

The Impact of Systematic Reviews on Health Care Policy in England

Frances Bunn

Submitted to the University of Hertfordshire in partial
fulfilment of the requirements of the degree of Doctor of
Philosophy

December 2010

Abstract

Background: The last few decades have seen a growing emphasis on evidence-based decision making in health care. Systematic reviews synthesising and evaluating research have been a key component of this movement. However, there is concern that such syntheses do not have the expected impact on policy and practice and more work is needed to enable us to maximise their potential. The aim of this study was to increase understanding of the likely impacts of systematic reviews on policy and identify factors that might facilitate their influence.

Methods: My own previously published work is integral to this study. I took ten systematic reviews on which I am an author and used established methods for the evaluation of research impact, including, bibliometrics and documentary review, to examine whether these reviews had influenced policy development. Data from these analyses were combined with an overview of the literature to identify factors that might increase impact.

Results: The reviews had influenced the development of national and international policy, although much of the impact was at a 'micro' level in the form of practice guidelines. There was considerable variation in the impact of the reviews. Reviews evaluating fluid resuscitation and road safety interventions showed the greatest evidence of impact and a review of qualitative studies on barriers to fall prevention the least. Differences might be explained by time since publication, type of question, importance to policy makers, the nature and strength of the evidence, the purpose of the review and the networks and strategies used for dissemination.

Conclusions: Systematic reviewers should consider the desired impacts of their work and include appropriate strategies for increasing impact, these should be detailed in the review protocol. This might include specifying methods to address applicability to particular contexts, and devising active strategies for dissemination.

Acknowledgements

I would like to thank my supervisors, Dr Geraldine Byrne and Professor Robyn Martin, for their valuable advice, support and encouragement; and Professor Sally Kendall, for her continuing support. This study is based on a number of publications and I would like to acknowledge the contribution of all my co-authors, in particular I would like to thank Professor Ian Roberts and my colleagues at the Cochrane Injuries Group. Finally I would like to thank my husband, Paul Kayley, for his tireless support.

Contents

Abstract.....	i
Acknowledgements.....	ii
List of tables and figures	vi
Abbreviations.....	vii
Chapter 1: Introduction	1
Introduction	1
The evolution of evidence based policy.....	3
Barriers to the development of evidence based policy	11
Explanatory frameworks and theories of the policy development process	13
Conclusion.....	20
Aims & objectives.....	21
Organisation of the submission	21
Chapter 2: Promoting the impact of research	25
Introduction	25
Defining research impact.....	25
Promoting the impact of research on health care policy	26
Promoting the impact of systematic reviews on health care policy.....	28
Facilitators for promoting the influence of research on health care policy	31
Collaborative approaches	31
Collaborative approaches and systematic reviews.....	34
Barriers to collaborative approaches.....	35
The role of dissemination in enhancing research use	36
Dissemination strategies and systematic reviews	38
Initiatives to increase the usefulness of systematic reviews.....	40
Conclusion.....	43
Chapter 3: Methods	51
Introduction	51
The role of my own research	51
Methods for determining impact	54
Evaluation frameworks	64
Selected framework.....	72

Other methodological considerations	73
Conclusion.....	74
Chapter 4: Strategies for fluid resuscitation: systematic reviews of RCTs	75
Introduction	75
Drivers behind review question/s.....	76
Systematic review methods.....	77
Results of impact analysis	85
Conclusion.....	96
Chapter 5: The prevention of traffic related injuries: systematic reviews of road safety interventions.....	104
Introduction	104
Drivers behind review question/s.....	104
Review methods	106
Results of impact analysis	110
Barriers to evidence-based road safety.....	119
Conclusion.....	124
Chapter 6: Identifying barriers and facilitators to fall prevention interventions: systematic review of qualitative research	126
Introduction	126
Drivers behind review question.....	126
The role of qualitative research in systematic reviews	127
Conceptual approaches/methodologies.....	129
Review methods	132
Results of impact analysis	140
Conclusion.....	145
Chapter 7: Systematic review to inform the development of NICE public health guidance	151
Introduction	151
Drivers behind review question.....	151
Rationale for guidelines	154
Potential problems with clinical guidelines	156
The development of NICE Guidelines	160
Methods for NICE reviews	161
The use of guidelines by practitioners.....	164
Results of impact analysis	167

Conclusion.....	172
Chapter 8: Evaluating the safety and efficacy of telephone consultation: a systematic review to inform service delivery & organization.....	173
Introduction	173
Drivers behind review question.....	173
Review methods	174
Results of impact analysis	178
Research impact and the social construction of policy	183
Conclusion.....	185
Chapter 9: Discussion & Conclusions	187
Introduction	187
Overview of the impact analysis.....	188
Factors that play a role in the extent to which systematic reviews can influence policy	197
Barriers & facilitators to research impact.....	201
Meeting the needs of policy makers whilst maintaining methodological rigour	204
Methodological approach and limitations of the study	205
Implications of the findings.....	210
Implications for future research	214
Conclusion.....	215
References	217
Appendix 1: Full list of authors publications.....	244
Appendix 2: Published papers included in submission	249

List of tables and figures

Tables

Table 2.1 Mechanisms to drive research impact	27
Table 2.2 Studies promoting the use of systematic reviews	46
Table 3.1 Published systematic reviews included in submission	52
Table 3.2 Search Strategies for identification of evidence of impact Chapters 4-8	62
Table 4.1 Downloads for Cochrane fluid resuscitation reviews in 2008 and 2009	89
Table 4.2 Evidence of fluid reviews' influence on guideline development	97
Table 5.1 Evidence of road safety reviews' influence on policy development	115
Table 6.1 Methods for systematically reviewing qualitative studies	147
Table 7.1 Results of internet searches to identify organisations citing the guidelines	170
Table 8.1 Secondary outputs from telephone triage review	181
Table 9.1 Summary of citations in policy documents: all reviews	190
Table 9.2 Number of downloads for Cochrane reviews	194

Figures

Figure 4.1 Results of citation analysis: fluid therapy reviews	87
Figure 5.1 Results of citation analysis: road safety reviews	112
Figure 8.1 Results of original and revised citation analyses: telephone consultation review	179
Figure 8.2 Policy analysis triangle	183

Abbreviations

BMA	British Medical Association
BMJ	British Medical Journal
CBA	Controlled before after study
CIG	Cochrane Injuries Group
CDSR	Cochrane Database of Systematic Reviews
DoH	Department of Health
EPOC	Effective Practice and Organisation of Care
HTA	Health technology assessment
ITS	Interrupted time series
KTE	Knowledge transfer and exchange
NICE	National Institute of Health and Clinical Excellence
NHS	National Health Service
NRS	Non randomised studies
RCT	Randomised controlled trial
SIGN	Scottish Intercollegiate Guidelines Network
STI	Sexually transmitted infection
WHO	World Health Organisation
WoS	Web of Science

Chapter 1: Introduction

Introduction

Over the last few decades there has been a growing emphasis on the use of evidence to inform decision making in health care (Dawes et al., 2005, Sackett et al., 1996). Whilst initially much of the focus was on the use of research evidence in medicine, and on which clinical interventions were most effective for patients, the movement has now expanded to include other aspects of health care practice including nursing and public health (Chalmers, 2003, Dawes et al., 2005, French, 1999, Kitson et al., 2008, Schuerman, 2002). There are a number of definitions of evidence-based practice but they commonly include an emphasis on using evidence that is the best available and that is up to date, valid and relevant (Dawes et al., 2005, Sackett et al., 1996).

One of the key aspects of evidence based practice has been the development of methods for the synthesis and integration of primary research. There are a number of terms for such syntheses but the most widely used and understood is 'systematic review'. A systematic review has been defined as 'a summary of the best available evidence that addresses a sharply defined question' (Sackett et al., 1996) which is 'prepared using a systematic approach to minimising biases and random errors and in which there is clear documentation of the process in a materials and methods section' (Chalmers, 1995). Systematic reviews have several advantages over other types of research that have led to them being regarded as particularly important tools for decision makers. For example, they are considered to be scientifically rigorous and can generally be conducted more rapidly than new primary research which may make them particularly useful for policy makers who may be called on to respond quickly to policy issues (Pawson, 2002). In addition, it inherently makes sense for decisions to be based on the totality of evidence rather than a single study (Sheldon, 2005).

However, despite this sound and scientific rationale for using systematic reviews for decision making, it has been pointed out that 'the policy and management

community in health care in general has been slow to use them for informing their decisions' (Sheldon, 2005). Indeed the usefulness of systematic reviews for aiding policy makers in the decision making process has come into question with authors suggesting a number of factors that might reduce their utility. These include a lack of good quality primary research for synthesis, a tendency for reviewers to focus on randomised controlled trials and controlled evaluations at the expense of other types of research, and inadequate evaluation of complex interventions with little recognition of the importance of contextual factors (Greenhalgh et al., 2003, Oliver, 2001, Pawson, 2002). Furthermore, it has been argued that researchers may view the relationship between policy and research in a simplistic linear fashion, whereby thinking proceeds sequentially through stages which are guided by rational analysis, without due consideration of the social, cultural, political, economic and ideological factors that contribute to the development of health policy. The factors which influence change are complex and further work is needed on the issues involved in ensuring systematic reviews make the appropriate contribution to the development of policy and practice.

In addition, recent years have seen a growing interest in the way in which research is used with researchers increasingly expected to consider the wider impacts of their work (HEFC, 2009). This may include the contributions research makes to society, culture, the economy, quality of life and public policy. The focus of this study is on systematic reviews and the impact they have on the development of health care policy in England. The study is based around ten systematic reviews (14 papers) (Alderson et al., 2000, Bunn et al., 2006a, Bunn et al., 2004a, Bunn et al., 2005, Bunn et al., 2003a, Bunn et al., 2003b, Bunn et al., 2008a, Bunn et al., 2004b, Bunn et al., 2008b, CIG Albumin Reviewers, 1998, Duperrex et al., 2002a, Duperrex et al., 2002b, Kwan et al., 2003, The Albumin Reviewers et al., 2004), on which I have made a contribution either as first author or co-author. The purpose of the study is to examine whether these reviews had influenced policy development and to identify mechanisms researchers can use to promote the use of systematic reviews in policy development.

Before I go further it is important to outline the definition of policy I have adopted for this study. Policy making can be viewed as involving the 'authoritative allocation of values' (Easton, 1953) and when interpreted broadly can include people making the policy as government ministers and officials, as local health service managers, or as representatives of a professional body. The results may take many forms ranging from national health policies made by the Government to clinical guidelines determined by professional bodies (Hanney et al., 2003).

In this chapter I set the context for the study by documenting the evolution of evidence based policy, exploring the development of systematic reviews, looking at barriers to the development of evidence based policy and examining some of the key frameworks and theories for explaining the relationship between research and policy. Finally I present the aims and objectives of the study and outline the organisation and contents of the submission in a chapter plan.

The evolution of evidence based policy

Evidence based practice

Although research has long been a factor in the development of health care the evidence based practice movement as we know it today is a relatively new phenomenon originating in the 1980s when the term evidence-based medicine emerged. A leading proponent of evidence based medicine described the process as "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients. It means integrating individual clinical expertise with the best available external clinical evidence from systematic research" (Sackett et al., 1996). The development of the evidence based medicine movement saw a shift from the paradigm of medicine as an 'art' to conceptualising it in more rational scientific terms. A lengthy medical education, personal experience and the guidance of colleagues was no longer seen as a sound basis for making decisions about health care. Instead proponents advocated that treatment decisions be based upon research evidence (Sackett et al., 1996). There were a number of drivers behind the

movement but key was the realisation that wide variation existed in the practice of individual clinicians (Wennberg, 1973), and that many were using treatments that colleagues considered inappropriate (Chassin et al., 1987). Indeed, there has been an accumulation of evidence that some ineffective, or possibly harmful, treatments have been overused whilst others shown to be effective were under-utilised (Chalmers, 2003, Hanney et al., 2005, Lau et al., 1992, Ottesen et al., 2001, Roberts and Bunn, 2002, Wennberg, 2002).

However, the ethical imperative to do no harm is only one of a number of factors involved in the development and growth of the evidence based practice movement. Other drivers which have been identified include: a better informed public, an increasing distrust of experts, the huge growth in information available, developments in information technology, a growth in the capacity and skills of the research community and increased emphasis on cost effectiveness, productivity and accountability (Davies, 2000b, Traynor, 2002). This agenda, and the calls for using the best available evidence have influenced all areas of health care and the term evidence based practice is now used to encompass the work of all health care professionals and not just doctors (Dawes et al., 2005). Whilst definitions and interpretations of the evidence-based movements in different professions reflect the differences in the contexts in which they operate they still appear to share key 'ingredients and tenets' (Traynor, 2002).

Key to any discussion of evidence based practice and policy is: what constitutes evidence? The Oxford English Dictionary describes evidence as 'the available body of facts or information indicating whether a belief or proposition is true or valid' (OED, 1998). In reality ideas of what constitutes evidence vary between research disciplines and between researchers and policy makers. Traditionally evidence based medicine has tended to take a narrow view of evidence with an elevation of quantitative research, in particular randomised controlled trials and meta-analyses, over other forms of evidence (Rycroft-Malone, 2006). Although ideas about what constitutes evidence have widened to encompass other types of research, such as uncontrolled studies and qualitative research, other forms of

non research based evidence such as experience, expert judgement, anecdote or theory are still largely discounted by researchers (Lomas, 2005).

Some critics have objected to the elevation of research evidence over other forms of knowledge and argued that advocates of evidence-based policy have failed to recognise the fallibility of scientific evidence, that reliable evidence can derive from other sources beside research (Hammersley, 2005, Rycroft-Malone, 2006) and that evidence is socially constructed and means different things to different people (Dopson et al., 2002, Dopson, 1999). In addition, the concept of a solid 'evidence-base' has been brought into question. In a recent paper on evidence-based policy Annette Boaz argued that the term evidence-base implies the existence of a solid foundation that policy or practice can build on whereas, in reality, the reliability and relevance of evidence changes over time and, therefore, a concept of an 'evidence pool' more accurately captures the dynamism and fluidity of evidence (Boaz, 2008b).

Systematic Reviews

One of the potential barriers to the widespread use of evidence in practice was the vast and unmanageable amount of information practitioners were faced with (Mulrow, 1994). As Iain Chalmers points out, clinicians had long recognised the challenge they faced in keeping up to date with relevant information (Chalmers and Trohler, 2000). He cites the Army surgeon John Rollo who in 1801 wrote 'life is too short for a conscientious physician to acquire –even with the most suitable education, unremitting observation, accurate investigation and unvaried reading–satisfactory confidence in the unreserved treatment of the sick committed to his charge' (Rollo, 1801). In addition, critics began to question the way researchers and practitioners failed to learn from, and build on, past research efforts. Archie Cochrane, a UK epidemiologist, encapsulated the idea that medical research should be catalogued and evaluated when he wrote 'it is surely a great criticism of our profession that we have not organised a critical summary, by specialty or subspecialty, adapted periodically, of all relevant randomised controlled trials' (Cochrane, 1979).

In response to this need for strategies for the management and evaluation of the evidence base researchers developed methodologies for synthesising primary studies. Central to this has been the development of statistical techniques for combining data from different studies. This work originated with the statistician Karl Pearson who in 1904 reported the use of formal statistical techniques to combine data from different medical studies. His rationale for combining data, that “many of the groups are far too small to allow of any definite opinion being formed at all, having regard to the size of the probable error involved” (Pearson, 1904), remains as one of the primary justifications for pooling studies today. Work on synthesis was carried on in the social sciences, rather than medicine, and it was the psychologist Glass who first used the term meta-analysis (Glass, 1976).

Systematic reviewing, however, developed as more than just a statistical method for pooling the results of different studies. It was a process which placed an emphasis on a scientific and systematic approach to reviewing research which was in stark contrast to the “subjective, scientifically unsound and inefficient reviews” often previously found in the medical literature (Light, 1984). As methods for synthesis developed commentators highlighted the difference between the poor scientific qualities of many non systematic reviews with the rigour of those produced using systematic review methodology. In a critique of review articles, published in four major medical journals between 1985 and 1986, Cynthia Mulrow concluded that ‘current medical reviews do not routinely use scientific methods to identify, assess, and synthesize information’. She goes on to say that the reviews evaluated tended to lack a clearly defined question, search methods were not defined, standardised criteria for assessing study quality were not used and synthesis tended to rely on informal qualitative descriptions rather than quantitative methods (Mulrow 1987). In contrast it was claimed that systematic reviews could integrate existing information; provide data for rational decision making; establish the consistency and generalisability of findings; and, if meta-analysis was used, increase the power and precision of treatment effects (Egger, 2001b). Crucially, the emphasis on a sound scientific

process reduces the risk of bias and, therefore, allows for more accurate and reliable conclusions to be drawn (Chalmers, 1995, Egger, 2001a, Mulrow, 1994). These perceived strengths mean that well conducted systematic reviews have come to be considered as one of the most important type of research in hierarchies of evidence for evaluating the effectiveness of interventions (Philips, 2001).

Evidence based policy

Despite some criticism of the evidence-based practice movement the concepts underlying evidence based practice, and the importance of systematic reviews, have gained widespread credence across all disciplines in health (Nutley, 2003b). Rycroft-Malone writes that the ideology of evidence based practice 'has penetrated the consciousness, discourse and working practices of professionals' (Rycroft-Malone, 2006). This widespread recognition of the need for practitioners to be aware of the importance of incorporating evidence into practice has, unsurprisingly, led to calls for managers and policy makers to consider the evidence when making policy decisions. One commentator declared 'if doctors are expected to base their decisions on the findings of research surely politicians should do the same ... the case for evidence-based policy making is difficult to refute' (Ham et al., 1995).

Many proponents of evidence-based policy have appeared to conceptualise it in terms that are very comparable to those used to describe evidence based practice. For example, Davies defines evidence based policy as "the integration of experience, judgement and expertise with the best available external evidence from systematic research" (Davies, 1999); a definition very like Sackett's description of evidence based practice (Sackett et al., 1996). However, it has been argued that policy making is fundamentally different from decisions in clinical practice and that evidence-based policy cannot be viewed as simply an extension of evidence-based medicine (Black, 2001). More recently terms such as 'evidence-influenced', 'evidence-informed' or 'evidence-aware' have begun to be used in place of the more traditional evidence-based to describe the

relationship between research and policy (Boaz, 2008b, Chalmers, 2005, Davies, 2000a). This change in language appears to represent a shift in thinking about, and expectations of, the contribution of evidence to policy.

At the same time as this growing interest among health care professionals in the use of evidence to make decisions, there were a number of political and organisational changes that created an environment which was particularly receptive to the burgeoning idea of incorporating evidence into the policy process. Throughout the 20th century there had been an increase in the number of organisations, such as pressure groups, researchers, think-tanks, professional bodies and statutory organisations, attempting to influence government actions often through the use of some form of evidence (Davies, 2000b). Although Governments in the UK had become increasingly responsive to the use of evidence, this was particularly marked when Labour were elected to power in 1997 with a philosophy of 'what matters is what works'. Their avowed rejection of political ideology in favour of a more pragmatic approach set a platform for the development of a more evidence-informed policy process (Davies, 2000b). This emphasis on the importance of knowing what works was further established in a number of government documents including the 1999 white paper on Modernising Government which called for policy makers to show a greater willingness to question inherited ways of doing things and to make 'better use of evidence and research in policy making' (Cabinet Office, 1999a) and the Cabinet Office report entitled 'Professional Policy Making for the 21st Century (Cabinet Office, 1999b) which stipulated using evidence as a core competency for good policy making. In addition, the 1991 national strategy 'Research for Health' (DOH 1991, 1993) led to a substantial increase in funding for health care research. This was accompanied by a shift in the way in which the research agenda was driven so that policy makers, and not just researchers or scientists, were involved in setting research priorities (Davies, 2000c).

Systematic reviews were central to this drive for creating an evidence base to inform policy. The Government initiatives described above involved the

development of an infrastructure to evaluate the effectiveness of health and public health interventions (Solesbury., 2001). This included the establishment of the National Institute for Health and Clinical Excellence and the NHS Centre for Reviews and Dissemination, and the provision of core funding support to the Cochrane Collaboration, an international organisation involved in preparing, maintaining and disseminating systematic reviews evaluating the effects of health care interventions. All of these bodies use systematic review methodology to evaluate evidence and to provide information on effectiveness.

So why are systematic reviews seen as so integral to decision making in health care and why have they received so much government support in England? This preference for systematic reviews is partly pragmatic in that systematic reviewers are thought to be able to respond more quickly, and more cheaply, to the demands of policy makers than researchers conducting primary studies (Pawson, 2002). Another key factor behind this emphasis on systematic reviews is an acceptance of one of the core tenets of evidence based practice, which is that it is vital to look at the totality of available evidence rather than at a single study in isolation. Indeed some commentators have claimed that 'generally the results of a single study are not worth disseminating' (Black, 2001) whereas, in contrast, systematic reviews can establish whether findings are consistent and generalisable across different populations and settings (Mulrow, 1994). Systematic reviews can also, where appropriate, incorporate statistical techniques for pooling data across studies which enhance the precision of treatment effects (Egger and Smith, 1997, Egger et al., 1997). A classic example of the advantage of pooling studies is incorporated into the Cochrane Collaboration logo which depicts the results of seven trials that evaluated the effects of a short course of corticosteroids given to women expected to give birth prematurely. Only two trials had clear cut, statistically significant effects, but when data from all of the studies were pooled, the power increased. This yielded a significant combined effect estimate which strongly indicated that the treatment reduced the risk of babies dying (Mulrow, 1995).

A more cynical interpretation of the prominence given to systematic reviews is that government and policy makers view systematic reviews as useful tools for rationing health care (Gingrich, 2009). Until recently decisions about the allocation of health care resources in England and Wales have been made by the National Institute of Health and Clinical Excellence (NICE). Their judgements, which often provoke a great deal of controversy, are based on evidence of effectiveness and cost-effectiveness which comes from systematic reviews. These issues are explored more fully in Chapter 7.

So where and how is it possible to envisage systematic reviews informing policy development? One well known public policy framework is the stages heuristic (Brewer and DeLeon, 1983, Lasswell, 1956) which divides the public policy process into four stages: agenda setting, policy formulation, the implementation of policy, and finally evaluation where the impact of policies is assessed. This framework has subsequently been criticised for being overly simplistic (Sabatier, 2007) but it does still provide a useful way for thinking about the policy process (Walt et al., 2008). Although it is possible to think that systematic reviews might impact on any stage of the process it is most likely that they could influence the first or last stages, agenda setting and evaluation. By bringing together all available research on a given topic systematic reviews can be used to develop an understanding of a problem, provide potential solutions, highlight gaps in the evidence and identify future research questions which may in turn be influential in determining the research agenda.

Indeed, systematic reviews can be used to help answer various types of questions. This is illustrated here using a variety of reviews which I have been involved in. Traditionally systematic reviews were used to answer clearly focused questions about the effectiveness of health care interventions, for example meta-analyses to evaluate the best type of fluid replacement for injured patients (Bunn et al., 2000b, Bunn et al., 2002, Bunn et al., 2008b), drug therapy for preventing infections in patients undergoing surgery for breast cancer (Bunn et al., 2006b), nutrition for head injured patients (Perel et al., 2006), pelvic floor

exercises for post natal continence problems (Wagg and Bunn, 2007) and care for patients with spinal injuries (Kwan and Bunn, 2005). Increasingly reviews are evaluating more complex community based interventions many of which focus on prevention. For example, the safety and efficacy of telephone triage (Bunn et al., 2005), strategies for preventing sexually transmitted infections and teenage pregnancies (Bunn et al., 2006a), the prevention of drugs and alcohol misuse in children (Petrie et al., 2007a) and the prevention of road related injuries (Bunn et al., 2003a, Duperrex et al., 2002a, Ker et al., 2005). In addition, systematic review methods are being used to synthesise qualitative research to explore questions about intervention development, delivery and acceptability rather than just effectiveness. For example, the potential barriers and facilitators to older people participating in falls prevention programmes (Bunn et al., 2008a), or the way parental perceptions of health behaviours may affect obesity prevention interventions for young children (Pocock et al., 2009). Reviews may also inform the development of guidelines at a national or local level (Bunn et al., 2006a).

Barriers to the development of evidence based policy

Despite the apparent rationale, and support, for using evidence, and in particular systematic reviews, to develop policy a number of commentators have argued that research has, in reality, had a minimal impact on government policy (Black, 2001, Feldman, 1999, Lomas, 2000b). This perceived inefficacy has been attributed to naivety on behalf of researchers about the policy process and, in the case of systematic reviews, to limitations in their methods that may restrict their usefulness to policy makers. Both of these ideas will be explored further; beginning with an examination of potential problems with systematic reviews and then moving on to issues and theories relating to policy development.

Limitations of systematic reviews as policy tools

As discussed earlier one of the distinguishing features of systematic reviews was the development of rigorous scientific methodologies to reduce the risk of bias. As randomised controlled trials (RCTs) are considered to be the 'gold standard' for evaluating the effectiveness of interventions much of the original

methodological endeavour focused on ways to synthesise them. For example, the Cochrane Collaboration, which has been highly influential in developing and promoting systematic review methods, has, to date, largely concentrated on the synthesis of RCTs. Many have argued, however, that this strict focus on RCTs has limited the utility and generalisability of systematic reviews and that other types of studies, such as less rigorous evaluations and qualitative studies, are needed to evaluate complex interventions and to give information about process, context, and the perspectives of users (Dixon-Woods et al., 2005, Harden et al., 2004, Oliver, 2005a, Speller et al., 1997). There is an increasing realisation that it is not enough to focus solely on 'what works' as policy makers have more complex questions and need to know 'what works for whom and in what circumstances' (Solesbury., 2001).

In the light of these criticisms researchers have developed new methods for synthesising non randomised (Reeves, 2008) and qualitative studies (Dixon-Woods et al., 2006b, Popay et al., 1998), for incorporating quantitative and qualitative studies together in the same review (Dixon-Woods et al., 2005, Harden et al., 2004, Harden and Thomas, 2005, Oliver et al., 2007) and for analysing the effects of complex social interventions in relation to context and settings (Pawson, 2002, Pawson et al., 2005). Most of these techniques are largely rooted in the same conceptual ideologies as conventional evidence based practice, incorporating or adapting many of the same systematic review methods, although some, such as Pawson's 'realist synthesis', use more radical methods and reject much of the orthodoxy of the systematic review movement (Pawson, 2002, Pawson et al., 2005). Although these methods may offer promising ways to make systematic reviews more relevant to policy, questions remain as to whether some approaches, such as realist synthesis, have the necessary methodological rigour and transparency which is central to systematic reviewing. The author's systematic reviews being considered in this study largely employ conventional systematic review methods for synthesising quantitative studies, in accordance with methods specified in the Cochrane handbook (Higgins, 2008). However, they also include examples of reviews incorporating

non randomized and qualitative studies. The methodological issues associated with different approaches are explored further in later chapters.

It has also been suggested that systematic reviews have not had the anticipated impact on policy because of a dearth of good quality primary research to review. Many interventions likely to be the subject of policy decisions have not been properly evaluated and this lack of good quality research affects the usefulness and applicability of systematic reviews (Juni et al., 2001). However, it can also be argued that the inherent emphasis on critical appraisal in evidence based practice and systematic reviews has highlighted many of the shortcomings of primary research and been instrumental in improving the conduct and reporting of studies.

Explanatory frameworks and theories of the policy development process

Although much credence has been given in recent years to the concept of evidence-based or informed policy it is generally accepted that the relationship between policy and research is complex, with government policies driven by many factors including 'ideology, value judgments, financial stringency, economic theory, political expediency, and intellectual fashion' (Davis and Howden-Chapman, 1996). Policy makers may be subject to many different influences including political imperatives and the media (Campbell, 2007) and non research evidence such as personal experience, local information and the opinions of eminent colleagues (Black, 2001, Campbell, 2007).

This idea that policy is influenced by many factors and is not a 'discrete' act or event but rather an "ethereal, diffuse, haphazard and somewhat volatile" process (Lomas, 2000a) is reflected in a number of policy theories. In the garbage can model of policy making the policy process is seen as dynamic, unpredictable, chaotic and often fundamentally irrational. Policy problems and solutions get 'dumped' together in a metaphorical garbage can and only become joined together when opportunities arise (Cohen et al., 1972). In the

enlightenment model (Weiss, 1976, Weiss, 1977) Weiss argues that policy makers dislike disruptions to the status quo and may ignore research that does not support their own values and beliefs. Although policy makers may sometimes be receptive to research findings, for example for large issues and at times of crisis, in general the effect of research is cumulative and indirect, rather than immediate, with concepts and ideas from research gradually filtering down into the policy process. Research is seen as only one of a number of knowledge sources and is 'considered less as problem solving than as a process of argument or debate to create concern and set the agenda' (Black, 2001). The nature of this process means that problems are not always recognized immediately and it may take considerable time for research to impact on policy (Buse, 2005). For example, as Buse points out, it took many years for the government to take the evidence on smoking and lung cancer seriously and even longer before any legislation was introduced (Buse, 2005). Weiss also argues that policy makers may have different values and goals from those of researchers. She sees researchers and policy makers as two distinct communities, with different conceptual frameworks, values and motivations, who often fail to understand each other's positions and language (Weiss, 1976).

One framework for understanding the way different factors influence policy is Walt and Gilson's health policy triangle (Walt and Gilson, 1994). This looks at the interaction between the policy context, the policy process, and the various actors, such as individuals, groups and organisations, involved in the formulation of policy. There are many actors and agencies beside the government and the research community who influence the policy agenda. This includes lobby groups, powerful 'producer groups' (such as the BMA), charities, the business community and of course the mass media. Buse argues that the media are able to 'raise, shape and determine issues and public opinion' which, in turn, influences the government to respond (Buse, 2005). This influence may be positive, for example the media raising awareness of important but neglected issues, or more negative such as giving topics levels of news coverage unrelated to their severity and the risks they pose to public health. This focus on particular

issues, because they are seen as dramatic or newsworthy, is a contributory factor in shaping individuals' perceptions of personal risk. For example, after a paper in the *Lancet* suggested a link between the combined measles, mumps and rubella (MMR) vaccine and autism (Wakefield et al., 1998) media coverage greatly increased levels of public anxiety and led to a decrease in vaccination. Although the *Lancet* paper has subsequently been heavily criticised, and further work has shown no evidence of a link between the MMR vaccine and autism, levels of vaccination remain lower than they were prior to the media furore (Health Protection Agency, 2009). The role of the media in the policy process is considered further in Chapter 4.

Researchers, and in particular proponents of evidence based policy, have been criticised for oversimplifying the relationship between research evidence and policy by assuming that there is a direct link between the two with policy makers identifying a problem which is then 'solved' by researchers (Harrison, 2001). Buse describes this as the Engineering or problem solving model which is based on the premise of a linear relationship between research and policy (Buse, 2005). Critics have argued that the underlying assumptions behind evidence based policy are naïve and flawed and that systematic reviews, and research in general, have failed to influence service policies because of an over-reliance on this simplistic model (Black, 2001, Boaz, 2008b, Buse, 2005, Marmot, 2004).

It has also been suggested that the opportunity for research to impact on policy is limited to small windows of time. In Kingdon's multiple streams model (Kingdon, 1984) the policy process is viewed as a mixture of three streams of actors and processes which generally operate independently of each other. These are: a problem stream consisting of various conditions that policy makers and citizens want addressed, a policy stream involving the ideas or proposed solutions to policy problems, and a politics stream consisting of elections and elected officials that is affected by the national mood, campaigning by pressure groups and administrative or legislative turnover. Kingdon argues that policy choices are made when the three streams are joined together at critical

moments of time. These opportunities for change are labelled as 'policy windows' and individuals or corporate actors who attempt to couple these three streams are 'policy entrepreneurs'. Policy entrepreneurs with greater access to policy makers and resources are more likely to be successful (Kingdon, 2003, Zahariadis, 2007).

In contrast to those who suggest that research has a limited influence on policy development others have argued that in recent years research has actually had a substantial impact on health policy. Indeed, it has been suggested that the recent political history in the UK, with the election of the previous Labour Government and its emphasis on 'what works', saw the re-emergence of a more evidence-informed policy process in the engineering mould (Nutley, 2000). Systematic reviews, with their focus on effectiveness have been seen as particularly key in this process. In the economising model Fox argues that a dominant free market ideology in many Western countries has led to researchers, funders and policy makers adopting an increasing emphasis on efficiency and value-for-money. Researchers have been able to influence policy because they are undertaking more policy relevant research and are producing research compatible with the dominant values and ideologies of those in positions of political power (Fox, 1990). Fox sees the development of methods to evaluate effectiveness (such as systematic reviews and meta-analysis) as key to this process (Fox, 2006). Whilst Fox's theory is based on conditions in the USA his arguments can also be generalised to England where priorities for research are increasingly being defined by funders and policy makers in contrast to the predominantly investigator led or science driven systems of the past (Davies, 2000c). The idea of policy driven research is explored further in Chapter 7 where I consider the impact of a review (Bunn et al., 2006a) commissioned by NICE to inform the development of public health guidance.

However, even if we accept that policy makers are willing to be influenced by evidence there are still a number of barriers to the use of evidence as a basis for policy making. One such challenge is the extent of scientific uncertainty that

often exists; due to either a lack of available relevant research (Macintyre et al., 2001) or what Weiss terms “the perennial problem of inconclusiveness” (Weiss, 1976). Harrison argues that effective implementation of research depends upon a number of factors including the existence of “comprehensive, authoritative statements based on systematic reviews of the research evidence” (Harrison, 2001). In reality it is uncommon for research to produce definitive conclusions with many studies showing little or no effect. This may be exacerbated by evidence based practice and its emphasis on well-designed evaluations and ‘gold standard’ study designs such as RCTs. Well designed controlled studies reduce the risk of bias and are, therefore, less likely to produce inflated treatment effects. As Buse points out, “it is generally accepted that the better designed the evaluation, the smaller the effect it is likely to demonstrate” (Buse, 2005). Even systematic reviews incorporating all the available evidence are often unable to deliver firm conclusions. In such instances policy makers have to decide if the research is merely inconclusive, if the intervention is ineffective in the situation or context where it was evaluated or if it is genuinely ineffective. This can be particularly difficult when complex health interventions are being evaluated.

Even where reliable research evidence is available its influence on policy development may still be limited. Research findings may be used selectively and are open to differing interpretations, misrepresentation, and controversy. In a paper looking at five different reviews on mentoring Boaz explored the question of how useful they were for guiding policy and practice (Boaz, 2005). She found that there were many ‘different viewpoints, incompatibilities and contradictory advice’ in the reviews and attributed this to a tendency by the reviewers to inflate their conclusions in an attempt to deliver clear policy recommendations. She argues that reviewers need to ‘rid themselves of the notion that there is a gold-standard method of research synthesis capable of providing unambiguous verdicts on programmes’ and that we need to be aware of, and accept, the limitations of systematic reviews.

Some theorists would argue that it is not just that research can be controversial, or open to differing interpretations, that is important but rather that policy makers see research purely in political terms and will interpret it to reinforce their own preferences. For example, in the strategic model research is seen as a tool used by governments to reinforce pre-determined positions or to hinder or postpone difficult political decisions (Weiss, 1979). Calls for further research before policy decisions can be made may be seen as a cynical attempt to remove an issue from the policy agenda (Buse, 2005). Another explanation which draws on some of the same ideas but employing a less cynical interpretation is the elective affinity model. This claims that research is more likely to have an effect on policy if members of the policy community have been involved in some way in the research process. Although this model suggests that research will only have an impact if the findings are congruent with the values and beliefs of the policy audience it does allow that research that is initially disregarded may play an 'enlightenment' role over a longer period of time (Short, 1997).

Another potential barrier to the use of research in the policy process is that researchers may lack the will, or the necessary communication and marketing skills, to ensure their work receives sufficient publicity to enter the policy arena. However, although they have been accused of naivety in their perceptions of the policy process, many researchers have become increasingly aware that policy is derived from a mixture of social, cultural, political, economic and ideological factors and that if they are to contribute to the policy agenda they need to consider the utility of their research, work to improve relationships with policy makers and accept that creating change may be a long and difficult process (Black, 2001, Maynard, 2006). Indeed the idea of collaborative research is becoming increasingly popular with both researchers (Lomas, 2000b, Patton, 1997) and government (National Audit Office, 2003, Stationary Office, 1999). Denis and Lomas define collaborative research as "a deliberate set of interactions and processes designed specifically to bring together those who study societal problems and issues (researchers) with those who act on or within those societal problems and issues (decision-makers, practitioners, citizens)"

(Denis and Lomas, 2003); and have coined the phrase 'linkage and exchange' to describe one model of collaborative working (Lomas, 2000b). Other processes to facilitate collaboration and describe collaborative research efforts include work on 'collaborative evaluation processes' (Patton, 1997), 'knowledge utilization' (Landry, 2001, Lavis et al., 2002) and 'knowledge transfer'. This trend towards collaboration is based on the belief that it is no longer appropriate to make a distinction between scientific experts and non-experts but that researchers, practitioners and lay people need to work together to produce scientific knowledge (Denis and Lomas, 2003). These ideas about knowledge transfer and exchange are explored further in Chapter 2.

The successful implementation of evidence-informed policy is of course a two way process requiring knowledge and understanding from both researchers and decision makers. However, it has been suggested that many policy makers lack the necessary 'research literacy' (Boaz, 2008c) to engage fully with research. Although the 'analysis and use of evidence' is now seen as a core skill for civil servants (Professional Skills for Government initiative <http://psg.civilservice.gov.uk>) an analysis by the Government Social Research Unit on evidence-based policy in practice (Campbell, 2007) found that 'the complexity and variety of different techniques were not well understood' and that 'specific techniques such as systematic reviews were even less well understood'. They identified a need for policy officials to have a clearer comprehension of the relative merits of different types of research. That policy makers were particularly unclear about what systematic reviews entailed or what they could offer, indeed many of the policy makers interviewed for the report had not even heard of systematic reviews, suggests that researchers still have a great deal more work to do to convince policy makers of the potential value of systematic reviews. However, it was not clear how many of the policy-makers interviewed were involved in health care where, I would suggest, it is likely that there would be greater awareness and use of systematic reviews than in other policy areas.

Many of the models described above, and the suggestions for how relationships between research and policy can be improved are dependent on the idea of researchers and policy makers as two distinct homogeneous groups engaged in different types of work and with different priorities, attitudes, career paths, rewards, training, knowledge and organisational constraints (Buse, 2005). However, it has been suggested that this notion of two distinct communities is overly simplistic with critics arguing that neither researchers nor policy makers are homogenous groups and that the translation, dissemination and communication of research findings are not the only barriers to translating research into policy. An alternative view is to see policy making and research utilisation in terms of networks and communities of actors who shape policy (Adam and Kriesi, 2007, Sabatier and Jenkins-Smith, 1993). These are formal and informal relationships that include all those who play a part in the generation, dissemination and evaluation of policy and can include researchers, politicians, journalists, civil servants, and members of the public. In the advocacy coalition framework Sabatier suggests research enters policy as much through political argument as through the transmission of knowledge (Sabatier and Jenkins-Smith, 1993).

Conclusion

There are then a variety of factors that need to be considered when assessing the influence of systematic reviews on policy. Certainly systematic reviews appear to be uniquely placed to contribute to the development of health care policy. Despite some criticism they hold an enviable position as a widely respected and used methodology, acknowledged by practitioners, funders and Government as central to the evaluation of evidence. However, there remains much debate about the extent of their influence and many commentators have been disappointed at what they see as a lack of impact on policy. This has been attributed to a number of causes including misconceptions about the nature of evidence, methodological shortcomings of systematic reviews, unrealistic expectations of what reviews can achieve, a fundamental misunderstanding of the way in which policy is developed and the role research plays in this, and a

lack of the necessary skills and competencies on the part of both researchers and policy makers. These issues are explored further throughout the study.

Aims & objectives

In this study I intend to examine the extent to which systematic reviews can inform or influence the development of health care policy. The study will be based on a significant amount of my own previously published research which I will use as illustrative examples to examine the following issues:

- Is there any evidence that any of my own systematic reviews have influenced policy at either a national or a local level?
- What factors play a role in the extent to which systematic reviews can influence policy e.g. review methods, funder, dissemination format, publicity, perceived importance of review question and 'fit' with values and ideologies of policy makers?
- What are the barriers and facilitators to systematic reviews informing policy?
- How can researchers produce systematic reviews that, whilst still methodologically rigorous, meet the needs of policy makers?

Organisation of the submission

In the following section I outline the organisation and content of the submission.

Chapter 1

In this first chapter I have set out the context and aims of the study by exploring the evolution of evidence based policy, the development of systematic reviews, barriers to the development of evidence based policy and key frameworks and theories for explaining the relationship between research and policy.

Chapter 2

Identifying ways of increasing the influence of systematic reviews on health care policy is a key aim of the study. In Chapter 1 I have touched on some of the

interventions designed to promote the impact of research on health care policy; such as knowledge transfer and exchange strategies. In Chapter 2 I examine in greater detail initiatives for promoting the use of research in the development of health care policy. In particular I look at literature on promoting the impact of systematic reviews.

Chapter 3

In Chapter 3 I describe the overall methodology for the submission. I begin by describing the integral part that my own published work plays in the submission and then go on to detail the methods employed to identify any evidence of impact that my reviews may have had on policy. I critically examine potential evaluation frameworks and outline the process by which I arrived at the framework I use in this study.

Chapters 4-8: Critique of methodological issues involved with systematic reviews and presentation of results of impact evaluation.

In Chapters 4-8 I present the published systematic review work that forms the basis of the submission. The rationale for the division of the reviews into these five chapters is based on differences in the review methods and in the subject matter of the reviews. For example, the reviews incorporate a variety of methodological approaches, including meta-analysis of randomised controlled trials (RCTs), meta-analysis of non-randomised studies, and synthesis of qualitative research, and cover a diverse range of health and public health related topics. The specific focus of Chapters 4-8 is presented below but each involves an evaluation of the impact of the review/s on health care policy. The papers on which the chapters are based can be found in Appendix 2.

Chapter 4

This chapter includes five systematic reviews concerned with the fluid resuscitation of critically ill patients (Bunn et al., 2000a, Bunn et al., 2000d, CIG Albumin Reviewers, 1998, Kwan et al., 2003, Roberts et al., 2004). All involved the statistical technique meta-analysis and they are used as illustrative examples to critically examine the methodological issues associated with this type of

review. In addition, the reviews are used as examples to critique some of the key issues surrounding the influence of reviews of clinical questions on healthcare policy and practice.

Chapter 5

This chapter includes two road safety reviews, one on traffic calming (Bunn et al., 2003a, Bunn et al., 2003b), and one on safety education of pedestrians (Duperrex et al., 2002a, Duperrex et al., 2002b). I discuss the methodological challenges associated with these reviews, such as the difficulties associated with identifying relevant studies in road safety research and the issues involved in the quality assessment and analysis of non-randomised studies. I also discuss some of the key barriers to the development of evidence-informed road safety policies.

Chapter 6

This chapter includes a review of qualitative studies evaluating barriers and facilitators to the successful delivery of falls prevention interventions (Bunn et al., 2008a). I use the review as a basis to explore the issues associated with the systematic review of qualitative research and critique the methods available. I also look more broadly at issues around the role of reviews of qualitative studies in informing policy decisions and examine methods that aim to increase the policy relevance of systematic reviews through the incorporation of qualitative research.

Chapter 7

This chapter is based on a systematic review evaluating the effectiveness of interventions for the prevention of teenage pregnancies and sexually transmitted infections (Bunn et al., 2006a). The review was commissioned by the National Institute of Health and Clinical Excellence (NICE) to inform the development of Public Health guidelines. I examine the extent to which the final guidelines were informed by the review and critically examine the role of guidelines in health care.

Chapter 8

This chapter includes a systematic review evaluating the effectiveness of telephone consultation (Bunn et al., 2004a, Bunn et al., 2005). This review differs from the others in the study because it is concerned with an issue relating to the way in which health services are organised and delivered. As well as addressing the methodological issues specific to this review I look at the complex political and socio-economic factors that are involved in policy changes involving service delivery and organisation.

Chapter 9

This chapter summarises the study findings, discusses the study's contribution to knowledge and looks at the implications of the findings. This includes a summary of the results of the impact evaluation, an analysis of how the results contribute to our knowledge of barriers and facilitators and a consideration of the place of systematic reviews in the policy process. It includes an overview of the methodological approach adopted and a discussion of the strengths and limitations of the study.

Chapter 2: Promoting the impact of research

Introduction

As discussed in the previous chapter there has been a growing interest in the role of evidence in the development of health care policy, a role advocated by Government, researchers and practitioners alike (Cabinet Office, 1999a, Cabinet Office, 1999b, Ham et al., 1995). It has, however, become increasingly apparent that there are many barriers to evidence-informed policy and that 'explicit and active strategies are required to ensure that research really does have an impact on policy and practice' (Walter, 2003b). In this chapter I begin by looking at what is meant by research impact and then move on to look at initiatives for promoting the use of research in the development of health care policy. In particular I focus specifically on literature that relates directly to initiatives or interventions for enhancing the influence of systematic reviews on the policy process.

Defining research impact

A variety of terms have been used to describe the impact of research on policy and practice. These include research impact, influence, outcomes, benefit, payback, translation, transfer, uptake and utilisation (Boaz, 2008a, Carden, 2004b). Research can be used either directly in decision-making related to policy or practice, or indirectly by contributing to the formulation of values, knowledge and debate. Commentators have pointed out that there is a key distinction to be made between 'conceptual' use, which brings about changes in levels of understanding, knowledge and attitude, 'symbolic use' which can lead to the mobilisation of support, and 'instrumental', or direct use, which results in changes in practice and policy making (Amara et al., 2004, Huberman, 1992, Nutley, 2003b, Weiss, 1976). Indeed, 'research impact forms a continuum, from raising awareness of findings, through knowledge and understanding of their implications, to changes in behaviour' (Nutley, 2003a). Strategies to enhance impact may be focused on any point along this continuum.

Of course it is difficult to consider the literature on promoting the influence of research without considering how we measure impact. In the field of health care there has been a fairly substantial amount of work exploring research impact but it has tended to focus on the impact on clinical practice rather than policy (Boaz, 2008a). The impact on practice may be more easily discerned than the impact on policy (Hanney, 2007); in contrast it may be more problematic to assess whether any changes in policy are attributable to research as it can be difficult to isolate the role research played in relation to the many other confounding factors that might be involved (Boaz, 2008a, Carden, 2004a, Hanney et al., 2000, Lavis et al., 2003a). Determining the impact of a specific piece of research is even more difficult as payback may come from an accumulation of research, that is the general 'stock or reservoir of knowledge', rather than from a single study (Hanney et al., 2000). Measuring impact is further complicated by the distinctions between direct or indirect influence. Conceptual and symbolic influence may be far harder to distinguish than instrumental or direct use of research. Methods for measuring research impact are considered further in Chapter 3.

Promoting the impact of research on health care policy

There is a distinction to be made between the natural uncontrolled spread of research and innovation (diffusion) and more active conscious efforts to spread research (dissemination) (Green et al., 2009). In this chapter I am concerned with active strategies to promote research influence. Isabel Walter and colleagues at the Research Unit for Research Utilisation (RURU) in Scotland developed a taxonomy of interventions which have been used to increase the impact of research (Walter, 2003b). After reviewing over 100 papers, that evaluated or described specific interventions to enhance the impact of research, they categorized them by the type of intervention and type of mechanism used to increase impact. They came up with 32 different categories of interventions which included: written materials (such as publication in journals), oral presentations, use of the mass media, research-based guidance, education, lobbying, research incentives, networks and collaboration. These were grouped

into eight categories to reflect the mechanisms considered to drive research impact; more detail of these can be seen in Table 2.1. These categories were designed to exist in parallel, rather than as a hierarchy, and are not mutually exclusive.

Table 2.1 Mechanisms to drive research impact (adapted from Walter 2003b)

Category	Underlying mechanism	Strategies	Underpinning theory
Dissemination	Research-based messages The extent to which dissemination interventions are tailored to potential users can vary considerably	Includes written materials, oral presentations	Adult learning theories which argue personal motivation important for behaviour change
Education	Learning: increasing knowledge and understanding of research findings	Includes lectures, interactive sessions, daily practice	Variety of educational theories
Social Influence	Social influence: changing norms and values as a route to changing behaviour	May include lobbying and patient-mediated interventions	Social influence and social learning theories
Collaboration	Communication: improving the flow of information and ideas between researchers and potential users	Aim to break down barriers between groups e.g. from differences in culture, timescales and values)	Social learning theory
Incentives	Motivation through reward: ways of acting will be more likely to recur when they are followed by positive consequences	Financial rewards, increase in professional status	Learning theories and economic models of rational behaviour
Reinforcement	Reinforcement through information	Audit, feedback and reminders	Diverse learning theories
Facilitation	Facilitation: providing the means to take action and removing barriers to that action	Financial, technical, organisational, or emotional assistance	Change management theories
Multifaceted initiatives	Generally target multiple mechanisms	Two or more interventions	Variety of theories but supported by transtheoretical model & social learning theories

Walter and colleagues also reviewed the literature on approaches to enhancing research use from the education, healthcare, social care and criminal justice sectors (Walter, 2003a). Of the literature they reviewed 60% were concerned with healthcare but, of those, many were focused on the impact of research upon practice rather than on policy. For example, studies were evaluating interventions such as training, audit and feedback, education, guidelines, incentives and computer support systems, all of which were aimed at health care professionals rather than policy-makers. However, they also consider some mechanisms relevant to increasing the influence of research on policy such as collaborative initiatives and active dissemination strategies. These will be considered in more depth later in the chapter.

Promoting the impact of systematic reviews on health care policy

The primary aims of this chapter were to review literature that focused specifically on ways of promoting the impact of systematic reviews on health care policy and to distinguish potential facilitators to the use of systematic reviews to inform policy. Although this was not a systematic review the aim was to identify as many relevant papers as possible. I searched for papers that either described or evaluated interventions designed to promote the influence of systematic reviews or that evaluated the impact of systematic reviews. Papers that appeared to be describing a process more akin to guideline development were excluded. I included all study designs including qualitative and descriptive studies.

To identify relevant studies I searched PubMed (2nd February 2009) with no date restrictions using the following search terms: policy OR policies AND knowledge transfer OR knowledge translation OR knowledge utilisation OR knowledge utilization OR knowledge broker* OR payback OR research impact OR research influence OR research uptake. These were combined with the MeSH term systematic review. In addition, I contacted experts in the area (Dobbins and Lavis at McMaster University in Canada) and used lateral search techniques such as checking reference lists, and using the 'related articles' function on Google

and 'cited by' function on Google Scholar for key papers. Such lateral search techniques are recommended when searching for studies in a complex area (Greenhalgh and Peacock, 2005). The searches were updated in February 2010.

Only nine papers were found which specifically addressed promoting the influence of systematic reviews on health care policy: (Atkins et al., 2005, Brussoni et al., 2006, Dobbins et al., 2001a, Dobbins et al., 2001b, Dobbins et al., 2009, Dobbins et al., 2004b, Keown et al., 2008, Lavis et al., 2006a, Lavis et al., 2005). Two of these (Dobbins et al., 2001a, Dobbins et al., 2001b) reported the same study. Three of the papers involved telephone surveys (Ciliska et al., 1999, Dobbins et al., 2001a, Dobbins et al., 2004a), two interviews (Lavis et al., 2005) and three described strategies used to increase the use and relevance of systematic reviews but included no formal evaluation of these methods (Atkins et al., 2005, Brussoni et al., 2006, Keown et al., 2008). I found only one formal evaluation of a strategy to increase systematic review impact. This was a randomised controlled trial to test the effectiveness of KTE strategies in Canadian public health decision making. The intervention compared three different strategies: access to an online registry of systematic reviews evaluating public health interventions, targeted evidence messages and knowledge brokering (Dobbins et al., 2009). More details of the included studies can be seen at the end of the chapter in Table 2.2.

Policy makers views on systematic reviews

There was some discrepancy in the literature as to the extent to which policy-makers used systematic reviews as a source of research evidence. In a study which used interviews with Canadian policy makers to explore how systematic reviews could better meet their needs (Lavis et al., 2005) it was found that none of the policy makers cited systematic reviews as a source of evidence but rather they were influenced by factors other than research including legal issues, pressure from stakeholders and public opinion. Policy makers tended not explicitly to place a high value on research evidence and certainly not on systematic reviews (Lavis et al., 2005). This was echoed in a recent UK report

that found few policy makers had heard of systematic reviews or meta-analyses nor were they aware of what these entailed or what they could offer (Campbell, 2007). However, more positive findings came from a number of other studies. In a telephone survey of public health policy makers in Ontario 57% of those surveyed had heard of systematic reviews and when prompted with a description of a review 86% said the description sounded familiar and 62% were able to give examples of reviews they knew about. It is not possible, however, to discern from their study whether policy makers actively used the reviews (Ciliska et al., 1999). In two later surveys, conducted to ascertain the role of recently completed systematic reviews in the development of public health policy, systematic reviews were seen as useful by policy makers and were valued above other types of research evidence (Dobbins et al., 2001a, Dobbins et al., 2001b, Dobbins et al., 2004a). These studies by Dobbins and colleagues were, however, conducted in an area of Canada where there have been considerable efforts to develop a policy-making culture that values research evidence. In addition, the nature of these studies, with their focus on systematic reviews, may have prompted policy makers to give more credence to systematic reviews than they otherwise might have done.

Although the evidence base in this area is weak a number of potential facilitators were identified that might increase the influence of systematic reviews on health care policy. These include collaboration between researchers and policy makers, incorporating examinations of context and applicability in reviews, appropriate dissemination strategies, responding to policy-makers' need for timely information and the development of methods to increase the usefulness of systematic reviews for policy makers. I will explore these in more detail in the sections below. In each section I begin with a general overview of literature and then focus on the literature relating to systematic reviews.

Facilitators for promoting the influence of research on health care policy

Collaborative approaches

A key mechanism to drive research impact, and one that has received much attention in the literature, is initiatives to increase collaboration between researchers and decision makers. Indeed, systematic reviews of barriers and facilitators to the use of research evidence by health policy makers have found that a key facilitator, and one supported by the most consistent and rigorous evidence, was interaction and personal contact between researchers and policy makers (Innvaer et al., 2002, Lavis et al., 2005). Dopson argues that interpersonal networks are crucial for the circulation of knowledge (Dopson et al., 2002). An example of this was the use of policy networks that brought policy-makers together with researchers through formally structured mechanisms. Initiatives such as this also helped to foster trust between policy makers and researchers which in turn promoted the use of research (Lavis et al., 2005).

In England a report by the National Audit Office states that ‘encouraging partnerships between researchers and users is a precondition of delivering evidence-based policy-making’ (National Audit Office, 2003). One of the roots of current collaborative approaches is based in academic work around knowledge translation (Denis and Lomas, 2003). Knowledge translation, which is also known as knowledge utilization, knowledge exchange, research transfer and research utilization, has been defined as ‘a process by which relevant research information is made available and accessible for practice, planning, and policy-making through interactive engagement with audiences and supported by user-friendly materials, and a communications strategy that enhances the credibility of the organization and, where relevant, reinforces key messages from the research’ (Program in Policy Decision-Making, 2003). Much of the recent work on research transfer has taken place in Canada where the promotion of research transfer and uptake and evidence-based decision making have been identified as

important priorities for the Canadian health-care system (National Forum on Health, 1997).

The transfer of information from researchers to decision makers has been a subject of interest since the 1950s (Huberman, 1990). The process was initially conceptualized as a logical flow of information from researchers to policy makers (Wingens, 1990) in line with the demand pull or problem solving model (Landry, 2001, Weiss, 1979). In the 1970s frustration over the ineffectiveness of knowledge transfer led to the dominance of the two communities model which attributed the lack of research impact to the cultural differences between the research and policy-making communities (Caplan, 1979). Although critics later claimed that this model was inadequate and overly pessimistic (Dunn, 1983, Wingens, 1990) subsequent approaches to improve collaboration and impact have once again embraced the two communities model emphasising the differences in culture, goals, information needs, timescales, power, reward systems and language between the research community and policy makers (Huberman, 1994, Lomas, 2000a).

These more recent approaches have, however, conceptualized knowledge translation not as a linear one-way transfer of information from researchers to policy makers but as a more collaborative process that involves interaction and exchange among researchers producing information and potential users such as policy makers and service providers (Huberman, 1994, Jacobson et al., 2003). Whilst proponents still see a distinction between the scientist and the non-scientist there is a 'mutual respect for the distinctive expertise that each brings to the research process' (Denis and Lomas, 2003). The producer-push and user-pull models of knowledge transfer and uptake have been supplanted by an interaction model to enhance knowledge transfer and uptake (Landry et al., 2001). Key to this process is that knowledge producers and users are personally known to each other and, therefore, familiar with each others' needs, priorities, aims and circumstances (Bogenschneider, 2000, Jacobson et al., 2003, Lomas, 2000a). For example, a UK study, describing a crime reduction initiative involving

researchers, policy makers and practitioners working closely together to develop and evaluate interventions, suggested that the co-location of research and policy teams was key to its success (Laycock, 2001). Denis and Lomas suggest that for some the purpose of collaboration is to ensure the local application of specific research results whereas for others the aim is for the researcher and practitioner to better understand each other's questions and approaches (Denis and Lomas, 2003). These aims echo the distinction between direct or instrumental use and conceptual or symbolic use.

Such strategies to develop and strengthen contact between the two groups have been labelled by some as 'linkage and exchange' (Lavis et al., 2002, Lomas, 2000a, Lomas, 2000b). Lomas describes how the Canadian Health Services Research Foundation has built principles of linkage and exchange into its activities. Health-sector researchers, managers and policy makers come together to set research priorities and review funding applications. To secure funding applicants must demonstrate both scientific merit and potential impact and investigative teams are required to include at least one decision maker actively engaged in management or policy in the area under study. He claims that there is much evidence that involving decision makers in research and bringing researchers into the policy making process is the best way to ensure the transfer of research findings (Lomas, 2000b). In a study of the role of health services research in public policy making Lavis and colleagues (Lavis et al., 2002) found that citable research was more likely to be a major influence in the policymaking process if policymakers had direct contact with researchers, for example through the establishment and maintenance of linkages with researchers. However, they also found that many efforts at interaction by research units yielded no tangible impact. They recommend that researchers create more opportunities for interaction with the potential users of their research, and that they should 'consider such activities as part of the 'real' work of research, 'not a superfluous add on' (Lavis et al., 2002).

One suggested mechanism for bridging the gap between researchers and policy makers is the use of knowledge brokers with an understanding of the culture of both decision maker and researcher environments and with the necessary communication, networking, problem solving and negotiation skills (Lomas, 2007). In a review of the evidence on promoting innovations in service organizations Greenhalgh and colleagues use the term 'boundary spanners' to describe individuals with significant social ties both inside and outside the organization (Greenhalgh et al., 2004). They suggest that such people can play a pivotal role in promoting the uptake of innovations.

Collaborative approaches and systematic reviews

Two studies looked specifically at collaborative approaches to promote the use of systematic reviews. One paper described strategies to create opportunities for stakeholder engagement throughout the stages of a systematic review (Keown et al., 2008). The authors describe how they have identified five potential opportunities for stakeholder engagement. These are:

- In an initial consultation between researchers and policy makers to identify suitable review topics
- At the beginning of the review process when questions are being defined and search strategies developed
- Throughout the process with a stakeholder involved as a team member
- At reaction meetings once preliminary findings are available
- In the dissemination of review results

From their observations of the process the authors have noted a number of benefits of stakeholder engagement which include an added depth to the review, better defined research questions, improved clarity of the final report and helpful input into recommendations. They claim that policy makers feel the process makes the reviews more useful. In the other study, the only RCT (Dobbins et al., 2009), it was found, somewhat counter-intuitively, that

knowledge-brokering was more effective in organisations that placed less value on research than those that already recognised the importance of evidence-based decision making. It could, however, be that there was less scope for improvement in organisations that already had a positive culture towards research use.

Barriers to collaborative approaches

Despite the increasing popularity of knowledge transfer strategies some commentators have been critical of what they see as an overly simplistic approach. Davies and colleagues suggest that the terms knowledge transfer and knowledge translation are misconceived terms for applied social research and that they do not adequately reflect the challenges, subtlety and complexity of research use (Davies et al., 2008). They argue that ideas of knowledge transfer are based in the traditional rational-linear models of research use and that 'the baseline assumption of two communities too easily leads to unsophisticated notions of knowledge and knowledge transfer' (Davies et al., 2008). They suggest the term 'knowledge interaction' may be more appropriate to 'describe the messy engagement of multiple players with diverse sources of knowledge'. In a systematic review of knowledge transfer and exchange (KTE) strategies (Mitton et al., 2007) the authors are also critical of the conceptualisation of KTE and argue that current ideas do not fit with the underlying complexities and politics of health policymaking.

A number of barriers to collaborative approaches have been identified (Lomas, 2000b, Walter, 2003a), both at the individual and the organizational level, and are linked to relationships between researchers and decision makers, modes of communication, time and timing, and context (Mitton et al., 2007). Developing effective partnerships requires time and energy and researchers may not have sufficient time or resources to develop linkage with policy makers (Lomas, 2000b, Walter, 2003a). Researchers have also cited problems identifying the point of entry into decision making organisations, difficulties balancing the competing agendas of partnerships and organisations and frustration with the

frequent restructuring and changes of personnel in decision making organisations. In addition, collaborative approaches may require some changes in the way researchers think, for example by broadening the range of methodologies used and rethinking the way problems are defined, (Denis and Lomas, 2003). They may also necessitate the researcher developing new skills in communication and negotiation (Lomas, 2000b), skills they may currently lack or do not wish to acquire (Carden, 2004a).

Barriers faced by policy makers included a lack of access to research findings, little time to read research, a poor understanding of research and inadequate skills to interpret it, and being overwhelmed by the sheer volume of research literature available (Nutley, 2003a, Walter, 2003a). In addition decision makers often need results faster than researchers can produce them (Campbell, 2007, Lomas, 2000b). Combined with these problems is the potential conflict between policy makers' desire to control findings for political reasons with researchers' need to publish (Lomas, 2000b). In the light of the additional resources required for incorporating KTE strategies into projects, and the fact that such funding is very often lacking, some researchers have raised the question of who should be responsible for developing and initiating KTE strategies. Should it be the role of researchers or policy makers (Mitton et al., 2007)?

Although collaborative approaches have been found to be key in promoting the use of research evidence, it has been questioned whether researchers should be seeking to promote these relationships or whether they may be responsible for compromising quality. It has been argued that if researchers do what is required for research to be used, then it may 'fail to fulfil one of its most important functions which is to be objective, reliable and unbiased' (Innvaer et al., 2002). Indeed, in the paper describing initiatives to involve stakeholders in the systematic review process the authors acknowledge that maintaining flexibility without compromising scientific rigour is a significant challenge (Keown et al., 2008).

The role of dissemination in enhancing research use

Another of the categories in Walter's taxonomy is that of dissemination (Walter, 2003b). For the majority of researchers dissemination is most commonly constituted by traditional approaches such as publication in peer-review journals and conference presentations. However, there is a lack of evidence as to whether such passive modes of dissemination are effective in driving research impact (Walter, 2003a). Indeed, a number of studies have found that the format of communication and presentation by researchers is often not considered 'user-friendly' by policy makers (Lomas, 2000b, Nutley, 2003a, Walter, 2003a). Research-use studies suggest that it is important to present research findings in formats that are tailored to potential customers (Lomas, 1991, Willison and MacLeod, 1999); systematic reviews of interventions to increase impact upon health care practice have found that active implementation strategies such as reminders, incentives, peer review, marketing and educational interventions are more effective in changing behaviour than passive distribution of recommendations, educational materials or guidelines (Walter, 2003a).

There is also evidence that seminars and workshops can encourage more direct use of research. Bogenschneider and colleagues used seminars to disseminate research findings to State-level policy makers in Wisconsin in the USA (Bogenschneider, 2000). They assessed effectiveness using questionnaires and follow-up telephone surveys and found the seminars were rated highly by participants and that policy makers reported using the information obtained in diverse ways. More recently alternative mechanisms for dissemination and knowledge transfer have been developed including the use of knowledge brokers trained specifically in information exchange (Law et al., 2004, Mitton et al., 2007). However, research in this area is limited and the effectiveness and cost-effectiveness of this strategy have been questioned (Pyra, 2003).

Dissemination strategies and systematic reviews

Five studies looked specifically at dissemination strategies in relation to systematic reviews (Atkins et al., 2005, Ciliska et al., 1999, Dobbins et al., 2004a, Dobbins et al., 2009, Lavis et al., 2005). In semi-structured interviews with managers and policy makers Lavis and colleagues found that barriers to research influence were the use of jargon and the fact that researchers often only publish for a scholarly audience in academic journals (Lavis et al., 2005). They explored the optimal way to present systematic review evidence and found that most policy makers supported a 1:3:25 format (i.e. one page of take home messages, a three-page executive summary that summarizes the full report, and a 25 page report, as well as a longer technical report if necessary). They argue that this format has the advantages of delivering research reports in a way that is more likely to be read, being tailored to meet the needs of different audiences, and helping researchers learn clarity and brevity. The up-front placement of take-home messages also reflects how many policy-makers actually read research reports (Lavis et al., 2005). However, there is some research to suggest that such a format may, in some instances, alienate those who are in disagreement or less receptive to the conclusions presented (Lavis, 2004, Lavis et al., 2006b).

In a telephone survey to determine whether recent systematic reviews of public health interventions were used in the development of new provincial public health policies in Canada it was found that policy makers valued systematic review evidence over other types of evidence. The executive summary was considered to be the most influential part of the review. The authors suggest, therefore, that researchers should put a great deal of effort into writing an executive summary (Dobbins et al., 2004b). In a recent RCT, of KTE strategies in public health decision making, the use of targeted messages was more effective in promoting evidence-informed decision making compared with alternatives such as a website and knowledge-brokering groups (Dobbins et al., 2009). In their guidance for undertaking reviews in health care the NHS centre for Reviews and dissemination suggest that dissemination is an integral part of the review process and should be considered from an early stage to allow time for planning,

development and implementation (CRD, 2009). They have developed a dissemination framework which incorporates a number of key attributes including the characteristics of the research message, the setting in which the message is received, the characteristics of the target audience(s), the source of the research message and the communication channel(s) used.

Another important point relating to dissemination is whether reviewers should make recommendations for policy and practice or if this is beyond their remit. One of the original tenets of systematic review methods, designed to maintain scientific rigour and reduce the risk of bias, was that reviewers should provide an accurate and unbiased assessment of what the data show rather than interpreting them in light of their own personal opinion. This has, however, led to some readers being frustrated by a lack of firm recommendations from systematic reviews, and disappointment when reports conclude that there is insufficient information to answer critical questions (Atkins et al., 2005). This tension is also reflected in the literature with managers and policy makers disagreeing about whether researchers should make recommendations (Lavis et al., 2005).

A number of studies have supported the idea of reviewers developing recommendations. For example, in a study of Canadian policy makers it was found that when reading systematic reviews they focused most on the results, conclusions and discussion sections (Ciliska et al., 1999) and, in a systematic review of evidence from interview studies of barriers and facilitators to the use of research evidence by policy makers, the inclusion of summaries with policy recommendations was a commonly reported facilitator (Innvaer et al., 2002). However, in his review of ways to improve the usefulness of systematic reviews for policy makers, Lavis suggests that researchers should avoid providing specific recommendations as research evidence alone is insufficient for making recommendations and researchers may not have a good sense of the values of those who will be affected by their decisions (Lavis et al., 2005). His reservations are supported by a UK researcher, Annette Boaz, who suggests that pressure to

deliver clear policy recommendations may lead systematic reviewers to inflate their conclusions and 'go beyond the evidence'. She argues we need to 'jettison the notion that a single review can deliver all-purpose policy advice' (Boaz, 2005).

Initiatives to increase the usefulness of systematic reviews

The influence of systematic reviews may also be increased by exploring ways of making them more relevant and useful to policy makers. Some suggestions for how this might be achieved are discussed in the following sections.

Context and applicability

A number of commentators have argued that key to making systematic reviews more useful to policy makers, and therefore more influential, is for the researcher to include an assessment of both the context in which the research was conducted and the applicability to a local setting (Gruen et al., 2005, Lavis et al., 2005). This echoes suggestions that it is important to go beyond simple questions of 'what works' to look at 'what works for whom and in what setting' (Solesbury, 2001). Gruen suggests that reviewers should consider the relative importance of the health problem under study, the relevance of outcome measures used, the practicality of the intervention, the appropriateness of the intervention and the cost-effectiveness of the intervention (Gruen et al., 2005). In reality, assessing context and applicability is not always straightforward. Although tools exist to assess applicability (Lavis et al., 2004, NICE, 2006) as a reviewer I have found these subjective and unsatisfactory. However, the biggest challenge that a reviewer faces when trying to assess applicability is that, in many instances, the content of interventions is poorly described and there is little detail of the providers, duration, intensity and setting, all of which is vital information if we are to make judgments about generalizability. It has been suggested that context may be more comprehensively examined by including published and unpublished (grey) research literature in a 'state-of-the-evidence' review that is broader than traditional systematic reviews (Benzies et al., 2006). However, such approaches may severely compromise review quality and validity

of the findings and there is, as yet, little evidence that incorporating 'evidence' of this nature improves assessments of local applicability. In addition, because this involves intensive searches for grey literature, reviews of this nature have significant resource implications.

Assessments of context are also dependent on the medium of dissemination and the audience. For example, Cochrane reviews are designed to be suitable for international audiences and so are deliberately not specific to one context. Lavis suggests that one way forward may be for researchers to produce reviews that can add to the international literature and that also include a locally adapted version (Lavis et al., 2005). Cochrane reviews are also, for the most part, focused on questions of effectiveness and, at present, there is no coordinated effort akin to the Cochrane Collaboration that has been undertaken to address questions other than 'what works?' (Lavis et al., 2006a). However, although the focus of the Cochrane Collaboration remains on randomised controlled evaluations of effectiveness, the latest version of the Cochrane handbook includes guidance on incorporating other study designs, such as qualitative research (Higgins, 2008). This might enable a greater breadth of questions to be considered in Cochrane reviews.

It has also been suggested that a key role of the systematic reviewer is to look at applicability of interventions for certain groups and to describe any differential effects by subgroups such as ethnicity and culture (Lavis et al., 2005). Although this is a laudable aim such analyses are often difficult to perform because of a lack of data from the primary studies. For example, in a review of interventions to prevent STIs and teenage pregnancies (Bunn et al., 2006a) the funders, NICE, asked for subgroup analyses by ethnicity and socio-economic status. This was understandable as the funders were keen to target interventions at those most at risk. However, we found little reliable evidence on the differential effects in these groups and a concern is that reviewers may feel under pressure to present information about applicability that is not supported by the data. Moreover,

there are a number of methodological concerns about the appropriateness of subgroup analyses (Higgins, 2008, Oxman and Guyatt, 1992).

Methodological initiatives

One of the criticisms of systematic reviews has been that a reliance on highly controlled research, such as RCTs, may limit their utility for policy makers and those in the 'real world'. Green argues that systematic reviews may 'deepen the chasm between research and practice' because 'most of the research qualifying as worthy of systematic reviews ... is highly controlled research under unrepresentative circumstances' and that this in turn means research is not adopted or implemented (Green et al., 2009). However, there is a growing literature on improving the usefulness of systematic reviews by incorporating a broader range of literature, including qualitative studies.

In addition, researchers are developing methods that allow them to go beyond answering questions merely about 'what works' and have developed techniques that allow them to address questions about cost-effectiveness, process, mechanisms and meanings (Dixon-Woods et al., 2005, Harden et al., 2004, Mays et al., 2005, Oliver, 2005b, Pawson et al., 2005, Thomas et al., 2004). The argument for such approaches is that they are more likely to yield reviews that are relevant to the decisions that policy-makers may face. The arguments against these methods are that they may introduce bias into what would otherwise be an approach that strives to minimise bias (Lavis et al., 2006a). To date, however, there is little evidence of whether such methodological initiatives improve research influence. Some have taken this further to produce syntheses that incorporate published literature, grey literature, decision makers' experience and researchers' knowledge and experience together into a 'policy synthesis' (Canadian Health Services Research Foundation, 2004). Such a technique is, however, relatively new and untested, presents formidable logistical challenges and should be used with caution (Lavis et al., 2005).

Timeliness

Timeliness has also been identified as a facilitator of research use (Ciliska et al., 1999, Lavis et al., 2005, Mitton et al., 2007). Systematic reviews, even though quicker to produce than primary research, may still take too long to be useful to policy makers (Campbell, 2007). In a review of evidence-based policy for practice for the Government Social Research Unit in the UK (Campbell, 2007) the authors suggest that to address the issue of timeliness there needed to be a modification of systematic reviews and the development of 'rapid evidence assessments' using similar rigorous techniques but in a much shorter time frame. How the authors envisage a process that maintains the same standards but takes less time is not specified. My own experience of producing a so called 'rapid' review for NICE (Bunn et al., 2006a) is that the funders wanted rapid delivery but did not want to compromise review quality, a situation that puts the researcher under a great deal of pressure. It may be that less rigorous rapid reviews and HTA assessments could be a bridge between science and policy making (Campbell, 2007, Rotstein and Laupacis, 2004, Watt et al., 2008). However, the scope of such syntheses may be more limited and they may not adhere to any single validated methodology (Watt et al., 2008). In addition, there is an obvious tension between the need to make reviews more useful, by exploring context, applicability, views and experiences and process, and the demands for timeliness. I have found that incorporating qualitative and quantitative research into the same review or searching for grey literature, whilst potentially valuable, are resource intensive and may lengthen the review process. This has implications for timely delivery to policy-makers concerned with a rapid response.

Conclusion

Although there is an increasing focus on research impact and the role of evidence in health care policy there is a lack of good quality evidence of interventions to promote the use of systematic reviews in the development of health care policy. This lack of established effective interventions is not limited

to systematic reviews. In a review of knowledge transfer and exchange strategies for research in general (Mitton et al., 2007) the authors noted that they found only one randomised controlled study. This lack of rigorous evaluation makes the transfer of findings to other or even similar, contexts difficult. The ability to generalise results to England is further restricted as many of the studies were conducted in Canada in areas with a policy culture that values research evidence and where initiatives at KTE are more common.

There is clearly some discrepancy in the literature over the extent to which policy makers actually make use of systematic review evidence in the policy development process. Although some of the Canadian studies showed a positive response to the use of systematic reviews, evidence from the UK suggests that systematic reviews are not well used or understood by many policy makers (Campbell, 2007). There are a number of well documented barriers that researchers and policy makers need to overcome before research can have an impact on policy. Some of the barriers that policy makers face, such as a lack of access to journals, lack of time to read and appraise articles, and a lack of critical appraisal skills, could feasibly be overcome by systematic reviews where reviewers have already done the work of finding relevant research, appraising their quality, and synthesizing the results (Ciliska et al., 1999). However, despite the potential advantages of reviews, barriers may remain as researchers may not be carrying out policy-relevant research, may be using inappropriate methods or may be using ineffective dissemination strategies. In addition, attempts to increase the usefulness of reviews, such as incorporating a variety of types of evidence and widening searches to include grey literature, may directly mitigate against timely delivery.

The literature does, however, point to a number of factors that may be key in promoting the use of systematic reviews. Lavis and colleagues make a number of recommendations for making systematic reviews more user friendly and increasing their impact. These include, involving policy-makers in setting questions and approach and interpreting the results, identifying benefits and

harms and describing any differential effects by subgroup (e.g. ethnicity and culture), looking at the context in which the research was conducted, developing a more user-friendly front end and adding additional local value to systematic reviews. They also suggest that reviews are included in The Cochrane Library or another source that provides 'one-stop shopping' for high quality reviews. It is clear that these strategies for increasing research impact require considerable investment on the part of the researcher in terms of time, resources and the development of new skills. However, despite a growing body of work on increasing research impact, the benefit of such developments is not established. In particular there is little empirical research on ways of increasing the influence of systematic reviews on health care policy and further work in this area is needed. In the next Chapter I describe the methods for the submission.

Table 2.2: Studies promoting the use of systematic reviews

Study ID	Study Type	Research aims & objectives	Main results	Comments
Atkins 2005 USA	Discussion paper	Describe lessons learned about how to increase the efficiency and impact of systematic reviews from network of evidence-based practice centres across North America	<p>Say that reviews must produce knowledge that is relevant to specific clinical and policy decisions and present information in concise and easily understood format.</p> <p>Planning for translation and implementation should be part of initial planning of the review.</p> <p>Lessons learned:</p> <ul style="list-style-type: none"> • Identify the right targets for evidence • Define appropriate questions and scope of a review • Working with partners important but reports need to address needs of partners without excessive tailoring to narrow interests of a single organisation • Balance consistency and flexibility in methods • Expand inclusion criteria to ‘best available evidence’ if higher quality evidence is lacking • Involving experts important but they need to be open minded enough to critically re-examine some of accepted conclusions in their field. 	No formal evaluation
Brussoni 2006 UK	Description of process to translate evidence into practice	<p>To bring together scientific evidence of what works in injury prevention with the knowledge and experience of practitioners.</p> <p>Uses a case study of reviews of evidence of effectiveness of smoke</p>	<p>Topics discussed:</p> <ul style="list-style-type: none"> • National policies and drivers seen as important influence on resources and staffing • Multi-agency partnerships seen as crucial • Meetings acted as valuable training tools and provided 	

		alarm installation Involved meetings with practitioners to discuss issues around implementation of interventions identified as effective.	mechanism for strengthening local partnerships Describes a process that they say considers local context and results in practical recommendations that reflect real-world practice	
Ciliska 1999 Canada	Telephone survey (before and after receiving systematic reviews) with public-health policy makers in Ontario Focus groups used to develop survey tool. Questionnaire pre-tested for content validity, revised and retested.	To gain an understanding of the research needs, perceptions of barriers to research utilization, and attitudes towards systematic reviews of decision-makers in public health at the levels of systems planning.	277 people eligible. 87% participated in first survey, 93% at follow up. 57% heard of systematic reviews, when prompted with a description 86% said the description sounded familiar, 62% able to give examples of reviews they knew about. When asked about priority reviews should be given in research agenda 62% said high and 9% top. For those who read the reviews most focused on the conclusions, discussion and results. Very few looked at tables. Barriers to research use: <ul style="list-style-type: none"> • Time • Availability of research results • Resources to implement research • Relevance • Policy climate • Timeliness 	Cannot discern from this survey whether policy makers actively use research
Dobbins 2001 (reported in two papers Dobbins 2001a & Dobbins)	Telephone survey of decision makers from all public health units in Ontario	To determine the extent to which 5 systematic reviews influenced public health decisions and policy development and to determine which characteristics of the innovation, organisation, environment, and individual	96% of decision makers participated in the survey, 63% reported using at least one SR in the past 2 yrs to make a decision, 50% perceived SR as having great deal of influence on program justification and 41% on planning decisions, 44% indicated SR has not influenced policy development at all Perception that one's organisation valued the use of research	Culture in Ontario that values research evidence. Results only generalisable to

2001b) Canada		predicted the influence the reviews had on those decisions	evidence for decision making and that ongoing training in critical appraisal was provided increased perception of the influence the systematic reviews had on public health decisions Most important predictors of use were position, expecting to use a review in the future, perceptions that reviews were easy to use and that they overcame barrier of limited critical appraisal skills.	organisations included in this study.
Dobbins 2004 Canada	Telephone survey	To determine whether the results of recently completed systematic reviews evaluating the effectiveness of public health interventions were used in the development of new provincial policies for public health practice	85% of decision makers participated in the survey, 96% of respondents reported that the systematic reviews played a part in developing new guidelines, 47% reported they contributed a great deal to the development of new recommendations for practice Decision makers valued the use of the systematic reviews to a greater extent than they did other types of information Significant predictor variables included the importance of the reviews in comparison to other sources of information and relevance of the reviews to policy decisions Majority of decision-makers rated executive summary as being the most important component of the systematic reviews	Review topics chosen in collaboration with provincial advisory group so relevant to public health policies under consideration. Culture in Ontario that values research evidence. Small sample size.
Dobbins 2009	RCT and qualitative interviews. 108 of 141 (77%) of public health organisations in Canada	To test the effectiveness of KTE strategies in Canadian public health decision making on programs related to the promotion of physical activity and healthy body weight in children. Involved three groups: 1) Control (who had access to an	Emerging findings from RCT include: <ul style="list-style-type: none"> • Having access to a registry of synthesised and translated research evidence (control grp) has no impact on evidence informed decision making (EIDM) • Targeted messaging significantly more effective in promoting EIDM than other strategies (p<.009) • A number of organisational factors modified the treatment effect • Simple KTE strategies may be as effective as complex ones (but 	

	participated in the study.	<p>online registry of systematic reviews evaluating public health interventions)</p> <p>2) Targeted message group (same as control plus direct mailing)</p> <p>3) Knowledge brokering (KB) group who received one-to-one input to build capacity of EIDM and assist in translating research evidence.</p> <p>Data collected baseline, post intervention, and 12 months.</p>	<p>need to be active rather than passive)</p> <p>Knowledge-brokering was more effective in organisations that placed less value on research evidence and less effective in those organisations that already recognised the importance of evidence-based decision making.</p> <p>Qualitative findings contradicted quantitative results. Participants in KB group perceived the KB to have significantly impacted EIDM capacity for them personally as well as for their organisation.</p> <p>Conclude that KB intervention may not have contained all the necessary components to produce a positive effect.</p>	
Keown 2008 Ontario, Canada	Discussion paper	<p>To describe various stakeholder engagement opportunities they employ throughout the stages of conducting a systematic review to increase knowledge utilization.</p> <p>Based on experience of 22 systematic reviews over 4 yr period.</p>	<p>They have identified 5 potential opportunities for stakeholder engagement:</p> <ul style="list-style-type: none"> • Stakeholder topic consultation • Stakeholder input meeting (e.g. input into setting question, literature searches) • Stakeholder as review team member • Stakeholder reaction meeting (discuss draft findings) • Stakeholder involvement in dissemination <p>Observed benefits</p> <ul style="list-style-type: none"> • Added depth to review • Help to define research question and add search terms • Improve clarity of final report and input into recommendations • Policy-makers feel it helps to make reviews more useful • Building capacity <p>Challenges</p> <ul style="list-style-type: none"> • Maintaining flexibility without compromising scientific rigour 	<p>Not a formal evaluation.</p> <p>Hard to generalise results to other settings</p>

			<ul style="list-style-type: none"> • Time and resource intensive • Difficult to find appropriate stakeholder to be review team member • Review may reveal controversial findings contrary to stakeholders expectations 	
Lavis 2005	<p>Systematic Review and interviews.</p> <p>Review included 17 studies 10 of which focused on health care policy-makers. In addition, they carried out interviews with policy makers in Canada and the UK to elicit views and experiences of using or commissioning systematic reviews.</p>	<p>To identify ways to improve the usefulness of systematic reviews for health care managers and policy makers.</p> <p>One of their main aims was to identify ways in which researchers could improve the usefulness of systematic reviews for health care managers and public policy-makers.</p>	<p>Ranked factors that influenced research-use by managers & policy makers from the most to the least rigorously demonstrated and consistent.</p> <p>Most important facilitators appeared to be:</p> <ul style="list-style-type: none"> • Interactions between researchers and health care policy makers • Timing and timeliness. <p>Other factors included:</p> <ul style="list-style-type: none"> • Policy networks that brought policy-makers & researchers together through formally structured mechanisms • Trust in the researcher <p>Barriers on the part of policy-makers included:</p> <ul style="list-style-type: none"> • Negative attitudes towards research evidence • Lack of necessary skills and expertise • Lack of perceived relevance • Use of jargon • Only publishing for a scholarly audience in academic journals <p>In those instances where policy makers reported using research evidence none cited systematic reviews. Many factors other than research influenced decision including legal issues, pressure from stakeholders and public opinion.</p>	<p>Overall there was a lack of research evidence in this area and it lacked rigour and consistency</p>

Chapter 3: Methods

Introduction

In this chapter I describe the overall methodology for the submission, and detail the strategies I employed to identify any evidence of impact or influence that my reviews may have had on policy. The core component of this submission is a body of my own published systematic review work and I begin by describing the role of that research and the way it is used and presented in the following chapters. This is followed by a discussion of the key issues involved in determining impact, the different techniques I used for evaluating the impact of my own reviews, and the framework I chose for the analysis.

The role of my own research

One of the aims of this submission was to explore the influence of systematic reviews on health care policy and to identify factors that may facilitate or restrict their impact. Central to this endeavour is a representation of my own previously published systematic reviews. These reviews form an integral part of the impact evaluation and are used as illustrative examples to enable an investigation of the influence of reviews on policy. The reviews are also used to critically explore the methodological issues associated with the systematic review process and to critique the different approaches adopted.

Details of these papers including the specifics of my contribution to each review can be seen in Table 3.1. The reviews form the basis for Chapters 4-8 with each chapter including a discussion of specific methodological issues and a presentation of any evidence of impact. The rationale for the division of the reviews into a number of chapters is based on differences in the systematic review methods employed and the review topics. For example, one chapter focuses on systematic reviews with meta-analysis of randomised controlled trials (RCTs) and another looks at meta-analysis of non-randomised studies. There is, of course, much overlap in the methods germane for different types of reviews. Even between reviews of RCTs and reviews of qualitative studies there are a

number of similarities, such as the importance of setting well defined questions and critically appraising studies. However, despite some overall methodological congruence there are significant differences between the various types of reviews and it is, therefore, appropriate to look at them separately. A list of all my published work can be found in Appendix 1 and copies of the papers included in this study can be found in Appendix 2. Five reviews were published in both the Cochrane library and a print journal and in these cases only one version of the review is included in the appendices.

Table 3.1 Published systematic reviews included in submission

Review	Published in:	First author Yes/No	Contribution to review process	Years published	Relevant chapter
Colloid Solutions for fluid resuscitation	Cochrane Database of Systematic Reviews (CDSR) (1)*	Yes	<ul style="list-style-type: none"> • Developing review protocol • Contributed to search strategy • Screened search records • Data extraction • Assessment of study quality • Data analysis • Writing up • Review updates 	<i>First published</i> 2000 <i>Updates:</i> 2001 2003 2008	4
Hypertonic versus isotonic crystalloid for fluid resuscitation in critically ill patients	CDSR (2)	Yes	<ul style="list-style-type: none"> • Developing review protocol • Contributed to search strategy • Screened search records • Data extraction • Assessment of study quality • Data analysis • Writing up • Review updates 	<i>First published</i> 2000 <i>Updates</i> 2002 2004	4
Human albumin solution for resuscitation and volume expansion in critically ill	CDSR (3a) BMJ (3b)	No	<ul style="list-style-type: none"> • Contributed to search strategy • Screened search records • Data extraction • Assessment of study 	<i>First published</i> 1998	4

patients			<ul style="list-style-type: none"> quality Commenting on the review 	<i>Updates</i> 2002 2004 2009	
Timing and volume of fluid administration for patients with bleeding following trauma	CDSR (4)	No	<ul style="list-style-type: none"> Protocol development Screened search records Data extraction Assessment of study quality Commenting on the review Review updates 	<i>First published</i> 2001 <i>Updates</i> 2003 2009	4
Colloids versus crystalloids for fluid resuscitation in critically ill patients	CDSR (5a) BMJ (5b)	No (involved in update not original review)	<ul style="list-style-type: none"> Data extraction Assessment of study quality 	<i>First published</i> 1998 <i>Updates</i> 2000	4
Area-wide traffic calming for preventing traffic related injuries	CDSR (6a) Injury Prevention (6b)	Yes	<ul style="list-style-type: none"> Developing review protocol Contributed to search strategy Screened search records Data extraction Assessment of study quality Data analysis Writing up Review updates 	<i>First published</i> 2003 <i>Updates</i> 2009	5
Safety education of pedestrians for injury prevention	CDSR (7a) BMJ (7b)	No	<ul style="list-style-type: none"> Contributed to search strategy Screened search records Data extraction Assessment of study quality Data analysis Writing up 	<i>First published</i> 2002	5
A systematic review of older people's perceptions of	Age and Ageing (8)	Yes	<ul style="list-style-type: none"> Developing review protocol Contributed to search strategy 	<i>First published</i> 2008	6

facilitators and barriers to falls prevention interventions			<ul style="list-style-type: none"> • Screened search records • Data extraction • Assessment of study quality • Data analysis • Writing up 		
Contraceptive advice and provision for the prevention of under 18 conceptions and STIs: a rapid review	NICE website (9)	Yes	<ul style="list-style-type: none"> • Contributed to search strategy • Screened search records • Data extraction • Assessment of study quality • Data analysis • Writing up • Input into development of guidance 	<i>First published</i> 2006	7
The effects of telephone consultation and triage on healthcare use and patient satisfaction: a systematic review	CDSR (10a) Br J Gen Pract (10b)	Yes	<ul style="list-style-type: none"> • Developed search strategy • Screened search records • Data extraction • Assessment of study quality • Data analysis • Writing up 	<i>First published</i> 2005	8

* each paper is given a number for ease of identification in tables and figures throughout the submission.

Methods for determining impact

As reported previously (Chapter 1) the definition of policy I have adopted involves not only national policies of the government but also policies agreed at national or local level by groups of health-care practitioners in the form of clinical or local guidelines, as well as policies developed by those responsible for training and education in various forms (Hanney et al., 2003). In addition, health policy covers both policy made in the public sector (by government) as well as policies in the private sector. Health is influenced by many determinants outside the health system (Buse et al., 2005) and I am, therefore, also interested in the actions of organizations external to the health system which have an impact on health (for example, the pharmaceutical or motoring industries).

As I noted in Chapter 2 many different terms have been used to define research impact. However, there is a general consensus of opinion that several types of research impact exist (Weiss, 1976, Huberman, 1992, Nutley, 2003b, (Estabrooks, 1999), including instrumental or direct impact, conceptual impact and symbolic impact (these are defined in Chapter 2 p 25, and again later in this Chapter). I was interested in exploring all these types of impact and for the purpose of my evaluation I took impact to mean any change that lay beyond the research process and its primary outputs (i.e. any change beyond the publication of the systematic review). However, as I highlighted in Chapter 1, the real world of policy making is complex and messy and because of this discerning the role of research in relation to all the other factors that might influence the policy making process is particularly challenging.

Although there is an increasing emphasis on research impact there are no agreed methods or instruments for measuring the impact of research (Boaz, 2008a). A variety of methods to assess research impact have been described and recommended in the literature. These include bibliometrics, documentary analysis, semi-structured interviews, case studies, panel review, surveys and network analysis (Boaz, 2008a, Hanney, 2007). The methods most frequently suggested for analysing the impact of research are bibliometrics, documentary review and interviews (Boaz, 2008a, Hanney et al., 2004b). Despite the potential for complex analysis using a range of techniques, the choice of methods for an evaluation of research impact should be realistic and fit with the scope of the task and the resources available (Boaz, 2008a).

However, choosing the appropriate methodology for this evaluation was not straight-forward as although there are a number of examples of research impact evaluations in the literature these tend to focus on the impact of research on practice rather than policy. In a recent review investigating the best way to evaluate the impact of research programmes on policy the authors found 35 papers that addressed the impact of research on health policy (Boaz, 2008a). However, many of those were focused on the relationship between research and

policy, rather than the impact of research on policy, and much of the work was reflective rather than documenting empirical examples (Boaz, 2008a).

Each of the methods has its own advantages and disadvantages and these needed to be considered when choosing methods fit for the purpose of my evaluation. Quantitative methods such as bibliometrics are widely used to quantify the impact of research. They can be used to manage large amounts of data, are transparent and reproducible and are suitable for comparisons across different research. In addition, they are cost-effective and can be applied to retrospective evaluations. However, bibliometric data usually focus on research outputs rather than outcomes and can only be considered a proxy for policy impact.

Qualitative methods such as semi-structured interviews, field visits and observations may generate rich descriptive data. However, it might be difficult to generalise their findings and they run the risk of being anecdotal and subjective (Boaz, 2008a). Such data may also be subject to a number of biases that lead policy makers to either overestimate or underestimate the influence of research. Asking a policy maker to focus on a specific systematic review might lead them to overestimate its significance, particularly if they are operating in a climate that purports to consider evidence when making policy decisions. On the other hand recall bias may mean that policy makers underestimate the influence of the research, particularly if there is a long gap between publication of the research and the interview. In addition, high turn-over of policy makers makes identifying suitable interviewees problematic, particularly for retrospective evaluations.

Although to date it has not commonly been used network analysis is a potential tool for impact evaluation. It involves analysis of relationships between different actors and can be used to map both formal and informal networks. The rationale for such an approach is that networks are important in disseminating research and shaping policy (Gray, 1973, Rogers, 1983, Sabatier and Jenkins-Smith, 1993). An advantage of such an approach is that it reflects the complex realities of

policy making. However, the disadvantage is that it focuses more on tracing research pathways or looking at the relationship between research information and policy rather than measuring impact. In addition, the identification of networks is likely to be more difficult in retrospective evaluations.

Owing to the limitations associated with most methods the use of multiple sources of evidence to identify research impact has generally been recommended (Croxson et al., 2001, Hanney et al., 2004a, Hanney et al., 2003, Lavis et al., 2003a). For example, in a study describing the methods used to determine the impact of research funded by the UK Arthritis Research Campaign, Hanney and colleagues used a mixture of documentary and literature review, semi-structured interviews with researchers and bibliometric analysis (Hanney et al., 2004b). However, overly complex approaches may be too time consuming and resource-intensive, and may not be justifiable relative to the potential benefits (Anderson, 2006, Beacham, 2005, Boaz, 2008a).

One of the challenges of this study was to identify appropriate methods for determining and comparing the impact of a group of reviews concerned with a range of topics and published over a ten year period. In addition, methods needed to be suitable for a retrospective analysis that involved tracking forwards from specific pieces of research rather than tracing backwards from policy. In light of these considerations I chose to use bibliometric analysis, documentary analysis and literature review, and some informal semi-structured email or telephone interviews with co-authors. Whilst use of these approaches may not facilitate in-depth exploration of the contextual factors -interpersonal, organisational and political- which influence the impact of systematic reviews on policy, they do provide a systematic and verifiable method of evaluating their contribution to a range of documents which inform policy. The use of more than one method enables richer data to be gathered and, also, allows for some triangulation. These methods are discussed further below.

Bibliometric analysis

One way of assessing research impact is to employ bibliometric methods. Bibliometrics employs quantitative analyses to measure patterns of scientific publication and citation, typically on journal papers (Ismail et al., 2009). One of the most important of these is citation analysis. This technique, which essentially involves counting the number of times a research paper or researcher is cited, works on the assumption that influential researchers and important works will be cited more frequently than others (Meho, 2007). Advantages of using a technique such as this is that citation rates are seen as an objective quantitative indicator for scientific success (Bornmann et al., 2008), they are robust and transparent and it is a relatively simple and cost-effective method. However, there are disadvantages which any researcher using this technique should be aware of. Whilst citation analysis may tell us the degree to which a piece of research has been useful to other researchers it does not give us an indication of the influence it may have had on decision makers; nor does it tell us how often reviews have been downloaded or read. Other criticisms of bibliometrics are that they focus on quantity rather than quality and measure the number of research outputs rather than research outcomes or impact (Boaz, 2008a).

In addition it has been suggested that citation counts are biased towards papers published in open access journals (Murali et al., 2004). The latter has been dubbed FUTON (full text on the internet) bias (Wentz, 2002). However, the extent to which open access may result in greater citation counts is not clear. In a longitudinal bibliometric analysis comparing cohorts of open access and non open access papers in one journal (Proceedings of the National Academy of Sciences) open access journals were twice as likely to be cited in the first 4-10 months after publication as non open access articles (Eysenbach, 2006). In contrast a recent RCT comparing open (free) full text access with subscription access found that although the open access articles were downloaded more often they were not cited more frequently than subscription access papers in the first year after publication (Davis et al., 2008). The later study may be more

reliable because it assigned papers to open or subscription access randomly, included more papers and covered papers in a greater range of journals.

As a result of these limitations citation analysis can only be considered as a surrogate or proxy measure for influence on policy and the results should be interpreted with some caution. Furthermore it is a proxy measure not yet satisfactorily proven to be linked to impact (Hanney, 2007). However, despite the potential drawbacks, bibliometric data were worth including in this study as they are a useful indication of the general influence of research.

Methods for citation analysis

The Thomson Scientific ISI citation databases which include the Science Citation Index (SCI) and the Social Sciences citation Index (SSCI) contain over 40 million records from more than 8700 journals and have traditionally been the main tool for citation analyses. However, the advent of the web has had a huge impact on citation analysis and in 2004 Scopus from Elsevier and Google Scholar from Google emerged to challenge the monopoly of the ISI citation index (Bakkalbasi et al., 2006, Meho, 2007). These bibliographic databases include additional document types such as books, chapters in books and conference proceedings that are not indexed in the ISI citation databases. Research has shown that Scopus offers about 20% more coverage than the Web of Science (Falagas et al., 2008) although it is limited to articles published after 1995.

Google scholar is a research orientated search engine that accesses conventional print material and web based material. It also extracts citation information and can be used as a citation index as well as a search engine. However, Google Scholar needs to be used with some caution as there is a lack of transparency about the sources and selection criteria (Jacsó, 2005, Smith, 2008) and the citation information can be flawed or inadequate (Falagas et al., 2008). Despite this, Google Scholar is a valuable tool as it contains significant resources not covered by other databases (Kousha and Thelwall, 2008), it can help in the retrieval of more obscure information, and is particularly useful for identifying

unpublished literature. Owing to the strengths and weaknesses of the different databases the use of multiple sources is generally recommended (Bakkalbasi et al., 2006, Meho, 2007). For this reason I used Scopus, Web of Science (WoS) (ISI citation indexes) and Google Scholar for my citation analysis.

Citation counts for Cochrane reviews may be artificially low because citing authors have incorrectly referenced Cochrane reviews (The Cochrane Library, 2008). Therefore, in addition to the citation searches I contacted Wiley Interscience the publishers of the Cochrane Library to ask if data on the number of times Cochrane reviews were accessed was available. They were able to provide me with data for 2008 and 2009. This data included the number of full text views and how this placed the review in the world ranking. They calculated world ranking by taking all the reviews on the Cochrane Library and ranking them based on the number of full text downloads. The total number of reviews on the Cochrane Library was 6,232 in 2008 and 6,840 in 2009.

Data from the citation analysis is considered in the context of the impact factor of the journal in which the research was published. The impact factor is a measure reflecting the average number of citations and is often used as a proxy for the relative importance of a journal within its field. Journals with higher impact factors are considered to be of greater importance than those with lower ones. There have been criticisms of the use of journal impact factors (Brown, 2007, Williams, 2007), including the claim that the use of impact factors conceals the large variation in the citation impact of individual papers within a journal (Seglen, 1997). However, it is still useful to consider the journal impact factor of as it has been shown to be strongly associated with the likelihood of subsequent citations (Callaham et al., 2002).

Documentary and literature review

Documentary analysis allows for the 'exploration and interpretation of existing documents and can elicit quantitative or qualitative findings' (Boaz, 2008a). Although the reliance on identifying and accessing existing records can be a

disadvantage, benefits of this technique are that it can be applied to a range of sources, provides contextual understanding and is cost-effective (Boaz, 2008a). Documentary and literature review can include identifying key citing papers and relevant clinical guidelines (Hanney et al., 2004b). As clinical guidelines fall within the definition of policy that I have adopted (Hanney et al., 2003), it was important to identify any guidelines that had incorporated evidence from any of the included systematic reviews. In addition, I attempted to identify literature that gave an indication that practice or policy may have changed as a result of a particular review.

To identify guidelines and relevant literature I undertook the following strategies:

- Handsearched the Cochrane document on the dissemination of Cochrane evidence (Canadian Cochrane Centre, 2004) to identify potential sources for further investigation. This document lists some of the bodies that use Cochrane reviews as a source of evidence for their guideline development.
- Screened the titles and abstracts of those papers identified by the citation analysis to see if any of them were guidelines or policy documents
- Searched The National Library of Guidelines and screened the list of guideline publishers on The National Health Service Library for Health (NHS NLH) (NHS NLH, 2009) and TRIP database using the guidelines filter
- Searched the following electronic databases: NHS NLH, PubMed and google.
- Searched the websites of relevant organisations or societies that might be involved in developing and issuing guidance
- Contacted Wiley Interscience, the publishers of the Cochrane Library, for information on the number of times reviews were downloaded

The details of specific search terms used and databases and websites searched for each review can be seen in Table 3.2. The searches on Google and Google Scholar were broad and generally resulted in many thousands of hits. As results

in these databases are presented by relevance (with those most relevant coming first) I screened at least the first 10 pages of hits returned. After that I stopped screening hits when it became clear they were no longer of any relevance.

Table 3.2 Search Strategies for identification of evidence of impact

Details of databases, search dates & search terms	Websites Searched
<p>Chapter 4</p> <p>PubMed (searched April 22nd 2009)</p> <p>Search for practice guidelines</p> <p>(Albumin* OR colloid* OR fluid therapy OR fluid resuscitation OR hypertonic OR volume supplementation) AND trauma OR injury OR intensive care</p> <p>Limits: Humans, Last 10 years, Practice Guidelines</p> <p>Searches for other papers related to impact on the use of fluids</p> <p>Google Scholar and NHS NLH</p> <p>(albumins/therapeutic use OR colloids/therapeutic use) AND (demand* OR prescribe* OR prescription OR trend* OR change* OR decrease* OR usage OR fall OR decline)</p> <p>Limits: Humans, date Jan 1998 to the present</p> <p>Bunn AND Fluid resuscitation or Fluid therapy OR hypertonic OR isotonic OR albumin OR colloid* OR crystalloid OR burn OR injury</p>	<ul style="list-style-type: none"> • American Association of Neurological Surgeons • Association of Surgeons of Great Britain and Ireland • American Association for the Surgery of Trauma • Brain Trauma Foundation (USA) • Canadian Agency for Drugs and Technologies in Health Care (CADTH) • Royal college of Surgeons of England • Royal College of Surgeons of Edinburgh • Royal Australian College of Surgeons • Scottish Intercollegiate Guidelines Network (SIGN) Society of Critical Care Medicine
<p>Chapter 5</p> <p>Pubmed (searched 12th June 2009)</p> <p>Traffic calming OR Road Safety OR RTC OR RTA OR pedestrian OR road OR traffic</p> <p>Limits: Humans, Last 10 years, Practice Guidelines</p> <p>Google Scholar</p> <p>(Traffic OR Road safety Or road) AND policy AND UK OR united Kingdom OR England</p> <p>Traffic calming AND bunn</p> <p>Evidence-based road safety</p>	<ul style="list-style-type: none"> • AAA foundation for road safety research (USA) • ARRB Australian Road Research Board • Australian Transport Safety Bureau • Department for transport • Danish council for road safety research • Road Peace (UK) • Royal Society for the Prevention of Accidents (ROSPA) • SWOV (institute for road safety research)
<p>Chapter 6</p> <p>Google and Google Scholar (29th March 2010)</p> <p>National Library of Guidelines (29th March 2010)</p>	<ul style="list-style-type: none"> • American Geriatric Society • British Geriatrics Society • National Osteoporosis Society

<p>Falls AND older people AND guidance Fall prevention AND older people Falls or Fall</p>	
<p>Chapter 7 Google Scholar and Google Guidelines AND public health Guidelines AND sexual health Guidelines AND teenage pregnancy NICE guidelines STI NICE sexually transmitted infections NICE one to one interventions</p>	<ul style="list-style-type: none"> • PubMed not searched because review published on NICE website not in journal • Also searched ERNIE database and Shared Learning Database on NICE website for evidence on implementation
<p>Chapter 8 Google and Google Scholar Guidelines AND telephone triage AND Bunn Guidelines AND telephone consultation AND Bunn National Library of Guidelines Guidance AND telephone triage AND Bunn Guidance AND telephone consultation AND Bunn Triage Telephone</p>	

Interviews with key informants

Many evaluations of impact are conducted by an independent party rather than the original investigator/s themselves. Such appraisals, therefore, need to access the ‘insider account’ which often involves questionnaires or interviews with the principal investigators on a study (Hanney, 2007). In this instance, however, I was assessing the impact of my own work and, therefore, already had the ‘insider account’. Whilst such an undertaking may be subject to problems with objectivity and bias, internal evaluators have the advantage of being in a better position to keep a record of impacts (Wooding, 2007). For this study I have supplemented my own recollections by undertaking informal telephone or email interviews with co-authors.

Having examined some of the available methods for impact evaluation I now want to consider ways in which such evaluations can be organised. Therefore, in the next section I look at some of the frameworks that have been developed to structure the assessment of research impact, and describe how I have drawn upon them to guide my own analyses.

Evaluation frameworks

In addition to the diversity in available methods for evaluating research impact there are also numerous frameworks available for structuring assessments of impact. In a systematic review of impact assessment Boaz found fourteen frameworks for structuring and interpreting data that were either discussed or used in the literature although, of those, only a small number had actually been used in impact evaluations (Boaz, 2008a). Although there are obviously areas of commonality between the domains included in the different frameworks there is no standard approach. Some, such as the payback model and the ESRC framework, were devised specifically for assessing the impact of either health research or social research and others were developed either during a study or after the data had been collected (Hanney, 2007). These frameworks have, in general, been designed to determine research impact in a range of domains, of which policy may just be one.

It has been suggested that it may not always be necessary to use a framework for impact evaluation as long as there are clear objectives and explicit techniques or methods to gather data (Hanney, 2007). However, frameworks can help to organize inquiry by identifying elements and relationships among elements that need to be considered for the generation of theory (Ostrom, 2007), although of course they do not of themselves predict behaviour or outcomes (Schlager, 2007). I felt, therefore, that it would be useful to refer to existing frameworks to structure my analysis. What is more using a common structured approach has the advantage that it allows more easily for comparison across the reviews (Wooding et al., 2004).

To decide on which framework/s would be most appropriate I began by looking at the information about frameworks provided by two recent reviews on impact assessment (Boaz, 2008a, Hanney, 2007). These reviews were used to identify frameworks and the key references associated with them. This was followed up by more detailed reading on those models that appeared to be most appropriate. I rejected frameworks for a number of reasons. For example, the ROAMEF framework (PREST, 2006) reflects key stages of the policy process however, it was concerned with the impact of whole programmes, was designed to be put in place before the beginning of a programme and does not specify methods for conducting the evaluation. The RAPID outcome assessment (ROA) is a visual mapping tool that evaluates the impact of a programme on policy and the policy environment. Although it has been suggested that it provides useful insights into the process of research impact (Boaz, 2008a) data collection methods associated with this model are too complex and time consuming and not fit for my purpose. In the end I looked in detail at three frameworks:

- The payback framework and model
- The Lavis knowledge transfer approach
- The research impact framework

These are described in more detail below.

The payback framework

I chose the payback framework (Buxton and Hanney, 1996, Buxton and Hanney, 1997, Hanney et al., 2004a, Hanney et al., 2000) for a number of reasons. It is the most commonly used framework in the evaluation of health research impact (Hanney, 2007), is well described in the literature and there are a number of publications detailing suggested methods for conducting evaluations. In addition, although it was not developed specifically for systematic reviews, it has been used to assess their impact (Soper and Hanney, 2007). This evaluation framework consists of a multidimensional categorization of the benefits, or

payback, from health research (Hanney et al., 2000). The five main categories of evaluation criteria for the framework are:

- (a) Knowledge production,
- (b) Research targeting, capacity building and absorption,
- (c) Informing policy and product development,
- (d) Health benefits
- (e) Broader economic benefits

It has been suggested that, whilst the dimensions of the framework are useful, it may not be necessary for all evaluations to include all the dimensions (Boaz, 2008a). Although I am mostly interested in the first three categories, and in particular the category on informing policy development, all five categories are described below briefly.

(a) Knowledge production

The first category of the model is knowledge production. Although this covers any type of publication peer-reviewed articles are considered to be most important. Factors such as the impact factor of the journal and the existence of an accompanying editorial are also seen as being of particular significance (Hanney et al., 2004b). Although as Hanney and colleagues point out there may be some journals that do not have an impact factor but that are still important vehicles for dissemination of the knowledge produced. For example, although the Cochrane Library had long been acknowledged as an important source of high quality information it only received an impact factor in 2007. What is more, the impact factor it now has may be an underestimation of the true impact of Cochrane Reviews worldwide and does not reflect their widespread use in guideline development, policy setting, and health care decision making by practitioners and consumers (The Cochrane Library, 2008).

(b) Research targeting, capacity building and absorption

A potential benefit of research is the better targeting of future research. Whilst this is, of course, relevant to all types of research it is particularly pertinent to systematic reviews which are often instrumental in the identification of areas for future research. This includes not only highlighting topics for research but also suggestions for the types of study design needed and the outcomes that should be measured. This category of the model also includes any impact upon research training and capacity building. Although my own analysis does not include an assessment of impact upon training and capacity building it is worth considering briefly here the potential role for systematic reviews in these areas. Systematic reviews, with their emphasis on multi-disciplinary teams, including clinicians and researchers, are in a key position to develop research capacity. The clear structure of a review and the guidance produced by organisations such as the Cochrane Collaboration and the Centre for Reviews and Dissemination also make systematic reviews feasible for those who do not have a great deal of previous research experience. The process of being involved in a review facilitates the development of skills in critical appraisal, analysis and report writing. Indeed, as a reviewer I have been involved in developing systematic review skills in health care professionals seconded to work on clinical questions relevant to their own area of practice (Cunningham et al., 2006, Petrie et al., 2007b, Pocock et al., 2010, Yanagawa et al., 2000).

(c) Informing policy development

Most pertinent to my work is the section of the framework which addresses the impact of research on policy development. As discussed earlier in the chapter I have adopted the definition of policy that is used by Hanney and colleagues in their framework; this is a broad interpretation of policy and includes the effect on policy locally and nationally and incorporates guidelines and the development of policies for education.

(d) Health benefits

Hanney and colleagues say that it is benefits in terms of health gains that might be viewed as the 'real' payback from health research. In this framework health benefits also include cost savings and improvements in the process of health care delivery. The measurement of health benefits is, however, rarely investigated in impact evaluations because of the difficulty of attributing particular health gains to specific pieces of research. As such it is largely beyond the remit of this work.

(e) Broader economic benefits

The last category in the framework is that of the economic benefits that might accrue from research: for example potential benefits to the economy through the commercial exploitation of research, or the benefits gained through reducing the economic burden of disease and creating a healthy workforce. The measurement of economic benefits is also beyond the remit of this submission.

Payback model

Hanney and colleagues have also developed a model for assessing payback which accompanies the evaluative framework described above (Hanney et al., 2000). The model includes seven stages which cover the whole research process including needs assessment, project specification, commissioning and outputs. Outputs are described as primary or secondary with primary outputs being the direct outputs of the research project, such as academic publications or presentations, and secondary outputs the wider impacts on policy and practice (for example citation in policy guidelines). In this study it is the secondary outputs I am concerned with as the evaluation concerns the impact of research papers (primary outputs) rather than projects as a whole. Stage 5 of the model is Applications which refers to actual changes in behaviour by practitioners and the public and the final stage involves impacts or final outcomes which include service and economic payback such as cost savings. Although the model is

described in stages they accept that the process may not be strictly linear (Hanney et al., 2000).

Knowledge transfer approach

The second framework that I looked at in detail is that by Lavis and colleagues (Lavis et al., 2003a). This is an assessment tool based on ideas of how best to transfer and facilitate the uptake of research knowledge. In contrast to the Payback Model this tool focuses particularly on the impact of research on the decision making process. The rationale for this approach is that it is too difficult to trace the complex pathways through which research might be translated into improved implementation or performance and from there into health gains. They argue, therefore, that it is the use of research knowledge to inform decision-making, rather than a change in health status, that constitutes the most appropriate generic measure of the impact of research (Lavis et al., 2003a). The tool includes the following four steps:

1. Identify the target audience for the research knowledge that has been funded or produced
 - a. general public
 - b. patients
 - c. clinicians
 - d. managers
 - e. research and development officers
 - f. public policy makers
2. Select the appropriate category of measures e.g.
 - a. Producer-push category if researchers led efforts
 - b. User-pull category if decision makers led efforts
 - c. Exchange categories if researchers and decision-makers have jointly led efforts
3. Select evaluation measures given the resources available
 - a. Process measures if limited resources
 - b. Intermediate if sufficient resources to conduct a survey

- c. Outcome measures if resources for case studies
4. Identify data sources and/or collect data, analyse the data, identify areas for future improvement and feed-back information to those involved

The advantages of this tool are that it focuses particularly on the decision making process and that it can be adapted to suit the resources available to the evaluator. For example, the choice of evaluation measure is dictated by available resources. They categorise outcome measures according to whether they capture a process associated with the pursuit of research impact (such as the number of papers published), an intermediate outcome (e.g. a change in awareness about a particular body of research knowledge) or an outcome (e.g. a decision to select one policy over another based on research evidence).

The Research Impact Framework

The third framework I considered was the Research Impact Framework. Developed by Kuruvilla and colleagues (Kuruvilla et al., 2006, Kuruvilla et al., 2007) this is a conceptual framework that uses a standardised way of describing a wide range of potential areas of health research impact and is designed to be used by researchers without any specific training in research impact assessment. It was created by identifying potential areas of health research impact and draws on a number of other models including the two I have already described: the Payback model of health research benefits (Buxton and Hanney, 1996) and Lavis's 'knowledge transfer' approach to assessing the impact of research (Lavis et al., 2003a). It includes four broad areas of impact: research-related impacts; policy impacts; service impacts; and societal impacts and looks at both positive and negative impacts. The authors of the framework acknowledge that there are many competing views on the role of science and its relationship with society, that there are various theories and models of the causal pathway of research impacts and that individual researchers may 'ascribe to different worldviews'. However, they say that the framework 'is not aligned with any particular philosophy, is not in itself evaluative and does not prioritise impacts or propose

causal pathways' (Kuruville et al., 2006). Whilst much of the framework covers areas that are outside the remit of my analysis I think it is worth considering in more detail the section of the framework concerning policy impacts as this is the focus of my own work. The framework includes several potential areas of policy impact which are further broken down into descriptive categories. These are:

- Levels of policy-making
 - e.g. international, national or subnational
 - Different groups, e.g. national and local politicians, health service administrators, managers/directors, representatives of local, national and international professional groups
- Type of policy
- Nature of policy impact. Further divided into four descriptive categories based on the work of Weiss (Weiss, 1998).
 - Instrumental use where research findings drive policy-making
 - Mobilisation of support where research provides support for policy proposals
 - Conceptual use where research influences the concepts and language of policy deliberations
 - Redefining/wider influence where research leads to rethinking and changing established practices and beliefs.
- Policy networks
- Political capital

This consideration of research impact on policy, with the distinction between direct and indirect use of research, attempts to deal with the complexity of the relationship between research and policy. However, in reality such distinctions may be difficult to make and judgments about which category to choose may be open to subjectivity.

Selected framework

None of the models described above were entirely suitable for my evaluation. Although the Payback model has much to recommend it, it has been criticised for not fully explaining the complex interface between research and policy (Boaz, 2008a). Therefore, I chose to adopt several of the domains from the Payback model as a basic framework but to incorporate aspects of the Research Impact Framework. This is shown in Box 1 below.

Box 1 Framework for research impact evaluation

1. Knowledge production
 - Publications (journal, impact factor)
 - Impact within research community (e.g. citation analysis)
 - Other methods of dissemination (e.g. oral presentations, press)
2. Research targeting
 - Influence on other research (e.g. follow-on research)
3. Informing policy development
 - Levels of policy-making
 - I. e.g. international, national or subnational
 - II. Different groups, e.g. national and local politicians, health service administrators, managers/directors, representatives of local, national and international professional groups
 - Type of policy
 - Nature of policy impact. Further divided into four descriptive categories based on the work of Weiss (Weiss, 1998).
 - I. Instrumental use where research findings drive policy making
 - II. Mobilisation of support where research provides support for policy proposals
 - III. Conceptual use where research influences the concepts and language of policy deliberations
 - IV. Redefining/wider influence where research leads to rethinking and changing established practices and beliefs.
 - Policy networks
4. Impact on practice

Although the Lavis tool is not incorporated in the framework, it is designed to be used prospectively; I have used it when considering the driver behind each review, for example who were the intended target audience for the review.

Other methodological considerations

There are several other aspects of a research impact evaluation that could affect the validity of the results. These include the timing and direction of an evaluation and the possibility of bias or conflicts of interest. These issues are addressed below.

Timing

The timing of an impact evaluation is also important. There needs to be sufficient time since the research was completed for change to have occurred but not be so long that the recall of individuals is affected. In this instance publication dates ranged from 1998 until 2008 and it is likely to be harder to discern evidence of impact for the more recently published reviews.

Direction of travel

Most evaluations of impact employ approaches that involve tracking forward from a piece of research or backward from a policy change or document (Boaz, 2008a). In this instance my evaluation involves using my reviews as a starting point and tracking forward from them. The major advantage of a study that starts with specific research and tracks forward is that, because of its more tightly defined focus, it may be more likely to identify impact (Hanney et al., 2003). However, a weakness of the approach is that it may be associated with a tendency to exaggerate the impact of the specific research under consideration (Hanney, 2007). Such studies may find it hard to distinguish between the specific research under investigation and other contributing factors.

Bias

It is likely that those undertaking investigations of research impact are doing so on behalf of those with a vested interest in proving the value of research, for example researchers, research funders or organisations involved in research. Indeed, most evaluations are sponsored, or funded, by the body that originally funded the programme of research (Hanney, 2007). As such many evaluations of research impact may be susceptible to bias or conflicts of interest with the risk of bias greater when a researcher is directly investigating their own research. Post-hoc analyses, such as mine, do not have the advantage of distinguishing the situation at baseline and determining before the project starts what measures of impact will be used.

Conclusion

There are a number of issues to consider when designing an investigation of research impact, including the choice of methods and framework, the direction and timing of the evaluation and the potential for bias. In addition there are several important caveats to consider when attempting to establish the impact of research on policy. For example, it is generally accepted that knowledge production activities are more easily discernable than impact on policy or health gain (Hanney, 2007). Then if we do identify evidence of impact we need to consider whether this is real or a result of bias and whether it is attributable to the research we are investigating or if it is possible that such effects may have occurred anyway (Molas-Gallart et al., 2000). Analyses of impact are further complicated by the convoluted nature of the relationship between research and policy. Whilst it may be challenging to identify the direct or instrumental impact of policy research this challenge is even greater if we are trying to show the way research may have been influential indirectly. For example, it may be particularly difficult to attribute indirect outputs, such as the conceptual use or diffusion of research ideas, to a particular piece of research or research programme.

Chapter 4: Strategies for fluid resuscitation: systematic reviews of RCTs

Introduction

In this chapter I present five systematic reviews (seven papers) that all addressed topics related to the fluid resuscitation of injured patients (Bunn et al., 2004b, Bunn et al., 2008b, CIG Albumin Reviewers, 1998, Kwan et al., 2003, Roberts et al., 2004). Full versions of the papers can be found in Appendix 2. The reviews cover clinical questions relating to the type of fluid to use and the method of administration. All of the reviews involved the employment of the statistical technique meta-analysis and they are used as illustrative examples to critically examine the methodological issues associated with this type of review. In addition, the reviews form the basis for an impact evaluation and critique of some of the key factors that shape health care policy. For example, by using these reviews as case studies, I examine the role that the pharmaceutical industry and the media play in decisions about health care policy.

On two of the included reviews I am first author (Bunn et al., 2004b, Bunn et al., 2008b) and on the other three (Alderson et al., 2004, Alderson et al., 2000, Kwan et al., 2003) I was a co-author. The rationale for including reviews on which I was a co-author, rather than just those on which I was first author, is that all of the reviews were concerned with some aspect of fluid resuscitation for injured patients and that together they form a significant body of work on the subject. Moreover, the methodological requirements for systematic reviews mean that they are often very collaborative in nature. All of the reviews were published in the Cochrane Library but two (Human albumin solution for fluid resuscitation and Colloids versus crystalloids for fluid resuscitation) were also published in the BMJ (CIG Albumin Reviewers, 1998, Schierhout and Roberts, 1998). All the reviews have been updated, most on several occasions. My specific contribution to each of the reviews is detailed in Chapter 3.

Drivers behind review question/s

The driver/s behind research has been identified as a factor that may affect the impact of research. Therefore, I begin by considering the drivers behind the systematic reviews included in this chapter. The reviews were part of a group of systematic reviews on the fluid resuscitation of critically ill patients that were conducted by members of the Cochrane Injuries Group. The Cochrane Collaboration consists of a number of collaborative review groups, such as the Injuries Group, that are concerned with preparing, maintaining and promoting high quality systematic reviews in a particular area of health care. Protocols for systematic reviews and completed reviews are peer reviewed and published in the Cochrane Library where they are open to public and professional scrutiny. The Cochrane Library is published on CD-ROM and the internet and electronic publication means that reviews can be updated or amended in response to post-publication criticism or new information. The intended audience for the reviews includes providers, practitioners and patients.

The scope of the Injuries Group covers the prevention, treatment and rehabilitation of traumatic injury; including the resuscitation of seriously injured and burned patients. The impetus for the choice of Cochrane review topics tends to be largely researcher or clinician driven. In this instance the choice of review topic was governed by several factors. In the first instance review topics were chosen after a prioritisation process conducted by the editorial team of the Injuries Group when the group was first established. However, the process then became more organic with one review question prompting another line of enquiry. For example, the albumin review came about because the first author on the colloids versus crystalloids review (Dr Gillian Schierhout) felt that the data suggested albumin might be harmful and should be the subject of further investigation. Likewise the finding that albumin was potentially harmful led to the review comparing the safety and efficacy of albumin with other synthetic colloids (Bunn et al., 2000a).

Systematic review methods

In the following section I enumerate the methods employed for the included reviews and critically examine the key methodological factors in systematic reviews and meta-analysis of randomised controlled trials (RCTs). These include an examination of issues around the following systematic review processes:

- Question development
- Identification of studies
- Data extraction
- Assessment of risk of bias
- Data synthesis including meta-analysis

Question development

The starting point of a systematic review is a protocol detailing inclusion criteria and methods. The use of a protocol provides researchers with the opportunity to define their question clearly and minimises the risk of ad hoc decisions that may introduce bias (Higgins, 2008). The importance of protocols to the Cochrane review process is underlined by the fact that all protocols are peer reviewed and published on the Cochrane Library. Central to the protocol is the development of a well-focused question which allows clear decisions to be made about what research to include in a systematic review and how to summarise it (Higgins, 2008). For a typical Cochrane review evaluating the effectiveness of an intervention the question is defined in terms of participants, intervention/s (and comparison/s), outcomes and types of study. The acronym PICO (participants, interventions, comparison and outcomes) is frequently used as a reminder of these criteria. One of the challenges for a reviewer is to construct a question that while focused is still meaningful. A question that is too broad becomes unmanageable but a question that is too narrow may have limited usefulness. In addition, any inclusion or exclusion criteria should have a rational justification

and be based on biological, sociological or other clearly warranted criteria (Bunn et al., 2001).

Part of the process of developing the review question includes specifying the outcomes to be included in the review. The outcomes chosen should include those that are likely to be meaningful to decision makers such as clinicians, patients and policy makers. However, frustratingly, those outcomes considered important by the reviewer may not necessarily be reported in individual studies. For example, in the fluid therapy reviews included in this chapter many authors reported surrogate outcomes such as intermediate physiological outcomes rather than those we considered relevant such as mortality, morbidity and adverse effects. Surrogate outcomes are flawed for a number of reasons. They are subject to intra and inter-observer variation, they may not be stable over time and importantly they have no face value to patients, their relatives or many decision makers. In addition, to be meaningful there would need to exist a strong predictive relationship between the variable and the primary outcome; a relationship that often is not proven.

Another key protocol decision is which study designs will be included in the review. As outlined in Chapter 1 RCTs are generally considered to be the optimal study design for unbiased estimates of intervention or treatment effects. Traditionally most Cochrane systematic reviews have focused on randomised controlled trials and all of the reviews included in this chapter are restricted to RCTs only. The inclusion of other types of studies in systematic reviews will be considered in subsequent chapters.

Identification of studies

As already outlined in Chapter 1 there are a number of methodological characteristics of systematic reviews that distinguish them from non-systematic traditional literature reviews. One of these is the emphasis on searching for all available literature regardless of publication status or country of origin. The rationale for the development of such highly sensitive search strategies is to

prevent or reduce the possibility of introducing bias into the review by missing relevant studies. One potential bias is publication bias where research with statistically significant findings is more likely to be published than is work with null or non-significant findings. For example, National Institutes of Health-Funded trials in the USA with significant results were more than twice as likely to be published than those showing non-significant results and in the UK among studies approved at the central Oxford research committee those with statistically significant results were more than twice as likely to be published as studies with null results (Easterbrook et al., 1991). More recent research has found that publication bias is still a threat to the validity of systematic reviews. In a systematic review investigating the impact of grey literature in meta-analyses of RCTs of health care interventions the authors found that published trials tended to be larger and show an overall greater treatment effect than grey trials such as abstracts and unpublished data (Hopewell et al., 2007b). In addition, trials with positive results tended to be published sooner than other trials (Hopewell et al., 2007a); which is sometimes known as 'time-lag bias'. Interestingly some investigators have found no association between the quality of study design and publication (Dickersin and Min, 1993, Easterbrook et al., 1991, Elvik, 1998) but rather the most common reasons why trials went unreported was that the investigators thought the results 'uninteresting' or else they 'did not have enough time'. Indeed, the publication strategy for the reviews included in this chapter reflects such biases. Those considered more 'interesting' were more likely to be submitted for dual publication in the Cochrane Library and another journal rather than just the Cochrane Library.

Therefore, to minimise the risk of introducing bias through failing to identify relevant studies systematic review search strategies are generally designed to be highly sensitive. In this context sensitivity refers to the ability of a search strategy to identify all relevant articles. A search strategy that missed relevant studies would be considered to have low sensitivity. However, the need for sensitivity must be balanced against the specificity or precision of a search. Specificity is a measure of the ability of the search strategy to exclude irrelevant

articles. A search strategy that retrieves many studies that are not relevant to the review would be considered to have low precision. Systematic review teams need to design search strategies that balance the requirement for maximising sensitivity against the practicalities of ensuring an adequate degree of specificity.

The most important sources of studies for systematic reviews, in particular for reviews of RCTs, are electronic databases such as PubMed (MEDLINE), EMBASE and The Cochrane Central Register of Controlled Trials (CENTRAL). The latter, which has been developed by the Cochrane Collaboration, is one of the largest repositories of controlled trials (Dickersin et al., 2002). In contrast to reviews that include non-randomised studies (as some of those in later chapters) searching for RCTs is relatively straight forward, there are a number of methodological filters that can be used to restrict searches to RCTs and controlled studies, thereby ensuring greater specificity without sacrificing sensitivity. As well as searches of the standard databases search strategies may also include the use of a range of specialist databases, handsearching of journals, searches of the internet for grey literature, contacting experts for information on possible studies and contacting drug companies or manufacturers to identify unpublished studies. In this context grey literature refers to research that is not formally published and that cannot be found easily through conventional channels such as electronic databases.

The augmentation of database searches with strategies such as handsearching and personal contacts can be time consuming and it is questionable whether the yields are always worth the effort involved. For example, for the albumin and colloids reviews we contacted the drug companies that manufactured the products to request any relevant information about unpublished studies. A number of the companies that manufactured colloids sent us information. This information, which in some cases was extensive, was handsearched but yielded no relevant studies. In an HTA report investigating the importance of comprehensive literature searches for systematic reviews of RCTs Egger and colleagues found that, although the worth of extensive searches to locate

difficult to find studies varied between specialities, in general trials which were harder to locate were often of poorer quality. They argue that, as such poor quality trials may increase the risk of bias, in situations where resources are limited detailed quality assessment should take precedence over extensive literature searches and translations of articles (Egger et al., 2003).

Data extraction

One of the distinguishing features of a systematic review is a comprehensive and scientific approach to extracting the necessary data from any primary studies that meet the inclusion criteria. This requires that a data collection tool, specific to each review, is developed and piloted. Such a tool visually represents the review question and planned appraisal of included studies, records decisions taken during the review process; and records outcomes data to be analysed (Higgins, 2008). The content and length of the form is largely dependent on the complexity of the review question with some reviews of complex interventions involving more detailed and lengthy data extraction procedures than the reviews in this chapter. For example, in a review of interventions for the prevention of sexually transmitted infections and teenage pregnancy (Bunn et al., 2006a) the funders, NICE, asked that we considered effectiveness in the context of how, by whom and where the intervention was delivered. We were also required to look at whether effectiveness was affected by socio-demographic characteristics such as sex, ethnicity and class. These factors, around the process and delivery of an intervention, are of less concern in reviews such as those in this chapter but are considered further in later chapters.

In the context of a systematic review, 'data' not only refer to results but also include any information about (or deriving from) a study, such as the methods, types of participants, the interventions, settings and context (Higgins, 2008). Owing to the high prevalence of data extraction errors (Gotzsche et al., 2007, Jones et al., 2005a), and the fact that independent double data extraction may result in fewer such errors (Buscemi et al., 2006) it is generally recommended that data extraction is performed by two reviewers independently (Higgins,

2008). For all reviews in this submission screening of records and data extraction was done by two reviewers independently.

Assessing the risk of bias

There is evidence that the quality of studies included in systematic reviews can influence effect estimates and that more rigorous studies are more likely to yield results that are 'closer to the truth' (Higgins, 2008). A bias is a systematic error, or deviation from the 'truth', in results or inferences and may lead to underestimation or overestimation of the true treatment effect (Higgins, 2008). Hence, it is vital to assess study quality. There are a number of elements that are associated with study quality and two of the most key are randomisation and allocation concealment.

Randomisation should involve both the generation of an unpredictable assignment sequence and the concealment of that sequence until allocation to study group occurs (Schulz et al., 1995). Of these, the concealment of allocation appears to be the most important. Analysing 250 trials from 33 meta-analyses, Schulz et al (Schulz et al., 1995) found that compared with trials with adequate allocation concealment, odds ratios were exaggerated by 41% for inadequately concealed trials and by 30% when the adequacy of concealment was unclear. In a similar study, inadequate allocation concealment exaggerated the effect estimate by 37% (Moher et al., 1998). More recent studies also substantiate these findings. A pooled analysis of seven methodological studies found that effect estimates from trials with inadequate concealment of allocation or unclear reporting of the technique used for concealment of allocation were on average 18% more 'beneficial' than effect estimates from trials with adequate concealment of allocation (95% confidence interval 5 to 29%) (Pildal et al., 2007). However, this exaggeration of effect estimates was found to be more of a threat when a subjective outcome (for example quality of life) was analysed with little evidence of bias in trials with objective outcomes (such as mortality) (Wood et al., 2008).

Allocation concealment is not the only factor associated with the risk of bias. Among adequately concealed trials and trials that did not clearly report about concealment, inadequate methods of sequence generation exaggerated effects by 25% and failure to double-blind by 17% (Schulz et al., 1995). Other potential sources of bias are incomplete outcome data, the selective reporting of outcomes (for example investigators only reporting data for outcomes that were statistically significant) and inadequate blinding of outcomes assessors.

There is debate about how to take account of quality assessment in a systematic review and, as more methodological research is conducted, recommendations for best practice have altered over time. Many scales and checklists designed to assess the validity and quality of RCTs have been identified (Moher et al., 1995). However, serious doubts remain about the usefulness of many of these scales as they may contain items not directly related to internal validity. At the time these reviews were originally conducted the Injuries Group policy was to judge quality primarily on allocation concealment as this was well known to be associated with bias (Schulz et al., 1995). More recently the Cochrane Collaboration has developed a more detailed method for assessing the risk of bias that involves the creation of a risk of bias table. This is a domain-based evaluation that looks at specific methodological factors known to be associated with bias, such as randomisation, allocation concealment and blinding of outcome assessment, and assesses studies on the level of bias likely to be present for each domain and for each study overall (Higgins, 2008).

Data synthesis

In a systematic review of RCTs data is often combined using the statistical technique of meta-analysis. To decide whether it is appropriate to combine data quantitatively in a meta-analysis researchers must judge whether there is sufficient homogeneity in terms of participants, interventions, study design and outcomes in the studies included in the review. This judgement may take into account, for example, known effect modifiers, biologic plausibility, methodological quality, relevance of outcomes and clinical experience or other

factors. Ideally these are all issues that will have been considered when formulating the research question and choosing the inclusion criteria. As well as the reviewer using their own judgement, prior to data synthesis, as to whether the use of meta-analysis is apposite heterogeneity can also be assessed once a meta-analysis has been conducted.

One way we can identify heterogeneity in a meta-analysis is through a visual inspection of the forest plot¹. For example if studies are homogenous, that is they are estimating the same thing, we would expect that there would be an overlap of confidence intervals. A poor overlap of confidence intervals, or the presence of obvious outliers², may suggest statistical heterogeneity. In addition, reviewers should use statistical tests to assess for heterogeneity. When these reviews were first conducted the test used in Cochrane reviews was the Chi-squared test which assesses whether the observed differences in studies are compatible with the play of chance alone. However, it has since been argued that as clinical and methodological diversity always occur in a meta-analysis statistical heterogeneity is inevitable (Higgins et al., 2003). Therefore, the focus has moved away from testing whether heterogeneity is present to measuring its impact or importance (Higgins and Thompson, 2002, Higgins et al., 2003). The statistic for assessing this is called the I-square statistic and is now included along with the Chi-square test on all forest plots in Cochrane reviews.

Of course even if tests show no statistical indication of heterogeneity this cannot be interpreted as evidence of homogeneity. This is not only because a non-significant test can never be interpreted as direct evidence in favour of the null hypothesis of homogeneity (Altman, 1991), but in particular because tests of homogeneity have low power, particularly where there are few trials, and may fail to detect as statistically significant even a moderate degree of genuine heterogeneity (Thompson and Pocock, 1991, Whitehead and Whitehead, 1991). Conversely, in meta-analyses with large number of trials the tests may be

¹ A graphical display designed to illustrate the relative strength of treatment effects. Shows the results of each individual study and the overall pooled result.

² An outlier is an observation that lies an abnormal distance from other values

oversensitive and detect heterogeneity that, while present, is not clinically significant. In all but one of the reviews included in this chapter (Kwan et al., 2003) there was no evidence of statistical heterogeneity, therefore, results for most outcomes were pooled in meta-analyses.

When judged to be appropriate, performing a meta-analysis can make an important contribution to a systematic review. Combining the results from similar randomised trials will increase the precision of the effect estimates and may allow the reliable estimation of even modest effects (Egger, 2001a). There are several methods for pooling study results, all of which take a weighted average of the study-specific effect estimates, with the weights being inversely proportional to the variance of the effect estimates. Clearly, in order to combine data, information on the effect estimate and its variance must be available and poor reporting of study results can therefore be an important obstacle to meta-analysis (Wagenaar, 1999). In the reviews in this chapter data for dichotomous outcomes, such as mortality, were pooled using the relative risk. This is one of a number of summary statistics available for the presentation of such data and was chosen over the odds ratio or risk difference (or absolute risk reduction) because it is less open to the possibility of misinterpretation.

Results of impact analysis

In this section I document any evidence that demonstrates ways in which the reviews included in this chapter may have impacted or influenced health care practice and policy. In addition I critically examine some of the issues related to policy impact. The results of the impact evaluation are presented using the framework described in Chapter 3. This includes the main headings of knowledge production, research targeting, informing policy development and impact on practice.

Knowledge production

Publications and other methods of dissemination

All the reviews were published in the Cochrane Database of Systematic Reviews (CDSR) on the Cochrane Library. There is no restriction to the word limit for Cochrane Reviews and the full final documents tend to be quite lengthy, particularly when, as with the colloids review, there are a large number of included studies. Reviews can be accessed and downloaded via subscription, although they are free to all UK residents via the NHS Health Information Resources (<http://www.library.nhs.uk/Default.aspx>), and PDFs of abstracts, standard or full versions of reviews are available. Reviews are accompanied by plain language summaries which are designed with consumers and the lay public in mind. Figures for 2007 show that CDSR has an impact factor of 4.654 which ranks it 14th out of 100 journals in the ISI category 'Medicine General and Internal'.

In addition to publication on CDSR the reviews on the use of human albumin and colloids versus crystalloids were also published in the BMJ (CIG Albumin Reviewers, 1998, Schierhout and Roberts, 1998). The impact factor for the BMJ was 12.827 in 2008 giving it a ranking of 5th in the category Medicine General and Internal (BMJ, 2010). The albumin review suggested that the use of human albumin was associated with an increase in mortality and the nature of the findings meant that as well as publication in the BMJ and Cochrane Library the results of the review were widely reported in the television and print media both in the United Kingdom and elsewhere. The reviews were also disseminated via conference papers, seminar and other forms of academic outputs.

Impact within research community

The citation analysis was initially performed between March and May 2009 and was updated in April 2010. Unsurprisingly, given the publicity that surrounded the publication of the review, the human albumin review (CIG Albumin Reviewers, 1998) was the most cited of any of the reviews in this chapter with

481 citations in WoS and 579 in Scopus. For some reason, possibly because of the way it was originally cited, data for the BMJ version was not available in Google Scholar. The other review that received a high number of citations was the review comparing crystalloids with colloids for fluid resuscitation, particularly the version published in the BMJ (Schierhout and Roberts, 1998). Of the other reviews the colloid solutions review (Bunn et al., 2008b) had 118 citations in Google Scholar, 23 in Scopus and 6 in WoS, the hypertonic versus isotonic review (Bunn et al., 2004b) had 104 in Google Scholar, 23 in Scopus and none in WoS, and the timing of fluid resuscitation review had 114 in Google Scholar, 25 in Scopus and none in WoS. A comparison of the numbers of citations for each review from the different databases can be seen in Figure 4.1.

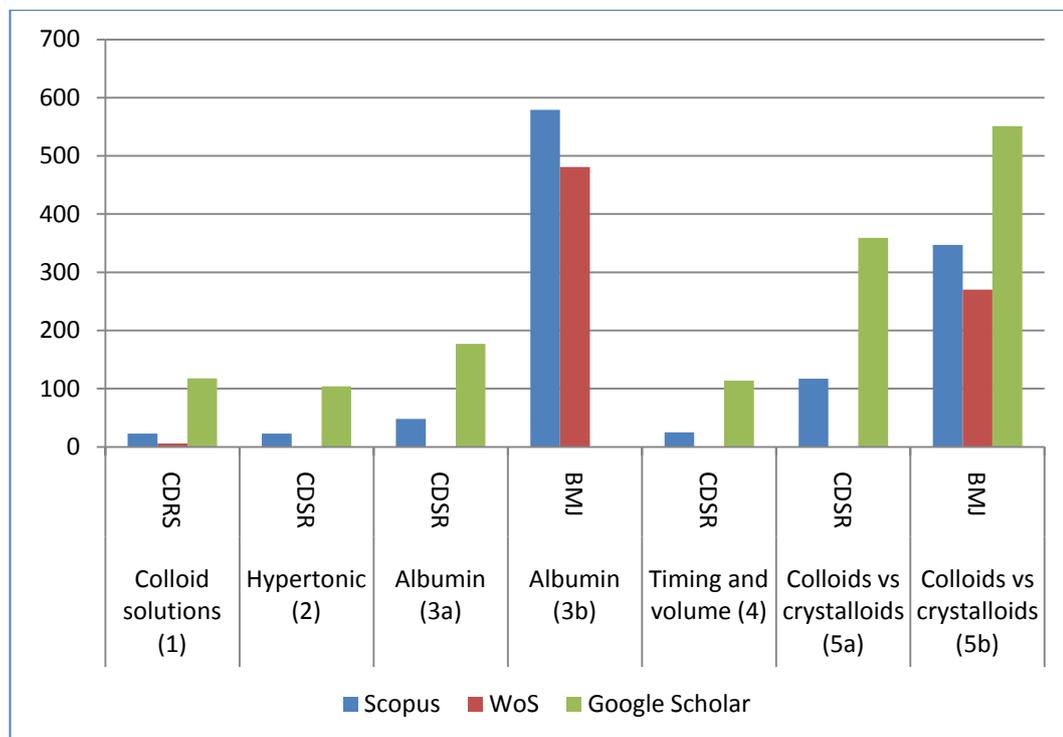


Figure 4.1 Results of citation analysis: Fluid therapy reviews³

There were a number of anomalies in the data, with considerable variation in citation counts for the same review in different databases. For example, the reviews evaluating colloid solutions hypertonic saline and timing and volume of fluid administration each had over 100 citations in Google Scholar but few or

³ Data from April 2010

none in WoS. In addition for the two papers that were published in two journals (albumin and colloids versus crystalloids) the BMJ version of the paper received substantially more citations than the version published in the Cochrane Library. Possible explanatory factors for this difference may be around the relative accessibility and visibility of the two reviews. Open access to the Cochrane Library is not universally available, it is only available electronically, and unlike the BMJ does not have an easily digestible weekly output. Furthermore, the format and length of Cochrane reviews may not be as user friendly for the reader. However, it has been suggested that the citation counts for Cochrane reviews are artificially low because citing authors have incorrectly referenced Cochrane reviews (The Cochrane Library, 2008). I investigate this further in Chapter 8 and my findings confirm that the citation data for Cochrane reviews in Scopus and WoS are inaccurate.

All the reviews in this chapter were cited at least once and even for the reviews that were less frequently cited the number of citations is above average. In a study of 493 published articles in emergency medicine the mean citations per year was 2.04 with 9.3% never cited (Callaham et al., 2002). However, it may not be appropriate to directly contrast such figures as there is evidence that review articles are generally highly cited in comparison with primary research (Bornmann et al., 2008, Meho, 2007). Other factors that may influence the probability of citations are impact, quality and prestige of the journal as well as journal accessibility, visibility and internationality (Bornmann et al., 2008, Callaham et al., 2002).

Data on the number of times the Cochrane reviews were downloaded were available for 2008 and 2009. The results of this can be seen in Table 4.1. The review that was downloaded the most was the review comparing colloids and crystalloids, this was ranked at 35 out of 6840 reviews. The review comparing hypertonic and isotonic crystalloid was downloaded the least but was still ranked in the top third of Cochrane review downloads for 2009.

Table 4.1 Downloads for Cochrane fluid resuscitation reviews 2008 & 2009

Review	Downloads 2008	Ranking 2008 (from a total of 6232)	Downloads 2009		Ranking 2009 (from a total of 6840)
	Full text*		Full text	Abstract	
Colloid solutions for fluid resuscitation	1516	170	1,387	2,457	327
Hypertonic versus isotonic solutions for fluid resuscitation	650	910	538	983	1620
Timing and volume of fluid resuscitation	691	823	694	1,153	1146
Colloids versus crystalloids for fluid resuscitation	2973	18	3,291	5,364	35
Human albumin solution for fluid resuscitation	1040	375	1,007	1,653	632

* data on abstract download not available for 2008

Research Targeting

The administration of intravenous fluids to maintain intravascular volume is a common health care intervention and yet controversy over whether to choose crystalloids or colloids has existed for over five decades (Rizoli, 2003).

Increasingly there have also been questions about when to administer fluids and in what volume (Revell et al., 2002). The citation analysis described above demonstrates that the systematic reviews discussed in this chapter have contributed to the knowledge pool in this area and stimulated debate. It has been found that many Cochrane reviews are a rich source of suggestions for further health-care research (Clarke et al., 2007) and there is certainly evidence that these reviews have been instrumental in influencing follow-on research (Cook and Guyatt, 2001).

For example, the original version of the albumin review highlighted the methodological weaknesses of the available studies and called for further well conducted RCTs in this area. Subsequently researchers in Australia have conducted a large multicentre double-blinded randomised controlled trial which set out to compare the effects of two fluid resuscitation regimens (human albumin or saline) on 28 day all-cause mortality in critically ill patients requiring intravascular volume resuscitation (Finfer et al., 2004a). The SAFE trial (saline versus albumin fluid evaluation) was conducted between 2001 and 2003 and included nearly 7000 participants. They found no difference in mortality between patients who were resuscitated with albumin and those given saline although there was a slight difference, in favour of saline, in mortality in patients who had trauma and head trauma. Although other factors may also have been involved it is clear that the review and surrounding controversy were instrumental in the initiation of this highly important trial.

Informing Policy Development

I begin this section by describing the impact of the human albumin review as this was clear and well documented at the time. Human albumin solution is a blood product that has been used for over half a century in the treatment of blood loss and burns. At the time, an estimated 100,000 patients were being treated in the UK alone with a cost to the NHS of close to 12 million (Roberts and Bunn, 2002). The Cochrane Injuries Group review found an increase in the risk of death in patients receiving albumin over those in the comparison group. Overall, the risk of death in patients receiving albumin was 14%, and the risk of death in patients not receiving albumin was 8%. The media were quick to pick up on a story that suggested that a widely used treatment may have harmed patients and most of the major newspapers in the UK carried headlines similar to this one in *The Guardian*: 'Blood Product May Have Killed Burn Victims'. As soon as the review had been completed the Cochrane Injuries Group wrote to the Chief Medical Office of Health giving early warning of the results. Five months later, the day before the review was published in the BMJ, an Expert Working Party was set up to consider the safety of human albumin administration. The Committee on the

Safety of Medicine took a whole year to reach a decision on albumin. They concluded that 'there is insufficient evidence of harm to warrant a withdrawal of albumin... the effect of albumin on mortality would only be answered by conducting a large RCT'.

This case study demonstrates how various actors or stakeholders with competing interests are involved in the policy process. In this instance stakeholders included the review authors, who had the support of the Cochrane Injuries Group and the Cochrane Collaboration, doctors involved in treating patients with human albumin, the pharmaceutical industry and the Government in the form of the medicines control agency. Of those actors the pharmaceutical companies were perhaps the most powerful and certainly the best funded. One of the reasons the Committee on the Safety of Medicines took a whole year to reach a decision was that the process involved lengthy negotiations with the plasma products industry. In addition, alarmed by the decline in sales of albumin, albumin manufacturers put substantial funds into a campaign to increase the use of albumin. In 2000 the Cochrane Injuries Group received a leaked document from the European Plasma Fractionation Association describing the 'Albumin Support Programme'. The albumin manufacturers had set aside \$2.2 million dollars for the programme whose main aim was to disseminate medical evidence that supported the use of albumin.

This attempt by pharmaceutical companies to boost sales of a product that at best was of doubtful value and at worst might be harmful is alarming but not that surprising. The fact that one of the primary concerns of the pharmaceutical industry is to increase sales and maximise profit is well recognised. Marcia Angell a former editor of the New England Journal of Medicine claims that the pharmaceutical industry has become 'primarily a marketing machine' (Angell, 2005) that 'co-opts every institution that might stand in its way'. Spielmans and Parry use the phrase 'marketing-based medicine' saying that 'science has largely been taken captive in the name of increasing profits for pharmaceutical firms (Spielmans and Parry, 2010). A systematic review investigating whether the

funding source of trials introduced bias found that studies funded by a drug company were four times more likely to have results favourable to the company than those funded by other sources (Lexchin et al., 2003). There is substantial evidence to suggest that this discrepancy is because drug companies suppress or manipulate data to make their products look more effective than they actually are (Angell, 2005, Smith, 2005, Spielmans and Parry, 2010).

Ultimately, however, as I go on to show, this example illustrates how a systematic review can have an impact despite strong opposition from those with a vested interest. Although whether such an influence would have occurred without the media furore that surrounded the review is unclear. In an editorial in the *Annals of Internal Medicine* Cook and Guyatt write 'the role of the intense reaction of the media to the Cochrane meta-analyses in influencing practice is difficult to prove, but easy to deduce' (Cook and Guyatt, 2001). A fuller account of the story of the human albumin review can be found elsewhere (Roberts and Bunn, 2002). Although the other reviews included in this chapter incited no media interest, and had no dramatic or obvious impact in the way that the albumin review did, there is some evidence that they have been instrumental in changing policy and practice. For example the review comparing colloids versus crystalloids (Schierhout and Roberts, 1998), which concluded that there was no evidence to support the continued use of colloids in the fluid resuscitation of critically ill patients, has been widely cited, included in many guidelines, and has been instrumental in fostering debate, and changing practice, in this area.

Levels and type of policy making

The impact of the reviews in this chapter can be seen on policy and practice at local, national and international levels. However, although I was able to discern influence on Government policy, it does appear that the majority of the impact has been at the level of professional bodies developing guidelines. Guidelines have been defined as 'directions of principles presenting current or future rules of policy. They may be developed by government agencies at any level, institutions, professional societies, governing boards, or by the convening of

expert panels (Biology Online, 2009). Guidelines are intended to ensure that health care recommendations are informed by the best available evidence (Oxman et al., 2006). The issues surrounding the development and use of guidelines are explored more fully in Chapter 7.

Inclusion in such a number of guidelines shows that the research has had a demonstrable impact on policy. The reviews discussed in this chapter have been widely cited in guidelines, with the albumin review (CIG Albumin Reviewers, 1998) and the review comparing colloids and crystalloids (Roberts et al., 2004) the most frequently cited. UK organisations drawing upon the reviews for the development of guidance include national Governmental bodies such as the HTA, NICE and SIGN, and professional bodies such as the Royal College of Surgeons Scotland, British Association for Immediate Care, National Blood Service and the British Committee for Standards in Haematology. The reviews have also been included in guidelines in the USA, Canada and Europe and by several professional bodies developing international guidelines. A list of guidelines that cite any of the fluid reviews presented in this chapter can be seen at the end of this chapter in Table 4.2. It should be noted that in reality this list of guidelines may, in fact, be more extensive as my searches were restricted to guidelines published in English or with an English abstract.

Nature of policy impact

The framework that I have adopted for this evaluation assesses the nature of the policy impact based on the categories devised by Weiss (Weiss, 1998). This attempts to distinguish between instrumental or direct impact as against conceptual or symbolic use. In this case there was evidence that all of the included reviews had some direct or instrumental impact in that they were used to inform practice guidelines. However, this evaluation only assessed whether the reviews were cited by guidelines and did not look in detail at whether those guidelines included recommendations that were in agreement with our review conclusions. In addition, most guidelines are based on a number of different publications and it is not easy to determine the contribution of individual

reviews. One of the categories in the framework is the extent to which research may have been influential in redefining, rethinking or changing established practices and beliefs. Although colloids, including human albumin, are still widely used in practice, there is no doubt that their use has declined (Roberts et al., 1999). In addition, the reviews, in particular the review comparing colloids and crystalloids, have stimulated the debate about fluid resuscitation and contributed to changes in practice.

Impact on practice

Although the focus of this submission is primarily on the influence of systematic reviews on policy it is also worth considering whether there is any evidence of impact on practice. To investigate the possible impact of the albumin review on sales of albumin in the UK one of the authors of the review (Ian Roberts) requested data on the monthly issues of albumin solutions between 1993 and 1998 from the Bio Products Laboratory (which serves England and Wales) and the Protein Fractionation Centre of the Scottish National Blood Transfusion Service (which serves Scotland and Northern Ireland) (Roberts and Bunn, 2002, Roberts et al., 1999). The Protein Fractionation Centre's data reflected virtually all albumin used in Scotland and Northern Ireland but data on Bio Products Laboratory's share of the market in England and Wales was not available (although it was believed to be substantial). In July 1998, after the publication of the review, issues from the Scottish National Blood Transfusion Service fell steeply from 180 kg in June to 62 kg in December. Issues from Bio Products Laboratory fell dramatically after July 1998 with a 40-45% drop in sales of 4.5% albumin and a 40% drop in 20% albumin (Roberts et al., 1999). This decline in sales was despite vigorous criticism of the review in the letters pages of the BMJ and was well in advance of advice of from the Committee on the Safety of Medicines. The decline in sales appears to be a clear indication that doctors took into account the evidence presented in the systematic review and that many doctors changed their practice in response.

In a report on the payback model (Wooding et al., 2004) Wooding and colleagues point out that behavioural change by practitioners is necessary for research findings to result in 'final outcomes'. These changes may be brought about by the influence of secondary outputs, such as policy decisions or guidance, or may be a direct result of the primary outputs of the research. In the case of the albumin review the changes in practice appeared to be a direct result of the primary output (the academic paper published in the BMJ) with the impetus for change coming from practice and practitioners rather than policy makers.

In 2004 I was asked to give a paper about the review, and its impact on policy and practice, at an HTA conference in Poland. For the presentation one of my co-authors, Ian Roberts, and I attempted to update the figures on the use of albumin in the UK and again contacted Bio Products Laboratory and the Scottish National Blood Transfusion Service. Bio Products Laboratory declined to give us updated data on albumin issues but figures from Scotland and Northern Ireland⁴ showed that the fall in the issue of 4.5% human albumin (HA) had been maintained. However, there had been a recent increase in sales of 20% HA. For the conference presentation I was also provided with data on the use of albumin in Sweden and Denmark by Professor Mona Britton from the Swedish Council for Technology Assessment in Health Care. This data showed there was a fall in albumin use in Sweden and Denmark that coincided with the publication of the review. In Denmark hospital costs for 1997-2002 fell from about 37 million Danish Crowns to under 15 million and in Sweden albumin sales fell from around 70 million Swedish Crowns to 35 million (personal communication).

Although the modified framework I have adopted for this evaluation does not include a consideration of the economic influence of the included reviews it is worth noting that there is evidence of considerable cost savings both to the health service in the UK and internationally. Colloids are more expensive than crystalloids, and albumin is considerably more expensive than other synthetic

⁴ Data provided by Neil Docherty 24th April 2004

colloids (McClelland, 1990). Therefore, the evidence provided that colloids are no more effective than crystalloids in reducing mortality, supports the adoption of a cheaper, but equally effective, treatment. The figures provided above, showing documented reductions in the use of albumin, indicate that substantial cost savings have already been made.

Conclusion

The systematic reviews included in this chapter are concerned with clinical questions around appropriate fluid resuscitation strategies; an area of clinical practice that has been subject to significant debate and controversy. They are exemplars of reviews employing rigorous Cochrane methodology to evaluate the effectiveness of clinical interventions and include the pooling of individual study results using the statistical technique meta-analysis. From the impact evaluation evidence emerged that all of the reviews have had some influence on policy and practice, for example being used to develop national and international clinical guidelines. In particular, two of the reviews (CIG Albumin Reviewers, 1998, Schierhout and Roberts, 1998) were found to have had a significant impact. For the first of these (CIG Albumin Reviewers, 1998) the impact of the review was heightened by the suggestion that the intervention being evaluated may have significantly increased mortality and the surrounding media furore increased the impact further. The other most influential review, comparing colloids and crystalloids (Schierhout and Roberts, 1998) has contributed significantly in creating debate about fluid resuscitation strategies and has contributed to changes in policy and practice. The reviews have also been instrumental in targeting future research.

Table 4.2 Evidence of Fluid reviews' influence on guideline development

1. Colloid solutions for fluid resuscitation. First published: 2000		
Guidelines Title	Organisation producing guidelines	Country
Clinical effectiveness and cost-effectiveness of prehospital intravenous fluids in trauma patients. (Dretske, 2004)	Health Technology Assessment (HTA)	UK
The use of pre-hospital intravenous fluid therapy in trauma 2004 (NICE, 2004)	National Institute for Health and Clinical Excellence (NICE)	UK
Fluid resuscitation in prehospital trauma care: a consensus view (Revell et al., 2002)	Consensus review including Royal college of surgeons Scotland, Ambulance service Association, The UK Military Defence Forces, British Association for Immediate Care (BASICS) and London Helicopter Emergency Medical Service	UK
Techniques For Assessing And Achieving Fluid Balance In Acute Renal Failure (Clark et al.)	Acute Dialysis Quality Initiative 2 nd International Consensus Conference	UK
The management of a child with a decreased conscious level. A nationally developed evidence-based guideline for hospital practitioners (The Paediatric Accident and Emergency Research Group)	The Paediatric Accident and Emergency Research Group	UK
Albumin vs synthetic colloids for fluid resuscitation May 2007	Canadian Blood Service (Provincial Blood Coordinating Office)	Canada
Canadian Association of Emergency Physicians Sepsis Guidelines: the optimal management of severe sepsis in Canadian emergency departments (Green et	Canadian Association of Emergency Physicians	Canada

al., 2008)		
Recommendations for diagnosis and treatment of forms of shock of the IAG Shock of the DIVI. Part 2: Hypovolemic shock (Adams et al., 2005)	German Interdisciplinary Association for Intensive and Emergency Medicine	Germany
Fluid resuscitation in neonatal and pediatric hypovolemic shock: a Dutch Pediatric Society evidence-based clinical practice guideline (Boluyt et al., 2006)	Dutch Pediatric Society	Netherlands
Fluid Resuscitation (Orlando Regional Medical Center, 2005)	Department of Surgical Education, Orlando Regional Medical Center	USA
2. Hypertonic versus near isotonic crystalloid for fluid resuscitation in critically ill patients: First published: 2000		
Guidelines Title	Organisation producing guidelines	Country
Clinical effectiveness and cost-effectiveness of prehospital intravenous fluids in trauma patients. (Dretske, 2004)	HTA	UK
The use of pre-hospital intravenous fluid therapy in trauma 2004 (NICE, 2004)	NICE	UK
Spanish Consensus Statement on Alternatives to Allogeneic Blood Transfusion (Leal et al., 2006)	Spanish Panel on alternatives to allogeneic blood transfusions	Spain
Guidelines for cardiopulmonary resuscitation (CPR) and emergency cardiovascular care (ECC) of pediatric and neonatal patients: Pediatric advanced life support (Atkins et al., 2006)	American Heart Association	USA

3. Human albumin solution for resuscitation and volume expansion in critically ill patients: First published: 1998		
Guidelines Title	Organisation producing guidelines	Country
Clinical effectiveness and cost-effectiveness of prehospital intravenous fluids in trauma patients. (Dretske, 2004)	HTA	UK
The use of pre-hospital intravenous fluid therapy in trauma 2004 (NICE, 2004)	NICE	UK
Resuscitation in hypovolaemic shock	Patient UK	UK
Guidelines on the management of massive blood loss (Stainsby et al., 2006)	British Committee for Standards in Haematology	UK
Transfusion guidelines for neonates and older children (Boulton, 2004)	British Committee for Standards in Haematology	UK
Management of massive blood loss: a template guideline (Stainsby et al., 2000)	National Blood service Northern Zone	UK
The management of a child with a decreased conscious level. A nationally developed evidence-based guideline for hospital practitioners (The Paediatric Accident and Emergency Research Group)	The Paediatric Accident and Emergency Research Group	UK
Consensus statement on the treatment of septic shock (Spapen et al., 1999)	The Belgian Society of Internal Medicine	Belgium
Management of bleeding following major trauma: a European guideline (Spahn et al., 2007)	Multidisciplinary Task Force for Advanced Bleeding Care in Trauma	Europe

Recommendations for diagnosis and treatment of forms of shock of the IAG Shock of the DIVI. Part 2: Hypovolemic shock (Adams et al., 2005)	German Interdisciplinary Association for Intensive and Emergency Medicine	Germany
Diagnosis and therapy of sepsis. S2 guidelines of the German Sepsis Society and the German Interdisciplinary Society of Intensive Care and Emergency Medicine (DIVI) (Reinhart et al., 2006)	German Interdisciplinary Society of Intensive Care and Emergency Medicine (DIVI)	Germany
Albumin vs synthetic colloids for fluid resuscitation May 2007	Canadian Blood service	Canada
Canadian Association of Emergency Physicians Sepsis Guidelines: the optimal management of severe sepsis in Canadian emergency departments (Green et al., 2008)	Canadian Association of Emergency Physicians	Canada
Fluid resuscitation in neonatal and pediatric hypovolemic shock: a Dutch Pediatric Society evidence-based clinical practice guideline (Boluyt et al., 2006)	Dutch Pediatric Society	Netherlands
Evidence-based colloid use in the critically ill: American Thoracic Society Consensus Statement (American Thoracic Society, 2004)	American Thoracic Society	USA
Guidelines for cardiopulmonary resuscitation (CPR) and emergency cardiovascular care (ECC) of pediatric and neonatal patients: Pediatric advanced life support (Atkins et al., 2006)	American Heart Association	USA
Newer Guidelines for neonatal resuscitation (Bajpai et al., 2001). 2000 conference on Cardiopulmonary Resuscitation & Emergency	Adapted from recommendations of the pediatric working group of the International Liaison Committee on Resuscitation (ILCOR)	India

Cardiovascular Care.		
4. Colloids versus crystalloids for fluid resuscitation in critically ill patients: First published: 1998		
Guidelines Title	Organisation producing guidelines	Country
Clinical effectiveness and cost-effectiveness of prehospital intravenous fluids in trauma patients. (Dretske, 2004)	HTA	UK
Fluid resuscitation in prehospital trauma care: a consensus view (Revell et al., 2002)	Consensus review including Royal college of surgeons Scotland, Ambulance service Association, The UK Military Defence Forces, British Association for Immediate Care (BASICS) and London Helicopter Emergency Medical Service	UK
The use of pre-hospital intravenous fluid therapy in trauma 2004 (NICE, 2004)	NICE	UK
Guidelines on the management of massive blood loss (Stainsby et al., 2006)	British Committee for Standards in Haematology	UK
Management of massive blood loss: a template guideline (Stainsby et al., 2000)	National Blood service Northern Zone	UK
British Consensus Guidelines on Intravenous Fluid Therapy for Adult Surgical Patients GIFTASUP (Powell-Tuck et al., 2008)	Endorsed by: The British Association for Parenteral & Enteral Nutrition (BAPEN), Association for Clinical Biochemistry, Association of Surgeons of Great Britain & Ireland and Society of Academic and Research Surgery, the Renal Association and the Intensive Care Society.	UK

Management of acute upper and lower gastrointestinal bleeding (SIGN, 2008)	Scottish Intercollegiate Guidelines Network (SIGN)	UK
The management of a child with a decreased conscious level. A nationally developed evidence-based guideline for hospital practitioners (The Paediatric Accident and Emergency Research Group)	The Paediatric Accident and Emergency Research Group	UK
Management of bleeding following major trauma> a European guideline (Spahn et al., 2007)	Multidisciplinary Task Force for Advanced Bleeding Care in Trauma	Europe
Consensus statement on the treatment of septic shock (Spapen et al., 1999)	The Belgian Society of Internal Medicine	Belgium
Recommendations for diagnosis and treatment of forms of shock of the IAG Shock of the DIVI. Part 2: Hypovolemic shock (Adams et al., 2005)	German Interdisciplinary Association for Intensive and Emergency Medicine	Germany
Diagnosis and therapy of sepsis. S2 guidelines of the German Sepsis Society and the German Interdisciplinary Society of Intensive Care and Emergency Medicine (DIVI) (Reinhart et al., 2006)	German Interdisciplinary Society of Intensive Care and Emergency Medicine (DIVI)	Germany
Fluid resuscitation in neonatal and pediatric hypovolemic shock: a Dutch Pediatric Society evidence-based clinical practice guideline (Boluyt et al., 2006)	Dutch Pediatric society	Netherlands
Evidence-based colloid use in the critically ill: American Thoracic Society Consensus Statement (American Thoracic Society, 2004)	American Thoracic Society	USA
Guidelines for cardiopulmonary resuscitation (CPR) and emergency cardiovascular care (ECC) of pediatric and neonatal patients: Pediatric	American Heart Association	USA

advanced life support (Atkins et al., 2006)		
Surviving Sepsis Campaign: International guidelines for management of severe sepsis and septic shock: 2004 (Dellinger et al., 2004) Surviving Sepsis Campaign: International guidelines for management of severe sepsis and septic shock: 2008 (Dellinger et al., 2008)	The International Surviving Sepsis Campaign Guidelines Committee	International
Canadian Association of Emergency Physicians Sepsis Guidelines: the optimal management of severe sepsis in Canadian emergency departments (Green et al., 2008)	Canadian Association of Emergency Physicians	Canada
5. Timing and volume of fluid administration for patients with bleeding: First published: 2001		
Guidelines Title	Organisation producing guidelines	Country
Clinical effectiveness and cost-effectiveness of prehospital intravenous fluids in trauma patients. (Dretske, 2004)	HTA	UK
The use of pre-hospital intravenous fluid therapy in trauma 2004 (NICE, 2004)	NICE	UK
Management of bleeding following major trauma: a European guideline (Spahn et al., 2007)	Multidisciplinary Task Force for Advanced Bleeding Care in Trauma	Europe
The new 2005 resuscitation guidelines of the European Resuscitation Council (Soar et al., 2006, Wenzel et al., 2006)	European Resuscitation Council	Europe

Chapter 5: The prevention of traffic related injuries: systematic reviews of road safety interventions

Introduction

The published work that I present in this submission includes systematic reviews covering a diverse range of subjects and employing a variety of methodological techniques. In the previous chapter the focus was on the treatment of traumatic injuries whereas this chapter concerns the prevention of injury; in particular the prevention of road traffic related injuries. The chapter includes two road safety reviews (four papers) one on area-wide traffic calming for preventing traffic related injuries (Bunn et al., 2003a, Bunn et al., 2003b) and one on safety education for pedestrians (Duperrex et al., 2002a, Duperrex et al., 2002b). Full versions of the papers can be found in Appendix 2.

I have already detailed the methods involved in undertaking systematic reviews of randomised controlled trials, much of which is also pertinent to these reviews. However, as there were particular methodological challenges associated with these road safety reviews, for example, the difficulties associated with identifying relevant studies in road safety research and the issues involved in the quality assessment and analysis of non-randomised studies (NRS), this chapter includes a critical discussion of these issues. As before, I also include information about the drivers behind the review questions and assess the impact of the reviews on policy. In addition, as there are differences in the culture of road safety research and policy compared to other areas of health care, particularly clinical research, I discuss some of the barriers to the development of evidence-based road safety policies.

Drivers behind review question/s

The reviews in the previous chapter were concerned with the treatment of trauma and injury. However, the remit of the Injuries Group also covers prevention and many of the review topics that were prioritised were concerned with the prevention of road traffic injuries. Death and disability resulting from

road traffic crashes is a global problem. At present 1.27 million people die each year and some 10 million people sustain permanent disabilities in road traffic crashes (WHO, 2008). For people under 44 years, road traffic crashes are one of the top three causes of death and disablement worldwide (WHO, 2009), and are the leading cause of death for people under 50 in the European Union (European Transport Safety Council, 2003). This is of particular concern to low and middle income countries where road traffic deaths are predicted to rise by 80% as these countries become increasingly motorised (Peden et al., 2004b). The identification of effective strategies for the prevention of road traffic injuries is, therefore, of global health importance.

Many of the casualties are vulnerable road users, such as pedestrians and cyclists, and the risks associated with walking, cycling and motorcycling remain very high in relation to those of car travel (Peden et al., 2004b, Sonkin et al., 2006). In Great Britain in 2008 there were 28,481 pedestrian casualties, 16,297 casualties for pedal cyclists and 21,549 for motorcyclists (Department for Transport, 2009). Although rates in the UK have fallen over the last few decades (Peden et al., 2004b), that downward trend may be because there has been a significant reduction in rates of people walking and cycling rather than due to improvements in safety (Roberts, 1993). World-wide, 46% of road traffic deaths are among vulnerable road users (WHO, 2009) and in some low and middle income countries the proportion of road deaths involving vulnerable road users is as high as 80% (Peden et al., 2004b).

Despite these alarming statistics many road-safety interventions are still focused on improving safety for car occupants. Therefore, in deciding upon review questions the co-ordinating editor of the Injuries Group, Professor Ian Roberts, was concerned that we gave priority to reviews with the potential to benefit vulnerable road users. In addition, review questions were guided by a wish to challenge the orthodoxy that often favours politically safe but relatively minor interventions, such as safety-education, over less popular but potentially more important population-based environmental or legislative interventions. Traffic

calming, for example, has the potential to impact not just upon injuries but on public health more widely. The creation of safer streets may increase levels of active transport, such as walking and cycling, and thereby impact on rates of obesity, diabetes and cardiovascular disease (Roberts and Arnold, 2007). In addition, as child pedestrian injuries and deaths disproportionately impact upon those from lower socio economic groups (Edwards et al., 2008, Lyons et al., 2003, Roberts, 1996), the creation of safer environments has the potential to reduce inequalities in health (Jones et al., 2005b, Liabo et al., 2003, Lyons, 2005). As in the previous chapter, the identification of review questions was researcher led rather than driven by funders or policy makers.

Review methods

Question development

As in the previous chapter one of the reviews presented here (Duperrex et al., 2002a) included randomised controlled trials only. Although RCTs are generally considered to be the optimum study design for evaluating the effectiveness of an intervention there are instances where it is appropriate to include non-randomised studies (NRS) in a systematic review. One justification for their inclusion is where the intervention being evaluated cannot be, or is unlikely to be, evaluated in a randomised trial for either ethical or practical reasons. For example, Cochrane Injuries Group researchers in the US evaluated the effectiveness of fencing around swimming pools to prevent children drowning (Thompson and Rivara, 2000). They included NRS as it was highly unlikely that it would ever be considered ethical to evaluate such an intervention in a randomised trial.

In theory it should be possible to conduct an RCT of a road safety intervention such as traffic calming. In reality, however, such interventions are rarely, if ever, evaluated in RCTs. This is partly owing to the logistical and practical issues involved but also because RCTs have not been considered as an integral part of the road-safety research culture. Indeed road-safety interventions have typically been evaluated using uncontrolled before/after designs. Another justification

for including NRS in a review is to provide evidence of effects that cannot be adequately studied in RCTs such as long-term or rare outcomes. For example, the primary outcomes of interest in the traffic calming review were injuries and deaths. These outcomes are relatively rare and studies need to be conducted over a number of years for any changes to be detected. Owing to these factors the inclusion criteria for the traffic calming review was widened to include controlled studies as well as RCTs. To reduce the risk of bias we specified that the control had to be concurrent and not historical. Despite the fact that we broadened our inclusion criteria to include some NRS it has been suggested that those inclusion criteria are still too strict and that the review would have been more informative if we had included uncontrolled before/after studies (Lyons, 2005).

Identification of studies

In the previous chapter I looked at the potential problems associated with publication bias in systematic reviews and how these form the basis for the rationale for searching for all available RCTs. Much less methodological work has been done on the importance of comprehensive searches for other study designs and it is not clear if the same publication biases are relevant to NRS (Higgins, 2008). Indeed, it is possible that those studies that are hardest to find are the poorest quality and, therefore, the most biased (Egger et al., 2003, Higgins, 2008). However, although the benefits of comprehensive searching for NRS are not established, the approach we took for the traffic calming review was to apply standard methods for searching and to attempt to identify all available studies. This was not an easy undertaking. NRS are harder to find than randomised studies as there are no registers of NRS, and many NRS are poorly indexed which makes developing reliable methodological search filters more difficult. These problems are further compounded by the difficulties of searching for road safety studies. For example, many studies are unpublished, are written in languages other than English and road safety databases are not user friendly.

The largest and most widely used road-related database is TRANSPORT. This database contains approximately 600,000 records including grey literature. To facilitate the identification of relevant studies in TRANSPORT we used word frequency analysis in an attempt to devise electronic search strategies with high sensitivity and positive predictive value (PPV) (Wentz et al., 2001). Despite our best efforts we were unable to devise search strategies that combined acceptable sensitivity and PPV for the retrieval of controlled evaluation studies of road safety interventions. Even strategies with a relatively low sensitivity (59%) had a low PPV (10%). In contrast search strategies for controlled studies in medical databases can achieve high sensitivity and PPV because terms describing the study methodology are included among the indexing (descriptor) terms. Road safety databases, however, have a limited range of indexing terms describing the study methodology. For example, one of the included studies in the traffic calming review is a prospective controlled study (Mackie et al., 1990), but nowhere in the TRANSPORT database record for this report is there an indication that it is a report of a controlled study. Owing to this inability to devise a strategy combining sensitivity and PPV our searches generated nearly 13,000 records which had to be screened for eligibility.

There are also difficulties associated with screening the results of searches. It is generally fairly easy to identify RCTs from the title or abstract but this is not the case with NRS where it is often not clear in either the title or abstract what study design was used. This may necessitate a much greater percentage of papers being obtained so that the hard copy could be screened. Again this was exacerbated by the shortcomings of the TRANSPORT database where limitations in the quality of abstracts (many of which were non-existent) meant that for the traffic calming review I had to obtain the hard copies of nearly 600 records, many more than would normally be retrieved for a comparable review of RCTs of a health care intervention. Of those only 16 were eventually included in the review. The need to obtain such a high proportion of hard copies was both time consuming and costly.

Critically Appraising Studies

Critical appraisal is a vital part of a review of NRS as the potential biases are likely to be greater than with a review of RCTs. In particular, studies are prone to selection bias and confounding. Selection bias occurs when there are differences in the two groups under study and confounding occurs when selection bias gives rise to imbalances between intervention and control groups on prognostic factors (Higgins, 2008). Confounding can have two effects in a meta-analysis, it can shift the estimate of the intervention effect (systematic bias) and increase the variability of the observed effects thereby introducing excessive heterogeneity among studies (Deeks et al., 2003). At the time that the traffic calming review was first conducted the Cochrane handbook did not include any advice on reviews of non randomised studies, although a section on this has subsequently been added. The handbook now advises looking at four areas for quality assessment:

- Was there a comparison?
- How were the groups created?
- Which parts of the study were prospective?
- On which variables was comparability (between the groups) assessed?

In the traffic calming review we assessed the quality of the studies by extracting data on how well the intervention and control areas were matched and the length of the before and after data collection periods. In addition, because of the risk of contamination, that is the potential for the effect on drivers' behaviour of driving in traffic calmed areas to remain in nearby non-traffic calmed streets, we also recorded the distance between the intervention and control areas.

Data synthesis

Because of the increased potential for selection bias, residual confounding, and the greater risk of other biases through poor design and execution, meta-

analyses of NRS are associated with an increased risk of heterogeneity. There is no way of controlling for these biases in the analysis of primary studies and no established methods for assessing how, or the extent to which, these biases affect primary studies (Higgins, 2008). Therefore, reviewers need to be aware of this increased potential for heterogeneity and take this into account when planning the analysis and interpreting the results. For the traffic calming review we used a random effects meta-analysis to take the increased heterogeneity into account. The assumption of random effects both allows for the anticipated heterogeneity between effects across studies and provides robustness if the assumption that events follow Poisson distributions is violated through overdispersion.

In the meta-analyses discussed in the previous chapter, and in the pedestrian safety education review, we used 'raw' unadjusted data as this is considered appropriate for RCTs. However, in NRS where there may be expected to be greater residual confounding it may be appropriate to use adjusted estimates. Meta-analysis of adjusted estimates can be performed using the generic inverse-variance outcome in RevMan. When we first conducted the traffic calming review this facility was not available within RevMan and the analyses were conducted in Stata version 7.0 (Stata corporation, College Station, Texas 77845 USA). The results for each study were expressed as rate ratios (the rate ratio is the ratio of event rates post and pre intervention in the intervention area divided by the corresponding post to pre-intervention event ratio in the control area).

Results of impact analysis

Knowledge production

Publication and other methods of dissemination

The pedestrian safety education review was published in the Cochrane Library (Duperrex et al., 2002b) and in an edition of the BMJ that focused on road safety (Duperrex et al., 2002a). The traffic calming review was published in the Cochrane Library (Bunn et al., 2003a) and in the journal Injury Prevention (Bunn

et al., 2003b). The Cochrane version of the traffic calming review was updated in 2009 and its publication on the Cochrane Library was accompanied by a press release. The press release resulted in the study being picked up by a number of media outlets both nationally and internationally. The review was also accompanied by a podcast summarising the key points of the review; this podcast is freely available on the Cochrane Website⁵. Podcasting is a method of publishing audio files via the internet in a format that allows users to download them and listen to them at their own convenience. They are increasingly being used to enhance the communication of research and of the top 100 general medical and international journals eight offer regular podcasts (Wilson et al., 2009). One such is the Cochrane Library which has been offering podcasts of selected reviews since 2008. It has been suggested though that the technical quality and listening experience of journal podcasts is variable (Wilson et al., 2009).

Impact within research community

The citation analysis for the reviews in this chapter was first conducted in June 2009 and updated in April 2010; results can be seen in Figure 5.1. This shows a comparison of citations in Scopus, WoS and Google Scholar for each paper, this includes the CDSR and Injury Prevention papers on traffic calming and the CDSR and BMJ versions of the safety education review.

The version of the traffic calming review published in Injury Prevention (Bunn et al., 2003b) had 20 citations in Scopus, 16 in Web of Science (WoS) and 36 in Google Scholar and the version published in the CDSR had no citations in either Scopus or WoS but 37 in Google Scholar. The BMJ version of the review on safety education for pedestrians (Duperrex et al., 2002a) received 42 citations in Scopus, 33 in WoS and 78 in Google Scholar whilst the Cochrane version received none in Scopus or WoS and 44 in Google Scholar. The average number of citations a paper receives may be affected by a number of factors including the impact factor of the journal and the type of speciality. For example, previous

⁵ <http://www.cochrane.org/podcasts/issue/Issue%204%2C%20October%202009/100>

research has found that the average number of citations in the five years after publication for a paper published in the BMJ is 32.5 (Perneger, 2004). The Safety Education review was cited more frequently than the Traffic Calming review but that may be accounted for by the fact that it was published in the BMJ which is a general journal with a higher impact factor than Injury Prevention which is a more specialised journal. As in the previous chapter, the versions in the Cochrane Library were not cited as often as their counterparts published in other journals, but this is likely to be because this citation information is flawed (the reliability of citation information for Cochrane reviews is investigated further in Chapter 8). Data from Wiley on review downloads revealed that the traffic calming review was downloaded 812 times in 2008 (620/6,232) and 334 in 2009 (2666/6840) and the pedestrian safety education review was downloaded 68 times in 2008 (5105/6,232) and 214 times in 2009 (3693/6840).

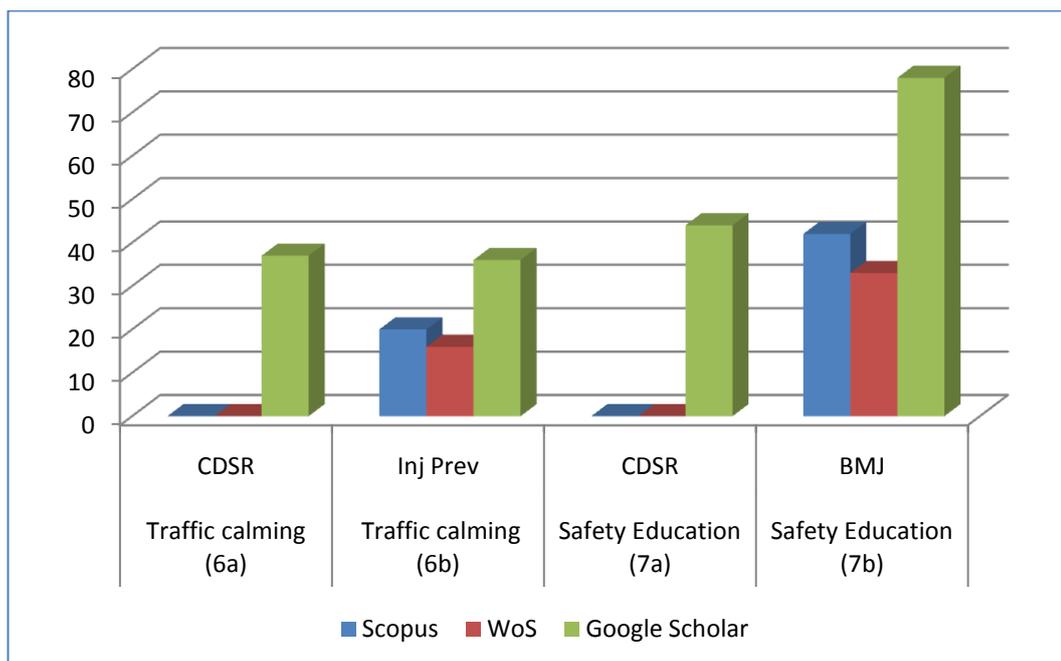


Figure 5.1 Results of citation analysis road safety reviews

Research targeting

In previous chapters I have described how the evidence-based practice movement, with its emphasis on rigour and insistence on high quality controlled evaluations of interventions, has become so dominant in health care research.

However, such attitudes, whilst beginning to filter into the road-safety discourse, are not as established or embedded in road-safety research. This is partly because the application of rigorous experimental research designs and the concept of systematically reviewing studies are less well established than in clinical research. Indeed, although a substantial amount of transport related research has been undertaken little of this has been directed towards ‘collating and evaluating evidence to inform policy or towards assessing effectiveness of policy’ (Terry, 2000). In addition, much research is market driven and aimed towards improving technical efficiency and safety of vehicle occupants (Terry, 2000) rather than improving the safety of pedestrians and cyclists.

Although it is hard to quantify how significant the role the Cochrane Injuries Group has been in promoting evidence-based road safety, there is no doubt it has been involved in championing the production and use of systematic reviews of road safety interventions, many of which are aimed at improving safety for vulnerable road users. In her essay entitled, ‘moving America towards evidence-based approaches to traffic safety’ Deborah Girasek writes:

‘Rigorous reviews of the state-of-our-art, such as those conducted by the Cochrane Injuries Group, must be consulted so that the science of the prevention strategies that are selected for promotion is beyond refute’ (Girasek, 2007).

However, any impact on research targeting appears to be attributable to the collective road-safety output of the Injuries Group⁶ rather than as a result of any individual review. Rather like the fluid reviews, much of their impact is because they form a coherent body of work that investigates many aspects of the same problem. Whilst this is not proven the body of work and its timing suggests that the Injuries Group has been instrumental in promoting the application of systematic review methods to road-safety research and contributed to a change in attitudes towards the conduct of road safety research.

⁶ There are 16 Injuries Group road-safety reviews published in the Cochrane Library in July 2010

Informing Policy Development

Levels and type of policy making

In the previous chapter I was able to present evidence of the way in which the reviews on fluid resuscitation strategies had influenced a range of clinical guidelines in both the UK and abroad. In road-safety, however, the development and use of guidelines is not as common and, therefore, I looked for the inclusion of the reviews in documents or reports by agencies (local, national and international) who might be involved in guiding and developing road-safety policy; for example, injury prevention reports by the World Health Organization (WHO). The results of this documentary review, including details about the report and organisation publishing it, are presented in Table 5.1.

Both reviews have been cited in a number of publications by UK organisations concerned with safety and injury prevention. These include charities, such as The Royal Society for the Prevention of Accidents and the National Children's Bureau, national and local Government departments such as Department of Health and Department of Transport, and evidence based resources such as 'best bets' which provides summaries of the best available evidence. Several of these reports were concerned with the role that interventions like safety education and traffic calming could play in reducing health inequalities.

Much of the identified impact of the reviews has been at an international level, for example in publications by the World Health Organization (WHO). The WHO aims to provide objective and reliable advice in the field of human health and to respond to the most pressing public health concerns around the world. Their outputs include guidelines, consensus reports, manuals and reviews which are aimed at both health workers and decision makers (Peden et al., 2004b).

Table 5.1 Evidence of reviews influence on policy development

Safety education for pedestrians (first published 2002)		
Guidelines/Report Title	Organisation	Country
A review of safety education: Principles for effective practice (McWhirter, 2008). This review then used to provide briefing paper on effective safety education for Department for Children, Schools and Families.	The Royal Society for the Prevention of Accidents (ROSPA)	UK
Traffic calming and childhood injury on the road (Liabo, 2004)	National Children's Bureau (produced from a more detailed document produced by what works for children)	UK
Can traffic calming measures achieve the children's fund objective of reducing inequalities in child health (Liabo et al., 2003) http://www.bestbets.org/bets/bet.php?id=218	Best Bets (best evidence topics). Concluded traffic calming had greater potential than safety education to reduce childhood inequalities	UK
Interventions to improve pedestrian skills in children : provisional statement (EuroSafe, 2007)	European Association for Injury Prevention and Safety promotion. Funded by European Union and aims to explore priority issues in injury prevention work	Europe
World report on road traffic injury prevention (Peden et al., 2004b)	World Health Organization	International
World report on child injury prevention (Peden et al., 2004a)	World Health Organization	International
Promoting evidence-informed decision making http://health-evidence.ca/articles/show/15392	Health Evidence Canada (funded by a number of research and government organisations to provide quality research evidence for decision makers)	Canada
Fact Sheet: Traffic education of children 4-12 years old (SWOV, 2009)	SWOV (Institute for Road Safety Research)	Netherlands
The effectiveness of road safety education (SWOV, 2009)	SWOV (Institute for Road Safety Research)	Netherlands

Traffic calming (first published 2003)		
Guidelines/Report Title	Organisation	Country
Traffic calming and childhood injury on the road (Liabo, 2004)	National Children's Bureau (produced from a more detailed document produced by what works for children)	UK
Tackling the wider social determinants of health and health inequalities: evidence from systematic reviews (Bambra et al., 2008)	Public Health Research Consortium on behalf of the Department of Health Policy Research Programme	UK
Joint Strategic Needs Assessment for Children in Kent (Kent PCT, 2008)	Report prepared for PCT and Local Authority	UK
Traffic calming, childhood pedestrian injury and inequalities and politics in: fifteenth seminar on behavioural research in road safety (Lyons, 2005)	Department for Transport (part of department of health accidental injury research initiative)	UK
How can injuries in children and older people be prevented? (Towner and Errington, 2004)	World Health Organization	International
World report on road traffic injury prevention (Peden et al., 2004b)	World Health Organization	International
World report on child injury prevention (Peden et al., 2004a)	World Health Organization	International
Youth and Road Safety (Toroyan and Peden, 2007)	World Health Organization	International
Prevention of chronic disease by means of diet and lifestyle changes (Willet et al., 2006)	The International Bank for Reconstruction and Development/The World Bank	International
Health Impacts of Transport a review (Kavanagh et al., 2005) (Kavanagh et al., 2005)	The Institute of Public Health in Ireland	Ireland
Strategies for Gain – the evidence on strategies to improve the health and well-being of Victorian children (Eagar, 2005)	Centre for Health Services Development	Australia
Promoting evidence-informed decision making http://health-evidence.ca/articles/show/18458	Health Evidence Canada (funded by a number of research and government organisations to provide quality research evidence for decision makers)	Canada
Urban Transportation, a Question of Health (Quebec Sante publique, 2006)	Agence de la Santé et des Services Sociaux de Montreal	Canada
Speed change management for New Zealand roads (Charlton, 2006)	Land Transport New Zealand (involved in allocating resources for transport)	New Zealand

Both reviews were cited in several reports from the WHO and in some instances contributed to the development of policy recommendations. For example, in the world report on traffic injury prevention (Peden et al., 2004b) the results of the traffic calming review feed into recommendation 5 (p162) on specific actions for the prevention of road traffic crashes. The recommendations in this report have been endorsed by member states through three United Nations General Assembly resolutions and one World Health Assembly resolution. The reviews had also been cited in a variety of practice and policy documents around the world including Europe, Australia, New Zealand and Canada.

The citation analysis presented here looks primarily at the number of citations and by whom they were cited. In general it does not consider how the research was presented and whether it was interpreted in the way originally intended. In a recent paper, using network analysis to explore how citation distortions can create unfounded authority for scientific ideas (Greenberg, 2009), Greenburg and colleagues use the phrase 'citation diversion'. By this they mean that the content of a paper is cited but that its meaning is altered in a manner that diverts its implications. An example of this can be seen in the interpretation of the findings of the Pedestrian Safety Education review. In our paper we wrote 'our review indicates that there is no reliable evidence supporting the effectiveness of pedestrian education for preventing injuries in children and inconsistent evidence that it might improve their behaviour, attitudes, and knowledge'. However, when cited in the WHO report on road traffic injury prevention (Peden et al., 2004b) it was used to support the statement that 'the use of information and publicity on their own is generally unsuccessful in reducing road traffic collisions'. This represents perhaps only a subtle distortion but none the less demonstrates how papers may not always be interpreted as the authors intended.

Nature of policy impact

The framework I used for this evaluation recognises that policy impact may take different forms and work on different levels. The section of the framework on

the nature of policy impact reflects this by including four categories based on the work of Weiss (Weiss, 1998): these are: 1) instrumental use, 2) mobilisation of support, 3) conceptual use and 4) redefining research practices and beliefs. There is, I believe evidence that the traffic calming and pedestrian safety education reviews have impacted to some extent at all of these levels. For example, as described above the traffic calming review had an instrumental effect on the development of road-safety policy by the WHO, policy that has the potential to impact upon many different countries. Although indirect impact is more difficult to measure, my own analysis, and informal interviews with co-authors and road safety experts, has suggested that the reviews have also influenced the policy process at a more conceptual level. There have been a number of shifts in the road safety paradigm including a change in emphasis from interventions that place the onus on the individual, such as education, towards more overarching population-based environmental or legislative interventions. The extent to which the reviews in this chapter have been instrumental to this process is hard to define. However, I believe that they, in conjunction with the whole road-safety output of the Injuries Group, have contributed to this process. In addition, the reviews, and the work of the Cochrane Injuries Group have impacted upon research thinking, practices and beliefs by promoting the use of systematic review methods and emphasising the value of using rigorous scientific methods for the evaluation of road safety interventions.

Impact on Practice

Although it was clear that the reviews had influenced policy, I found no real evidence that they had impacted upon road-safety practice. However, as the focus of my investigation was on policy rather than practice this is perhaps not surprising; there may well have been impacts that my evaluation methods were unable to identify. In the previous chapter we saw how much clinical practice is now directed by guidelines issued locally and nationally. The same is not true in road safety where many initiatives, such as traffic calming, may be implemented locally by local authorities rather than as a result of a central policy (Terry, 2000).

Terry argues that local authorities may have implemented traffic calming because it was seen as a relatively cheap way of improving noise abatement, air quality and pedestrian safety and one with a greater impact than some more expensive large scale construction projects. Although the traffic calming review was found to be on the international policy agenda there is no guarantee that this ensures it will impact on actual practice. In relation to other interventions traffic calming may be more expensive and logistically difficult to implement, particularly for low and middle income countries. It may be easier, for example, to put in place legislation rather than make whole scale changes to the road environment. Indeed, Cochrane reviews on the use of helmets for the prevention of injuries (Liu et al., 2008, Macpherson and Spinks, 2007, Thompson et al., 2000) have influenced the development of WHO policy on helmet laws (WHO, 2006); legislation that is now being widely implemented in Asia and Africa.

Barriers to evidence-based road safety

An unscientific approach to road safety?

In the previous sections I have demonstrated that systematic reviews of road-safety can influence policy. However, that impact may be diluted or obstructed by a number of barriers (Breen, 2004, Girasek, 2007, Roberts et al., 2006, Terry, 2000). Many of these are, of course, the universal barriers that hinder knowledge translation and research impact upon policy. For example, in her enlightenment model Weiss suggested that researchers may ignore research that does not support their own values (Weiss, 1976, Weiss, 1977) and there is evidence of this in road safety. In a Dutch study into the use of scientific road safety knowledge in decision making (Bax, 2006), the author concluded that policy makers primarily use knowledge if it fits in with an existing policy line, if they have asked for the information themselves, or if the knowledge is relatively new. Experts in the US have also commented on the use of road-safety interventions that are not based on scientific evidence (Hedlund, 2007) arguing

that there is a need to 'make intuition-based road-safety delivery unacceptable' (Hauer, 2005).

The selective use of research evidence in the development of road-safety policy has also been critiqued by researchers in the UK (Achara et al., 2001). In March 2000 the previous UK Government launched its road safety strategy which aimed to achieve a 40% reduction in road deaths and serious injuries (DETR., 2000). One of the initiatives included in this plan was the introduction of school-based driver education. This move to introduce school-based driver education was despite evidence from a previous systematic review that it may be associated with an increase in deaths and injuries in teenagers, possibly because it enables them to obtain a driving licence sooner than they otherwise might have (Vernick et al., 1999). An update of this review by a team of Cochrane Injuries Group reviewers also found no evidence of any benefit from school-based driver education and a suggestion that it may be linked to an increase in the proportion of teenagers involved in traffic crashes (Roberts and Kwan, 2001). In a paper in the *Lancet* (Achara et al., 2001) the Cochrane reviewers point out that the government had not taken previous randomised controlled trials into account but that instead the policy was supported by a poor quality uncontrolled before/after study. This research, from the British Institute of Traffic Education Research, looked at attitudes, knowledge and intended behaviour. These are surrogate outcomes and there is no reliable evidence that changes in self-reported knowledge and attitudes lead to changes in behaviour or reduced crash rates. Indeed, surrogate outcomes have been found to be misleading or dangerous (Grimes and Schulz, 2005). In addition, the survey had a poor response rate of 36%. This example reflects both the selective use of research and a lack of understanding on the part of policy-makers of the need for rigorous controlled evaluations in road-safety research.

The role of industry

In the same way that the pharmaceutical industry is able to influence healthcare policy, those with a vested financial interest appear to be able to influence road-safety policy (Roberts et al., 2006). For example, the alcohol industry has been involved in lobbying against the introduction of random breath tests and reductions in legal blood alcohol limits for driving (Rutherford, 2002, WHO, 2001). The WHO recommendations for safe blood alcohol levels for driving are 0.05 mg/ml whereas in the UK they remain at 0.08 mg/ml (ICAP, 2009). Although the UK safety legislation record is generally good in relation to other European countries, in this instance the UK compares unfavourably with many other European countries. This difference may, in part, be due to the power of the alcohol industry in the UK.

Another group with a commercial interest in road-safety policy is car manufacturers. Car manufacturers have become involved in road safety through the creation of the Global Road Safety Partnership (GRSP), an alliance between business, non-governmental organizations, and governments that aims to improve road safety. However, there is a concern that such involvement is inappropriate as 'companies may only support their own commercial interests rather than implementing evidence-based interventions' (Mohan and Roberts, 2001). In a word frequency analysis, comparing road safety documents produced by the GRSP and the WHO, researchers found important differences in the emphasis given to different road safety strategies. The GRSP documents were found to be more likely to emphasise driver training and safety education over speed reduction despite a lack of strong scientific evidence for such a position (Roberts et al., 2006). This discrepancy may be because the GRSP are not scientists and are unable to critically evaluate information in the way that researchers at the WHO can, or it may be that they are still stuck in the old paradigm that sees road-safety as the responsibility of the individual. However, it is not hard to see how car manufacturers may prefer to promote educational strategies as these are less likely to impact upon them commercially. In addition,

education may be attractive for policy makers who see it as politically safe and non controversial (Stone, 1989).

The role of lobby groups

However, industry and manufacturers are not the only powerful interest groups able to influence road safety policy. Indeed car owners and the motoring public have a substantial impact in this arena. Francis Terry argues that the government is reluctant to instigate road safety policy that is likely to invoke negative public reaction. For example, although the benefits of wearing seat belts were conclusively demonstrated during tests in the early 1960s, ministers did not feel that they could introduce a seat belt law until 50% of the population were wearing them (Terry, 2000). The first attempt to get seat belt legislation passed was in 1973 and there were many other unsuccessful attempts until the bill was finally passed in the Transport Act of 1981 (Transport Bill, 1981). As seatbelts prevent around 25,000 serious injuries annually the delay had serious health and economic implications (Breen, 2004). Similarly both previous and current UK governments have been reluctant to give priority or funding to 20mph zones despite experiments showing a dramatic reduction (typically 60%) in the number of pedestrians killed or seriously injured. Although tests were first conducted in 1990 it was at least five years before the Department of Transport would consider giving any priority to spending on 20 mph zones more generally (Terry, 2000). Public acceptability may be a more important factor in determining road-related policy than the evidence.

Indeed it would appear that certain vociferous minority groups have had a disproportionate effect on road safety policy, such as in the case of the use of speed cameras (Breen, 2004, Pilkington, 2003). Despite surveys demonstrating widespread public support and the support of the select committee on transport (Select Committee on Transport, 2002), the use of speed cameras is still being questioned in the media (Breen, 2004). This is largely as a result of campaigns against speed cameras waged by groups such as The Association of British Drivers (ABD) and Motorists Against Detection (MOD) (Breen, 2004, Pilkington,

2003). Despite relatively small membership (ABD only has 1000 members) they have managed to get disproportionate TV and media coverage. Pilkington argues that 'such minority radical groups are the product of a dominant car culture' and that they reject population measures to increase road safety such as speed cameras and traffic calming. He argues that lobbying by pro-motorist groups has influenced government policy on speed cameras; for example, the decision by the previous Government that all speed cameras would have to be highly visible (bright yellow) despite there being no evidence on how increased visibility might impact on effectiveness.

A large part of the socio-economic burden caused by road injuries is borne by the health sector, and health care professionals have been shown to influence the policy debate. For example, although seat-belt legislation was slow in coming Breen argues that lobbying by health care groups such as the British Medical Association (BMA), the Royal College of Surgeons, The British Paediatric Association, and the Child Prevention Committee was key in getting seat-belt legislation finally passed (Breen, 2004). However, although health care groups can have some influence in the policy arena they are often up against powerful opposition from motoring groups and those with a vested commercial interest. The BMA, for example has campaigned unsuccessfully for the implementation of random breath testing since the 1980s (Peden et al., 2004b).

Individualism versus collectivism

There is another barrier to road safety and one that is at the heart of much of the protest against safety measures, such as traffic calming and speed cameras. In western cultures, with their emphasis on individualism over collectivism, people may see the cause and solution of problems such as road traffic deaths and injuries in terms of individual behaviour rather than a collective response (Girasek, 2007). Breen, for example, suggests that it was politicians opposed to what they saw as the 'nanny state' who blocked the introduction of seat-belt legislation (Breen, 2004). One of the reasons that education may be so popular as an intervention, despite little evidence of effectiveness, is that it 'appeals to

values of personal freedom and individual responsibility' (Girasek, 2007).

Commentators have argued that child pedestrian injuries have historically been seen as the responsibility of the child and that the solution lies with them rather than the environment (Stone, 1989).

Conclusion

In this chapter I have demonstrated how two road-safety systematic reviews have impacted on policy both directly and indirectly. The reviews have been particularly influential on international policy which is in many ways understandable as road related deaths and injuries are a global phenomenon that is a rapidly increasing problem in low and middle income countries. In addition, traffic calming may be seen as important as improvements to the road environment may have far reaching public health and environmental impacts that go beyond the prevention of traffic injuries (Roberts and Arnold, 2007). However, although these reviews appear to be on the policy agenda that, in itself, is no guarantee that they will be taken up and affect practice. Some reviews may ultimately be more instrumental because they evaluate an intervention that is simpler to implement. Indeed, one of the attractions of safety education is that it is relatively simple and inexpensive to implement.

Although there is evidence of direct impact upon policy, much of the influence of the reviews may be more in line with Weiss's ideas about the enlightenment role of research (Weiss, 1976) with the reviews influencing ideas of decision makers in a way that may eventually filter down to policy. For example, the reviews may have challenged the over-reliance on politically safe interventions such as safety education and highlighted the need for more comprehensive population-based strategies. The reviews may also have played a part in changing the thinking of road-safety researchers by emphasising the need for rigorous research designs and promoting the use of systematic reviews in injury prevention.

It was, however, difficult to consider the impact of one review in isolation. The Cochrane Injuries Group has now produced many road safety systematic reviews

and it appears that much of their influence may be as a collective body of work on road-safety. These systematic reviews have highlighted the need for greater rigour in road-safety research and challenged and changed the existing research and policy culture in road-safety. In addition, the focus on 'what works' has drawn attention to discrepancies between the evidence and existing road-safety policy and challenged the over-reliance on politically expedient but ineffective interventions.

What is clear is that changing policy to increase road safety is often a long and arduous process. Industry, motoring lobby groups and the general car-dependent public are all influential in the policy process. In such an arena the production and publication of research may not be enough to change policy. Researchers may need to be more vocal and pro-active in lobbying for changes to road-safety policy. There also needs to be a recognition that to reduce the theory/practice gap requires 'long-term commitment among researchers, practitioners and policy makers' (Sussman et al., 2006).

Chapter 6: Identifying barriers and facilitators to fall prevention interventions: systematic review of qualitative research

Introduction

In the previous two chapters I focused on reviews that evaluated the effectiveness of interventions. However, systematic review methodology can also be used to ask different types of questions; for example, questions that concern contextual factors such as beliefs and attitudes, or that examine the processes of implementing successful interventions. Review questions such as these are not best answered using quantitative studies, such as randomised controlled trials, but instead may require qualitative approaches. In this chapter I present a review which, because it was exploring factors around the implementation of interventions, focused on qualitative studies. The review was designed to investigate the barriers and facilitators to the successful delivery of falls prevention interventions (Bunn et al., 2008a).

Although reviews of qualitative studies in general adhere to the same structure and methodological approaches as reviews of effectiveness, there are a number of key differences. This chapter, therefore, includes a critical discussion of the methodological issues associated with systematic reviews of qualitative studies. As the place of qualitative research in systematic reviews has been more controversial and is less well-established I include a critical examination of the role of qualitative research in evidence based practice and systematic reviews. As before, I discuss the drivers behind the review question and assess the impact of the review on policy.

Drivers behind review question

The review was conducted as part of a Department of Health funded project exploring barriers and facilitators to fall prevention in older people. Although previous systematic reviews had evaluated the effectiveness of fall prevention interventions (Gillespie et al., 2001, Gillespie et al., 2009) they had also identified

that compliance with these interventions was not always good. There was, therefore, a need to examine more fully issues around uptake and participation. The main aim of the review was to appraise the existing research evidence of barriers and facilitators which influence compliance with, and adherence to, interventions to reduce falls and fall-related fractures in older people.

The role of qualitative research in systematic reviews

Much of the discourse around evidence-based practice in health care and public health is based on the assumption that when we talk about evidence we mean quantitative research, in particular randomised controlled trials (Rycroft-Malone et al., 2004b). Systematic reviewers, for example, have prioritised the randomised controlled trial over any other type of evidence. One of the reasons for this is that evidence-based medicine had its roots in the discipline of epidemiology (Swanson, 2001) and epidemiologists predominately use quantitative methods, with an underlying positivistic paradigm (Jack, 2006). Proponents of such approaches have perceived qualitative research as less rigorous and, in traditional hierarchies of research, qualitative research has often been absent or poorly ranked (Upshur, 2001). Black argues that health services research is only considered 'real and serious' when it uses quantitative approaches (Black, 1994). It should be noted though that many of these research hierarchies were based primarily on methods for addressing questions of effectiveness or causation, for which qualitative approaches are not designed. Many hierarchies now include qualitative research and the emphasis has shifted to choosing the appropriate methodological approach for the research question (Petticrew and Roberts, 2003).

Systematic reviewers in the fields of health and social care have begun to recognise that for many interventions, in particular complex interventions, we need to go beyond merely addressing questions of 'what works'. Instead we need to seek an understanding of the influence of contextual factors and the social, political and economic determinants of health and health-related behaviours (Hills, 2000). To gain a greater understanding of these factors it is

increasingly recognised that researchers may need to 'conduct research within a naturalistic paradigm using qualitative research approaches (Jack, 2006).

Qualitative research can help us go beyond questions purely about what works and help us to explore questions such as: why does an intervention work and for whom does it work and in what circumstances (Barbour, 2000)? It can help us plan and interpret quantitative research, allow us to explore contradictions in the evidence, access sensitive or hard to reach settings where RCTs would be unethical or unfeasible, explore barriers and facilitators to the successful implementation of interventions and programmes and allow for the development of theory.

Qualitative research, therefore, has a number of roles within systematic reviews, both in conjunction with quantitative research and in its own right. A chapter on the role of qualitative research within Cochrane reviews has recently been added to the Cochrane Handbook (Noyes et al., 2008). The authors of this chapter list a number of ways in which qualitative research might contribute to reviews of effectiveness. This includes:

- Helping to define the question
- Enhancing reviews by synthesizing evidence from qualitative research identified whilst looking for evidence of effectiveness
- Extending reviews by undertaking a search to specifically seek out evidence from qualitative studies to address questions directly related to the effectiveness review
- Supplementing reviews by synthesizing qualitative evidence within a stand-alone, but complementary, qualitative review to address questions on aspects other than effectiveness

The above are all concerned with how qualitative research might enrich traditional reviews of quantitative studies. Many qualitative systematic reviews, however, are not associated with particular evaluations of effectiveness although by gaining a greater understanding of human behaviour or beliefs they may further the development of appropriate interventions.

However, although the use of qualitative research in systematic reviews has become relatively commonplace there are those who have argued that it is inappropriate to attempt to synthesise qualitative research. The basis for such criticisms is that systematic reviews are reductionist, that you cannot make any form of generalisation from qualitative research, and that the uniqueness and complexities of qualitative studies resist 'summing up' (Light, 1984). Dixon Woods is a qualitative researcher who has done much work around systematic reviews. Whilst she is an advocate of systematic reviews of qualitative research she does acknowledge that a basic tension exists between qualitative research, which relies on interpretation and reflection, and the more rigid structured processes associated with systematic reviews (Dixon-Woods and Fitzpatrick, 2001). Despite such concerns about the place for systematic reviewing in qualitative research it has been suggested that a failure to synthesise means qualitative researchers are working in isolation and are in danger of reinventing the wheel (Sandelowski et al., 1997). As with quantitative research there is a danger of duplication and there may be benefits to situating qualitative research within larger programmes of work so that it forms part of the whole research body and contributes to the development of policy and practice in a meaningful way.

Conceptual approaches/methodologies

The methods for systematic review of quantitative studies are well established and widely acknowledged. Although various guidelines exist for the conduct of systematic reviews, such as those produced by the Cochrane Collaboration (Higgins, 2008), the NHS Centre for Reviews and Dissemination (CRD, 2009) and the National Institute for Health and Clinical Excellence (NICE, 2006), these manuals by and large propose similar methods. The conceptual basis of these methods is that the adoption of a structured approach is essential to minimise bias and error. The need for a well-defined question, comprehensive searches and critical appraisal of included studies are well accepted, and methods for synthesis widely agreed upon. In contrast such concord does not exist amongst those synthesising qualitative research. Indeed, there are a variety of methods

for synthesis guided by a number of different conceptual approaches. Methods can broadly be grouped into those designed for use with quantitative and qualitative data, such as thematic analysis, realist synthesis, and Bayesian meta-analysis, and those such as meta-ethnography which are intended for the synthesis of qualitative data only. These and other approaches are summarised in Table 6.1. This table combines previous overviews (Barnett-Page and Thomas, 2009, Dixon-Woods et al., 2006a) with my own observations.

Qualitative systematic reviews have been conceptualised as ‘interpretive’ as distinct from the more ‘integrative’ approach taken in quantitative reviews (Noblit and Hare, 1988). Integrative reviews primarily involve the synthesis of quantitative studies, are concerned with combining or amalgamating data and may involve techniques such as meta-analysis. Interpretive reviews, on the other hand, are suitable for synthesising qualitative research and involve both induction and interpretation. Synthesis is achieved through subsuming the concepts identified in the primary studies into a higher order theoretical structure. Dixon-Woods expands on this by saying that ‘the defining characteristic of an interpretive synthesis is its concern with the development of concepts, and with the development and specification of theories that integrate these concepts... the main concern is not aggregations of data but theory’ (Dixon-Woods et al., 2005). However, despite their differences, there is overlap between the methods with integrative reviews involving elements of interpretation and interpretive syntheses involving elements of aggregation of data (Dixon-Woods et al., 2005). Indeed, not all qualitative reviews fall into the interpretive paradigm with some taking a more quantitative integrative approach than others. For example, Dixon-Woods argues that some methods such as narrative summary, meta-ethnography and meta-synthesis lie more at the interpretive end of the spectrum whilst others such as content analysis and comparative analysis are more integrative. In addition, some methods are designed purely for the synthesis of qualitative research whereas others are intended to be used for the synthesis of quantitative and qualitative research, often integrated into the same review.

In a critical review of methods for the synthesis of qualitative research Barnett-Page and Thomas suggest that many of the differences in approaches are explained by epistemological positions; in particular whether they can be categorised as falling into 'realist' or 'idealist' epistemologies (Barnett-Page and Thomas, 2009). They used categories based on those of Spencer and colleagues (Spencer et al., 2003) which are organised into the following spectrum:

- Subjective idealist: there is no shared reality independent of multiple alternative human constructions
- Objective idealism: there is a world of collectively shared understandings
- Critical realism: knowledge of reality is mediated by our perceptions and beliefs
- Scientific realism: it is possible for knowledge to approximate closely an external reality
- Naive realism: reality exists independently of human constructions and can be known directly

Barnett-Page and Thomas classify different methods by where they fit on this spectrum and argue that iterative methods such as critical interpretive synthesis and meta-study are at the 'subjective idealist' end of the spectrum. For example when describing meta-study Patterson argues that as primary studies are themselves constructs then syntheses are constructs of constructs (Patterson et al., 2001) and that there is no such thing as an 'absolute truth'. Meta-ethnographic approaches and grounded theory they categorise as methods informed by objective idealism. In contrast thematic synthesis and framework synthesis, which are designed to directly inform policy and practice, take more realist approaches to synthesis (Barnett-Page and Thomas, 2009).

Whilst there may be some justification for such a diversity of methodological approaches there is also a real danger of duplication. Dixon-Woods warns that 'in seeking methodological developments, existing methods will be overlooked, and there will be a proliferation of methods that risk re-inventing the wheel

(Dixon-Woods et al., 2005). In addition, the profusion of terms can be confusing and may mask some of the basic similarities in methodological approaches (Barnett-Page and Thomas, 2009). Indeed, many reviews of qualitative studies, such as the one presented in this chapter, do not strictly adhere to one methodology but rather incorporate aspects from several.

Review methods

Question Development

One of the first things we teach those new to systematic reviewing is how to develop a well defined question. This generally involves specifying the type of intervention, participants, outcomes and study designs we intend to include. We stress the importance of getting this part of the review process right before we begin our searches as post-hoc changes to the question are discouraged because they are considered to increase the risk of bias. However, a number of researchers have suggested that with qualitative reviews a more iterative approach to question development may be more appropriate. Dixon-Woods, for example, argues that there are 'strong epistemological reasons for adopting a position closer to that of primary qualitative research in formulating a question for a review that includes qualitative research (Dixon-Woods et al., 2006a). Unlike a purely quantitative review the definition of what the phenomenon of interest is may be generated through the process of searching for and identifying papers. Indeed, it has been suggested that the question should be treated as 'a compass rather than an anchor, and as something that is not finally settled until the end of the review' (Eakin and Mykhalovskiy, 2003). My own experiences would support the view that more iterative processes are appropriate for systematic reviews of qualitative studies. However, it is still important to have a well defined question before significant data extraction, analysis and interpretation of results is undertaken.

Identification of studies

There are a number of factors that can make the identification of qualitative studies problematic. One of these is the lack of reliable search filters that filter out those studies using the research methods of interest. Such methodological filters are 'predetermined search strategies that use terms related to research design to identify all those studies using the research methods of interest to the reviewer' (Evans, 2002); and methodological filters have been used to great effect for the identification of RCTs. The development of search filters for qualitative studies is difficult partly because at present most databases do not have reliable indexing terms for qualitative research (Barroso et al., 2003, Dixon-Woods et al., 2001, Evans, 2002) and formal indexing terms may perform poorly for identifying relevant studies (McKibbin et al., 2006). For example the descriptor for qualitative research in Medline is of limited value as it was only introduced in 2003. Before that it came under the heading of 'nursing-methodology research'. In contrast CINAHL introduced the term 'Qualitative Studies' in 1988 which reflects a greater interest in qualitative research for nursing researchers. In addition developing search filters is complicated by the fact that qualitative research encompasses many different research methods and approaches, such as ethnography, phenomenology and grounded theory (Evans, 2002), and that the term 'qualitative' is used and understood in a variety of ways (Dixon-Woods et al., 2006a, Grant, 2004).

Another reason that qualitative research is hard to identify is that qualitative research reports are renowned for producing poor descriptors of the research methods, with neither the title nor abstract stating the methods used (Evans, 2002). For example, the title may be more likely to reflect some aspect of the findings rather than the methods used in the study (Evans, 2002). In comparison the methods would almost always be reflected in the title of an RCT. Furthermore, many abstracts for qualitative studies may be missing. In an evaluation of different search strategies Shaw and colleagues found that 23 per cent of relevant qualitative studies on breastfeeding did not have an abstract in the database (Shaw et al., 2004).

However, there has been significant progress in identifying and developing methodological filters for qualitative studies. The Cochrane handbook (Noyes et al., 2008) points out that there are now empirically-tested methodological filters for MEDLINE (Wong et al., 2004), CINAHL (Wilczynski et al., 2007), Psychinfo (McKibbon et al., 2006) and EMBASE (Walters et al., 2006). Research suggests that of the major databases CINAHL may be the best source of qualitative studies as the methodological indexing terms they use will identify a greater number of qualitative studies than in other databases such as medline or Psychinfo (Evans, 2002, Flemming and Briggs, 2007).

So far the discussion in this section has been based on the assumption that we should take a similar approach to searching for reviews of qualitative studies as we would in a quantitative review; i.e. we should attempt to identify all the available research and our search strategies should be transparent and reproducible. Such comprehensive search strategies are one of the cornerstones of systematic reviewing and considered crucial for minimising bias. This is the approach that we adopted for the paper included in this chapter. However, it has been suggested that the justification for exhaustive search strategies is not as great in systematic reviews of qualitative research and that it may be appropriate to take a more purposive sampling approach, 'aiming to provide a holistic interpretation of a phenomenon, where the extent of searching is driven by the need to reach theoretical saturation and the identification of the disconfirming case may be more appropriate' (Dixon-Woods et al., 2006b). After all qualitative studies themselves do not claim to be representative or generalisable.

Critically Appraising Studies

Traditionally a key component of systematic reviews of quantitative studies, and one that distinguishes them from conventional literature reviews, is the critical appraisal of included studies. As discussed in Chapter 4, quality assessment is considered essential as it enables us to make judgments about the potential level of bias in a review. In comparison, the role of critical appraisal in systematic

reviews of qualitative research is less well established, with much debate about whether reviewers should critically appraise qualitative research at all. Arguments against critical appraisal are generally constructed around the uniqueness of qualitative research. As Mays and Pope point out extreme relativists may take the position that all research is unique and valid in its own terms (Mays and Pope, 2000). In addition, it has been argued that quality assessment stifles the interpretive and creative aspects of qualitative research (Schwandt, 1996) and that qualitative research represents a distinctive paradigm and as such should not be judged by conventional measures of validity, generalisability and reliability. However, it has been suggested that such a position would find little support among health researchers (Murphy et al., 1998). Despite such concerns qualitative researchers involved in evidence synthesis generally agree that some form of quality assessment is important (Campbell et al., 2003, Dixon-Woods and Fitzpatrick, 2001, Noyes et al., 2008, Popay et al., 1998), as poor quality studies may distort findings and lead to difficulties in interpretation.

Although there is now a growing consensus that quality assessment of qualitative studies is important there is, as yet, no agreement about the form such appraisals should take. Qualitative researchers may have very different disciplinary, philosophical, and theoretical ideas and researchers from different paradigms may hold very diverse opinions about the characteristics that define a good qualitative study (Dixon-Woods et al., 2004, Sandelowski et al., 1997). Dixon-Woods points out that a key problem is that 'much of the work on developing appraisal criteria for qualitative research has been the tendency to treat it as a unified field, both at the level of data collection (such as focus groups) and at the level of the methodological approach (such as grounded theory)' (Dixon-Woods et al., 2004). In qualitative research the methodological approach is generally closely linked to the theoretical perspective adopted by the researcher, and reflecting such differences in quality assessment can be a considerable challenge. What might be considered essential for one methodology or perspective may not be deemed important for another.

Quality assessment checklists and guidelines

Even if a reviewer accepts the position that critical appraisal is a necessary part of the review process they are still faced with the question of how they go about quality assessment and whether or not they should use a formal framework or checklist as a guide. Reviewers of quantitative studies also face these dilemmas but despite a proliferation of quantitative checklists there is good agreement on what the important quality domains are; for example randomisation, allocation concealment, blinding and attrition rates. There are now over 100 sets of quality criteria for qualitative research but many of these take 'non-reconcilable positions' on a number of issues (Dixon-Woods et al., 2004) and fail to distinguish between different study designs or theoretical approaches (Dixon-Woods et al., 2006a). No single set of guidelines is definitive (Mays and Pope, 2000). In response to this dilemma a project by the UK National Centre for Social Research attempted to further a consensus on quality criteria by producing a framework for assessing qualitative research. The development of this framework was guided by a number of existing frameworks, interviews with those active in the area and a workshop (Spencer et al., 2003). This framework is one that is recommended in the Cochrane handbook and is the one that we used for the review presented in this chapter. However, like Dixon-woods, we found the framework lengthy and unwieldy (Dixon-Woods et al., 2004) and had to shorten it to make it fit for our purpose. Even the shortened version remains lengthy and time-consuming to complete.

Whilst all quality assessment is to some extent subjective, even for quantitative studies, my experience of assessing qualitative studies is that subjectivity is even more of a concern. We overcame this to some extent by two reviewers independently assessing quality and then meeting to discuss the criteria and agree final decisions. It may be though that quality assessment of qualitative studies will always be inherently subjective. Dixon-Woods argues that some aspects of qualitative research, particularly those relating to insight and interpretation, are difficult to appraise and may remain very difficult to measure except through the subjective judgement of experienced qualitative researchers

(Dixon-Woods et al., 2004, Dixon-Woods et al., 2007). In a comparison of three methods for quality appraisal (Dixon-Woods et al., 2007) researchers found that structured approaches, such as the use of frameworks, did have some advantages but that they did not result in higher agreement between reviewers than unprompted judgment. However, the use of judgement alone may require a reviewer to have a greater depth of expertise than a more structured approach. As a reviewer with a quantitative rather than qualitative background I found the checklist a useful guide despite its flaws.

How should quality information be incorporated into a review?

In addition to concerns about the objectivity of quality assessment a reviewer must also consider which quality domains to include and if they should be weighted in any way. For example, with RCTs allocation concealment is regarded as a key marker of quality because of empirical evidence that poor allocation concealment can lead to distortion of treatment effects (Schulz et al., 1995). As with the appraisal of quantitative studies, there is also a need to distinguish between those criteria which are about quality of process and analysis and those which concern transparency of reporting (Dixon-Woods et al., 2004). In recent years the advent of guidelines for reporting RCTs , diagnostic and observational studies (CONSORT (Begg et al., 1996), STARD (Bossuyt et al., 2003) and STROBE (Vandenbroucke et al., 2007)) has meant that reports of quantitative studies are more likely to be reported in a standardised way. In addition, I would argue that systematic reviews with their emphasis on quality assessment of individual studies have highlighted the need for improved quality and reporting in primary studies. In contrast reports of qualitative studies tend to vary in writing styles and publication formats with detailed descriptions of methods and data analysis procedures typically lacking (Harden et al., 2004).

Finally, and very importantly, the reviewer is faced with the question of how to use information on quality within the review. For reviews of quantitative studies that include a meta-analysis we can explore the impact of quality by performing sensitivity analyses. Such a statistical option is not available to us when

reviewing qualitative studies. Alternatives may be to use quality as a criteria for excluding studies (Campbell et al., 2003, Estabrooks et al., 1994), or to grade the studies so that those considered better quality are given more weight in the analysis than those of poorer quality (Attree, 2004). Researchers are undertaking more methodological work in this area and in a paper describing the methods used for thematic synthesis of qualitative studies (Thomas and Harden, 2008) the authors describe how they investigated the contribution studies made to the synthesis in light of their quality. They found that better studies appeared to contribute most to the synthesis. In the review included in this chapter we did not undertake sensitivity analyses although this is something that I would consider for future review of qualitative studies. Such strategies are, however, complicated by the lack of consensus on what constitutes a 'good' quality qualitative study.

Data synthesis

Methods for synthesis

As discussed earlier in the chapter there has been a proliferation of methodological approaches to the systematic review of qualitative studies. A number of commentators have attempted to describe the similarities and differences between methods for synthesis of data (Barnett-Page and Thomas, 2009, Dixon-Woods et al., 2005). Clearly some methods reflect a more quantitative logic than others with the main concern being to summarise, map or catalogue the findings often through the identification of common themes and concepts. In meta-summary, for example, Sandelowski describes an aggregative method that involves summarising rather than 'transforming' findings and that even allows for the calculation of 'effect sizes' (Sandelowski and Barroso, 2007). In contrast other methods are much more concerned with the development of theory. Dixon-Woods describes her method of critical interpretive synthesis and says that the 'defining characteristic of an interpretive synthesis is its concern with the development of concepts, and with the development and specification of theories that integrate those concepts' (Dixon-Woods et al., 2006a). The main

product is not aggregations of data but theory. Similarly Britten talks about the development of first, second and third order constructs suggesting that systematic reviews may allow the development of new higher order theoretical ideas (Britten et al., 2002). Some methods, such as meta-ethnography and meta-synthesis are designed for use with qualitative studies only, whereas the instigators of methods such as critical interpretive synthesis and realist synthesis claim that they can be used to synthesise all types of evidence. It is also important to remember that the interpretative emphasis of qualitative reviews may mean that, unlike quantitative reviews which are in theory reproducible, reviewers will not necessarily come to the same conclusions (Dixon-Woods et al., 2006a).

It has been suggested that variations in the extent to which methods attempt to 'go beyond' the primary study is influenced by the epistemological position where those from a more idealist perspective 'seek to push beyond the original data to a fresh interpretation of the phenomena under review'. In contrast those taking a more 'realist' perspective may focus on describing and summarising their data. Indeed, as a reviewer with a predominantly quantitative background the approach I took in our review could be described as more 'realist' than 'idealist'.

Sampling

In the earlier critique of issues around the identification of studies I noted that some researchers do not feel that it is necessary to identify all available studies but that a purposive approach, whereby the reviewer chooses papers they think are relevant, could be appropriate. The issue of sampling rather than a comprehensive approach is also relevant when we talk about analysis. Reviews of qualitative studies may identify many studies, and assessment and analysis may be time consuming and costly. It has been suggested that it is not necessary to include all identified studies but that instead the reviewer could adopt an approach similar to that of sampling for primary qualitative research (Dixon-Woods et al., 2006a, Schreiber et al., 1997). Such an approach, however, would

rely on the reviewer deciding which papers to include, and this could be problematic for a number of reasons. As Dixon-Woods acknowledges, such an approach could leave the review open to the accusation that it is no longer objective or 'systematic'. Indeed, one of the major criticisms of the traditional literature review was the subjective way in which studies were chosen for inclusion. In addition, there is concern that sampling would mean missing relevant studies which could have usefully contributed to the phenomenon under study (Jensen and Allen, 1996). Our approach for this review was to include all relevant studies in the analysis.

Results of impact analysis

Knowledge production

Publication and other methods of dissemination

The review was published in the journal *Ageing and Society* in 2008 (Bunn et al., 2008a) and was also disseminated via conference papers and seminars.

Impact within research community

When the citation analysis for the paper was carried out on the 18th August 2009 there were no citations in Scopus or Web of Science and only one in Google Scholar. The latter was a citation in a thesis from the University of Hong Kong. This lack of citations may be explained by the fact that the paper was only published in 2008 and had, therefore, had little time to make an impact. It has been found that on average it takes three years for a paper to be cited (Grant and Lewison, 1997) and the median time lag between the publication of a paper and its inclusion in published guidelines is eight years (Grant et al., 2000). Indeed, when I repeated the citation analysis in April 2010 there was one citation in Scopus and Web Of Science, a Danish paper looking at barriers to participation in a hospital falls clinic, and five citations in Google Scholar (one of which was a self-citation).

Informing policy development

I was unable to identify any discernable effect on health care policy in response to the publication of this review. As already suggested one explanation for this may be the relatively short time since publication. In his work on the policy cycle Sabatier argues that the process of policy change 'requires a time perspective of a decade or more' (Sabatier, 1988). This ties in with Weiss's ideas of the enlightenment function of research (Weiss, 1977) that argues that a focus on short-term impact will underestimate the influence of research because one of the functions of research is to alter perceptions and concepts of policy makers and researchers over time. In addition, it is worth considering whether it is appropriate for all research to impact on policy. The impact of a review must, to some extent, be related to the aim or purpose of the study. In this instance the review was part of a larger project and was designed to inform the development of a qualitative study. Indeed the purpose of systematic reviews is often to inform the development of a piece of primary research, or to justify the need for primary research, and its impact on policy may be through the primary study it informed.

As I found so little evidence of any influence upon policy this chapter does not include a detailed evaluation of impact. Instead in the following section I look more broadly at issues around systematic reviews of qualitative research and their role in informing policy decisions.

Do reviews of qualitative research impact upon policy?

In an examination of the role of qualitative research in influencing the policy process Rist argues that qualitative research may impact upon policy at a number of levels (Rist, 1994, Rist, 2003). He takes the position, as do others (Chelimsky, 1985, Guba E, 1984, Nakamura R and Smallwood, 1980), that the policy making process is a cycle that involves three phases; policy formulation, policy implementation and policy accountability. Each phase has its own order and logic, own information requirements and own policy actors (Rist, 1994). At the

policy formulation stage policy makers need to define clearly and understand the problem or condition that they are facing. He suggests that qualitative research may be highly relevant at this stage; for example looking at the social construction of problems, and community and organizational receptivity. However, he argues that because 'the window for policy formulation is small' it may not allow for the conduct of new qualitative research. Therefore, 'the information that can be passed through has to be ready and in a form that enhances quick understanding' (Rist, 1994). Risk also suggests that qualitative research may inform policy implementation arguing that the 'ground-level view of implementation is best done through qualitative research'. In the final stage of policy accountability he says that qualitative research may allow for an examination of the strengths and weaknesses of a programme'.

Rist suggests, however, that qualitative research has not fulfilled its potential contribution to the policy process because 'there is no broad-based and sustained tradition within contemporary social science of focusing on qualitative research specifically on policy issues' (Rist, 1994). Others have argued though that qualitative studies have been routinely excluded from evidence review and policy development (Graham and McDermott, 2006). However, there is now a growing support for the inclusion of qualitative research in systematic reviews. Many commentators have claimed that reviews of quantitative research may be too narrow to help policy-makers and practitioners make decisions (Greenhalgh et al., 2003, White, 2001) and suggested that incorporating qualitative research will increase their utility and relevance (Oliver, 2005b, Oliver, 2001). Indeed, discussions with policy makers have found that qualitative research, and the stories of individuals, are an important part of what they see as evidence (Petticrew et al., 2004, Whitehead et al., 2004).

Several methods have been developed that aim to increase the policy relevance of systematic reviews through the incorporation of qualitative research. Oliver and colleagues (Nilsen et al., 2006, Oliver, 2005b, Oliver et al., 2008) describe how they have developed a framework to enable them to synthesise findings

from both qualitative and quantitative studies in the same review. The framework, which includes a theoretical component, allowed them to examine evidence for effectiveness alongside qualitative evidence on barriers and facilitators. Data is displayed in a way that enables readers to navigate the charts presented to find evidence relevant to their circumstances (Oliver et al., 2008). Such an approach they suggest enabled them to make explicit recommendations for policy and practice based on 'transparent evidence' (Oliver, 2005b) and allows policy makers to draw out relevant implications for themselves (Oliver et al., 2008). They cite how a review on promoting healthy eating and physical activity for young people (Shepherd et al., 2001) was able to directly influence health care policy such as the National Service Framework for coronary heart disease (DOH., 2000).

An alternative approach is Bayesian meta-analysis. From a Bayesian perspective (Spiegelhalter et al., 2000, Sutton and Abrams, 2001) synthesis of research evidence is a decision-making process and as such 'pre-existing beliefs, subjective judgments and access to external sources of evidence will all shape interpretation' (Roberts et al., 2002). In a review of factors affecting the uptake of childhood immunisation, qualitative research was used to revise the pre-existing subjective beliefs of the reviewers; beliefs that were then used to inform the analysis. The reviewers concluded that 'the use of either qualitative or quantitative research alone might not identify all relevant factors, or might result in inappropriate judgments about their importance' (Roberts et al., 2002).

Another technique for incorporating qualitative research, and one that claims to increase relevance to policy makers, is Realist Synthesis. This is concerned with explanations of the mechanisms of programmes in particular contexts and settings rather than overall judgements of effectiveness. It has its roots in philosophy and Pawson and colleagues suggest that it is an appropriate methodological approach for complex policy interventions (Pawson et al., 2005). Their support of the realist approach is based on the complex, active and dynamic nature of complex service interventions; interventions that are subject

to reinvention and adaptation in different circumstances and contexts. They suggest that under a realist evaluation the basic question changes from 'what works' to 'what is it about this programme that works for whom in what circumstances' (Pawson et al., 2005). Realist synthesis does not distinguish between, or prioritise, different research designs but instead takes the position that 'multiple methods are needed to illuminate the richer picture'. Like framework synthesis existing theoretical standpoints form the basis of the review, with prevailing theories mapped, tested and refined.

Whilst the realist review may use conventional systematic review headings as a starting point it is methodologically very different. The process is far more iterative and non-linear than a traditional systematic review with steps of the process overlapping each other. Owing to the emphasis on refining theories they argue that 'second thoughts can (and should) occur at any stage' (Pawson et al., 2005). They claim this method is particularly relevant to policy makers because 'realist review raises the status of linkage from a recommendation to a methodological requirement' (Pawson et al., 2005). In addition, they take the standpoint that research influences policy according to Weiss's theory of enlightenment and that realist review is particularly suited to informing policy through the 'drip drip of enlightenment'.

The methods for systematic review described above all make claims for increasing policy relevance. However, one of the dilemmas is whether some of these approaches are compatible with the basic systematic review principles of a rigorous scientific approach. For example, although I can see the rationale of realist synthesis it is not a methodology that I would choose. Although realist synthesis can supposedly incorporate all types of research, including quantitative research, it does not, to my mind, satisfactorily address questions around the effectiveness of interventions. Although 'what works' is not the only question we should be asking I still believe it is a fundamental question and one that we need to approach with rigour. I would also argue that it is appropriate to prioritise some research designs over others. This does not mean prioritising

quantitative research over qualitative research but rather it means recognising that they might contribute to different parts of the evidence 'jigsaw' (Whitehead et al., 2004). In addition, I believe it is right to regard some quantitative designs, such as RCTs, as superior to other quantitative designs, such as uncontrolled studies, because RCTs are less susceptible to bias.

The approaches above all are concerned with reviews that include both quantitative and qualitative research in the same review. Whether reviews, such as ours, that only include qualitative research can be as useful for policy makers is not clear. Indeed, in an echo of Rist's claim that qualitative researchers have not focused on influencing policy I was able to find little literature that specifically addressed the question of the impact of qualitative systematic reviews upon policy. It may be that qualitative research is primarily seen by decision makers as useful for explaining the results of quantitative research or for developing quantitative research (Olson, 2001); indeed this is how it is conceptualised in Cochrane handbook.

Conclusion

Although there is increasing support for the inclusion of qualitative research in systematic reviews there remains some debate and confusion about the most appropriate methods and whether systematic reviews of qualitative research can be conducted using the same structures as those offered by conventional systematic review methodology (Dixon-Woods et al., 2006a). These concerns arise both because of technical issues, such as how to identify studies, but also due to epistemological questions about what constitutes qualitative research, what constitutes a good quality study and should searches be comprehensive. Unlike conventional systematic reviews, processes for reviews including qualitative research may need to be more iterative and organic (Greenhalgh et al., 2005). Despite the diversity of techniques none, as yet, seem to have secured a dominant position.

I found no evidence that the systematic review presented in this Chapter (Bunn et al., 2008a) has had any impact or influence upon health care policy in England. This lack of impact is perhaps not particularly surprising when we consider the fact that the review is only recently published and has not yet had time to influence the policy process. In addition, it should be noted, that the purpose of the review was to inform the development of a primary study and, thus, fulfilled the purpose for which it was designed. Most of the work that I found on improving the policy relevance of qualitative reviews concerned reviews incorporating both quantitative and qualitative studies; I found little that addressed reviews that only included qualitative research. It is possible that qualitative reviews may have greater impact if they are formally linked to a review of quantitative studies evaluating effectiveness. It would appear that more work is needed to explore whether reviews that only include qualitative research can impact upon policy and to identify ways of facilitating their influence.

Finally it is worth considering how the approach chosen to synthesise qualitative research might impact upon its relevance or utility for policy makers, and therefore upon its potential to impact upon the development of health care policy. Rist argued that qualitative research needed to be in a form that 'enhances quick understanding' (Rist, 1994) and this must be as pertinent to systematic reviews as it is to primary research. It has been suggested that some methodological approaches to the synthesis of qualitative research may be more useful to policy makers than others. Some of the more interpretive approaches may be more complex, and may operate at a conceptual or symbolic level that may require further interpretation by policy makers before they can usefully inform practice (Barnett-Page and Thomas, 2009).

Table 6.1 Methods for Systematically Reviewing Qualitative Evidence

Method	Author/examples	Main features	Advantages	Disadvantages
Bayesian Meta-analysis	Example of this approach (Roberts et al., 2002)	<ul style="list-style-type: none"> Qualitative evidence used to identify variables to be included in meta-analysis 	<ul style="list-style-type: none"> Prevents partial synthesis by including both quantitative and qualitative 	<ul style="list-style-type: none"> Not easy to implement Techniques still under development
Content analysis		<ul style="list-style-type: none"> Technique for categorising data and determining frequencies of categories 	<ul style="list-style-type: none"> Transparent processes Converts qualitative data into quantitative form and makes it easier to manipulate within quantitative frameworks 	<ul style="list-style-type: none"> Reductive Results may be oversimplified Diminishes complexity and context
Critical interpretive synthesis	<p>Dixon-Woods (Dixon-Woods et al., 2006b)</p> <p>Adaptation of meta-ethnography and also borrows techniques from grounded theory</p>	<ul style="list-style-type: none"> Adaptation of traditional meta-ethnographic methods Aim is generation of a synthesising argument (lines of argument) Sampling highly iterative, sample rather than identify all literature Development of theoretical categories based on analysis of conceptual similarities and differences that identified in the literature, and constant comparison across these Critique rather than critical appraisal – treats literature as an object of inquiry Recognises the partial nature of any account of the evidence but is explicit and reflexive about this 	<ul style="list-style-type: none"> Provides guidance on sampling and critical appraisal Can be used to synthesis all types of evidence including quantitative evidence Puts the author back in 	<ul style="list-style-type: none"> Only suitable for experienced qualitative researchers More work is needed on the impact of appraisal on synthesis Not inherently reproducible Many issues remain to be resolved

Framework synthesis	Brunton and Oliver (Brunton et al., 2006, Oliver et al., 2008) Based on framework analysis as outlined by Pope and colleagues (Pope et al., 2000)	<ul style="list-style-type: none"> • Rationale is that qualitative research produces large amount of textual data • Framework synthesis provides highly structured approach to organising and analysing data • Uses a priori framework • Indexing using numerical codes • Can be used to map nature and range of concepts under study 	<ul style="list-style-type: none"> • Potentially useful for policy makers 	
Grounded theory approach	Based on work by Glaser and Strauss. Developed as method for synthesis e.g. (Eaves, 2001, Kearney, 2001)	<ul style="list-style-type: none"> • Over-riding concern is generation of theory • Inductive approach to analysis – allowing theory to emerge from data • Uses constant-comparison method • Generation of higher order themes • Theoretical sampling and saturation means number of papers for review can be limited 	<ul style="list-style-type: none"> • Potential for generation of new theory 	<ul style="list-style-type: none"> • Not well established as a method for synthesis • Few examples • Lacks transparency • No guidance for how to critically appraise studies
Meta-ethnography	Based on work by Noblit and Hare (Noblit and Hare, 1988)	<p>Involves 3 major strategies:</p> <ul style="list-style-type: none"> • Reciprocal translational analysis (RTA) – key themes or concepts are identified • Refutational synthesis – contradictions between reports are examined • Lines of agreement synthesis (LOA) – involves building a general interpretation grounded in the findings of the separate studies • Developed to include qualitative research using diverse methodological approaches 	<ul style="list-style-type: none"> • Systematic approach – preserves interpretive approaches of primary data • Greater methodological work has been done than for some other methods 	<ul style="list-style-type: none"> • Is solely a method for synthesis with no guidance on sampling or critical appraisal. • Demanding and laborious

Meta-narrative	Greenhalgh (Greenhalgh et al., 2005) Approach informed by Thomas Kuhn's ideas that knowledge produced within particular paradigms (Kuhn, 1962)	<ul style="list-style-type: none"> • Designed to synthesise evidence to inform complex policy-making questions • Involves multi-disciplinary team • Synthesised research from wide range of disciplines • Unit of analysis was unfolding storyline • Key dimensions or themes identified 	<ul style="list-style-type: none"> • Potentially useful for policy makers 	<ul style="list-style-type: none"> • Time consuming and complex requiring multi-disciplinary team • Few examples
Meta-summary	Sandelowski (Sandelowski and Barroso, 2007)	<ul style="list-style-type: none"> • Aggregative • Findings accumulated and summarised rather than 'transformed' • Reflects a quantitative logic 	<ul style="list-style-type: none"> • Can be used to discover patterns or themes 	<ul style="list-style-type: none"> • No theory generation • Calculation of 'effect sizes' may be inappropriate
Meta-synthesis	Sandelowski (Sandelowski and Barroso, 2007)	<ul style="list-style-type: none"> • Interpretive integration of qualitative findings • Goes beyond summaries to offer novel interpretations of findings • More penetrating than meta-summary 		
Meta-Study	(Paterson et al., 2001)	<ul style="list-style-type: none"> • Interpretive approach grounded in constructivist orientation to epistemology • Involves 3 components: meta-data-analysis, meta-method, and meta-theory • No such thing as absolute truth • Involves analysis of theory, methods and findings of qualitative research and synthesis 	<ul style="list-style-type: none"> • Allows for theoretical development • Involves consideration of theoretical/conceptual basis of included studies 	<ul style="list-style-type: none"> • Time consuming & lengthy • Although critical appraisal recommended criteria not established • May only be suitable for experienced qualitative researchers

		<p>into new ways of thinking</p> <ul style="list-style-type: none"> • Requires adoption of theoretical framework • Involves identifying all relevant studies • Advanced method for seasoned researchers 		
Narrative summary	Examples of this approach (Fairbank et al., 2000, Popay et al., 2006)	<ul style="list-style-type: none"> • Varies from simple description of findings to more interpretive and reflexive accounts 	<ul style="list-style-type: none"> • Can be used to integrate quantitative and qualitative research 	<ul style="list-style-type: none"> • May be criticised for its lack of transparency
Realist synthesis	Pawson (Pawson et al., 2005)	<ul style="list-style-type: none"> • Theory driven • Integrates diverse forms of evidence by using them for proof or refutation of theory 	<ul style="list-style-type: none"> • Can be used for synthesis of quantitative and qualitative data together 	<ul style="list-style-type: none"> • Treats all forms of evidence as equally authoritative • May lack rigour in evaluation of effectiveness
Thematic analysis	Example of this approach (Garcia et al., 2002, Thomas and Harden, 2008)	<ul style="list-style-type: none"> • Some overlap with methods for narrative summary • Involves identification of prominent or recurrent themes • Summarises findings of different studies under thematic headings 	<ul style="list-style-type: none"> • Flexible and allows integration of quantitative and qualitative research • Allows for transparency • Process has been well described 	<ul style="list-style-type: none"> • Failure to distinguish between data or theory driven approaches • Not clear how different themes weighted • More work needed on how to use quality data within the review

Chapter 7: Systematic review to inform the development of NICE public health guidance

Introduction

In this chapter I present a systematic review of interventions for the prevention of sexually transmitted infections and teenage pregnancies (Bunn et al., 2006a). The review was funded by the National Institute of Health and Clinical Excellence (NICE) to inform the development of public health guidance. As in previous chapters I will examine the drivers behind the review, discuss the methods used and evaluate the impact of the review. In addition, as the review was designed to inform guidelines, I critically explore issues around the development and implementation of guidelines. This includes the rationale for guidelines, a consideration of potential limitations of guidelines, a description of the guideline development process and an examination of the impact of guidelines on practice and the potential barriers to their successful implementation.

The nature of the impact evaluation in this chapter differs from the others in this submission because the review was specifically commissioned to inform Government guidelines; and therefore the link between the review and policy is already established. In addition, the review was not published in a journal but rather is published on the NICE website; where it is included as a background document to the guidance. Therefore, although the review is in the public domain, it is the guidance, rather than the review, that is most likely to be accessed. In evaluating the impact of the review and guidance I begin by examining the extent to which the recommendations in the guidelines are concordant with the evidence presented in the systematic review; then I present evidence relating to the impact of the review or the guidance.

Drivers behind review question

The review in this chapter was commissioned by the National Institute of Health and Clinical Excellence (NICE). NICE was created in 1999 as an independent body to appraise the clinical and cost-effectiveness of health technologies referred to

it from the Department of Health; its guidance covers England and Wales. It produces guidance in three areas: public health (the promotion of good health and prevention of ill health), health technologies (guidance on new and existing medicines, treatments and procedures within the NHS) and clinical practice (specific conditions and treatments). Despite ongoing controversy over its role and criticisms of a number of its decisions NICE is now well established and has become 'a core policy driver within the national health service in England and Wales' (Collier, 2008).

The review included here (contraceptive advice and provision for the prevention of under 18 conceptions and STIs) was one of three reviews commissioned to inform the development of public health guidance on the prevention and treatment of sexually transmitted infections and under 18 conceptions. The others, produced by researchers from different institutions, considered evidence for the effectiveness of screening for chlamydial infection in sexually active young men and women, and effectiveness of partner notification for sexually transmitted infections including HIV. The three effectiveness reviews, along with an economic evaluation, were used to develop guidelines which were published in February 2007⁷.

Topics for NICE reviews generally come directly from the Department of Health. However, topics for consideration can come from a number of sources including clinical and public health professionals, patients, carers and the general public, and the Department of Health's national clinical directors. In addition, suggestions can come from within NICE itself. A number of criteria are taken into account when selecting topics. This includes the following:

- burden of disease (population affected, morbidity, mortality)
- resource impact (i.e. the cost impact on the NHS or the public sector)
- policy importance (i.e. whether the topic falls within a government priority area)

⁷ <http://www.nice.org.uk/PHI003>

- whether there is inappropriate variation in practice across the country
- factors affecting the timeliness or urgency for guidance to be produced.

In this case guidance on sexual health was felt to be needed because there had been large increases in many sexually transmitted infections in the UK over the previous 10 years, with diagnoses of chlamydia, gonorrhoea and infectious syphilis having more than doubled since 1995. The UK also had the highest rate of teenage pregnancy in Western Europe (UNICEF, 2001). In response to these problems the Government of the time set out a number of targets concerning sexual health and the prevention of unwanted teenage pregnancies which the guidance was designed to promote. This included a joint Public Service Agreement (PSA) target for the Department of Health (DH) and the Department for Education and Skills (DfES) to reduce under eighteen conceptions by 50% by 2010. In addition, the guidance was intended to support the delivery of a range of measures for improving sexual health as set out in the public health white paper *Choosing Health* (DOH., 2004), and the implementation of the *National Service Framework for Children, Young People and Maternity Services* (DOH, 2004) which set standards for health promotion and prevention with young people to reduce the risk of both teenage pregnancy and acquiring a sexually transmitted infection. These measures also all supported the national teenage pregnancy strategy (Social Exclusion Unit, 1999) and the national strategy for sexual health and HIV (DOH, 2001) which had a number of aims including reducing the transmission of HIV and STIs, and reducing the rates of unintended pregnancies.

The full version of the review produced for NICE is over 250 pages long. Therefore, only the executive summary of the report is presented in Appendix 2. The full version of the report can be found on the NICE website at the following address:

http://www.nice.org.uk/nicemedia/pdf/STI_University_Of_Hertfordshire_Final.pdf

Rationale for guidelines

Clinical guidelines have been defined as ‘an attempt to distil a large body of medical expertise into a convenient, readily usable format’ (Cook et al., 1997)’ and as ‘systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances’ (Woolf et al., 1999). NICE claims that their guidelines can be used to:

- Provide recommendations for the treatment and care of people by health professionals
- Develop standards to assess the clinical practice of individual health professionals
- Assist in the education and training of health professionals
- Help patients to make informed decisions
- Improve communication between patient and health professional

Although guidelines are not a new technology, in recent years they have attracted increasing attention and the use of guidelines is now widespread not just in the UK but also internationally. The increasing emphasis on developing and using guidelines is undoubtedly associated with the rise of the evidence-based practice movement (Timmermans and Mauck, 2005). However, there are other factors, rooted in issues that most health care systems face, that have also contributed to their ascendancy (Woolf et al., 1999). One of the most significant of these is the rising cost of health care which is fuelled by an increasing demand for care, more expensive technologies and an ageing population. Before interventions are recommended for widespread use NICE assesses both clinical effectiveness and cost-effectiveness. However, the extent to which guidelines are seen as mandatory regulations rather than recommendations varies from body to body and from country to country. NICE is the most influential guideline development body in England, although the current Government has recently reduced its power by stipulating that its guidance is no longer legally binding in England and Wales. The public health guidance has never been mandatory.

Of course one of the aims of guidelines is to improve the quality of health care. In Chapter 1 I discussed how variations in service delivery and treatment were a factor behind the growth of the evidence-based practice movement. Guidelines are designed to reduce variations in services delivery among different providers, hospitals and geographical regions. This is to make care consistent and provide equal access to health care and because it is presumed that some of this variation is attributable to inappropriate care. Woolf and colleagues suggest a number of potential health and service related benefits of clinical practice guidelines (Woolf et al., 1999). These can be seen in box 2 below.

Box 2. Potential benefits of guidelines

The use of guidelines may:

- Improve health outcomes
- Improve quality of clinical decisions
- Improve consistency of care
- Empower patients and the public by improved access to information
- Promote fair distribution of care and resources
- Improve efficiency, free up resources and optimise value for money
- Support quality improvement activities
- Identify gaps in the evidence
- Identify interventions of proven worth and call attention to ineffective, dangerous and wasteful practices
- Through critical appraisal of evidences identify design flaws in existing studies
- Influence public policy by calling attention to under recognised health problems, clinical services and preventive interventions and to neglected populations and high risk groups

Potential problems with clinical guidelines

However, despite claims that guidelines improve patient care, their widespread use is not accepted by everyone as a good thing. Arguments against their use include the assertions that recommendations in guidelines may be inaccurate, that they may not take into account the needs of individual patients, that they may be used to ration or divert resources inappropriately and that their use may be detrimental to practitioners. These issues are explored more fully below.

Recommendations may be wrong

Commentators have argued that a key limitation of guidelines is that they may often be wrong. They suggest that one major reason for this is that scientific evidence about what to recommend is often lacking, misleading or misinterpreted and that even when the evidence is good recommendations will still involve subjective value judgements when the benefits of an intervention are weighted against the potential harms (Woolf et al., 1999). The sort of evidence considered to be 'gold standard' by many proponents of evidence-based practice, such as RCTs, is often not available (Timmermans and Mauck, 2005). Even when RCTs are available they may be flawed; for example follow-up may be too short and relevant outcomes not measured (Spence, 2009). It has also been suggested that as the evidence base for many technologies and practices is often ambiguous and contested that evidence must be 'continually interpreted and reframed in accordance with the local context and priorities' (Ferlie et al., 2001).

An examination of the use of effectiveness and cost-effectiveness reviews by Government bodies, including NICE, to make decisions about health care provision in Britain, Australia and Canada found that all had ongoing issues with the quality and strength of the experimental evidence (Clement et al., 2009). In another paper examining the data used for pharmacoeconomic analyses for a publicly funded insurance program in Australia the authors found that for a number of questions no randomised trials were available and that even when they were available they often had serious methodological flaws and that there

was uncertainty around the estimates of effectiveness. In addition, they felt that some of the assumptions on which economic models had been based could be open to challenge (Hill et al., 2000). Although both these papers concern decisions about clinical therapies, in particular drug therapy, this lack of evidence applies to public health guidance as well. Indeed, the lack of 'gold standard' evidence is likely to be even more of a problem for complex public health interventions. When undertaking the review included in this chapter (Bunn et al., 2006a) I found limited evidence and, of that which was available, little was from the UK.

Another reason why guidelines may be wrong is because they are unduly influenced by the opinions, clinical experience and composition of the group developing the guidelines (Kane, 1995, Woolf et al., 1999). An overreliance on expert opinion may be particularly dangerous (Kane, 1995) and factors such as professional role and status may distort group processes (Pagliari and Grimshaw, 2002, Pagliari et al., 2001). In addition, guidelines may be wrong because patient's needs may not be the only factor considered when making recommendations and practices may be recommended to help control costs, serve societal needs or protect the special interests of clinicians, policy makers or managers (Woolf et al., 1999).

Guidelines too inflexible for the needs of individual patients

Another criticism of guidelines is that they are inflexible and do not take into account the circumstances, medical history and values and preferences of individual patients (Woolf, 1997); what is best for patients overall may not be appropriate for individuals. Guidelines tend to focus on individual conditions whereas, in reality, many patients have complex health care needs with multiple health care problems. These concerns have led to doctors, faced with non-uniform clinical problems, criticising guidelines as 'cookbook medicine' (Woolf et al., 1999). Researchers in the USA have suggested that there is the potential for conflict between medical humanism, which seeks to understand the patient as a person and emphasises 'shared decision making' between health care

professionals and patients, and evidence-based practice which aims to put medicine on a firm scientific footing and emphasises the use of standardised guidelines. They argue that an outright collision between these two movements can only be avoided if guidelines remain as 'recommendations rather than mandates'(Hartzband and Groopman, 2009).

In addition, it has been suggested that guidelines are often flawed because they are derived from studies carried out in selected groups of patients who may not be representative of all populations (Spence, 2009), particularly if they have coexisting conditions (Hartzband and Groopman, 2009). Certainly in my experience RCTs often have strict inclusion criteria which mean that patients with comorbidities or complex medical conditions are excluded from studies. This may limit the generalisability of the findings beyond the particular group of participants in the trial.

Guidelines used to ration health care

One of the greatest areas of controversy surrounding guideline use in the UK is whether they should be used to make mandatory recommendations about health care provision. Until recently health technology guidance has been legally binding in England and Wales which has led to the accusation that they were used as a tool to ration health care (Gingrich, 2009). Critics of NICE have argued that its decisions are not based on effectiveness so much as cost-effectiveness and that even if a treatment is shown to be effective it may be rejected if it is too expensive (Smith, 2000). However, others have pointed out the dilemma faced by the majority of health care systems, namely that resources are scarce but competing demands are infinite (Walker et al., 2007). Sir Michael Rawlins the Chairman of NICE argues that 'some form of rationing is inevitable in every health care system and that the issue is not whether to ration but how to go about it'. He suggests that some countries, such as the USA, ration on the basis of individual wealth or income whereas in Europe 'no citizens are denied access to basic health care merely because they are poor'(Rawlins, 2009).

It has also been argued that the development of formal guidelines may channel resources inappropriately or unfairly. For example, Woolf and colleagues argue that 'imprudent recommendations for costly interventions may displace limited resources that are needed for other services of greater value to patients'. They suggest that guidelines may be used by advocacy groups (and health professionals) to give inaccurate impressions about the relative importance of diseases and the effectiveness of interventions (Woolf et al., 1999). In an editorial in the BMJ Richard Smith also makes this point (Smith, 2000) when he says that although patients may receive a drug which could benefit them they may end up worse off overall because the cost of the drug may mean that there will be cuts to other services, such as nursing and social care, instead.

Disadvantages for practitioners

It has also been claimed that guidelines may have negative effects on health care practitioners. Practitioners may find guidelines inconvenient and time consuming to use (Woolf et al., 1999) and be confused by conflicting guidelines on the same topic. In addition, it has been suggested that a widespread reliance on guidelines may deskill practitioners and that 'simplistic algorithms may not take into account the complexity of clinical practice and the parallel and iterative thought processes inherent in clinical judgement' (Woolf et al., 1999). Instead of using clinical judgement practitioners will be encouraged to follow guidelines that treat all patients as the same and, therefore, will be 'poorly equipped to contend with the variations between patients they will encounter in actual clinical circumstances' (Timmermans and Mauck, 2005). Clinicians may also resent guidelines because they feel they are a threat to their personal autonomy and because algorithms and guidelines may result in health care professionals being replaced by less expensive, less skilled workers (Timmermans and Mauck, 2005).

The development of NICE Guidelines

One of the criticisms levelled at NICE is that its decision making process is too lengthy (Martin, 2007). Some critics have compared it unfavourably with the Scottish Medicines Consortium which they claim is able to make decisions more quickly and cheaply. However, as Minhas and Patel point out the process used by NICE includes greater mechanisms for scoping, public consultation, revision and appeals (Minhas and Patel, 2008). Despite some criticisms of NICE and their processes several reviews from independent agencies have reported favourably on their work. These include two reviews from the House of Commons Health Committee (House of Commons Health Committee, 2002, NICE, 2008) and a review by the World Health Organization (Hill et al., 2003). It has also been acknowledged in a select committee report that 'the job of NICE is much more difficult than we first imagined' (NICE, 2008).

The development of NICE guidelines is a process that involves a number of stages. Once the Department of Health has referred a topic to NICE then national organisations representing patients, carers, and health professionals involved in their care, can register as stakeholders. These stakeholders are consulted throughout the guideline development process. The next step involves the preparation of a scope statement which sets out what the guideline will, and will not, cover. NICE, registered stakeholders and an independent guideline review panel can all contribute to the development of the scope. Then a guidelines development group made up of health professionals, representatives of patient and carer groups and technical experts is established. Once the review evidence is produced the guideline development group assesses the available evidence and produce a draft guideline. There is at least one public consultation period for registered stakeholders to comment on the draft guideline. An independent guideline review panel reviews the guideline to check that stakeholder comments have been taken into account. After the guideline development group finalises the recommendations, the collaborating centre produces the final guideline; NICE formally approves the final guideline and issues its guidance to the NHS. As a reviewer my involvement in the process

began only once the scope statement had been finalised. My role included producing the review to time, presenting the findings to PHIAC (Public Health Interventions Advisory Committee), revising the review in light of PHIACs comments, and commenting on the draft guidance.

Methods for NICE reviews

The methods used for the review presented in this chapter were those stipulated in the NICE handbook (NICE, 2006). As the methods were generally similar to those recommended for Cochrane Reviews I do not enumerate the methods of the review in detail but instead discuss aspects of the review process or methods which were different from those discussed in previous chapters.

Question development

This review differs significantly from any of the others presented in this submission because the way it was commissioned meant that the question was driven by policy makers rather than researchers. The title, subject and inclusion criteria all reflect the scope statement⁸ which, as is usual practice, was decided upon by NICE and the various stakeholders involved in the guideline development process. The review question included the prevention of sexually transmitted infections in all age groups and all populations and the prevention of conceptions in the under 18s. NICE conceptualised this as one review, whereas it may have been better as two separate reviews.

The breadth of the question was narrowed somewhat by a focus on one-to-one (i.e. individual rather than group-based) interventions only. This excluded many studies which were delivered to groups rather than individuals. Whether it was appropriate to exclude group based interventions is debatable. Field work to test the review findings showed that practitioners would have liked more evidence about the relative effectiveness of one-to-one interventions compared with other interventions, such as group based sessions (Cook et al., 2007). It is possible that the guidance may have been more useful if it had included group-

⁸ <http://www.nice.org.uk/guidance/index.jsp?action=download&o=31896>

based interventions although this would have made the review very large, less focused and would have taken much longer to complete. In addition, it could be argued that the focus on one-to-one interventions made the guidance of direct relevance to practitioners in clinical settings who would typically see patients on an individual, rather than a group, basis.

The scope of the review was broader than Cochrane reviews because it also included systematic reviews, and controlled before/after studies as well as RCTs. The rationale for including systematic reviews is that they provide a synthesis of the evidence and may give more information than a single study. However, reviewing systematic reviews can be problematic. For example, other systematic reviews may have addressed slightly different questions and may not have presented the required detail from the primary studies. In this case it is necessary to go back to those studies for further information. In this instance, however, we found no systematic reviews that met our inclusion criteria and so the dilemma of how to deal with them did not arise. The review also included qualitative studies that either looked at the process of the interventions (e.g. how and why they do/do not work) and/or focused on the user perspective of potential barriers and facilitators. In Chapter 6 I looked in depth at the rationale and methods for including qualitative studies. The inclusion of qualitative studies was, therefore, intended to allow us to address some of these more complex questions and go beyond simple questions about 'what works'.

Data extraction

The processes for data extraction were similar to those used for Cochrane reviews and involved independent double data extraction. However, the complexity of the interventions under study, the need to go beyond simple questions about 'what works' and the need to address effectiveness in different groups of participants meant that the data extraction form was lengthy and complex.

Critical appraisal

Critical appraisal for NICE reviews is done using quality assessment checklists. The domains included in the checklist are similar to those recommended in the Cochrane handbook. However, unlike Cochrane reviews, which generally don't include summary scores, NICE quality assessment requires that studies are given an overall summary score. Studies are graded as (++), (+) or (–), with (++) meaning studies are judged to be of high quality and (–) that they are of poor quality. Summary scores are not recommended in the Cochrane handbook because they involve assigning weights to different items in the scale which can be difficult to justify (Higgins, 2008). Furthermore they have been shown to be unreliable as assessments of validity (Juni et al., 1999).

Analysis

Owing to the wide scope of the research question, and the heterogeneity in interventions, participants and outcomes, I did not consider meta-analysis to be appropriate. Instead data were presented in the text of the review and in the results column of the evidence tables. Although the guideline development group agreed with the decision not to pool studies in a meta-analysis they did find it difficult to interpret the results without an overall summary statistic. Therefore, I suggested that in the final draft of the review, the results be presented graphically in forest plots but without pooled summary statistics. The chair of PHIAC felt this was very useful and something that they should consider for future reviews. Indeed, the updated NICE methods manual (NICE, 2009) now includes a recommendation that

'Forest plots should be used to show effect estimates and confidence intervals for each study (when available, or when it is possible to calculate them). If possible, they should be used even when it is not appropriate to do a metaanalysis and present a pooled estimate'.

This is accompanied in the handbook (p81) by an example of a forest plot taken from the review presented in this chapter (Bunn et al., 2006a).

This debate about how best to present the results is indicative of the difficulties reviewers can face presenting results in systematic reviews where it has not been appropriate to pool studies in a meta-analysis. In complex reviews with large number of studies it can be particularly difficult to present the results in a coherent and easily understood way. NICE have tried to address this to some extent by presenting results in tables with an indication of whether the intervention had a positive effect (+), a negative effect (-) or no statistically significant effect (0). Although this does give a snapshot of whether or not a result was statistically significant it could be misleading as a non significant result is not necessarily an indication that the intervention is not effective; many trials lack the power to detect moderate treatment effects (Tarnow-Mordi and Healy, 1999). NICE also require that each review question is accompanied by at least one evidence statement that reflects the strength (quality, quantity and consistency) of the evidence and makes a statement about its applicability. These statements can also be used to highlight the lack of evidence (NICE, 2009). Reducing large amounts of complex data into clear, self-contained statements is challenging. The accuracy of these statements was particularly important as they were used by the guidelines development group as a basis for their recommendations.

The use of guidelines by practitioners

It is important to look at the impact of the public health guidelines on sexual health within the context of current thinking about guideline implementation. Therefore, before I look at the impact of the review I will critically explore some of the current research around the implementation of guidelines, including the impact of NICE guidelines. This includes the extent to which guidelines are used by practitioners and the potential barriers to their successful implementation.

There is evidence that clinical guidelines can be effective in changing the process and outcome of care (Grimshaw et al., 1995, Grimshaw and Russell, 1993, Thomas et al., 2000), although it has been found that some are adhered to more closely than others (Grol, 2001). These differences in concordance could be

explained by a variety of factors such as the type of health problem involved, the method of development used, the content of the recommendations, the source of dissemination, or the format and layout (Davis and Taylor-Vaisey, 1997). Compliance has been found to be better for guidelines of acute care than chronic care and is associated with a better quality of evidence supporting the recommendations, compatibility of the recommendations with existing values, less complexity of the decision making needed, more concrete description of the desired performance, and fewer new skills and organisational changes needed to follow the recommendations (Burgers et al., 2003, Foy et al., 2002, Grilli and Lomas, 1994, Grol et al., 1998).

The mode of dissemination may have a bearing on the uptake of guidelines. A number of studies have found passive dissemination of guidelines ineffective (Lomas, 1991, Oxman et al., 1995) or of only limited effectiveness (Farmer et al., 2008). However, other studies have suggested that passive dissemination of guidelines does have an effect. Two separate studies looked at the effect of an Effective Health Care Bulletin for glue ear issued in 1992 (Black and Hutchings, 2002, Mason et al., 2001). Effective Health Care Bulletins were produced by the National Centre for Reviews and Dissemination funded by the Department of Health. In both studies they found that the passive dissemination of the guidelines was associated with a significant decline in rates of surgery for glue ear and conclude that 'distributing printed recommendations to decision makers may influence surgery rates' (Mason et al., 2001). However, determining how much of the change is attributable to the guidelines and how much to other confounding factors is not easy. The authors acknowledge that there may be a number of contextual factors that contributed to this decline. This includes pre-existing professional concerns about the value of surgery, the introduction of an internal market into the NHS, and growing apprehension among patients fuelled by scepticism in the mass media (Black and Hutchings, 2002).

Sheldon and colleagues undertook an evaluation to assess the extent of implementation of NICE guidance (Sheldon et al., 2004). This involved

interrupted time series analysis, review of case notes, a survey and interviews to evaluate the implementation of 12 sets of guidelines. They found that implementation of NICE guidance was variable; some had been associated with changes in practice and others had not. However, the retrospective observational nature of their study made it difficult to separate the effect of the guidelines from other factors that may have influenced professional practice.

The effect of NICE guidelines on practice is also assessed in another paper which looked at the effect of NICE guidance on wisdom tooth extraction and primary total hip replacement (Ryan et al., 2004). NICE guidance recommended that pathology free impacted wisdom teeth should not be operated upon. The authors looked at hospital activity data for 88 trusts that could provide three years data from before guidance published and two years afterwards. Although there was a downward trend in extraction, this appeared to have started before the NICE guidance and may be a response to previously issued professional guidelines. They found no significant change in behaviour relating to total hip replacement. They suggest that one reason for this lack of impact is the relatively passive dissemination of the guidelines (Ryan et al., 2004). However, Michael Rawlins the Chairman of NICE suggests that implementation of NICE guidance is better than some commentators have suggested. In a letter defending NICE's role in implementing their guidelines (Rawlins and Dillon, 2005) he cites a report by Abacus International (Howard and Harrison, 2005) which looked at the impact of 28 appraisals. The results showed that 12 appraisals were implemented fully, 12 were incompletely implemented, and four were over-implemented.

The quality of the evidence supporting recommendations has also been found to be a facilitator in the implementation of guidelines. In a pilot study, of a national initiative funded by the DOH to demonstrate how practice could successfully be changed to bring it into line with research evidence, they found the most influential factors were strong evidence, supportive opinion leaders and interaction within a committed organization. Unclear evidence was a real barrier

(Dopson et al., 2001). In addition, guidelines that are easy to understand, can be tried out, and do not require specific resources, have a greater chance of implementation (Francke et al., 2008). Health-care professionals were less likely to use guidelines for patients who had complex conditions or co-morbidities.

The difficulties associated with changing the behaviour of health care professionals may be a significant barrier to guideline implementation. Practitioners may be resistant to guidelines because they are seen as a threat to the use of autonomy and discretion in professional work (Timmermans and Mauck, 2005). An extensive review of the evidence suggests that changing behaviour is possible, but that this change generally requires comprehensive approaches at different levels (practitioner, team practice, hospital, wider environment), tailored to specific settings and target groups (Grol and Grimshaw, 2003). There is also substantial literature to suggest that the successful implementation of most clinical guidelines requires organizational as well as individual change (Grimshaw et al., 2004, Rycroft-Malone et al., 2004a) and that effective strategies generally have multiple components (Francke et al., 2008). Cultural and environmental characteristics such as poor leadership, a lack of support from peers, little emphasis on continuing education and insufficient staff and time can be major barriers to the implementation of evidence (Francke et al., 2008, Kitson et al., 1998).

Results of impact analysis

Were the guidelines evidence-based?

Before I present the results of the impact evaluation I consider to what extent the guidelines are 'evidence-informed'. In other words I wanted to assess whether the recommendations in the guidelines were consistent with the evidence presented in the review itself. The final guidelines are, after all, influenced by a number of factors, including evidence from the other reviews and economic evaluation, the results of the fieldwork, and the views of the stakeholders and guideline development group. In addition, it is possible that

guideline developers are subject to financial and political pressures that may influence their recommendations.

To evaluate the extent to which the guidelines were informed by the review I compared the recommendations with the evidence from the review. Overall, the recommendations appeared to be in line with the evidence from the review. However, there were some slight discrepancies. For example, in recommendation two in the guidelines it recommends that health care practitioners should:

'have one-to-one structured discussions with individuals at high risk of STIs. The discussions should be structured on the basis of behaviour change theories. They should address factors that can help reduce risk-taking and improve self-efficacy and motivation. Ideally, each session should last at least 15-20 minutes. The number of sessions will depend on individual need' (NICE, 2007b).

Although there was evidence in the review that 15-20 minute sessions could be effective, the review suggested that the most effective interventions were those with sessions that were longer than that; for example 60 or 90 minutes. However, a recommendation suggesting sessions of such length is unlikely to be practical or possible in a typical UK health setting. Guideline developers have to take into account the context and feasibility of intervention delivery.

Having been involved in the review and seen the process of guideline development at first hand has made me cognisant of the difficulties involved in creating guidelines when the evidence base is poor. Although there were some good quality studies they were all done in the USA and there was little UK-based evidence. The original scope specified a number of sub-questions looking at effectiveness in high risk groups, such as looked after children and sex workers. However, for many of these groups, there was little or no relevant evidence.

Knowledge production

Publication and methods of dissemination

The guidance was published on the NICE website in February 2007. A number of documents were available including the full guidance, a summary of the guidance and all background documents including the full reviews, the scope statement, and fieldwork evaluations. In addition the guidance was accompanied by a press release (NICE, 2007a) and implementation advice. The implementation advice includes information on identifying a project lead, carrying out a baseline assessment, assessing cost, and building and implementing an action plan.

Impact within research community

Because the review was an unpublished report, and therefore would not be indexed in WoS or Scopus, I only performed the citation analysis in Google Scholar. The review had only been cited once and this was a self-citation in a book chapter on which I was an author (Trivedi et al., 2007a). I did, however, search the internet for research papers and organisations that had cited the guidance. There were a number of commentaries or discussion papers relating to the guidance and one of these referred to the review but did not formally cite it (Ward, 2007). Considering that the review is only published on the NICE website, that review authors are not specifically mentioned in the guidance, and that the reviews are not as easily accessible as the guidance, this lack of citation is hardly surprising. It was also not possible to do a formal citation analysis on the published guidance but internet searches found a number of references to the guidance.

The organisations or websites where the guidance was discussed or cited can be seen in Table 7.1.

Table 7.1 Results of internet searches to identify organisations citing the guidelines

Organisation or place where guidance cited
NICE guidance cited in guidelines produced by British Association for Sexual Health and HIV (Rogstad et al., 2008)
Improvement and Development Agency http://www.idea.gov.uk/idk/core/page.do?pageId=6034364 (website that mentions guidance)
East midlands Public Health Observatory http://www.empho.org.uk/THEMES/teenagepregnancy/tp9.aspx (website link to guidelines)
NHS Clinical Knowledge Summaries http://www.cks.nhs.uk/contraception_emergency/evidence/references (cites guidelines)
The National Guideline Clearinghouse™ (NGC). US based database of evidence-based clinical practice guidelines and related documents. http://www.guideline.gov/summary/summary.aspx?doc_id=10789&nbr=005617&string=chlamydia

Research targeting

In the concluding section of the review (Bunn et al., 2006a) I highlighted some of the gaps in the evidence. These included a need for studies that were based in the UK and were large enough to detect primary outcomes such as a reduction in STIs or conceptions. In addition, I suggested there was a lack of information on the effectiveness of peer-led and school-based interventions, and little evidence on one-to-one interventions in vulnerable groups such as young people in or leaving care, young people from some ethnic backgrounds (primarily Pakistani & Bangladeshi), sex workers, refugees and asylum seekers. The gaps in the evidence that I identified are largely reflected in the final NICE guidance (Appendix B p36). This information on gaps in the evidence was then used by PHIAC to make recommendations about future research questions. These included recommendations for further research about effectiveness of interventions delivered in different settings and by different professionals and

for more work in high-risk groups. However, I was unable to ascertain to what extent these recommendations had been acted upon and whether the review or guidance had been instrumental in directing future research.

Informing policy development

As already mentioned, public health guidance is not mandatory and local service providers are under no obligation to follow it. The extent to which the recommendations have been adopted is difficult to ascertain and is, in reality, beyond the scope of this work. Certainly a number of commentaries published in response to the guidance highlighted the difficulties of implementing recommendations in busy and hard pressed health services. Helen Ward, a specialist in infectious disease epidemiology, points out that because sexual health services in the UK have been streamlined, with a move towards rapid testing, single visits and, in some cases, self-completed sexual histories and self-collected specimens, adding one or two 20 minute counselling sessions into such a service provides a major challenge (Ward, 2007). In a critique of how the guidance might apply to general practice Oakeshott and Graham identify a number of barriers to implementation (Oakeshott and Graham, 2007). These include a lack of time to identify high risk individuals, a lack of personnel trained in sexual health counselling and the challenges of working with young people who may not return for follow up appointments or may not wish to attend counselling sessions. To what extent service providers have attempted to overcome these barriers or the extent to which they may have been successful in implementing guidance is unclear.

NICE was not initially explicitly responsible for promoting and monitoring the implementation of their guidelines. However, this has now changed and since 2004 the institute has had an implementation support strategy. The NICE website now includes two databases which include details of implementation. One is the Evaluation and Review of NICE implementation evidence database (ERNIE) which is a source of information on the implementation and uptake of NICE guidelines. The other is The Shared Learning database which provides

examples from organisations implementing NICE guidance and the lessons they have learnt. I searched both of these databases but was unable to find any information relating to the implementation of the sexual health guidance with which this chapter is concerned. Indeed, although there was data on a number of health technology and clinical practice guidance there was very little information on any of the public health guidance.

Conclusion

In this chapter I have examined the impact of a systematic review on the development of NICE guidance. As the review was specifically commissioned as part of the guideline development process the link between the review and policy is clearly established. It is apparent, therefore, that the review had a direct impact on Government policy in England. However, there are a number of barriers to the implementation of guidelines and there has been some debate in the literature about the extent to which NICE guidance has been successfully implemented. I found little information to indicate whether or not the guidance considered in this chapter (NICE, 2007b) had been successfully implemented.

Chapter 8: Evaluating the safety and efficacy of telephone consultation: a systematic review to inform service delivery & organization

Introduction

In previous Chapters the focus of the included reviews has largely been on interventions for preventing or treating ill health; for example the prevention of injuries or sexually transmitted infections or the treatment of critically ill patients. In contrast the review presented in this chapter, which looks at the effectiveness of telephone consultation and triage (Bunn et al., 2004a, Bunn et al., 2005), is primarily concerned with an issue relating to the way in which health services are organised and delivered. The review, which was published in the Cochrane Library and the British Journal of General Practice, was conducted using methods which were largely similar to those already described in previous chapters, in particular Chapters 4 and 5. However, there are some methodological issues specific to this review, such as the inclusion of interrupted time series designs, which I examine further in this chapter.

As before, I also include information about the drivers behind the review question and assess the impact of the review on policy. The systematic review presented in this chapter relates to the effectiveness of telephone consultation, a topic that was subject to considerable controversy and debate at the time we conducted the review. Therefore, I end the chapter by looking at the impact of the review in light of the political and socio-economic context of the time and consider the relationship between evidence, political agendas and professional interests. I am first author on the papers and was responsible for all stages of the review including: developing a review protocol, screening search records, data extraction, critical appraisal, data analysis and writing up.

Drivers behind review question

Many systematic reviews are undertaken to inform the development of primary research and, indeed, that was the main impetus behind this review. A colleague

was planning a study to evaluate the impact of the advice given by NHS direct staff (Byrne et al., 2007) and, as we were unable to find a previous systematic review on the subject, we felt that it was important to review the existing evidence. NHS Direct is a 24-h telephone advice service, based in England and Wales, which is designed to help callers to self-manage problems and reduce unnecessary demands on other National Health Service (NHS) provision (Munro et al., 2000). Set up in December 1997 NHS Direct was promoted as part of the previous Governments modernising strategy which included improving access and providing patient centred, technologically sophisticated care in the NHS (Executive, 1998, Executive, 2000). NHS Direct is staffed by nurses and reflected a move towards extending the role of nurses so that they take on responsibility for tasks previously done by doctors. At the time NHS direct was a relatively new initiative and, despite its rapid expansion, there was a lack of studies evaluating its effectiveness or safety (Byrne et al., 2007).

Although the focus of the primary research study was on NHS Direct we chose to set the inclusion criteria for the review somewhat wider and included all forms of telephone consultation or triage delivered by any type of health care worker. The rationale for this was that telephone consultation was being used in a number of settings (not just NHS Direct) and an overview of any information about its safety and efficacy was required. However, despite broadening our inclusion criteria in this way, we found only 11 relevant studies.

Review methods

The Cochrane Collaboration is structured so that different reviews fall under the remit of different Collaborative Review Groups who each have their own editorial team. The reviews in Chapters 4 and 5 fall under the scope of the Cochrane Injuries Group, of which I am an Editor, whereas the review in this chapter was conducted for the Cochrane Effective Practice and Organisation of Care Group (EPOC). EPOC differs from most other Cochrane Groups in that its scope is concerned not with a particular medical condition or area of health care but instead is focused around the way services are organised and delivered. The

focus of EPOC is on reviews of interventions designed to improve professional practice and the delivery of effective health services. This includes various forms of continuing education, quality assurance, informatics, financial, organisational and regulatory interventions that can affect the ability of health care professionals to deliver services more effectively and efficiently. Organisational interventions are defined as those which involve a change in the structure or delivery of health care. In other words, an organisational intervention is a change in who delivers health care, how care is organised, or where care is delivered'⁹; telephone consultation may involve any or all of those changes.

Types of studies

Although each Collaborative Review Group is largely guided by the methods set down in the Cochrane handbook there are some differences between the processes and methods of different Review Groups. This is largely a result of the differing Group scopes although, of course, it is also influenced by the backgrounds, beliefs and attitudes of the different editorial teams. The EPOC editorial team have developed methods and procedures that they consider to be appropriate for the kind of reviews that they produce. One of the key differences is that EPOC argue that it may not be feasible to evaluate many organisational or professional interventions in an RCT and that it may be necessary to consider other study designs. For example, although the introduction of small scale telephone consultation, such as in one particular GP practice, may be amenable to evaluation in an RCT the introduction of more widespread organisational changes (such as nationwide introduction of telephone consultation) maybe difficult to test in an RCT. Of course EPOC are not alone in including non randomised studies in their reviews. In Chapter 5 I have already described a systematic review of traffic calming that included controlled before and after studies (CBAs). However, unlike the traffic calming

⁹ <http://www.epoc.cochrane.org/en/scope.html>

review, the review in this chapter also includes Interrupted time series designs (ITS).

ITS are quasi-experimental studies that may or may not involve a control group and involve multiple data collection points before and after an intervention or natural event. The trend in pre-tests is then compared to the trend in post-tests. The problem with ITS, and the reason why many Cochrane Groups do not include them, is that they are more susceptible to bias than RCTs and controlled evaluations. For example, ITS may be subject to threats to internal validity such as maturation bias, where there is a pattern of improvement in the group prior to the intervention, or instrumentation bias where there are changes in the way records are kept or outcomes measured (EPOC, 1998). In addition, it is difficult to be sure if any change is a result of the intervention or whether it can be attributed to some other confounding factor.

In response to such concerns EPOC have developed detailed quality assessment criteria for non controlled studies. Some of the quality criteria for ITS are those that are applicable to other study designs, such as blinded assessment of primary outcomes and whether the primary outcome measures are reliable, but others are specific to ITS. For example, an assessment of whether there were sufficient data collection points before and after the intervention (EPOC guidance specifies at least three points before and three points after), whether the intervention is independent of other changes (e.g. changes to policy and practice which may have contributed to any change in outcome) and whether the data was analysed appropriately (formal test for trend using appropriate methods).

Interrupted time series can be classified as long time series, which require at least 20 observation points pre-intervention, and short time series which require at least three observation points pre-intervention and three post-intervention. The review presented here included three short time series. In general, the studies were felt to be of reasonable quality. In all three the intervention appeared to be independent of other changes, they had blinded assessment of the primary outcome and the number of data collection points before and after

the intervention was considered to be sufficient. One study was judged to be at risk of instrumentation bias as methods of data collection changed from manual to electronic recording after the start of the intervention. In addition, the data in this study was not analysed appropriately and the analysis was redone, by an EPOC statistician, using time series regression techniques. However, despite a generally favourable assessment of their methodological quality, the increased risk of bias associated with ITS means that their results need to be considered with some caution.

Types of outcomes

When deciding on appropriate outcomes to include reviewers need to consider whether they should include adverse effects and, if so, how they should go about this. Although this is relevant to all reviews it was particularly important to consider possible adverse outcomes for this review as concerns exist about the safety of telephone consultation, particularly when it is undertaken by nurses substituting for doctors (Crouch and Dale, 1998, Florin and Rosen, 1999, Salk et al., 1998).

Data on adverse events may be sparse, and indeed it was in this review, but the absence of information does not mean that the intervention is safe. Reviews of RCTs may not pick up adverse events as they may be too uncommon or too long term to be observed within RCTs (Higgins, 2008). Two of the RCTs included in the review reported mortality but both were underpowered for this outcome. It might be argued that if review criteria were broadened to include observational studies such as cohort studies or case-control studies they would be more able to detect adverse events. However, in an article looking at issues associated with assessing harmful effects in systematic reviews, McIntosh and colleagues caution against an unquestioning assumption that observational studies are the best source of harmful effects data. They suggest that the difficulty of interpreting observational study data may outweigh the benefits (McIntosh et al., 2004).

Analysis

As with a number of the reviews presented in this submission heterogeneity meant that it was not appropriate to pool studies in a meta-analysis; in this case there was heterogeneity in study design, interventions, outcomes and participating health professionals. Therefore, we presented a narrative and tabular summary of findings which, where possible, included an assessment on the size of the effect observed and statistical significance of the studies. For each study where possible we reported the main results in natural units in the results table and post-intervention differences and 95% confidence intervals or P values. For interrupted time series where possible we calculated a change in the level of outcome at the first point after the introduction of the intervention, and estimated a change in the slopes of the regression line (calculated as post-intervention minus pre-intervention slope).

Results of impact analysis

Knowledge production

Impact within research community

The citation analyses were conducted in January 2010. When I first ran the citation analysis the Cochrane Library version of the review had no citations in Web Of Science (WoS), two in Scopus and 48 in Google scholar. The version of the review published in the British Journal of General Practice had six citations in WoS, 11 in Scopus and 18 in Google Scholar. In the citation analysis for this chapter, as for some of the others, the citation count in Google Scholar was far higher than for the other databases. The discrepancy was particularly pronounced for the Cochrane version of the review; there were 48 citations in Google Scholar and only two in Scopus and none in WoS. Even taking into account the fact that Google Scholar includes grey literature whereas Scopus does not this difference between the two seemed very great. I wanted to explore this further to see what this disparity could be attributed to.

It has been suggested that the citation information in Google Scholar is flawed (Falagas et al., 2008) and it could, therefore, be that the citation count from Google Scholar is inaccurate. However, it has also been noted that citation counts for Cochrane reviews may be artificially low because citing authors have incorrectly referenced Cochrane reviews (The Cochrane Library, 2008). To investigate the accuracy of the citation information in Google Scholar I began by checking the 48 citing documents in Google Scholar to see whether they were correct. This involved looking at each of the citing papers or reports and checking that it had cited the Cochrane version of the review. Of the 48 citations 46 were found to be correct. I then checked to see which of those 46 papers or reports were indexed in Scopus and Web of Science. Of those 46 papers or reports from Google Scholar 20 were indexed in Scopus and 17 in Web of Science. Therefore, the total citation counts for the review should have been 22 for Scopus (there were an additional two citations in Scopus not picked up by Google Scholar) and 17 for Web Of Science. The results of both the original citation analysis and the revised analysis can be seen in Figure 8.1.

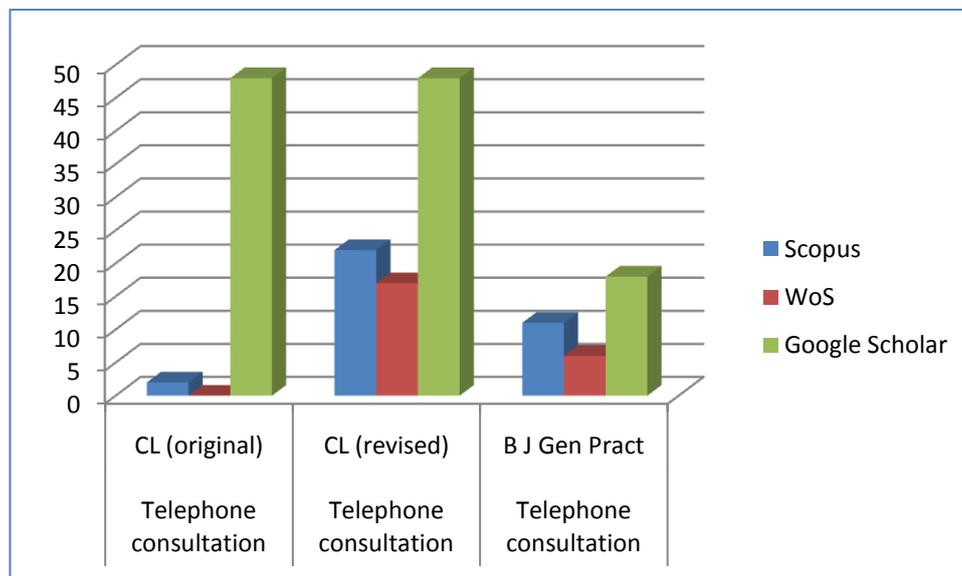


Figure 8.1 Results of original and revised citation analyses: telephone consultation review

Information from Wiley Interscience¹⁰ suggests that the review was of interest to the research and practice communities. Data on full text downloads was available for 2008 and 2009. In 2008 the review was accessed 681 times which gave it a world ranking of 841 out of 6,232, and in 2009 it was accessed 886 times which gave it a world ranking of 778 out of 6840.

Research targeting

Telephone consultation and triage continues to be an issue receiving attention on the research agenda. To what extent the review presented in this chapter may have contributed to this is difficult to gauge. However, a cluster randomised trial of telephone triage in general practice that has recently been funded under the NIHR Health Technology Assessment Programme cites the review in the study protocol (Campbell J, 2009). In setting out the rationale for their choice of intervention and outcome measures the authors refer to the review when they say 'the identification and choice of the most relevant outcomes have been a contentious issue in previous evaluations of triage systems in primary care'.

Informing policy development

Levels and type of policy development

In the payback model outputs from research are classified as primary or secondary (Wooding et al., 2004). Primary outputs are the direct outputs of the research project, such as academic publications or presentations, and secondary outputs are the wider impacts on policy and practice. Table 8.1 shows the details of the secondary outputs from the review (results from both papers are combined).

Once duplicate citations had been discounted there were 39 citations in peer reviewed journals and a further seven papers (most of which were not in English) where it was unclear whether it was a peer review journal or other type of publication. There was evidence that the review had had an impact both

¹⁰ Personal communication from Laura Sampson Associate Editor Wiley-Blackwell 26/2/2010

nationally, locally and internationally. The review was cited in national policy documents, including Scottish and Welsh policy documents, and in local policy documents in England. The review was also cited in a number of international reports and papers, some of which appeared to be feeding into policy. For example, a review of patient safety in primary health care produced by the Australian Commission of Safety and Quality in Health Care (Australian Commission safety and quality in health care, 2009).

Table 8.1 Secondary outputs from telephone triage review

Type of output	Number	Example/s
Peer reviewed article	39	British Medical Journal
Other type of article (e.g. not clear if peer reviewed journal, not English Language)	7	Bollettino dei medicine Svizzeri Australian Family Physician
National policy	3	Shifting the Balance of Care (NHS Scotland) Rapid review of hospital element of unscheduled care services in North Wales
Local policy	2	Strategy for commissioning of Urgent Care Services for NHS Birmingham and North East NHS Trust
National practice guideline	3	Rapid review of brief interventions and referral for smoking cessation (NICE 2006) Telephone advice lines for people with long term conditions (Royal College of Nursing)
Grey literature	8	Salisbury – An Evaluation of Advanced Access in General Practice (NIHR SDO 2007) Singh – Making the Shift (NHS Institute for Innovation and Improvement 2006)
Thesis	2	Swedish medical dissertation
Book chapter	2	Moore 2008 in From General Practice to Primary Care: The industrialization of family Medicine (Illiffe S)
Total	66	

Nature of policy impact

The citation of the review in various policy documents provided some evidence that the review had had a direct impact on policy. However, interestingly the review was interpreted differently in different documents with some using it to support telephone consultation and others using the review to suggest that the safety and efficacy of telephone consultation were unproven. For example, in the document shifting the balance of care by NHS Scotland (NHS Scotland) the review is used to support the use of tele-medicine and tele-health whereas in a rapid review for the National Public Health Services for Wales (Webb, 2009) the authors conclude that there is a lack of evidence to support telephone consultation. My own feeling would be that the latter interpretation would be more accurate and that, whilst it may be a promising intervention, data on some important outcomes, such as patient satisfaction and adverse events, is lacking. Whether the review had any impact at a more indirect level was hard to judge. However, this review is the only piece of work I have been involved in that is related to telehealth and it is possible that this makes it more unlikely that I would be aware of indirect impacts. For example, in areas where I have done more work, such as the prevention and treatment of injury, I am more familiar with the field and, therefore, more able to identify examples of indirect impact.

Type of policy

Black identifies three types of policy including practice policies (use of resources by practitioners), service policies (resource allocation and pattern of services) and governance policies (organisational and financial structures) (Black, 2001). He argues that the influence of research on the latter has generally been limited as decisions about governance are driven by many factors including ideology, value judgements, financial and economic considerations and political expediency. Changes to service policies may be influenced by evidence. However, whether the move to introduce nationwide telephone consultation, in the form of NHS Direct, was primarily driven by evidence or by ideology or economic or political expediency is debatable.

Research impact and the social construction of policy

Evidence is, of course, only one factor in determining policy. Numerous commentators have pointed out that diverse political, economic and social factors contribute to the formulation of health care policy. This systematic review of telephone triage provides an opportunity to examine more fully the interplay between research evidence, political and social expediency and professional agendas.

A variety of 'actors' including individuals, groups and organizations contribute to the policy process. Walt and Gilson (Walt and Gilson, 1994) depict this interaction in their policy analysis triangle which focuses on the content, context, process and actors involved in the policy process (see Figure 8.2).

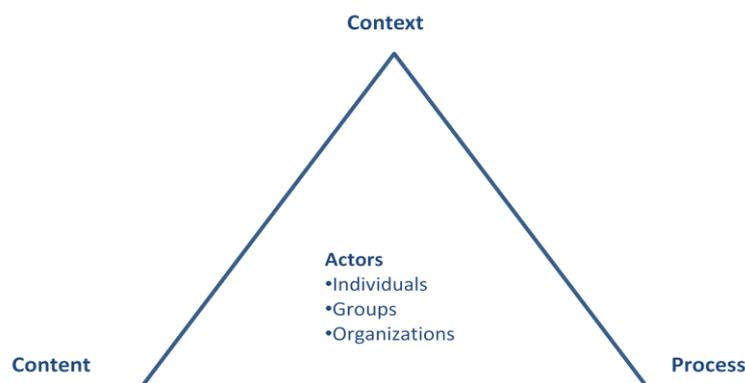


Figure 8.2 Policy analysis triangle (source: Walt and Gilson 1994)

Although this triangle is an over-simplification of the complex relationships and inter-relationships involved in the policy process it does provide a useful way of looking at the different factors that might affect policy. It shows how actors are at the centre of health policy. In the pluralist model of policy development power is dispersed throughout society and health policy emerges as the result of conflict and bargaining among large numbers of groups organized to protect the specific interests of its members (Buse et al., 2005). In the example we are focusing on here, NHS Direct, the main actors were the Labour Government, who saw NHS Direct as a key part of their modernising plan, and professional interest groups such as the medical and nursing professions. Many doctors were

opposed to the scheme, questioning its safety, effectiveness, and cost-effectiveness (Ferriman, 2005, Hayes, 2000), with the British Medical Association claiming that the service had expanded too fast before being adequately evaluated (Glasper, 2000).

Traditionally the medical profession have been a powerful interest group and health care systems have tended to be organized in deference to their preferences (Buse et al., 2005). Although the medical profession is still a powerful interest group in the UK, recent years have seen a number of challenges to their status. Some of these challenges, such as a growing recognition of the expertise of patients in relation to their own health (particularly for chronic conditions) and increasing the skills of nurses so that they can take over clinical tasks and responsibilities that were previously the remit of the doctor, were core components of the previous governments modernising agenda and NHS direct. Although some of the criticisms of NHS Direct from doctors may have been valid they may also have been opposed to NHS Direct because of the challenge it represented to their authority and their position as 'gatekeepers' of the NHS (Chiam, 2000). In such a context professional agendas might outweigh the desire for, or recognition of, the evidence however good it may be.

Although it is certainly true that many actors may influence the policy process, and in previous chapters I have provided examples of the power of industry and other lobbying groups, the role of the state cannot be ignored. Pluralism has been criticised as a model because it portrays the state as a neutral umpire arbitrating between the needs of other interest groups. In contrast public choice theorists suggest that the state is in itself an interest group which wields considerable power over the policy process, often to further the interests of bureaucrats and policy makers. Those espousing Elitism theories go further and argue that key economic and political decisions tend to be taken by a small elite in order to preserve the existing economic regime (Buse et al., 2005). In Chapter 1 I documented how the previous Government had given increasing credence to the notion of evidence-informed policy, a notion reinforced by a number of

policy documents (Cabinet Office, 1999a, Cabinet Office, 1999b). However, in reality, political and economic expediency may, and I would suggest often does, take precedence over evidence.

The way many major changes to service delivery and organisation come about illustrates how political and economic factors and intellectual fashion are often far more powerful than evidence. Although frameworks for the design and evaluation of complex interventions suggest that evaluation should be sequential, moving from theory to modelling, explanatory trials, pragmatic trials, and ultimately long term implementation (Campbell et al., 2000) in reality this sequence is rarely followed. The Government often introduce new services before evaluation can take place and subsequent evaluations may have to use unreliable methods such as uncontrolled before and after studies (McDonnell et al., 2006). When conducting the review of telephone consultation the only studies we found evaluating NHS Direct did not meet our inclusion criteria even though these had been broadened to include controlled studies and ITS's. This use of research to reinforce and support positions that Government have already adopted is in line with ideas proposed by Weiss in her strategic model of the policy process which was discussed in Chapter 1 (Weiss, 1979). Of course once a service, such as NHS Direct, is widely established not only is it more difficult to evaluate but it is difficult to withdraw the service without substantial reorganisation and disruption of other services (McDonnell et al., 2006).

Conclusion

I did find examples that the review had had some impact on policy, both in the UK and internationally. Interestingly the review had been interpreted and used in different ways with some policy makers using it to support the use of telephone consultation and others suggesting that the safety and efficacy of telephone consultation is unproven. This suggests that evidence may be used to reinforce positions already adopted. There was no evidence the review had impacted on policy relating to NHS Direct. However, this was not particularly surprising as NHS direct was well established before we conducted the review,

and the evaluations of NHS Direct that we found were uncontrolled studies and did not meet our inclusion criteria.

The impact of the review should be considered in light of the political and socio-economic context of the time. In previous chapters I have discussed how complex social, cultural, political, economic and ideological factors may contribute to the development of health policy (Davis and Howden-Chapman, 1996). The way in which services are organised and delivered may be even more subject to such factors than a more discrete medical intervention. For example, in the case of an intervention like telephone consultation it is easy to see that political and economic factors may take precedence over research evidence.

Chapter 9: Discussion & Conclusions

Introduction

The aim of this study was to examine the extent to which systematic reviews can influence the development of health care policy in England, to identify factors that might be important in increasing the impact of reviews and to examine how researchers can produce systematic reviews that meet policy makers' needs without sacrificing methodological rigour. Integral to this submission are ten of my own previously published systematic reviews (15 reports or papers). These have been used as illustrative examples to investigate the questions addressed in this study. The reviews cover a range of health and public health related topics. They provide information on:

- Fluid resuscitation for critically ill patients (Alderson et al., 2000, Bunn et al., 2000b, Bunn et al., 2004b, Bunn et al., 2008b, CIG Albumin Reviewers, 1998, Kwan et al., 2003)
- The prevention of road-traffic injuries (Bunn et al., 2003a, Bunn et al., 2003b, Duperrex et al., 2002a, Duperrex et al., 2002b)
- Barriers to participation in fall prevention interventions (Bunn et al., 2008a)
- The prevention of sexually transmitted infections and teenage pregnancies (Bunn et al., 2006a)
- The effect of telephone consultation on service use and patient satisfaction (Bunn et al., 2004a, Bunn et al., 2005)

Systematic reviews by nature are collaborative endeavours that ideally involve more than one author in order to ensure that the review team has both methodological and clinical or subject expertise. In addition, 'gold standard' methods for reviews (Higgins, 2008) stipulate that data extraction and critical appraisal should be undertaken by two reviewers independently. The particular contribution that I made to each review is documented in Chapter 3. Although the level of my input varied between reviews, in general I have chosen to use

reviews to which I made a significant contribution. I am first author on six of the ten reviews included. This final chapter includes:

- An overview of the results of the impact-analysis
- An examination of the factors that might play a role in the extent to which systematic reviews can influence policy
- An analysis of how the results contribute to knowledge about barriers and facilitators
- An examination of how researchers can produce systematic reviews that meet the needs of policy makers but are still methodologically rigorous
- An overview of the methodological approach and the strengths and limitations of the study
- Conclusions about the implications of the study including ramifications for systematic reviewers and future research

Overview of the impact analysis

A primary aim of this study was to examine whether a cohort of my own previously published work had had any influence on the development of health care policy. In this section I present an overview of impact and provide some evidence to suggest that systematic reviews can influence policy. The results of the impact evaluation are structured using a framework that combined domains from two separate frameworks, the Payback model and the Research Impact Framework. The framework includes the domains: knowledge production, research targeting, informing policy development and impact on practice. The outcome of interest in this study was any evidence of direct or indirect impact on policy, with policy making understood in a broad rather than narrow sense.

I adopted a conceptual approach to the study which was based on Weiss's work on the relationship between research and policy (Weiss, 1976, Weiss, 1977, Weiss, 1998). In her work she makes the distinction between impact that is direct, i.e. research is instrumental in driving policy making, and that which is

more indirect, for example involving the use of research to mobilise support, influence concepts and language, or question established practices and beliefs. This conceptual approach is generally accepted by policy scholars as an accurate reflection of the way research outputs feed into the policy process. A summary of the findings from the impact analysis in Chapters 4-8 can be seen in Table 9.1. It is structured to reflect the domains of the framework used for the analysis and includes any evidence of impact on knowledge production, including total number of citations in Google Scholar for each review; any evidence of research targeting; levels of policy making, including the number of guidelines and policy documents citing the paper; the type of policy; and the nature of the policy impact, for example evidence of direct or indirect impact, and any evidence of impact on practice.

Knowledge production

The papers were published over a ten year period, between 1998 and 2008. Of the ten reviews included in the analysis nine were published in peer-reviewed journals, the other (Bunn et al., 2006a) was published on the website of the National Institute for Health and Clinical Excellence. Some of the most frequently cited papers were the five reviews on fluid therapy (Chapter 4). All had an impact on the research community although this was particularly significant for two (Alderson et al., 2000, CIG Albumin Reviewers, 1998). The least cited of all the papers was the review of qualitative studies (see Chapter 6) looking at barriers and facilitators to the uptake of fall prevention interventions (Bunn et al., 2008a). However, this was also the most recently published paper.

Table 9.1: Summary of main impacts on health care policy

Review and year first published	Knowledge production	Research targeting	Levels of policy making	Type of policy	Nature of policy impact (e.g. direct or instrumental or indirect or conceptual)
1. Colloid solutions 2000	Outputs Published in CDSR Citations GS= 118	No evidence of research targeting identified	Impact largely at level of practice guidelines. Guidelines/policy documents UK=5 Non UK = 5 International =1	Influence of reviews appears to be largely related to practice policies (e.g. use of resources by practitioners)	Instrumental use: informing the development of practice guidelines (e.g. by NICE and HTA) Indirect impact of review unclear
2. Hypertonic vs. Isotonic 2000	Outputs Published in CDSR Citations GS = 104	No evidence of research targeting identified	Impact largely at level of practice guidelines. Guidelines/policy documents UK =2 Non UK =2 International =0	Influence of reviews appears to be largely related to practice policies	Instrumental use: informing the development of practice guidelines (e.g. by NICE and HTA) Indirect impact of review unclear
3. Albumin 1998	Outputs Published in CDSR/BMJ Significant impact in research community Citations: CDSR version received 177 citations in GS	Review played a role in the initiation of large multi-centre RCT conducted in Australia	Guidelines/policy documents UK=7 Non UK = 8 International =2	Influence of reviews appears to be largely related to practice policies	Instrumental use: informing the development of practice guidelines (e.g. by NICE and HTA) Direct impact on practice with 40% reduction in use of human albumin in UK; also evidence of reduction in use in other countries Indirect impact by stimulating

	BMJ version received 579 citations in Scopus (data from GS not available)				debate and changing thinking and established practices around fluid resuscitation
4. Timing and volume 2001	Outputs Published in CDSR Citations: GS=114	No evidence of research targeting	Impact largely at level of practice guidelines. Guidelines/policy documents UK =2 Non UK =2 International = 2	Influence of reviews appears to be largely related to practice policies	Instrumental use: informing the development of practice guidelines (e.g. by NICE and HTA) Indirect impact of review unclear
5. Colloids vs crystalloids 1998	Outputs Published in CDSR/BMJ Significant impact in research community Citations: GS=359/551	Review suggested albumin might be harmful which led to CIG undertaking review focusing on albumin.	Guidelines/policy documents identified UK =8 Non UK=7 International=2	Influence of reviews appears to be largely related to practice policies	Instrumental use: informing the development of practice guidelines (e.g. by NICE and HTA) Indirect impact by stimulating debate and changing thinking and established practices around fluid resuscitation
6. Traffic calming 2003	Outputs Published in CDSR/Inj Prev Citations: GS=54/46	No evidence of research targeting identified but may have impacted on research thinking in road safety.	Guidelines/policy documents identified UK=4 Non UK=5 International=5	Greatest evidence of impact was found on international policy in form of WHO road safety guidance. Review fed directly into development of policy recommendations which were endorsed by member states through UN resolution.	Instrumental use: informing the development of international policy Indirect impact: may have contributed to a shift in road safety paradigm (e.g. highlighting need for more population based strategies)

7. Pedestrian safety education 2002	Outputs Published in CDSR/BMJ Citations: GS= 54/95	No evidence of research targeting identified but may have impacted on research thinking in road safety.	Guidelines/policy documents identified UK=3 Non UK=3 International=3	Informed the development of EU and WHO policy documents	Indirect impact: may have contributed to a shift in road safety paradigm (e.g. highlighting lack of evidence to support road safety education for children as a road safety strategy)
8. Fall prevention: Barriers/facilitators	Outputs Published in Age & Ageing Citations: GS = 5	Review informed the design and conduct of primary study	No evidence of impact at any level of policy. Not cited in any guidelines/policy documents.	No evidence of impact on policy	No evidence of direct or indirect impact
9. Contraceptive advice 2006	Outputs Published on NICE website Citations GS= 0	No evidence of research targeting identified	Review directly informed practice guidelines for England and Wales. No evidence of international impact.	Directly informed the development of NICE guidance	Instrumental use: clear link between review and development of NICE guidance Unclear if NICE guidance has been implemented or impacted on practice
10. Telephone consultation 2005	Outputs Published in CDSR/BJGP Citations: GS =55/23	Review informed the design and conduct of primary study	Guidelines/policy documents identified UK=8 Non UK=0 International=0	Little evidence of impact on policy	No evidence of direct impact on policy or practice. Indirect impact of review unclear.

CDSR= Cochrane Database of Systematic Reviews, GS = Google Scholar

Eight of the reviews were published in the Cochrane Library, five of which were also published in an additional peer review journal. On first inspection it appeared that the Cochrane versions were cited far less than the versions published elsewhere, for example the BMJ albumin paper was more highly cited than the Cochrane version. This discrepancy may be a reflection on the prominence, accessibility or readability of Cochrane reviews in comparison to those published in other journals. However, it may also be because citation counts for Cochrane reviews are artificially low because they have been incorrectly cited (The Cochrane Library, 2008). My investigations (see Chapter 8) substantiate the latter theory. I found that the citation information for Cochrane reviews in Scopus and WoS is highly flawed and does not reliably reflect the impact of the papers on the research community.

The impact of the Cochrane reviews within the research and practice community was further assessed by an analysis of the frequency with which the reviews were downloaded. Data for 2008 and 2009 can be seen in Table 9.2. Although there was considerable variation between the reviews the data do suggest that all of the reviews have been of interest to decision makers. The figures, however, only give an indication of the popularity of the reviews in comparison with other Cochrane reviews and do not provide any information on who downloaded them or how they were subsequently used. Interestingly despite the fact that the albumin review had had the highest overall citation figures of any of the reviews it is now downloaded less than some of the other fluid therapy reviews. It is possible to speculate that this recent lack of interest is because the review, and subsequent RCT on the subject, have affected practice to such an extent that albumin is no longer widely used in clinical practice.

Table 9.2 Number of downloads for Cochrane reviews

Review	Downloads 2008	Ranking 2008 (from a total of 6,232)	Downloads 2009	Ranking 2009 (from a total of 6840)
1. Colloid solutions	1516	170	1387	327
2a. Albumin	1040	375	1653	632
3. Hypertonic vs. Isotonic	650	910	538	1620
4a. Colloids vs crystalloids	2973	18	3291	35
5. Timing and volume	691	823	694	1146
6a. Traffic calming	812	620	334	2666
7a. Pedestrian Safety Education	68	5105	214	3693
9a. Telephone Consultation	681	841	886	777

Research targeting

I also assessed whether research had contributed to the setting of future research agendas. Such 'research targeting' is not always easy to determine or prove but there was one example of a systematic review leading to further research. The human albumin review discussed in Chapter 4 (CIG Albumin Reviewers, 1998) was a significant driver in a subsequent large RCT undertaken in Australia (Finfer et al., 2004b).

Informing policy development

The main purpose of the impact evaluation was to assess whether the reviews had influenced the formulation of health care policy in any way. I found a number of examples where the reviews had appeared to influence policy. In particular the reviews had had an impact on the development of policies in the

form of guidelines at a local level or for professional bodies, with nine of the ten reviews contributing to the development of practice guideline (see Table 9.1). There was also some evidence of impact on Government policy with six reviews cited in NICE guidelines, and one used to develop NICE guidance. In addition, a number of the reviews had impacted on international policy. This was particularly the case for the road safety reviews (Chapter 5) that had fed directly into the development of policy recommendations by the WHO (Peden et al., 2004b), recommendations which have been endorsed by member states through United Nations General Assembly resolutions. This is especially notable because, in an examination of the development of recommendations by WHO departments, Oxman and colleagues found that systematic reviews were rarely used for developing recommendations but instead the process usually relied heavily on experts (Oxman et al., 2007).

Levels and type of policy impact

This evidence relating to the impact of the reviews should be considered in the context of how, where and at what level we can realistically expect reviews to support the development of health care policy. In theory systematic reviews might support the policy process in a number of ways, including setting the policy agenda by providing evidence that a problem exists, giving examples of the impacts of policies on people and organisations, evaluating policy initiatives and providing feedback from experiments (Whitehead et al., 2004). However Lomas suggests that the functional role of research in setting agendas may be relatively small because civil servants dislike inputs being pushed at them (Lomas and Brown, 2009). In addition, some levels of policy may be more difficult to influence than others. 'Macro' policies emanating from the Department of Health may be particularly difficult to influence as they tend to be the result of a closed system of decision making (Walt, 1994). Indeed, much of the impact of the reviews in this study was in the development of a range of 'micro' policies, such as guidelines for professional bodies. The reviews which appeared to have the most influence were those related to clinical practice questions around appropriate fluid resuscitation strategies. This is perhaps unsurprising as it has

been suggested that some policy and practice areas lend themselves more readily than others to being informed by research. For example, it may be easier for research to influence the development of practice policies, such as the use of resources by practitioners, rather than governance policies that relate to organisational and financial structures.

It has also been suggested that we have misguided expectations about the extent to which scientific evidence can influence behaviour (Green et al., 2009, Weiss, 1979). Green argues that the literature on diffusion and dissemination tells us that people, whether policy makers, practitioners or the public, will filter the information they receive, and selectively choose information that fits with their perceived needs, priorities, and circumstances (Green et al., 2009).

Nature of policy impact

In a recently published study exploring ways of measuring research utilization Estabrooks and colleagues attempt to clarify the construct of research utilization and identify observable indicators of research use (Estabrooks, et al 2011). Although their focus is on research utilization by practitioners, in particular the use of research in nursing practice, some of their work is applicable to the use of research by policy makers. Like others, they construct research use in terms of instrumental, conceptual and symbolic impact and suggest that the use of evidence to develop or update educational material, policy and guidelines is an indicator of instrumental research impact. Similarly in this study I took inclusion of one of the reviews in a guideline or other policy document as an indicator of instrumental research impact, and, as previously reported found a number of examples of such impact.

However, many have argued that much of the impact of research on policy is likely to be conceptual or symbolic rather than instrumental (Black, 2001, Lavis et al., 2003b, Lomas and Brown, 2009, Weiss, 1976) and it is well documented that such indirect impacts are difficult to determine. Estabrooks and colleagues suggested that questioning of current practice, revising ideas and engagement with relevant research are all indicators of conceptual use of research. In

Chapter 4 I suggested that there was evidence that the reviews relating to human albumin and colloids versus crystalloids created debate and discussion about established practices and beliefs. Indeed, the high number of citations associated with these reviews is potentially an indicator of their importance in this debate. In addition, the reviews in Chapter 5 may have influenced thinking about the evaluation of road safety interventions and been used to mobilise support in favour of environmental and legislative interventions. However, in both cases this indirect impact is difficult to substantiate. For the remaining reviews I found no discernable indication of conceptual or symbolic impact.

Impact on practice

The main focus of the evaluation was the extent to which the reviews had impacted on policy. However, there was an obvious example of a review impacting directly on practice. The albumin review (CIG Albumin Reviewers, 1998) had a significant impact on practice leading to a 40% decrease in the use of human albumin in the UK and documented falls in other countries. Interestingly, practitioners made changes to their clinical behaviours well before the protracted deliberations of policy makers had been concluded. In this instance improvements in health care delivery were achieved directly in response to the review rather than because of a policy directive. However, policy in the form of guidelines may have subsequently played a role in further embedding these changes into clinical practice.

Factors that play a role in the extent to which systematic reviews can influence policy

It is apparent from the results of the impact evaluation that some reviews have had a greater impact than others. In this section I draw upon findings from the impact evaluation and current literature to examine the factors that play a part in the influence of systematic reviews on policy.

Dissemination

One factor that may have contributed to the degree of impact that different reviews have had was the mode of dissemination. Most of the reviews were disseminated traditionally via publication in peer-review journals. Eight reviews were published on the Cochrane library and, although Cochrane reviews have been found to be of comparable or better quality to those published in print journals (Jadad et al., 1998), the format may not be particularly user friendly (Grimshaw, 2004). Many of the reviews are lengthy and decision makers may find the detailed tables, forest plots, figures and quality assessment information difficult to interpret. Of course reviews in print journals, while generally shorter, are also often full of academic language and tables and statistics that policy makers may find daunting (Lavis et al., 2005). It is possible that the impact of the reviews in this study may have been increased by more active or user-friendly dissemination strategies. From the overview of the literature on the dissemination of reviews (Chapter 2) it was clear that the inclusion of short summaries of review findings is particularly important for policy makers (Dobbins et al., 2004b, Lavis et al., 2005). However, few studies compared different review formats and it is clear that more work is needed on the development of appropriate methods for disseminating reviews. Although the Cochrane Collaboration has made efforts to improve dissemination, such as through lay summaries and podcasts, more work is needed to improve the readability of Cochrane reviews.

The dissemination of research results through the media appeared to be a significant factor in determining the degree to which the reviews influenced policy. Gold talks about how rare it is for a single paper to make an impact on its own (Gold, 2009). When it does she describes this as a 'big bang' where the results of a single study frame the debate in new ways, or drive initiatives. She argues that in reality most studies by themselves will not have the scope, robustness or focus for a 'big bang' but suggests that publication in a prestigious journal can generate press coverage that propels the communication and uptake

of key messages from research. Of the ten reviews included in this study there was one example of a review making a 'big bang'. This was the study on human albumin (CIG Albumin Reviewers, 1998), and the fact that it was widely reported in the media both in the UK and internationally undoubtedly increased the impact of the review. The 2009 update of the Cochrane traffic calming review was also picked up by a number of media outlets. This was largely because the publication of the update was accompanied by a press release. Of course it is difficult to quantify the influence of a press release. It is generally only reviews that are considered 'newsworthy', for example showing clear evidence of benefit or harm, that tend to get accompanied by a press release and such reviews may be more likely to have an impact anyway. However, the updated review received more press coverage than the original version that had no press release.

Despite the undoubted influence of the media some commentators have questioned whether traditional media still has the same power to spark trends and create opinion. Green suggests that the explosion of the Internet and the saturated media environment have generated a new type of audience that is more critical of mass media and more reliant on other people's opinions at the interpersonal level (Green et al., 2009). In such an environment innovations in connectivity such as two way electronic devices, text messaging, blogging and twitter may be as influential as more traditional forms of media. Researchers should explore how such technology can be used to communicate the findings of systematic reviews in a format that would make them more accessible for decision makers.

Drivers behind review question

The drivers behind a review may also determine how influential reviews are; including who initiated the question and why, the type of question being asked, the perceived importance of the question and whether it fits with the values and ideologies of policy makers. It has been suggested that researcher-initiated syntheses run the risk of addressing questions that are of interest to other researchers rather than policy makers (Gold, 2009). Eight reviews were

Cochrane reviews, seven of which were undertaken while working for the Cochrane Injuries Group. The set up of the Cochrane Collaboration has meant that Cochrane reviews have typically reflected the interests of individual researchers rather than being priority-driven (Grimshaw, 2004). In this instance the choice of questions for the Injuries Group reviews (Chapters 4 & 5) was influenced by a prioritisation process undertaken by the group when it was first established. This process drew on the expertise of an international group of researchers and practitioners but, for the most part, did not include policy makers. The impact of many of these reviews indicates that such a collaborative prioritisation process can be a successful way of identifying review questions of importance to decision makers and practitioners. Of the other three reviews, two were specifically undertaken to inform the development of primary research (Bunn et al., 2005, Bunn et al., 2008a) (Chapter 6 & 8) and one was initiated by policy makers to inform the development of NICE guidance (Bunn et al., 2006a) (Chapter 7). Although the latter was initiated by policy makers its impact on practice may have been limited because the recommendations did not fit well with current practice and service organisation.

Review methods

The reviews in this study incorporate a range of study types and involve a variety of techniques for analysis including meta-analysis, narrative presentation of results and thematic analysis of qualitative studies. In general, however, the majority are conducted according to Cochrane principles and primarily ask questions about 'what works' with controlled studies prioritised over other study designs. The exceptions to this were the review commissioned by NICE, in which I looked at context and applicability and attempted to answer questions about 'what works for whom in what circumstances' (Bunn et al., 2006a, Bunn et al., 2008a), and the review of qualitative studies exploring attitudes to fall prevention interventions (Bunn et al., 2008a).

The choice of study types to include in each review was affected by a number of considerations including rigour, suitability and the complexity of the question.

For example, although it was appropriate to restrict clinical questions about fluid resuscitation to RCTs, more complex public health questions, such as the review concerning sexual health, may need to involve more diverse study types including qualitative studies. Despite the inclusion of non randomised studies in a number of the reviews, my approach may still be considered by some to be overly positivist and biomedical (Morgan and Ziglio, 2007). However, I would argue that although reviewers may need to look beyond questions about what works it is still of fundamental importance to establish the effectiveness of interventions. We should not be neglecting the question 'what works' but in some instances should also be asking additional questions, such as 'what works for whom and why'.

Barriers & facilitators to research impact

Policy makers and determinants of research use

There is evidence to suggest a positive attitude towards research is an important determinant of research use (Estabrooks et al., 2003). In Chapter 2 I reviewed a number of studies that explored policy makers' attitudes towards the use of systematic reviews. The extent to which policy makers knew about or valued systematic reviews as a source of evidence was the subject of some discrepancy. However, overall it appeared that policy makers were not always familiar with systematic reviews and did not necessarily place a high value on systematic review evidence. This is likely to be a barrier to research impact. There are, of course, other potential determinants of research use and in Chapter 7 I explored some of the factors that might influence the use of guidelines by health care professionals. The nature and strength of the evidence has been found to be a facilitator for research use (Dopson et al., 2001) and this might explain why some of the reviews in this study had a greater impact than others. In addition, the cultural and organisational environment in which policy makers operate is likely to influence their use of research (Grimshaw et al., 2004, Rycroft-Malone et al., 2004a). In Chapter 2 some of the studies that reported the most positive attitudes towards systematic review evidence were conducted in Canada where

there have been considerable efforts to develop a policy-making culture that values research evidence.

Social networks

An important factor in determining impact may be the interactions and connections between researchers and other actors in the policy process. There is evidence to suggest that the social networks of researchers play a vital role in the communication and dissemination of research (Gray, 1973, Rogers, 1983) and the degree of impact. Although there is a variety of definitions and ideas about what constitutes a network, most agree that it involves groups of actors linked together, either loosely or closely structured, but capable of working together collectively (Walt et al., 2008). Although initially policy analysis focused on the role of the state, more recent work acknowledges the involvement of a much larger number of actors in the policy process with policy decisions less top down and taking into account the values and beliefs of expanded networks (Buse et al., 2005, Hajer and Wagenaar, 2003).

Although difficult to quantify, the impact of some reviews may have been facilitated by informal and formal social networks. For example, links between the Cochrane Injuries Group and the Unintentional Injury Prevention (UIP) Team at the WHO may have increased the impact of the road safety reviews on WHO policy. These links may be stronger because the UIP team at WHO operates from a similar epistemological position (many researchers are epidemiologists) to the CIG and it may be easier to disseminate research within networks where messenger and audience are similar. Indeed, some organizations may be more 'permeable' to outside knowledge because of their institutional philosophies, the positioning of influential staff members, or financial strength (Greenhalgh et al., 2004). In addition it should be noted that the Cochrane Collaboration is itself an influential organization involved in many networks in the UK and internationally.

Knowledge transfer and exchange strategies

Although many networks involving researchers are informal and not specifically intended to facilitate dissemination, there is an increasing emphasis on the development of more formal networks whose primary aim is the transfer of research knowledge to decision makers. The adoption of such knowledge transfer and exchange (KTE) strategies appears to make inherent sense. Gold says that networking and trust are important in knowledge transfer and so researchers 'willing to engage in communications that go beyond publication are likely to find a greater payoff' (Gold 2009, p1131).

Chapter 2 provides some evidence to support the use of KTE strategies to improve the dissemination and uptake of systematic reviews. However, much of the research in this area is anecdotal or descriptive and the benefits of KTE strategies need to be further established in robust evaluations. There is also some concern that the adoption of collaborative or interactive approaches might compromise the scientific rigour of research (Innvaer et al., 2002, Keown et al., 2008). I was unable to draw any conclusions about the benefits of KTE from the impact evaluation in this study. This was because only one review, the review conducted to inform NICE guidance, included any formal KTE strategies. This KTE was led by the policy makers at NICE rather than initiated by the researchers and it is not clear what impact this had on dissemination. It is possible that the impact of other reviews might have been greater if more formal KTE strategies had been employed.

A key consideration is that substantial investment may be required in order to understand the audience and its needs', build credibility, develop actionable audience-specific messages, and continue the exchange of ideas (Lavis et al., 2005, Lavis et al., 2002). For KTE to be effective activities may need to be sustained for a considerable length of time which may be costly and beyond the resources of many researchers. Although researchers have a responsibility for KTE this must be shared with policy makers, and funding bodies need to be aware of the resources required to support such initiatives. In addition

researchers may need to be supported in the development of appropriate KTE skills.

Meeting the needs of policy makers whilst maintaining methodological rigour

Throughout this study I have grappled with the question of how we produce reviews that are useful to policy makers whilst still methodologically rigorous. Of course 'rigour' may mean different things to different people and my idea of rigour may well differ from someone coming from a less 'positivist' background. From my perspective, rigour means adhering closely to Cochrane standards for evaluating effectiveness. However, I am aware of the limitations of such an approach and some reviews, particularly those asking complex public health questions, may require a broader approach; for example incorporating qualitative research.

I discussed earlier how a consideration of context, for example does an evaluation work in particular settings and for particular populations, may be important in order to make reviews relevant to policy makers. However, asking more questions and including a greater diversity of study designs within a review increases the complexity of the task that in turn means that more resources are required. Increasing the cost and time taken to do a review has implications for whether the review might impact on policy. It has been suggested that there are only small windows of time where research may impact on the policy process (Kingdon, 1984) and the window of opportunity may have passed by the time a synthesis is commissioned and completed (Gold, 2009). One alternative is for researchers to do a synthesis of previous reviews in a 'review of reviews' (Bunn et al., 1999, Bunn et al., 2000c, Trivedi et al., 2007b). This is an increasingly viable option as the burgeoning interest in systematic reviews means there is now a substantial reservoir of reviews to draw upon. However, it is only feasible if previous good quality reviews have addressed the question of interest. The resource constraints under which policy makers often operate have also led to the suggestion that some parts of the review process could be made less time

consuming. For example, there is some interest in more iterative methods for searching (personal communication NICE). In addition, there is scope for further work on methods for qualitative reviews and whether 'sampling' of the literature, rather than including all available studies, might be appropriate. Such initiatives will need to be thoroughly evaluated to ensure that we are not sacrificing methodological rigour in favour of expediency.

Methodological approach and limitations of the study

Much previous work on the impact of research on policy is reflective rather than documenting empirical examples (Boaz, 2008c) and there are few impact evaluations that focus on the impact of research on policy. In addition, although received wisdom suggests that systematic reviews are useful tools for policy makers, there is a lack of data to support this. In this study I have explored these issues further by taking methods previously used for evaluating the impact of research and applying them to systematic reviews. I found only one other example of an impact analysis concerned with systematic reviews (Soper and Hanney, 2007). This study, therefore, provides further information on the application of impact evaluation methods for systematic reviews. My analysis also differs from many in that the focus is on specific outputs in the form of published papers rather than the impact of whole projects or programmes.

There are a variety of methods available for evaluating the impact of research. I used primarily quantitative methods, including bibliometrics and literature and documentary review. I chose this approach because they are methods that could be used for comparison across reviews, are transparent and reproducible, and are suitable for retrospective evaluations. There may, however, be an inherent contradiction in the use of quantitative methods such as bibliometrics to measure the messy, complex process of policy making and it has been suggested that qualitative methods such as observations and semi-structured interviews may generate useful descriptive and explanatory data for examining the policy context and the pathways to research use. However, as methods they run the risk of being anecdotal and subjective and it can be difficult to generalise their

findings (Boaz, 2008a). Although I conducted some informal interviews with co-authors, and with other colleagues who might be aware of evidence of impact, this was limited as I already had the 'insider account'. Interviews with policy makers might have added useful additional information to the study but were not considered feasible given the breadth of subject areas covered in the analysis. Moreover, interviews would be subject to recall bias and it is unlikely that it would have been possible to identify individuals involved in policy decisions relating to specific reviews, some of which were published over a decade ago. High turnover of policy makers is a problem even for prospective evaluations (Dobbins 2009).

Although a number of frameworks for structuring research impact exist none were entirely suitable for the purposes of this study. Therefore, I took components from two established frameworks (the Payback Model and the Research Impact Framework) to create a framework for this study. This had the advantage of using a well respected framework (The Payback Model) but making it more relevant for a study focused on policy.

Limitations

Limitations of methods for impact analysis

Evaluating the impact of research is complex and difficult and there are a number of methodological issues that might have a bearing on the validity of the results of this study. At present there are no agreed instruments or methods for determining impact (Peckham et al., 2008) and many of the methods currently used focus on outputs rather than outcomes or processes. My primary outcome was impact on policy and although I took citation in a policy document as an indicator of influence on policy, this was a proxy for impact and, furthermore, can only be considered an intermediate outcome. It does not tell us if a policy was implemented. However, the implementation of policy is difficult to measure and the farther you travel away from the policy or policy document the harder it is to attribute such changes to specific pieces of research. In addition, although

the focus of my research was on policy impact that, in itself, is only ultimately of interest if that policy leads to improvements in patient or system related outcomes. Such outcomes were beyond the remit of this study.

Limitations of bibliometrics and documentary review

Another potential limitation of this study is that the methods I chose to use have not been developed for the specific purpose of evaluating the impact of research on policy (Boaz, 2008a). It could be argued that as bibliometrics and documentary review do not involve enquiry in the real world of policy making such methods fail to capture the complexities of the policy making process.

Indeed, there are a number of limitations associated with the use of bibliometrics to assess the impact of research in general. A major criticism of bibliometric techniques is that they tend to focus on quantity rather than quality, measuring the number of research outputs rather than research outcomes or impact (Boaz, 2008a). In my analysis I took citation in a policy document (such as guidelines) as an indicator of evidence of impact. However, as I was generally not party to the process by which such guidelines were developed it was unclear how significant a role the reviews played in the guideline development process and what influence they might have had in relation to other competing factors.

Another possible limitation of the citation analysis in this study was that I focused on whether a paper was cited rather than the context in which it was cited. There are many reasons why people cite work and, as well as acknowledging the contribution of previous work, the purpose of the citation may be to dispute the claims of previous work or criticise it (Garfield, 2006). For example, many of the papers citing the albumin review were critical of the methods of the review or conclusions drawn. In addition, citations may be distorted or manipulated so that the findings are not used as the researchers originally intended (Greenberg, 2009, Hoyt and Garrison, 1997).

Risk of bias

Another potential threat to the validity of this study is the risk of bias. This was a retrospective analysis and as such may be at greater risk of bias than one where data is collected prospectively. Evidence from the citation analysis and documentary review was supplemented by my own recollections of instances of impact. Whilst information provided by researchers themselves may provide examples of impact not available through documentary review or bibliometric analysis (Bunn and Kendall, 2011, Kalucy et al., 2009) such information may be subject to recall bias. In addition, the risk of bias in an impact evaluation is greater when the researcher, as is the case in this study, is directly investigating her own research, and it has been suggested that strong evaluations would include elements of both internal and external review (Hovland, 2007). In reality, however, most impact evaluations are carried out by those with a vested interest and, as such, are subject to the risk of bias. Furthermore, internal evaluations have the advantage of the 'insider account' which means they may have greater access to information that allows them to unravel the complexity of the policy making process (Walt et al., 2008).

The direction of travel is another potential source of bias in an impact evaluation. I tracked forward from research papers and, although this has the advantage of a more tightly defined focus, which may help to identify any contribution that the research has made, such an approach may be more likely to inflate the impact of research than tracing backwards from policy (Hanney et al 2007). Hanney and colleagues suggest that starting with research may imply a linearity that obscures the complex reality of how policy changes occur. However, they acknowledge that it is difficult to see how the approach of tracing backwards could be applied to diverse research programmes (Hanney, 2007).

The papers included in the impact analysis were self-selected and do not represent all the systematic reviews I have authored. Indeed, the choice of two of the papers (Alderson et al., 2000, CIG Albumin Reviewers, 1998) was governed by the knowledge that they had had an impact, and would, therefore, make good

case studies. However, although I have not included all my publications, those that I have chosen represent a substantial body of my published systematic review work (see Appendix 1 for a full list of publications). The breadth of systematic reviews included in this study was both a strength and a limitation. Whilst, the range of topics covered, allowed for wide-ranging consideration of potential barriers and facilitators, and of the extent to which such barriers might be subject specific or more generic, the scope of the evaluation limited the choice of data collection methods.

In this study I have compared the impact of a variety of reviews. However, direct comparisons of the reviews are problematic for a number of reasons. One consideration is the variation in citation windows. The reviews were published at different times over a ten year period and those reviews published recently may not yet have had time to impact on policy. In his work on the policy cycle Sabatier notes that it may take up to a decade for the full impact of research to be apparent (Sabatier, 1988). In addition, citation volume typically peaks in the third or fourth year post-publication and, therefore, a window of five years has been suggested as most appropriate for research assessment (Ismail et al., 2009, van Leeuwen et al., 1999). Furthermore, the diverse nature of the review topics means that one should be cautious when comparing the impact of reviews because, as I have previously discussed, some policies and policy-making processes may lend themselves more readily than others to being informed by research (Lavis et al., 2002).

It has been suggested that comparative assessment of scientometric indicators, such as citation rates, may be hindered by different standards valid in different science fields and subfields, and that direct comparison is only possible after standardization or normalisation (Schubert and Braun, 1996). As I did not undertake formal normalization, direct comparisons of citation rates across the different chapters should be interpreted in light of the differing journals and their respective impact factors. In the ISI WoS database, the BMJ, Cochrane Library, and British Journal of General Practice are all categorised as 'Medicine,

General and Internal'. As these journals are classified in the same subcategory it is reasonable to make some comparisons across these journals. Of the other two journals in which I published papers, Ageing and Society is classified in the subcategory 'Gerontology' and Injury Prevention is in the subcategory 'Public, environmental and occupational health'. The impact factors of the various journals (ISI WoS 2009) ranged from 13.660 (BMJ) to 1.770 (Ageing and Society). Furthermore, normalisation is more important when trying to make direct comparisons of the impact of individual researchers rather than the impact of different papers; in this study the major purpose of the citation analysis was to trace the flow of knowledge, for example looking to see if reviews were cited in particular policy documents, rather than to draw any conclusions from direct comparisons of citation counts.

Implications of the findings

This submission contributes to knowledge in several ways. It is based on a number of previously published reviews which offer a unique interpretation of knowledge in a range of subject areas. In addition the study offers insight into the application of methods for evaluating the impact of a diverse group of systematic reviews. Whilst there has been much written on the relationship between research and policy there are few examples of impact evaluations (Boaz 2008a) and this study represents a documented example of an empirical evaluation. Finally the submission contributes to our understanding of how researchers might increase the impact of systematic reviews.

The results of the impact analysis presented in this study provide some evidence that systematic reviews influence health care policy. However, there are significant difficulties associated with determining the impact of specific pieces of research on health care policy (Hanney 2007). The real world of policy making is complex and messy and questions remain about the validity of current methods for evaluating impact of research on policy. As with many impact evaluations I used bibliometric indicators to measure research impact. Whilst they were a reliable measure that could be used across a diverse group of

reviews questions remain about what meaning we can attach to them in terms of policy impact (Hanney 2005). Although documentary review suggested that a number of the reviews had impacted on practice guidelines the methods I used did not enable me to quantify the extent of their contribution. Neither was I able to assess whether policies were implemented or if they impacted on practice.

Identifying suitable methods for the evaluation of impact on policy is particularly difficult. It is possible that qualitative methods, such as interviews with policy makers or direct observation of decision making processes may be more appropriate than quantitative methods for evaluating research impact on policy. However, such methods can be difficult to generalise and may not fit with retrospective evaluations of specific pieces of research. The retrospective nature of the study made it more susceptible to recall bias. Nevertheless, whilst prospective evaluations may have many advantages they also run the risk of underestimating research influence as it may take a number of years for research to impact on the policy process.

The conceptual approach that I used for my study was based on three main types of research impact; instrumental, conceptual and symbolic. This classification of research utilisation is well established and commonly used (Weiss 1976, Stetler., 1985, Estabrooks., 1999). However, although it is generally agreed that much research impact is conceptual or symbolic these are far harder to distinguish than instrumental impact (Hanney et al., 2000). This proved to be the case in my study where indirect impacts were difficult to substantiate.

My study included ten systematic reviews which were concerned with a variety of topics. The rationale for including such a diverse group of reviews was to evaluate the impact of a body of my own work and explore how researchers can increase the impact of systematic reviews. My focus was on the impact of systematic reviews more generally rather than being subject specific. However, it is possible that a narrower focus on reviews in one subject area (i.e. road safety) may have made identifying evidence of impact easier. In such an evaluation it would be more feasible to employ a greater mix of quantitative and

qualitative methods and could involve both tracking forwards from specific reviews and tracking backwards from policy documents.

Although I found evidence of impact there was substantial variation in the influence of individual reviews. The extent to which systematic reviews contribute to the development of health care policy is dependent on a variety of methodological, organisational, political and economic factors. Some of these imperatives that drive policy formulation may be more difficult to influence than others, in particular the political and economic ones may often be beyond the control of individual researchers. However, researchers generally have greater governance over the methodological and organisational elements of the systematic review process, and manipulating these may allow them to increase the influence of their work on health care policy.

The methodological approach adopted by a researcher undertaking a review might be a contributing factor to its overall impact and, as such, the approach should be chosen in the light of the desired impact of the review. Is the primary purpose to impact on practice and/or policy or to inform research? Such considerations will have a bearing on the way the question is formulated and the review undertaken. For example, some review questions, particularly those concerned with public health or complex interventions, will need to be framed to include a consideration of contextual factors and applicability to particular settings or populations. This may necessitate mixed methods reviews including quantitative and qualitative studies. However, for many clinical questions, such as those considered in Chapter 4, a focus on RCTs remains appropriate.

Throughout this study the importance of using appropriate dissemination strategies has been a recurring theme. Researchers need the skills and knowledge necessary to disseminate their work and developing these should be a part of their professional development. Some dissemination strategies are reasonably simple, for example the adoption of user friendly review formats and inclusion of executive summaries and key messages summarising salient points of the review. Researchers should consider other formats for dissemination

instead of relying simply on dissemination via peer review journals. In Lavis's knowledge transfer tool (see Chapter 3) he suggests that it is important to identify the audience for whom the research is planned (Lavis et al., 2003a). This might include the general public, patients, clinicians, managers and public policy makers. Considering for whom the research is intended may make it easier for researchers to target and evaluate their dissemination activities. In addition it is worth addressing at the outset what processes are available for transferring the knowledge and the desired impact of the transfer of knowledge (Gold, 2009, Lavis et al., 2005, Mitton et al., 2007). For instance, in a review I am currently involved in we are holding focus groups involving policy makers, service providers and patient representatives to disseminate and discuss the findings and develop recommendations relevant to our local setting. Investing time and effort into developing appropriate networks, for example through building links with local service providers, commissioners and representatives from the voluntary sector, may be key to dissemination.

Even when researchers have the skills and knowledge to disseminate their work they may not have the time or financial resources required. Knowledge transfer and exchange strategies, such as the use of knowledge brokers, have been put forward as a means of improving impact but these may require substantial investment from both researchers and policy makers. Simpler methods such as targeted messaging have been found to be effective (Dobbins et al., 2009) and may be a more realistic alternative. Ultimately however, if KTE is considered an essential part of the review process then funding bodies and research institutions need to provide support for such initiatives.

It is frequently argued that systematic reviews are of limited use when they lack firm conclusions or recommendations; Gold suggests that although syntheses identifying gaps in the evidence may be useful those drawing more concrete conclusions may be more likely to be valued by policy makers (Gold, 2009). There is, however, a tension between the established tenet of systematic reviewing that emphasises an unbiased presentation of the evidence and the

need to make reviews of interest to policy makers. Researchers have been criticised for going 'beyond the evidence' in an effort to increase the impact of their reviews (Boaz, 2005) and, indeed, ultimately it is unhelpful to policy makers, practitioners or service users to make inflated claims that are not substantiated by the evidence. I would argue that calls for 'further research' whilst frustrating may sometimes be appropriate. In such cases decision-makers, researchers and service users may work together to produce recommendations whilst acknowledging the limitations of the recommendations and the need for further work.

Implications for future research

This study has identified a number of gaps in the evidence base. Areas for future research identified by the study include a need for the following:

- The development and testing of qualitative methods for measuring the impact of research on policy
- Evaluation of different methods of dissemination, including more user friendly presentation and the use of technologies such as podcasts and interactive media
- Evaluation of methods for improving the presentation of reviews so that they are easier for decision makers to understand
- Evaluation of the most appropriate formats for reviews including how to present complex scientific information in a format that is accessible for decision makers.
- Development and evaluation of different methods for search strategies including work on the use of iterative rather than comprehensive methods
- Qualitative work looking at how decision makers interpret and understand systematic reviews and how this could be facilitated

Conclusion

The last few decades have seen a growing emphasis on evidence-based decision making in health care, with research-syntheses such as systematic reviews seen as key sources of evidence. In addition, more recently, there has been burgeoning interest in the way in which research is used; researchers are increasingly expected to demonstrate that their work has contributed to society in some way, whether this is through an impact on the economy, culture, quality of life or public policy. This study brings these two trends together in an exploration of the impact of systematic reviews on health care policy.

Those seeking to influence policy through the findings of systematic reviews need to be aware of the social, cultural, political, economic and ideological factors that contribute to the development of health care policy. Evidence may play its part in decision making but researchers must be realistic about what evidence-syntheses can achieve. However, despite this competition from other influences, some of which are very powerful, this study provides some evidence that systematic reviews can influence health care policy at a variety of levels. Although much of the impact is at a 'micro' level systematic reviews can influence national and international policy. These impacts are, however, not inevitable and need to be routinely and actively encouraged.

As a researcher who has undertaken a number of systematic reviews across a range of subject areas I have come to realise that producing a rigorous review of the evidence is not in itself always enough to influence policy making. The process of dissemination must be seen as part of the review process and reviewers must adopt suitable methods for conducting and disseminating their research. In addition, it is important to have a clear idea of the desired impacts of a review as there is now a substantial literature to suggest that it is naïve to conceptualise the impact of research on policy and practice in the same way. The needs and methods of working of policy-makers may be fundamentally different from those of practitioners, and systematic reviewers seeking to influence policy may need to adapt their approach to reflect the needs of their

target audience. Finally, whilst it is important for researchers to consider how they might increase the influence of their work, such impacts are difficult to measure. Questions remain about how we define and measure policy impact and more work is needed to develop suitable methods for impact analysis.

References

- ACHARA, S., ADEYEMI, B., DOSEKUN, E., KELLEHER, S., LANSLEY, M., MALE, I., et al. (2001) Evidence based road safety: the Driving Standards Agency's schools programme. *Lancet*, 358, 230-2.
- ADAM, S. & KRIESI, H. (2007) The network approach. IN SABATIER, P. (Ed.) *Theories of the Policy Process*. Cambridge, Westview Press.
- ADAMS, H. A., BAUMANN, G., CASCORBI, I., EBENER, C., EMMEL, M., GEIGER, S., et al. (2005) Recommendations for diagnosis and treatment of forms of shock of the IAG Shock of the DIVI. Part 2: Hypovolemic shock. *Intensivmed*, 42, 96-109.
- ALDERSON, P., BUNN, F., LEFEBVRE, C., LI, W. P., LI, L., ROBERTS, I., et al. (2004) Human albumin solution for resuscitation and volume expansion in critically ill patients. *Cochrane Database Syst Rev*, CD001208.
- ALDERSON, P., SCHIERHOUT, G., ROBERTS, I. & BUNN, F. (2000) Colloids versus crystalloids for fluid resuscitation in critically ill patients. *Cochrane Database Syst Rev*, CD000567.
- ALTMAN, D. (1991) *Practical statistics for medical research*, London, Chapman and Hall.
- AMARA, N., OUIMET, M. & LANDRY, R. (2004) New Evidence on Instrumental, Conceptual, and Symbolic Utilization of University Research in Government Agencies.
- AMERICAN THORACIC SOCIETY (2004) Evidence-based colloid use in the critically ill: American Thoracic Society Consensus Statement. *Am J Respir Crit Care Med*, 170, 1247-59.
- ANDERSON, R. (2006) The policy impact of population health surveys: an illustration of the measurement challenges using the NSW Health Survey. *Evidence & Policy*, 2, 167-183.
- ANGELL, M. (2005) *The Truth about the Drug Companies* Random House.
- ATKINS, D., FINK, K. & SLUTSKY, J. (2005) Better information for better health care: the Evidence-based Practice Center program and the Agency for Healthcare Research and Quality. *Ann Intern Med*, 142, 1035-41.
- ATKINS, D. L., BERG, M. D., BERG, R. A., BHUTTA, A. T., BIARENT, D. & BINGHAM, R. (2006) 2005 American Heart Association (AHA) guidelines for cardiopulmonary resuscitation (CPR) and emergency cardiovascular care (ECC) of pediatric and neonatal patients: Pediatric advanced life support. *Pediatrics*, 117, E1005-E1028.
- ATTREE, P. (2004) Growing up in disadvantage: a systematic review of the qualitative evidence. *Child Care Health Dev*, 30, 679-89.
- AUSTRALIAN COMMISSION SAFETY AND QUALITY IN HEALTH CARE (2009) Patient safety in Primary healthcare: a review of the literature.
- BAJPAI, A., AGGARWAL, R. & DEORARI, A. K. (2001) Newer guidelines for neonatal resuscitation. *Perinatology*, 3, 297-306.
- BAKKALBASI, N., BAUER, K., GLOVER, J. & WANG, L. (2006) Three options for citation tracking: Google Scholar, Scopus and Web of Science. *Biomed Digit Libr*, 3, 7.
- BAMBRA, C., GIBSON, M., PETTICREW, M., SOWDEN, A., WHITEHEAD, M. & WRIGHT, K. (2008) Tackling the wider social determinants of health and health inequalities: evidence from systematic reviews. Public Health Research Consortium.
- BARBOUR, R. S. (2000) The role of qualitative research in broadening the 'evidence base' for clinical practice. *J Eval Clin Pract*, 6, 155-63.
- BARNETT-PAGE, E. & THOMAS, J. (2009) Methods for the synthesis of qualitative research: a critical review. *BMC Medical Research Methodology*, 9, 59.
- BARROSO, J., GOLLOP, C. J., SANDELOWSKI, M., MEYNELL, J., PEARCE, P. F. & COLLINS, L. J. (2003) The challenges of searching for and retrieving qualitative studies. *West J Nurs Res*, 25, 153-78.

- BAX, C. A. (2006) Besluitvorming over verkeersveiligheid in het Nationaal Verkeers- en Vervoersplan NVVP : de inspraakfase, het kabinetsstandpunt en de behandeling in de Tweede Kamer. D-2006-4. SWOV. Leidschendam.
- BEACHAM, B. K., L. MCINTYRE, E. (2005) Understanding & measuring research impact http://www.phcris.org.au/phplib/filedownload.php?file=/elib/lib/downloaded_files/publications/pdfs/phcris_pub_3236.pdf. *focus on*, 2, 12pp.
- BEGG, C., CHO, M., EASTWOOD, S., HORTON, R., MOHER, D., OLKIN, I., et al. (1996) Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *Jama*, 276, 637-9.
- BENZIES, K. M., PREMJI, S., HAYDEN, K. A. & SERRETT, K. (2006) State-of-the-evidence reviews: advantages and challenges of including grey literature. *Worldviews Evid Based Nurs*, 3, 55-61.
- BIOLOGY ONLINE (2009) Definition of guidelines <http://www.biology-online.org/dictionary/Guidelines> accessed 19th June 2009.
- BLACK, N. (1994) Why we need qualitative research. *J Epidemiol Community Health*, 48, 425-a-426.
- BLACK, N. (2001) Evidence based policy: proceed with care. *Bmj*, 323, 275-9.
- BLACK, N. & HUTCHINGS, A. (2002) Reduction in the use of surgery for glue ear: did national guidelines have an impact? *Qual Saf Health Care*, 11, 121-4.
- BMJ (2010) About BMJ <http://resources.bmj.com/bmj/about-bmj>.
- BOAZ, A., FITZPATRICK, S., SHAW, B (2008a) Assessing the impact of research on policy: A review of the literature for a project on bridging research and policy through outcome evaluation.
- BOAZ, A., PAWSON, R (2005) The Perious Road from Evidence to Policy: Five Journeys Compared. *Jnl Soc. Pol*, 34, 175-194.
- BOAZ, A. G., L. LEVITT, R. SOLESBURY, W. (2008b) Does evidence-based policy work? Learning from the UK experience. *Evidence & Policy*, 4, 233-53.
- BOAZ, A. G., L; LEVITT, R; SOLESBURY, W. (2008c) Does evidence-based policy work? Learning from the UK. *Evidence & Policy*, 4, 233-53.
- BOGENSCHNEIDER, K. O., J. LINNEY, K. MILLS, J. (2000) Connecting Research and Policymaking: Implications for Theory and Practice from the Family Impact Seminars. *Family Relations*, 49, 327-339.
- BOLUYT, N., BOLLEN CASPER, W., BOS ALBERT, P., KOK JOKE, H. & OFFRINGA, M. (2006) Fluid resuscitation in neonatal and pediatric hypovolemic shock: a Dutch Pediatric Society evidence-based clinical practice guideline. *Intensive Care Medicine*, 32, 995-1003.
- BORNMANN, L., MUTZ, R., NEUHAUS, C. & DANIEL, H. D. (2008) Citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics*, 8, 93-102.
- BOSSUYT, P. M., REITSMA, J. B., BRUNS, D. E., GATSONIS, C. A., GLASZIOU, P. P., IRWIG, L. M., et al. (2003) Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Bmj*, 326, 41-4.
- BOULTON, F. (2004) Transfusion guidelines for neonates and older children. *British Journal of Haematology*, 124, 433-453.
- BREEN, J. (2004) Road safety advocacy. *Bmj*, 328, 888-90.
- BREWER, G. D. & DELEON, P. (1983) *The foundations of policy analysis*, Monterey CA, Brooks/Cole.
- BRITTEN, N., CAMPBELL, R., POPE, C., DONOVAN, J., MORGAN, M. & PILL, R. (2002) Using meta ethnography to synthesise qualitative research: a worked example. *J Health Serv Res Policy*, 7, 209-15.

- BROWN, H. (2007) How impact factors changed medical publishing and science. *BMJ*, 334, 561-564.
- BRUNTON, G., OLIVER, S., OLIVER, K. & LORENC, T. (2006) A Synthesis of Research Addressing Children's, Young People's and Parents' Views of Walking and Cycling for Transport. London, EPPI-Centre, Social Science Research Unit, Institute of Education.
- BRUSSONI, M., TOWNER, E. & HAYES, M. (2006) Evidence into practice: combining the art and science of injury prevention. *Inj Prev*, 12, 373-377.
- BUNN, F., ALDERSON, P. & HAWKINS, V. (2000a) Colloid solutions for fluid resuscitation. *Cochrane Database Syst Rev*, CD001319.
- BUNN, F., BROOKS, F., APPLETON, J., MEAD, M., MAGNUSSON, J., NORTON, S., et al. (2006a) Contraceptive advice and provision for the prevention of under 18 conceptions and STIs: a rapid review. On behalf of the National Institute for Health and Clinical Excellence.
- BUNN, F., BYRNE, G. & KENDALL, S. (2004a) Telephone consultation and triage: effects on health care use and patient satisfaction. *Cochrane Database Syst Rev*, CD004180.
- BUNN, F., BYRNE, G. & KENDALL, S. (2005) The effects of telephone consultation and triage on healthcare use and patient satisfaction: a systematic review. *Br J Gen Pract*, 55, 956-61.
- BUNN, F., COLLIER, T., FROST, C., KER, K., ROBERTS, I. & WENTZ, R. (2003a) Area-wide traffic calming for preventing traffic related injuries. *Cochrane Database Syst Rev*, CD003110.
- BUNN, F., COLLIER, T., FROST, C., KER, K., ROBERTS, I. & WENTZ, R. (2003b) Traffic calming for the prevention of road traffic injuries: systematic review and meta-analysis. *Inj Prev*, 9, 200-4.
- BUNN, F., CUNNINGHAM MELANIE, E. & HANDSCOMB, K. (2006b) Prophylactic antibiotics to prevent surgical site infection after breast cancer surgery. *Cochrane Database of Systematic Reviews: Reviews 2006 Issue 2 John Wiley & Sons, Ltd Chichester, UK DOI: 10.1002/14651858.CD005360.pub2*.
- BUNN, F., DICKINSON, A., BARNETT-PAGE, E. & HORTON, K. (2008a) A systematic review of older people's perceptions of facilitators and barriers to participation in falls prevention interventions. *Ageing & Society*, 28, 449-472.
- BUNN, F., DIGUISEPPI, C. & ROBERTS, I. (1999) Health of Londoners Injury Prevention. A rapid review prepared for the NHS Executive. London.
- BUNN, F., DIGUISEPPI, C. G. & ROBERTS, I. (2001) Systematic Reviews of Injury Studies. IN RIVARA, F. P. C., P; GROSSMAN, D; MAIER, R; KEOPSELL, T. (Ed.) *Injury Control A Guide to Resarch and Program Evaluation*. Cambridge University Press.
- BUNN, F. & KENDALL, S. (2011) Does nursing research impact on policy? A case study of health visiting research and UK health policy. *Journal of Research in Nursing*, 16, 169-191.
- BUNN, F., LEFEBVRE, C., LI WAN PO, A., LI, L., ROBERTS, I. & SCHIERHOUT, G. (2000b) Human albumin solution for resuscitation and volume expansion in critically ill patients. The Albumin Reviewers. *Cochrane Database Syst Rev*, CD001208.
- BUNN, F., ROBERTS, I. & DIGUISEPPI, C. (2000c) A National Contract on Accidents. In Evidence from systematic reviews of research relevant to implementing the 'wider public health' agenda. NHS Centre for Reviews and Dissemination <http://www.york.ac.uk/inst/crd/wph.htm>.
- BUNN, F., ROBERTS, I., TASKER, R. & AKPA, E. (2000d) Hypertonic versus isotonic crystalloid for fluid resuscitation in critically ill patients. *Cochrane Database Syst Rev*, CD002045.

- BUNN, F., ROBERTS, I., TASKER, R. & AKPA, E. (2002) Hypertonic versus isotonic crystalloid for fluid resuscitation in critically ill patients. *Cochrane Database Syst Rev*, CD002045.
- BUNN, F., ROBERTS, I., TASKER, R., AKPA, E. & TRIVEDI D (2004b) Hypertonic versus near isotonic crystalloid for fluid resuscitation in critically ill patients. *Cochrane Database Syst Rev*, CD002045.
- BUNN, F., TRIVEDI, D. & ASHRAF, S. (2008b) Colloid solutions for fluid resuscitation. *Cochrane Database Syst Rev*, CD001319.
- BURGERS, J. S., GROU, R. P., ZAAT, J. O., SPIES, T. H., VAN DER BIJ, A. K. & MOKKINK, H. G. (2003) Characteristics of effective clinical guidelines for general practice. *Br J Gen Pract*, 53, 15-9.
- BUSCEMI, N., HARTLING, L., VANDERMEER, B., TJOSVOLD, L. & KLASSEN, T. P. (2006) Single data extraction generated more errors than double data extraction in systematic reviews. *J Clin Epidemiol*, 59, 697-703.
- BUSE, K., MAYS, N. & WALT, G. (2005) Power and the policy process. *Making Health Policy*. Open University Press.
- BUSE, K. M., N; WALT, G. (2005) *Making Health Policy*, Open University Press.
- BUXTON, M. & HANNEY, S. (1996) How can payback from health services research be assessed? *J Health Serv Res Policy*, 1, 35-43.
- BUXTON, M. & HANNEY, S. (1997) Assessing Payback from Department of Health Research and Development: Second Report: Main Report, Research Report No 24. Uxbridge, HERG, Brunel University.
- BYRNE, G., MORGAN, J., KENDALL, S. & SABERI, D. (2007) A survey of NHS Direct callers' use of health services and the interventions they received. *Primary Health Care Research and Development*, 8, 91-100.
- CABINET OFFICE (1999a) Modernising government. Stationery Office: London, Ch 2, para 6 Cm 4310. Available at <http://archive.cabinetoffice.gov.uk/moderngov/whtpaper/index.htm>.
- CABINET OFFICE (1999b) Professional policy making for the twenty-first century. A report by the Strategic Policy Making Team. London HMSO.
- CALLAHAM, M., WEARS, R. L. & WEBER, E. (2002) Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. *Jama*, 287, 2847-50.
- CAMPBELL J (2009) Effectiveness and cost-effectiveness of telephone triage of patients requesting same day consultations in general practice: a cluster randomised controlled trial comparing nurse-led and GP-led management systems. NIHR Health Technology Assessment Programme www.hta.ac.uk/1865.
- CAMPBELL, M., FITZPATRICK, R., HAINES, A., KINMONTH, A. L., SANDERCOCK, P., SPIEGELHALTER, D., et al. (2000) Framework for design and evaluation of complex interventions to improve health. *Bmj*, 321, 694-6.
- CAMPBELL, R., POUND, P., POPE, C., BRITTEN, N., PILL, R., MORGAN, M., et al. (2003) Evaluating meta-ethnography: a synthesis of qualitative research on lay experiences of diabetes and diabetes care. *Soc Sci Med*, 56, 671-84.
- CAMPBELL, S. B., S. COATES, E. DAVIES, P. PENN, G. (2007) Analysis for policy: Evidence-based policy in practice. London, Government Social Research Unit.
- CANADIAN COCHRANE CENTRE (2004) The Dissemination of Cochrane Evidence. An inventory of resources that use Cochrane Reviews. The Cochrane Collaboration.
- CANADIAN HEALTH SERVICES RESEARCH FOUNDATION (2004) Policy Synthesis: Program Overview. Ottawa, Canadian Health Services Research Foundation.
- CAPLAN, N. (1979) The Two-Communities Theory and Knowledge Utilization. 22, 459-470.

- CARDEN, F. (2004a) Issues in assessing the policy influence of research. *International Social Science Journal*, 56, 135-151.
- CARDEN, F. (2004b) Issues in assessing the policy influence of research. *International Social Science Journal*, 56, 131-151.
- CHALMERS, I. (2003) Trying to Do More Good than Harm in Policy and Practice: The Role of Rigorous, Transparent, Up-to-Date Evaluations. *Annals of the American Academy of Political and Social Science*, 589, 22-40.
- CHALMERS, I. (2005) If evidence-informed policy works in practice, does it matter if it doesn't work in theory? *Evidence & Policy*, 1, 227-42.
- CHALMERS, I. & TROHLER, U. (2000) Helping physicians to keep abreast of the medical literature: Medical and Philosophical Commentaries, 1773-1795. *Ann Intern Med*, 133, 238-43.
- CHALMERS, I. A., D (EDS) (1995) *Systematic Reviews*, London, BMJ Publishing Group.
- CHARLTON, S., . (2006) Speed change management for New Zealand. *Land Transport New Zealand Research Report*. University of Waikato & TERNZ Ltd.
- CHASSIN, M. R., KOSECOFF, J., SOLOMON, D. H. & BROOK, R. H. (1987) How coronary angiography is used. Clinical determinants of appropriateness. *Jama*, 258, 2543-7.
- CHELIMSKY, E. (1985) Old patterns and new directions in program evaluation. IN CHELIMSKY, E. (Ed.) *Program evaluation: Patterns and directions*. Washington DC, American Society for Public Administration.
- CHIAM, M. (2000) Divided opinions over benefits of NHS Direct. *Br J Nurs*, 9, 389.
- CIG ALBUMIN REVIEWERS (1998) Human albumin administration in critically ill patients: systematic review of randomised controlled trials. Cochrane Injuries Group Albumin Reviewers. *Bmj*, 317, 235-40.
- CILISKA, D., HAYWARD, S., DOBBINS, M., BRUNTON, G. & UNDERWOOD, J. (1999) Transferring public-health nursing research to health-system planning: assessing the relevance and accessibility of systematic reviews. *Can J Nurs Res*, 31, 23-36.
- CLARK, W. C., MEHTA, R. L. & SCHETZ, M. Acute Dialysis Quality Initiative 2nd International Consensus Conference Workgroup 5.
- CLARKE, L., CLARKE, M. & CLARKE, T. (2007) How useful are Cochrane reviews in identifying research needs? *J Health Serv Res Policy*, 12, 101-3.
- CLEMENT, F. M., HARRIS, A., LI, J., YONG, K., LEE, K. & MANNIS, B. (2009) Using Effectiveness and Cost-effectiveness to Make Drug Coverage Decisions: A Comparison of Britain, Australia, and Canada. *JAMA*, 302, 1437-1443.
- COCHRANE, A. (1979) A critical review, with particular reference to the medical profession. In: *Medicines for the Year 2000*. London: Office of Health Economics.
- COHEN, M. D., MARCH, J. G. & OLSEN, J. P. (1972) A Garbage Can Model of Organizational Choice. *Administrative Science Quarterly*, 17, 1-25.
- COLLIER, J. (2008) Parliamentary review asks NICE to do better still. *BMJ*, 336, 56-57.
- COOK, D. & GUYATT, G. (2001) Colloid use for fluid resuscitation: evidence and spin. *Ann Intern Med*, 135, 205-8.
- COOK, D. J., GREENGOLD, N. L., ELLRODT, A. G. & WEINGARTEN, S. R. (1997) The relation between systematic reviews and practice guidelines. *Ann Intern Med*, 127, 210-6.
- COOK, P., CORBETT, K., DOWNING, J., CROSSLEY, M. & BELLIS, M. (2007) A fieldwork evaluation of NICE guidance on sexual health interventions. Liverpool, Centre for Public Health, John Moores University.
- CRD (2009) CRD's guidance for undertaking reviews in health care. University of York, Centre for Reviews and Dissemination.

- CROUCH, R. & DALE, J. (1998) Telephone triage--how good are the decisions? (Part 2). *Nurs Stand*, 12, 33-9.
- CROXSON, B., HANNEY, S. & BUXTON, M. (2001) Routine monitoring of performance: what makes health research and development different? *J Health Serv Res Policy*, 6, 226-32.
- CUNNINGHAM, M., BUNN, F. & HANDSCOMB, K. (2006) Prophylactic antibiotics to prevent surgical site infection after breast cancer surgery. *Cochrane Database Syst Rev*, CD005360.
- DAVIES, H., NUTLEY, S. & WALTER, I. (2008) Why 'knowledge transfer' is misconceived for applied social research. *J Health Serv Res Policy*, 13, 188-90.
- DAVIES, H. N., S. SMITH, P. (2000a) Introducing evidence-based policy and practice in public services. *What works? Evidence-based policy and practice in public services*. Bristol, The Policy Press.
- DAVIES, H. N., S. SMITH, P. (2000b) *What Works? Evidence-based policy and practice in public services*, Bristol, The Policy Press.
- DAVIES, H. N., S. TILLEY, N (2000c) A strategic approach to research and development. IN DAVIES, H. N., S. SMITH, P. (Ed.) *What works? Evidence-based policy and practice in public services*. Bristol, The Policy Press.
- DAVIES, P. (1999) What is evidence-based education? *British Journal of Educational Studies*, 47, 108-21.
- DAVIS, D. A. & TAYLOR-VAISEY, A. (1997) Translating guidelines into practice. A systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines. *Cmaj*, 157, 408-16.
- DAVIS, P. & HOWDEN-CHAPMAN, P. (1996) Translating research findings into health policy. *Soc Sci Med*, 43, 865-72.
- DAVIS, P. M., LEWENSTEIN, B. V., SIMON, D. H., BOOTH, J. G. & CONNOLLY, M. J. (2008) Open access publishing, article downloads, and citations: randomised controlled trial. *Bmj*, 337, a568.
- DAWES, M., SUMMERSKILL, W., GLASZIOU, P., CARTABELLOTTA, A., MARTIN, J., HOPAYIAN, K., et al. (2005) Sicily statement on evidence-based practice. *BMC Med Educ*, 5, 1.
- DEEKS, J. J., DINNES, J., D'AMICO, R., SOWDEN, A. J., SAKAROVITCH, C., SONG, F., et al. (2003) Evaluating non-randomised intervention studies. *Health Technol Assess*, 7, iii-x, 1-173.
- DELLINGER, R. P., CARLET, J. M., MASUR, H., GERLACH, H., CALANDRA, T., COHEN, J., et al. (2004) Surviving Sepsis Campaign guidelines for management of severe sepsis and septic shock. *Crit Care Med*, 32, 858-73.
- DELLINGER, R. P., LEVY, M. M., CARLET, J. M., BION, J., PARKER, M. M., JAESCHKE, R., et al. (2008) Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock: 2008. *Crit Care Med*, 36, 296-327.
- DENIS, J. L. & LOMAS, J. (2003) Convergent evolution: the academic and policy roots of collaborative research. *J Health Serv Res Policy*, 8 Suppl 2, 1-6.
- DEPARTMENT FOR TRANSPORT (2009) Road Casualties in Great Britain - main results 2008
<http://www.dft.gov.uk/pgr/statistics/datatablespublications/accidents/casualtiesmrcgbmainresults2008> accessed 3rd July 2009.
- DETR. (2000) Tomorrow's roads: safer for everyone: the Government's road safety strategy and casualty reduction targets for 2010. London
 Department of Environment, Transport and the Regions.

- DICKERSIN, K., MANHEIMER, E., WIELAND, S., ROBINSON, K. A., LEFEBVRE, C., MCDONALD, S., et al. (2002) Development of the Cochrane Collaboration's Central Register of Controlled Clinical Trials. *Eval Health Prof*, 25, 38-64.
- DICKERSIN, K. & MIN, Y. I. (1993) Publication bias: the problem that won't go away. *Ann N Y Acad Sci*, 703, 135-46; discussion 146-8.
- DIXON-WOODS, M., AGARWAL, S., JONES, D., YOUNG, B. & SUTTON, A. (2005) Synthesising qualitative and quantitative evidence: a review of possible methods. *J Health Serv Res Policy*, 10, 45-53.
- DIXON-WOODS, M., BONAS, S., BOOTH, A., JONES, D. R., MILLER, T., SUTTON, A. J., et al. (2006a) How can systematic reviews incorporate qualitative research? A critical perspective. *Qualitative Research*, 6, 27-44.
- DIXON-WOODS, M., CAVERS, D., AGARWAL, S., ANNANDALE, E., ARTHUR, A., HARVEY, J., et al. (2006b) Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Med Res Methodol*, 6, 35.
- DIXON-WOODS, M. & FITZPATRICK, R. (2001) Qualitative research in systematic reviews. Has established a place for itself. *Bmj*, 323, 765-6.
- DIXON-WOODS, M., FITZPATRICK, R. & ROBERTS, K. (2001) Including qualitative research in systematic reviews: opportunities and problems. *J Eval Clin Pract*, 7, 125-33.
- DIXON-WOODS, M., SHAW, R. L., AGARWAL, S. & SMITH, J. A. (2004) The problem of appraising qualitative research. *Qual Saf Health Care*, 13, 223-5.
- DIXON-WOODS, M., SUTTON, A., SHAW, R., MILLER, T., SMITH, J., YOUNG, B., et al. (2007) Appraising qualitative research for inclusion in systematic reviews: a quantitative and qualitative comparison of three methods. *J Health Serv Res Policy*, 12, 42-7.
- DOBBINS, M., COCKERILL, R. & BARNESLEY, J. (2001a) Factors affecting the utilization of systematic reviews. A study of public health decision makers. *Int J Technol Assess Health Care*, 17, 203-14.
- DOBBINS, M., COCKERILL, R., BARNESLEY, J. & CILISKA, D. (2001b) Factors of the innovation, organization, environment, and individual that predict the influence five systematic reviews had on public health decisions. *Int J Technol Assess Health Care*, 17, 467-78.
- DOBBINS, M., DECORBY, K. & TWIDDY, T. (2004a) A knowledge transfer strategy for public health decision makers. *Worldviews Evid Based Nurs*, 1, 120-8.
- DOBBINS, M., HANNA, S., CILISKA, D., MANSKE, S., CAMERON, R., MERCER, S., et al. (2009) A randomized controlled trial evaluating the impact of knowledge translation and exchange strategies. *Implementation Science*, 4, 61.
- DOBBINS, M., THOMAS, H., O'BRIEN, M. A. & DUGGAN, M. (2004b) Use of systematic reviews in the development of new provincial public health policies in Ontario. *Int J Technol Assess Health Care*, 20, 399-404.
- DOH (2001) The national strategy for sexual health and HIV. London: Department of Health. .
- DOH (2004) National Service Framework for Children, Young People and Maternity Services. London: Department of Health.
- DOH. (2000) National Service Framework for Coronary Heart Disease: Modern Standards and Service Models. London, DoH.
- DOH. (2004) Choosing Health Making healthy choices easier. London: Department of Health.
- DOPSON, S., FITZGERALD, L., FERLIE, E., GABBAY, J. & LOCOCK, L. (2002) No magic targets! Changing clinical practice to become more evidence based. *Health Care Manage Rev*, 27, 35-47.

- DOPSON, S., LOCOCK, L., CHAMBERS, D. & GABBAY, J. (2001) Implementation of evidence-based medicine: evaluation of the Promoting Action on Clinical Effectiveness programme. *J Health Serv Res Policy*, 6, 23-31.
- DOPSON, S. G., J. LOCOCK, L. CHAMBERS, D. (1999) Evaluation of the PACE Programme: Final Report. Oxford, Healthcare Management Institute and Wessex Institute for Health Services Management.
- DRETSKE, J. S., S. BAYLISS, S. BURLS, A. (2004) Clinical effectiveness and cost-effectiveness of prehospital intravenous fluids in trauma patients. Health Technology Assessment.
- DUNN, W. N. (1983) Measuring Knowledge Use. *Science Communication*, 5, 120-133.
- DUPERREX, O., BUNN, F. & ROBERTS, I. (2002a) Safety education of pedestrians for injury prevention: a systematic review of randomised controlled trials. *Bmj*, 324, 1129.
- DUPERREX, O., ROBERTS, I. & BUNN, F. (2002b) Safety education of pedestrians for injury prevention. *Cochrane Database Syst Rev*, CD001531.
- EAGAR, K., BREWER, C ET AL (2005) Strategies for Gain - the evidence on strategies to improve the health and wellbeing of Victorian children, March 2005, 172p. Centre for Health Service Development, University of Wollongong.
- EAKIN, J. M. & MYKHALOVSKIY, E. (2003) Reframing the evaluation of qualitative health research: reflections on a review of appraisal guidelines in the health sciences. *J Eval Clin Pract*, 9, 187-94.
- EASTERBROOK, P. J., BERLIN, J. A., GOPALAN, R. & MATTHEWS, D. R. (1991) Publication bias in clinical research. *Lancet*, 337, 867-72.
- EASTON, D. (1953) *The Political System*, New York, Alfred A Knopf.
- EAVES, Y. D. (2001) A synthesis technique for grounded theory data analysis. *J Adv Nurs*, 35, 654-63.
- EDWARDS, P., GREEN, J., LACHOWYCZ, K., GRUNDY, C. & ROBERTS, I. (2008) Serious injuries in children: variation by area deprivation and settlement type. *Arch Dis Child*, 93, 485-9.
- EGGER, M., DAVEY-SMITH, G., ALTMAN, D (EDS) (2001a) *Systematic Reviews in Health Care: Meta-Analysis in Context*, London, BMJ Publishing.
- EGGER, M., DAVEY SMITH, G., O'ROURKE, K. (2001b) Rationale, potentials, and promise of systematic reviews. IN EGGER, M., DAVEY SMITH, G., ALTMAN, DG. (Ed.) *Systematic Reviews in Health Care: Meta-Analysis in Context*. London, BMJ Publishing Group.
- EGGER, M., JUNI, P., BARTLETT, C., HOLENSTEIN, F. & STERNE, J. (2003) How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technol Assess*, 7, 1-76.
- EGGER, M. & SMITH, G. D. (1997) Meta-Analysis. Potentials and promise. *Bmj*, 315, 1371-4.
- EGGER, M., SMITH, G. D. & PHILLIPS, A. N. (1997) Meta-analysis: principles and procedures. *Bmj*, 315, 1533-7.
- ELVIK, R. (1998) Are road safety evaluation studies published in peer reviewed journals more valid than similar studies not published in peer reviewed journals? *Accid Anal Prev*, 30, 101-18.
- EPOC (1998) Draft EPOC Methods Paper: Including Interrupted Time Series (ITS) Designs in a EPOC Review
<http://www.epoc.cochrane.org/Files/Website/Reviewer%20Resources/inttime.pdf>.
- ESTABROOKS, C. A., FIELD, P. A. & MORSE, J. M. (1994) Aggregating Qualitative Findings: An Approach to Theory Development. *Qual Health Res*, 4, 503-511.

- ESTABROOKS, C. A., FLOYD, J. A., SCOTT FINDLAY, S., O'LEARY, K. A. & GUSHTA, M. (2003) Individual determinants of research utilization: a systematic review. *Journal of Advanced Nursing*, 43, 506-520.
- ESTABROOKS, C. E. (1999) The conceptual structure of research utilization. *Research in Nursing and Health*, 22, 203-216.
- EUROPEAN TRANSPORT SAFETY COUNCIL (2003) Transport safety performance in the EU: a statistical overview. Brussels, ETSC.
- EUROSAFE (2007) Interventions to improve pedestrian skills in children. IN EUROPEAN ASSOCIATION FOR INJURY PREVENTION AND SAFETY PROMOTION (Ed.) [http://www.eurosafe.eu.com/csi/eurosafe2006.nsf/wwwVwContent/I3effective measures-11.htm](http://www.eurosafe.eu.com/csi/eurosafe2006.nsf/wwwVwContent/I3effective%20measures-11.htm).
- EVANS, D. (2002) Database searches for qualitative research. *J Med Libr Assoc*, 90, 290-3.
- EXECUTIVE, N. H. S. (1998) The new NHS: modern and dependable; a national framework for assessing performance; consultation document. Department of Health.
- EXECUTIVE, N. H. S. (2000) The NHS Plan: a plan for investment, a plan for reform. London: Department of Health.
- EYSENBACH, G. (2006) Citation advantage of open access articles. *PLoS Biol*, 4, e157.
- FAIRBANK, L., O'MEARA, S., RENFREW, M. J., WOOLRIDGE, M., SOWDEN, A. J. & LISTER-SHARP, D. (2000) A systematic review to evaluate the effectiveness of interventions to promote the initiation of breastfeeding. *Health Technol Assess*, 4, 1-171.
- FALAGAS, M. E., PITSOUNI, E. I., MALIETZIS, G. A. & PAPPAS, G. (2008) Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *Faseb J*, 22, 338-42.
- FARMER, A. P., LEGARE, F., TURCOT, L., GRIMSHAW, J., HARVEY, E., MCGOWAN, J. L., et al. (2008) Printed educational materials: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev*, CD004398.
- FELDMAN, S. (1999) Strangers in the night: research and managed mental health care. *Health Aff (Millwood)*, 18, 48-51.
- FERLIE, E., GABBAY, L., FITZGERALD, L., LOCOCK, L. & DOPSON, S. (2001) Evidence-based Medicine and Organisational Change: An Overview of Some Recent Qualitative Research. IN ASHBURNER, L. (Ed.) *Organisational Behaviour and Organisational Studies in Health Care: Reflections on the Future*. Basingstoke, Palgrave.
- FERRIMAN, A. (2005) NHS Direct and NHS 24 condemned as flawed by BMA conference. BMA Annual Representative Meeting Manchester 27-30 June *BMJ* 331, News Extra.
- FINFER, S., BELLOMO, R., BOYCE, N., FRENCH, J., MYBURGH, J. & NORTON, R. (2004a) A comparison of albumin and saline for fluid resuscitation in the intensive care unit. *N Engl J Med*, 350, 2247-56.
- FINFER, S. R., BOYCE, N. W. & NORTON, R. N. (2004b) The SAFE study: A landmark trial of the safety of albumin in intensive care. *Medical Journal of Australia*, 181, 237-238.
- FLEMMING, K. & BRIGGS, M. (2007) Electronic searching to locate qualitative research: evaluation of three strategies. *J Adv Nurs*, 57, 95-100.
- FLORIN, D. & ROSEN, R. (1999) Evaluating NHS direct. Early findings raise questions about expanding the service. *Bmj*, 319, 5-6.
- FOX, D. M. (1990) Health policy and the politics of research in the United States. *J Health Polit Policy Law*, 15, 481-99.

- FOX, D. M. (2006) The 'long and lonely road': translating evidence into policy reply. *Health Econ Policy Law*, 1, 423-5.
- FOY, R., MACLENNAN, G., GRIMSHAW, J., PENNEY, G., CAMPBELL, M. & GROL, R. (2002) Attributes of clinical recommendations that influence change in practice following audit and feedback. *J Clin Epidemiol*, 55, 717-22.
- FRANCKE, A., SMIT, M., DE VEER, A. & MISTIAEN, P. (2008) Factors influencing the implementation of clinical guidelines for health care professionals: A systematic meta-review. *BMC Medical Informatics and Decision Making*, 8, 38.
- FRENCH, P. (1999) The development of evidence-based nursing. *Journal of Advanced Nursing*, 29, 72-78.
- GARCIA, J., BRICKER, L., HENDERSON, J., MARTIN, M. A., MUGFORD, M., NIELSON, J., et al. (2002) Women's views of pregnancy ultrasound: a systematic review. *Birth*, 29, 225-50.
- GARFIELD, E. (2006) The History and Meaning of the Journal Impact Factor. *JAMA*, 295, 90-93.
- GILLESPIE, L. D., GILLESPIE, W. J., ROBERTSON, M. C., LAMB, S. E., CUMMING, R. G. & ROWE, B. H. (2001) Interventions for preventing falls in elderly people. *Cochrane Database Syst Rev*, CD000340.
- GILLESPIE, L. D., ROBERTSON, M. C., GILLESPIE, W. J., LAMB, S. E., GATES, S., CUMMING, R. G., et al. (2009) Interventions for preventing falls in older people living in the community. *Cochrane Database Syst Rev*, CD007146.
- GINGRICH, N. (2009) Economist Debates: Health care: Statements <http://www.economist.com/debate/days/view/394>. *The Economist*.
- GIRASEK, D. (Ed.) (2007) *Moving America towards evidence-based approaches to traffic safety*, Washington, AAA Foundation for Traffic Safety.
- GLASPER, A. (2000) Does NHS Direct deserve its recent criticism? *British journal of nursing* 9, 1034.
- GLASS, G. (1976) Primary, secondary and meta-analysis of research. *Educat Res* 5, 3-8.
- GOLD, M. (2009) Pathways to the use of health services research in policy. *Health Serv Res*, 44, 1111-36.
- GOTZSCHE, P. C., HROBJARTSSON, A., MARIC, K. & TENDAL, B. (2007) Data extraction errors in meta-analyses that use standardized mean differences. *Jama*, 298, 430-7.
- GRAHAM, H. & MCDERMOTT, E. (2006) Qualitative Research and the Evidence Base of Policy: Insights from Studies of Teenage Mothers in the UK. *Journal of Social Policy*, 35, 21-37.
- GRANT, J., COTTRELL, R., CLUZEAU, F. & FAWCETT, G. (2000) Evaluating "payback" on biomedical research from papers cited in clinical guidelines: applied bibliometric study. *Bmj*, 320, 1107-1111.
- GRANT, J. & LEWISON, G. (1997) Government funding of research and development. *Science*, 278, 878-80.
- GRANT, M. J. (2004) How does your searching grow? A survey of search preferences and the use of optimal search strategies in the identification of qualitative research. *Health Info Libr J*, 21, 21-32.
- GRAY, V. (1973) Innovation in the states: A diffusion study. *American Political Science Review*, 67, 1174-1185.
- GREEN, L. W., OTTOSON, J. M., GARCIA, C. & HIATT, R. A. (2009) Diffusion theory and knowledge dissemination, utilization, and integration in public health. *Annu Rev Public Health*, 30, 151-74.
- GREEN, R. S., DJOGOVIC, D., GRAY, S., HOWES, D., BRINDLEY, P. G., STENSTROM, R., et al. (2008) Canadian Association of Emergency Physicians Sepsis Guidelines: the

- optimal management of severe sepsis in Canadian emergency departments. *Cjem*, 10, 443-59.
- GREENBERG, S. A. (2009) How citation distortions create unfounded authority: analysis of a citation network. *BMJ*, 339, b2680-.
- GREENHALGH, T. & PEACOCK, R. (2005) Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *Bmj*, 331, 1064-5.
- GREENHALGH, T., ROBERT, G., MACFARLANE, F., BATE, P. & KYRIAKIDOU, O. (2004) Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q*, 82, 581-629.
- GREENHALGH, T., ROBERT, G., MACFARLANE, F., BATE, P., KYRIAKIDOU, O. & PEACOCK, R. (2005) Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review. *Soc Sci Med*, 61, 417-30.
- GREENHALGH, T., TOON, P., RUSSELL, J., WONG, G., PLUMB, L. & MACFARLANE, F. (2003) Transferability of principles of evidence based medicine to improve educational quality: systematic review and case study of an online course in primary health care. *Bmj*, 326, 142-5.
- GRILLI, R. & LOMAS, J. (1994) Evaluating the message: the relationship between compliance rate and the subject of a practice guideline. *Med Care*, 32, 202-13.
- GRIMES, D. A. & SCHULZ, K. F. (2005) Surrogate End Points in Clinical Research: Hazardous to Your Health. *Obstetrics & Gynecology*, 105, 1114-1118.
- GRIMSHAW, J. (2004) So what has the Cochrane Collaboration ever done for us? A report card on the first 10 years. *Canadian Medical Association Journal*, 171, 747.
- GRIMSHAW, J., FREEMANTLE, N., WALLACE, S., RUSSELL, I., HURWITZ, B., WATT, I., et al. (1995) Developing and implementing clinical practice guidelines. *Qual Health Care*, 4, 55-64.
- GRIMSHAW, J. M. & RUSSELL, I. T. (1993) Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet*, 342, 1317-22.
- GRIMSHAW, J. M., THOMAS, R. E., MACLENNAN, G., FRASER, C., RAMSAY, C. R., VALE, L., et al. (2004) Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess*, 8, iii-iv, 1-72.
- GROL, R. (2001) Successes and failures in the implementation of evidence-based guidelines for clinical practice. *Med Care*, 39, 1146-54.
- GROL, R., DALHUIJSEN, J., THOMAS, S., VELD, C., RUTTEN, G. & MOKKINK, H. (1998) Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. *Bmj*, 317, 858-61.
- GROL, R. & GRIMSHAW, J. (2003) From best evidence to best practice: effective implementation of change in patients' care. *Lancet*, 362, 1225-30.
- GRUEN, R. L., MORRIS, P. S., MCDONALD, E. L. & BAILIE, R. S. (2005) Making systematic reviews more useful for policy-makers. *Bull World Health Organ*, 83, 480.
- GUBA E, G. (1984) The effect of definitions of policy on the nature and outcomes of policy analysis. *Educational Leadership*, 42.
- HAJER, M. A. & WAGENAAR, H. (2003) *Deliberative policy analysis*, Cambridge, Cambridge University Press.
- HAM, C., HUNTER, D. J. & ROBINSON, R. (1995) Evidence based policymaking. *Bmj*, 310, 71-2.
- HAMMERSLEY, M. (2005) Is the Evidence-based practice movement doing more good than harm? Reflections on Iain Chalmers' case for research-based policy making and practice. *Evidence & Policy*, 1, 85-100.

- HANNEY, S., GRANT, J., WOODING, S. & BUXTON, M. (2004a) Proposed methods for reviewing the outcomes of health research: the impact of funding by the UK's 'Arthritis Research Campaign'. *Health Research Policy and Systems*, 2, 4.
- HANNEY, S., MUGFORD, M., GRANT, J. & BUXTON, M. (2005) Assessing the benefits of health research: lessons from research into the use of antenatal corticosteroids for the prevention of neonatal respiratory distress syndrome. *Soc Sci Med*, 60, 937-47.
- HANNEY, S., PACKWOOD, T. & BUXTON, M. (2000) Evaluating the Benefits from Health Research and Development Centres: A Categorization, a Model and Examples of Application. *Evaluation: The International Journal of Theory, Research and Practice*, 6, 137-160.
- HANNEY, S. R., GONZALEZ-BLOCK, M. A., BUXTON, M. J. & KOGAN, M. (2003) The utilisation of health research in policy-making: concepts, examples and methods of assessment. *Health Res Policy Syst*, 1, 2.
- HANNEY, S. R., GRANT, J., WOODING, S. & BUXTON, M. J. (2004b) Proposed methods for reviewing the outcomes of health research: the impact of funding by the UK's 'Arthritis Research Campaign'. *Health Res Policy Syst*, 2, 4.
- HANNEY, S. R. G., J; WOODING, S; BUXTON, M.J. (2007) An assessment of the impact of NHS Health Technology Assessment Programme. IN NHS HEALTH TECHNOLOGY ASSESSMENT PROGRAMME. VOL 11, N. (Ed.).
- HARDEN, A., GARCIA, J., OLIVER, S., REES, R., SHEPHERD, J., BRUNTON, G., et al. (2004) Applying systematic review methods to studies of people's views: an example from public health research. *J Epidemiol Community Health*, 58, 794-800.
- HARDEN, A. & THOMAS, J. (2005) Methodological issues in combining diverse study types in systematic reviews. *International Journal of Social Research Methodology*, 8, 257-271.
- HARRISON, S. (2001) *Implementing the results of research and development in clinical and managerial practice*, Abingdon, Radcliffe Medical Press.
- HARTZBAND, P. & GROOPMAN, J. (2009) Keeping the Patient in the Equation -- Humanism and Health Care Reform.
- HAUER, E. (2005) The Road Ahead. *Journal of Transportation Engineering*, 131, 333-339.
- HAYES, D. (2000) The case against NHS Direct. *Doctor*, 13, 36-39.
- HEALTH PROTECTION AGENCY (2009) Completed primary courses at two years of age: England and Wales, 1966-1977, England only 1978 onwards. 2006.
- HEDLUND, J. (2007) Improving traffic safety culture in the United States: The Journey Forward. Washington, AAA Foundation for Traffic Safety.
- HEFC (2009) Research Excellence Framework Second consultation on the assessment and funding of research. Higher Education Funding Council for England.
- HIGGINS, J. G., S (EDITORS) (2008) *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.*
- HIGGINS, J. P. & THOMPSON, S. G. (2002) Quantifying heterogeneity in a meta-analysis. *Stat Med*, 21, 1539-58.
- HIGGINS, J. P., THOMPSON, S. G., DEEKS, J. J. & ALTMAN, D. G. (2003) Measuring inconsistency in meta-analyses. *Bmj*, 327, 557-60.
- HILL, S., GARATTINI, S., VAN LEONHOUT, J., O'BRIEN, B. & DE JONCHEERE, K. (2003) Technology appraisal programme of the National Institute for Clinical Excellence: a review by the WHO. WHO.
- HILL, S. R., MITCHELL, A. S. & HENRY, D. A. (2000) Problems with the interpretation of pharmaco-economic analyses: a review of submissions to the Australian Pharmaceutical Benefits Scheme. *Jama*, 283, 2116-21.

- HILLS, M. (2000) Human science research in public health: the contribution and assessment of a qualitative approach. *Can J Public Health*, 91, 1-4-12.
- HOPEWELL, S., CLARKE, M., STEWART, L. & TIERNEY, J. (2007a) Time to publication for results of clinical trials. *Cochrane Database Syst Rev*, MR000011.
- HOPEWELL, S., MCDONALD, S., CLARKE, M. & EGGER, M. (2007b) Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database Syst Rev*, MR000010.
- HOUSE OF COMMONS HEALTH COMMITTEE (2002) Review of the National Institute for Clinical Excellence. London, Stationery Office.
- HOVLAND, I. (2007) Making a difference: M&E of policy research. *Working Paper 281*. London, Overseas Development Institute
- HOWARD, S. & HARRISON, L. (2005) NICE Guidance Implementation Tracking. Data Sources, Methodology and Results. ABACUS International.
- HOYT, P. & GARRISON, J. (1997) Political manipulation within the Small Group: Foreign Policy Advisors in the Carter Administration. IN HART, P., STERN, E. & SUNDELIUS, B. (Eds.) *Beyond Groupthink: Political Group Dynamics and Foreign Policy-Making*. MI, The University of Michigan Press.
- HUBERMAN, M. (1990) Linkage Between Researchers and Practitioners: A Qualitative Study. 27, 363-391.
- HUBERMAN, M. (1992) 'Linking the practitioner and researcher communities for school improvement'. *Address to the International Congress for School Effectiveness and Improvement*. Victoria, B.C.
- HUBERMAN, M. (1994) Research utilization: The state of the art. *Knowledge, Technology, and Policy*, 7, 13-33.
- ICAP (2009) 16. Blood Alcohol Concentration Limits. *ICAP Blue Book - Practical Guides for Alcohol Policy and Prevention Approaches*. International Centre for Alcohol Policies
<http://www.icap.org/PolicyTools/ICAPBlueBook/BlueBookModules/16BloodAlcoholConcentrationLimits/tabid/176/Default.aspx>.
- INNVAER, S., VIST, G., TROMMALD, M. & OXMAN, A. (2002) Health policy-makers' perceptions of their use of evidence: a systematic review. *J Health Serv Res Pol*, 7, 239-244.
- ISMAIL, S., NASON, E., MARJANOVIC, S. & GRANT, J. (2009) Bibliometrics as a tool for supporting prospective R&D decision-making in the health sciences: Strengths, weaknesses and options for future development. RAND Corporation.
- JACK, S. M. (2006) Utility of qualitative research findings in evidence-based public health practice. *Public Health Nurs*, 23, 277-83.
- JACOBSON, N., BUTTERILL, D. & GOERING, P. (2003) Development of a framework for knowledge translation: understanding user context. *J Health Serv Res Policy*, 8, 94-9.
- JACSO, P. (2005) Google Scholar: the pros and the cons. *Online Information Review*, 29, 208-214.
- JADAD, A. R., COOK, D. J., JONES, A., KLASSEN, T. P., TUGWELL, P., MOHER, M., et al. (1998) Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. *Jama*, 280, 278-80.
- JENSEN, L. A. & ALLEN, M. N. (1996) Meta-Synthesis of Qualitative Findings. *Qual Health Res*, 6, 553-560.
- JONES, A. P., REMMINGTON, T., WILLIAMSON, P. R., ASHBY, D. & SMYTH, R. L. (2005a) High prevalence but low impact of data extraction and reporting errors were found in Cochrane systematic reviews. *J Clin Epidemiol*, 58, 741-2.

- JONES, S. J., LYONS, R. A., JOHN, A. & PALMER, S. R. (2005b) Traffic calming policy can reduce inequalities in child pedestrian injuries: database study. *Inj Prev*, 11, 152-6.
- JUNI, P., ALTMAN, D. G. & EGGER, M. (2001) Systematic reviews in health care: Assessing the quality of controlled clinical trials. *Bmj*, 323, 42-6.
- JUNI, P., WITSCHI, A., BLOCH, R. & EGGER, M. (1999) The hazards of scoring the quality of clinical trials for meta-analysis. *Jama*, 282, 1054-60.
- KALUCY, E. C., JACKSON-BOWERS, E., MCINTYRE, E. & REED, R. (2009) The feasibility of determining the impact of primary health care research projects using the Payback Framework. *Health Res Policy Syst*, 7, 11.
- KANE, R. L. (1995) Creating practice guidelines: the dangers of over-reliance on expert judgment. *J Law Med Ethics*, 23, 62-4.
- KAVANAGH, P., DOYLE, C. & METCALFE, O. (2005) Health Impacts of Transport a review. The Institute of Public Health in Ireland.
- KEARNEY, M. H. (2001) Enduring love: a grounded formal theory of women's experience of domestic violence. *Res Nurs Health*, 24, 270-82.
- KENT PCT (2008) Joint Strategic Needs Assessment for Children in Kent.
- KEOWN, K., VAN EERD, D. & IRVIN, E. (2008) Stakeholder engagement opportunities in systematic reviews: knowledge transfer for policy and practice. *J Contin Educ Health Prof*, 28, 67-72.
- KER, K., ROBERTS, I., COLLIER, T., BEYER, F., BUNN, F. & FROST, C. (2005) Post-licence driver education for the prevention of road traffic crashes: a systematic review of randomised controlled trials. *Accid Anal Prev*, 37, 305-13.
- KINGDON, J. W. (2003) *Agendas, alternative and public policies (2nd ed)*, New York, Longman.
- KINGDON, J. W. (1984) *Agendas, Alternatives, and Public Policies* Boston, Little Brown.
- KITSON, A., HARVEY, G. & MCCORMACK, B. (1998) Enabling the implementation of evidence based practice: a conceptual framework. *Qual Health Care*, 7, 149-58.
- KITSON, A., RYCROFT-MALONE, J., HARVEY, G., MCCORMACK, B., SEERS, K. & TITCHEN, A. (2008) Evaluating the successful implementation of evidence into practice using the PARIHS framework: Theoretical and practical challenges. *Implementation Science*, 3, 1.
- KOUSHA, K. & THELWALL, M. (2008) Sources of Google Scholar citations outside the Science Citation Index: A comparison between four science disciplines. *Scientometrics*, 74, 273-294.
- KUHN, T. (1962) *The Structure of Scientific Revolutions*, Chicago, University of Chicago Press.
- KURUVILLA, S., MAYS, N., PLEASANT, A. & WALT, G. (2006) Describing the impact of health research: a Research Impact Framework. *BMC Health Serv Res*, 6, 134.
- KURUVILLA, S., MAYS, N. & WALT, G. (2007) Describing the impact of health services and policy research. *J Health Serv Res Policy*, 12 Suppl 1, S1-23-31.
- KWAN, I. & BUNN, F. (2005) Effects of prehospital spinal immobilization: a systematic review of randomized trials on healthy subjects. *Prehosp Disaster Med*, 20, 47-53.
- KWAN, I., BUNN, F. & ROBERTS, I. (2003) Timing and volume of fluid administration for patients with bleeding. *Cochrane Database Syst Rev*, CD002245.
- LANDRY, R., AMARA, N. & LAMARI, M. (2001) Utilization of social science research knowledge in Canada. *Research Policy*, 30, 333-349.
- LANDRY, R. A., N; LAMARI, M. (2001) Climbing the ladder of research utilization. *Science Communication*, 9, 171-182.

- LASSWELL, H. D. (1956) *The decision process: Seven categories of functional analysis*, University of Maryland, Maryland.
- LAU, J., ANTMAN, E. M., JIMENEZ-SILVA, J., KUPELNICK, B., MOSTELLER, F. & CHALMERS, T. C. (1992) Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med*, 327, 248-54.
- LAVIS, J., DAVIES, H., GRUEN, R., WALSHE, K. & FARQUHAR, C. (2006a) Working Within and Beyond the Cochrane Collaboration to Make Systematic Reviews More Useful to Healthcare Managers and Policy Makers. *Healthcare Policy / Politiques de Santé*, 1, 21-33.
- LAVIS, J., DAVIES, H., OXMAN, A., DENIS, J. L., GOLDEN-BIDDLE, K. & FERLIE, E. (2005) Towards systematic reviews that inform health care management and policy-making. *J Health Serv Res Policy*, 10 Suppl 1, 35-48.
- LAVIS, J., ROSS, S., MCLEOD, C. & GILDINER, A. (2003a) Measuring the impact of health research. *J Health Serv Res Policy*, 8, 165-70.
- LAVIS, J. N. (2004) Linking research to action. In: World report on knowledge for better health: strengthening health systems. Geneva, World Health Organisation.
- LAVIS, J. N., LOMAS, J., HAMID, M. & SEWANKAMBO, N. K. (2006b) Assessing country-level efforts to link research to action. *Bull World Health Organ*, 84, 620-8.
- LAVIS, J. N., POSADA, F. B., HAINES, A. & OSEI, E. (2004) Use of research to inform public policymaking. *The Lancet*, 364, 1615-1621.
- LAVIS, J. N., ROBERTSON, D., WOODSIDE, J., MCLEOD, C. & ABELSON, J. (2003b) How can research organizations more effectively transfer research knowledge to decision makers? *The Milbank Quarterly*, 81, 221 - 248.
- LAVIS, J. N., ROSS, S. E., HURLEY, J. E., HOHENADEL, J. M., STODDART, G. L., WOODWARD, C. A., et al. (2002) Examining the role of health services research in public policymaking. *Milbank Q*, 80, 125-54.
- LAW, M., KERTROY, M., TEPLICKY, R., KING, S., PLEWS, N., HANA, S., et al. (2004) Knowledge Transfer in Health Care. McMaster University.
- LAYCOCK, G. (2001) Hypothesis-Based Research: The Repeat Victimization Story. *J Criminology and Criminal Justice*, 1, 59-82.
- LEAL, R., MUÑOZ, M., PÁRAMO, J. A., GARCÍA-ERCE, J. A. & LLAU, J. A. (2006) Spanish Consensus Statement on Alternatives to Allogeneic Blood Transfusions: The Seville Document. *Transfusion Alternatives in Transfusion Medicine*, 8, 35-40.
- LEXCHIN, J., BERO, L. A., DJULBEGOVIC, B. & CLARK, O. (2003) Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*, 326, 1167-1170.
- LIABO, K. (2004) Traffic calming and childhood injury on the road. IN BUREAU, N. C. S. (Ed.) *Highlight no 207*.
- LIABO, K., LUCAS, P. & ROBERTS, H. (2003) Can traffic calming measures achieve the children's fund objective of reducing inequalities in child health. IN [HTTP://WWW.BESTBETS.ORG/BETS/BET.PHP?ID=218](http://www.bestbets.org/bets/bet.php?id=218), B. B. (Ed.).
- LIGHT, R., PILLEMAR, DB (1984) *Summing Up: the Science of Reviewing Research*. Cambridge, Massachusetts: Harvard University Press; 1984.
- LIU, B. C., IVERS, R., NORTON, R., BOUFOUS, S., BLOWS, S. & LO, S. K. (2008) Helmets for preventing injury in motorcycle riders. *Cochrane Database Syst Rev*, CD004333.
- LOMAS, J. (1991) Words without action? The production, dissemination, and impact of consensus recommendations. *Annu Rev Public Health*, 12, 41 - 65.
- LOMAS, J. (2000a) Connecting research and policy. *Can J Policy Res*, 1, 140-1.
- LOMAS, J. (2000b) Using 'linkage and exchange' to move research into policy at a Canadian foundation. *Health Aff (Millwood)*, 19, 236-40.
- LOMAS, J. (2007) The in-between world of knowledge brokering. *Bmj*, 334, 129-32.

- LOMAS, J. & BROWN, A. D. (2009) Research and advice giving: a functional view of evidence-informed policy advice in a Canadian Ministry of Health. *Milbank Q*, 87, 903-26.
- LOMAS, J. C., T. MCCUTCEON, C. MCAULEY, L. LAW, S. (2005) Conceptualizing and Combining Evidence for Health System Guidance. IN WWW.CHSRF.CA, C. H. S. R. F. C. A. F. (Ed.).
- LYONS, R. (2005) Traffic calming, childhood pedestrian injury and inequalities and politics. IN TRANSPORT, D. F. (Ed.) *Behavioural research in road safety*. Her Majesty's Stationery Office.
- LYONS, R. A., JONES, S. J., DEACON, T. & HEAVEN, M. (2003) Socioeconomic variation in injury in children and older people: a population based study. *Inj Prev*, 9, 33-7.
- MACINTYRE, S., CHALMERS, I., HORTON, R. & SMITH, R. (2001) Using evidence to inform health policy: case study. *Bmj*, 322, 222-5.
- MACKIE, A. M., WARD, H., A., & WALKER, R. T. (1990) Urban Safety Project: 3: Overall evaluation of area wide schemes. TRRL Research Report, RR 263
- MACPHERSON, A. & SPINKS, A. (2007) Bicycle helmet legislation for the uptake of helmet use and prevention of head injuries. *Cochrane Database Syst Rev*, CD005401.
- MARMOT, M. G. (2004) Evidence based policy or policy based evidence? *Bmj*, 328, 906-7.
- MARTIN, D. (2007) Absence of policy on how to use statins in England and Wales: do not let JBS2 fill the void. *Int J Clin Pract*, 61, 1967-1970.
- MASON, J., FREEMANTLE, N. & BROWNING, G. (2001) Impact of effective health care bulletin on treatment of persistent glue ear in children: time series analysis. *Bmj*, 323, 1096-7.
- MAYNARD, A. (2006) The 'long and lonely road': translating evidence into policy. *Health Econ Policy Law*, 1, 415-21.
- MAYS, N. & POPE, C. (2000) Qualitative research in health care. Assessing quality in qualitative research. *Bmj*, 320, 50-2.
- MAYS, N., POPE, C. & POPAY, J. (2005) Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. *J Health Serv Res Policy*, 10 Suppl 1, 6-20.
- MCCLELLAND, D. B. (1990) ABC of transfusion. Human albumin solutions. *Bmj*, 300, 35-7.
- MCDONNELL, A., WILSON, R. & GOODACRE, S. (2006) Evaluating and implementing new services. *BMJ*, 332, 109-112.
- MCINTOSH, H., WOOLACOTT, N. & BAGNALL, A.-M. (2004) Assessing harmful effects in systematic Reviews. *BMC Medical Research Methodology*, 4, 19.
- MCKIBBON, K. A., WILCZYNSKI, N. L. & HAYNES, R. B. (2006) Developing optimal search strategies for retrieving qualitative studies in PsycINFO. *Eval Health Prof*, 29, 440-54.
- MCWHIRTER, J. (2008) A review of safety education: Principles for effective practice. Royal Society for the Prevention of Accidents (ROSPA).
- MEHO, L. I. (2007) The Rise and Rise of Citation Analysis. *Physics World*, 20, 32-36.
- MINHAS, R. & PATEL, K. (2008) From rationing to rational: the evolving status of NICE. *J R Soc Med*, 101, 436-442.
- MITTON, C., ADAIR, C., MCKENZIE, E., PATTEN, S. & WAYE PERRY, B. (2007) Knowledge Transfer and Exchange: Review and Synthesis of the Literature. *Milbank Quarterly*, 85, 729-768.
- MOHAN, D. & ROBERTS, I. (2001) Global road safety and the contribution of big business. *BMJ*, 323, 648-.

- MOHER, D., JADAD, A. R., NICHOL, G., PENMAN, M., TUGWELL, P. & WALSH, S. (1995) Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials*, 16, 62-73.
- MOHER, D., PHAM, B., JONES, A., COOK, D. J., JADAD, A. R., MOHER, M., et al. (1998) Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet*, 352, 609-13.
- MOLAS-GALLART, J., TANG, P. & MORROW, S. (2000) Assessing the non-academic impact of grant-funded socio-economic research: results from a pilot study. *Research Evaluation*, 9, 171-182.
- MORGAN, A. & ZIGLIO, E. (2007) Revitalising the evidence base for public health: an assets model. *Promotion & Education*, 14, 17-22.
- MULROW, C. D. (1994) Rationale for systematic reviews. *Bmj*, 309, 597-9.
- MULROW, D. (1995) Rationale for systematic reviews. IN CHALMERS, I. A., D. (Ed.) *Systematic Reviews*. London, BMJ Publishing Group.
- MUNRO, J., NICHOLL, J., O'CATHAIN, A. & KNOWLES, E. (2000) Evaluation of NHS Direct first wave sites: second interim report to the Department of Health. *Medical Care Research Unit, University of Sheffield, Sheffield*.
- MURALI, N. S., MURALI, H. R., AUETHAVEKIAT, P., ERWIN, P. J., MANDREKAR, J. N., MANEK, N. J., et al. (2004) Impact of FUTON and NAA bias on visibility of research. *Mayo Clin Proc*, 79, 1001-6.
- MURPHY, E., DINGWALL, R., GREATBATCH, D., PARKER, S. & WATSON, P. (1998) Qualitative research methods in health technology assessment: a review of the literature. *Health Technology Assessment*, 2, 1 - 274.
- NAKAMURA R, T. & SMALLWOOD, F. (1980) *The politics of policy implementation*, New York, St Martins.
- NATIONAL AUDIT OFFICE (2003) Getting the evidence: using research in policy making. Report by the Comptroller and Auditor General. London: HMSO.
- NATIONAL FORUM ON HEALTH (1997) Canada health action: Building on the legacy. Ottawa: National Forum on Health.
- NHS NLH (2009) The National Library of Guidelines, The NHS National Library of Health. NHS.
- NHS SCOTLAND Shifting the balance of care
<http://www.shiftingthebalance.scot.nhs.uk/improvement-framework/high-impact-changes/use-tele-medicine-and-tele-health-to-support-care-delivery/>.
- NICE (2004) Pre-hospital initiation of fluid replacement therapy in trauma. National Institute of Health and Clinical Excellence. Technology Appraisal 74
<http://www.nice.org.uk/nicemedia/pdf/ta074guidance.pdf>.
- NICE (2006) The guidelines manual: London: National Institute of Health and Clinical Excellence. Available from NICE.org.uk.
- NICE (2007a) Press Release. NICE guidance calls for more to be done to halt the rise of sexually transmitted infections and prevent under 18 conceptions
<http://www.nice.org.uk/guidance/index.jsp?action=download&o=31902>.
- NICE (2007b) Prevention of sexually transmitted infections and under 18 conceptions. PH3 February 2007 <http://guidance.nice.org.uk/PH3>.
- NICE (2008) Second report of the health committee 2007-2008, HC27-1. London, Stationery Office.
- NICE (2009) Methods for the development of NICE public health guidelines (second edition). London, National Institute of Health and Clinical Excellence. Available from NICE.org.uk.
- NILSEN, E. S., MYRHAUG, H. T., JOHANSEN, M., OLIVER, S. & OXMAN, A. D. (2006) Methods of consumer involvement in developing healthcare policy and

- research, clinical practice guidelines and patient information material. *Cochrane Database Syst Rev*, 3, CD004563.
- NOBLIT, G. W. & HARE, R. D. (1988) *Meta-ethnography: synthesising qualitative studies*, Newbury Park, California, Sage.
- NOYES, J., POPAY, J., PEARSON, A., HANNES, K. & BOOTH, A. (2008) Chapter 20: Qualitative research and Cochrane Reviews. IN HIGGINS, J. & GREEN, S. (Eds.) *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008]*. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.
- NUTLEY, S. P.-S., J. SOLESBURY, W. (2003a) Models of research impact: a cross sector review of literature and practice, Building Effective Research 4. London, learning and Skills Research Centre.
- NUTLEY, S. W., I. DAVIES, HTO. (2003b) From knowing to doing: a framework for understanding the evidence-into practice agenda. *Evaluation*, 9, 125-148.
- NUTLEY, S. W., J. (2000) Evidence and the policy process. IN DAVIES, H. N., S. SMITH, P. (Ed.) *What works? Evidence-based policy and practice in public services*. Bristol, The Policy Press.
- OAKESHOTT, P. & GRAHAM, A. (2007) NICE guidance on one-to-one interventions to reduce sexually transmitted infections and under-18 conceptions: a view from general practice. *Sex Transm Infect*, 83, 171-a-172.
- OED (1998) *The new Oxford dictionary of English*, Oxford, Clarendon Press.
- OLIVER, S., HARDEN, A., REES, R., SHEPHERD, J., BRUNTON, G. & OAKLEY, A. (2007) Young people and mental health: novel methods for systematic review of research on barriers and facilitators. *Health Educ Res*.
- OLIVER, S. H., A; REES, R; SHEPHERD, J; BRUNTON, G; GARCIA, J; OAKLEY, A (2005a) An emerging framework for including different types of evidence in systematic reviews for public policy. *Evaluation*, 11, 428-446.
- OLIVER, S. H., A; REES, R; SHEPHERD, J; BRUNTON, G; GARCIA, J; OAKLEY, A (2005b) An emerging framework for including different types of evidence in systematic reviews for public policy. *Evaluation*, 11, 428-446.
- OLIVER, S. P., G (EDS) (2001) *Using Research for Effective Health Promotion*, Buckingham, Open University Press.
- OLIVER, S. R., REES, R. W., CLARKE-JONES, L., MILNE, R., OAKLEY, A. R., GABBAY, J., et al. (2008) A multidimensional conceptual framework for analysing public involvement in health services research. *Health Expect*, 11, 72-84.
- OLSON, K. (2001) Using qualitative research in clinical practice. IN MORSE, J. M., SWANSON, J., M & KUZEL, A. J. (Eds.) *The nature of qualitative evidence*. Thousand Oaks, CA, Sage.
- ORLANDO REGIONAL MEDICAL CENTER (2005) Fluid Resuscitation. Department of Surgical Education, Orlando Regional Medical Center.
- OSTROM, E. (2007) Institutional rational choice: an assessment of the institutional analysis and development framework. IN SABATIER, P. A. (Ed.) *Theories of the policy process 2nd Edition*. Boulder, Westview Press.
- OTTESEN, M. M., KOBER, L., JORGENSEN, S. & TORP-PEDERSEN, C. (2001) Consequences of overutilization and underutilization of thrombolytic therapy in clinical practice. TRACE Study Group. TRAndolapril Cardiac Evaluation. *J Am Coll Cardiol*, 37, 1581-7.
- OXMAN, A. D. & GUYATT, G. H. (1992) A consumer's guide to subgroup analyses. *Ann Intern Med*, 116, 78-84.
- OXMAN, A. D., LAVIS, J. N. & FRETHEIM, A. (2007) Use of evidence in WHO recommendations. *The Lancet*, 369, 1883-1889.

- OXMAN, A. D., SCHUNEMANN, H. J. & FRETHEIM, A. (2006) Improving the use of research evidence in guideline development: 7. Deciding what evidence to include. *Health Res Policy Syst*, 4, 19.
- OXMAN, A. D., THOMSON, M. A., DAVIS, D. A. & HAYNES, R. B. (1995) No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *Cmaj*, 153, 1423 - 1431.
- PAGLIARI, C. & GRIMSHAW, J. (2002) Impact of group structure and process on multidisciplinary evidence-based guideline development: an observational study. *J Eval Clin Pract*, 8, 145-53.
- PAGLIARI, C., GRIMSHAW, J. & ECCLES, M. (2001) The potential influence of small group processes on guideline development. *J Eval Clin Pract*, 7, 165-73.
- PATERSON, B., THORNE, S., CANAM, C. & JILLINGS, C. (2001) *Meta-Study of Qualitative Health Research - A practical Guide to Meta-Analysis and Meta-Synthesis*, Thousand Oaks, CA, Sage.
- PATTON, M. (1997) *Utilization-focused evaluation*, Newbury Park, California, Sage Publications.
- PAWSON, R. (2002) Evidence-based policy: the promise of 'Realist Synthesis'. *Evaluation*, 8, 340-358.
- PAWSON, R., GREENHALGH, T., HARVEY, G. & WALSHE, K. (2005) Realist review--a new method of systematic review designed for complex policy interventions. *J Health Serv Res Policy*, 10 Suppl 1, 21-34.
- PEARSON, K. (1904) Report on certain enteric fever inoculation statistics. *BMJ*, 1243-1246.
- PECKHAM, S., WILLMOTT, M., ALLEN, P., ANDERSON, S. & GOODWIN, N. (2008) Assessing the impact of the NHS Service Delivery and Organisation Research and Development Programme. *Evidence & Policy: A Journal of Research, Debate and Practice*, 4, 313-330.
- PEDEN, M., OYEBITE, K., OZANNE-SMITH, J., HYDER, A., BRANCHE, C., FAZLUR RAHMAN, A., et al. (2004a) World report on child injury prevention. Geneva, World Health Organization.
- PEDEN, M., SCURFIELD, R., SLEET, D., MOHAN, D., HYDER, A., JARAWAN, E., et al. (Eds.) (2004b) *World report on road traffic injury prevention* <http://whqlibdoc.who.int/publications/2004/9241562609.pdf>, WHO, Geneva.
- PEREL, P., YANAGAWA, T., BUNN, F., ROBERTS, I., WENTZ, R. & PIERRO, A. (2006) Nutritional support for head-injured patients. *Cochrane Database Syst Rev*, CD001530.
- PERNEGER, T. V. (2004) Relation between online "hit counts" and subsequent citations: prospective study of research papers in the BMJ.
- PETRIE, J., BUNN, F. & BYRNE, G. (2007a) Parenting programmes for preventing tobacco, alcohol or drugs misuse in children <18: a systematic review. *Health Educ Res*, 22, 177-91.
- PETRIE, J., BUNN, F. & BYRNE, G. (2007b) Parenting programmes for preventing tobacco, alcohol or drugs misuse in children under 18: a systematic review. *Health Educ Res*, 22, 177-91.
- PETTICREW, M. & ROBERTS, H. (2003) Evidence, hierarchies, and typologies: horses for courses. *J Epidemiol Community Health*, 57, 527-529.
- PETTICREW, M., WHITEHEAD, M., MACINTYRE, S. J., GRAHAM, H. & EGAN, M. (2004) Evidence for public health policy on inequalities: 1: the reality according to policymakers. *J Epidemiol Community Health*, 58, 811-6.

- PHILIPS, B., SACKETT, D., BADENOCH, D., STRAUS, S., HAYNES, B., DAWES, M. (2001) Oxford Centre for Evidence Based Medicine Levels of Evidence <http://www.cebm.net/index.aspx?o=1025> Accessed August 2008.
- PILDAL, J., HROBJARTSSON, A., JORGENSEN, K. J., HILDEN, J., ALTMAN, D. G. & GOTZSCHE, P. C. (2007) Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *Int J Epidemiol*, 36, 847-57.
- PILKINGTON, P. (2003) Speed cameras under attack in the United Kingdom. *Inj Prev*, 9, 293-294.
- POCOCK, M., TRIVEDI, D., WILLS, W., BUNN, F. & MAGNUSSON, J. (2009) Parental perceptions regarding healthy behaviours for preventing overweight and obesity in young children: a systematic review of qualitative studies. *Obes Rev*.
- POCOCK, M., TRIVEDI, D., WILLS, W., BUNN, F. & MAGNUSSON, J. (2010) Parental perceptions regarding healthy behaviours for preventing overweight and obesity in young children: a systematic review of qualitative studies. *Obesity Reviews*, 11, 338-353.
- POPAY, J., ROBERTS, H., SOWDEN, A., PETTICREW, M., ARAI, L., RODGERS, M., et al. (2006) Guidance on the Conduct of Narrative Synthesis in Systematic Reviews A product from the ESRC Methods Programme.
- POPAY, J., ROGERS, A. & WILLIAMS, G. (1998) Rationale and standards for the systematic review of qualitative literature in health services research. *Qual Health Res*, 8, 341-51.
- POPE, C., ZIEBLAND, S. & MAYS, N. (2000) Qualitative research in health care. Analysing qualitative data. *Bmj*, 320, 114-6.
- POWELL-TUCK, J., GOSLING, P. & LOBO, D. E. A. (2008) British consensus guidelines on intravenous fluid therapy for adult surgical patients. IN GIFTASUP (Ed.).
- PREST (2006) Smart innovation: a practical guide to evaluating innovation programmes. Manchester, Policy Research in Engineering, Science and Technology, University of Manchester.
- PROGRAM IN POLICY DECISION-MAKING (2003) Develop a systematic approach to knowledge transfer. IN UNIVERSITY, M. (Ed.).
- PYRA, K. (2003) Knowledge Translation: A review of the literature. Prepared for the Nova Scotia Health Research Foundation
- QUEBEC SANTE PUBLIQUE (2006) Urban Transportation a Question of Health. IN MONTREAL, A. D. L. S. E. D. S. S. D. (Ed.).
- RAWLINS, M. (2009) Economist Debates: Health care: Statements <http://www.economist.com/debate/days/view/394>. *The Economist*.
- RAWLINS, M. & DILLON, A. (2005) What's the evidence that NICE guidance has been implemented?: More recent data on NICE implementation show different picture. *BMJ*, 330, 1086-.
- REEVES, B., DEEKS, JJ., HIGGINS, PT., WELLS, A., ON BEHALF OF THE COCHRANE NON-RANDOMISED STUDIES GROUP (2008) Including Non randomised studies (Chapter 13). In the Cochrane Handbook Version 5.0.0 <http://www.cochrane-handbook.org/>.
- REINHART, K., BRUNKHORST, F. M., BONE, H. G., GERLACH, H., GRUNDLING, M., KREYMAN, G., et al. (2006) Diagnosis and therapy of sepsis. S2 guidelines of the German Sepsis Society and the German Interdisciplinary Society of Intensive Care and Emergency Medicine (DIVI). *Clinical Research in Cardiology*, 95, 429-454.
- REVELL, M., PORTER, K. & GREAVES, I. (2002) Fluid resuscitation in prehospital trauma care: a consensus view. *Emerg Med J*, 19, 494-8.

- RIST, R. (1994) Influencing the policy process with qualitative research. IN DENZIN, N. & LINCOLN, Y. (Eds.) *Handbook of qualitative research*. Thousand Oaks, Sage.
- RIST, R. (2003) Influencing the policy process with qualitative research. IN DENZIN, N. & LINCOLN, Y. (Eds.) *collecting and Interpreting Qualitative Research*. Sage.
- RIZOLI, S. B. (2003) Crystalloids and colloids in trauma resuscitation: a brief overview of the current debate. *J Trauma*, 54, S82-8.
- ROBERTS, I. (1993) Why have child pedestrian death rates fallen? *Bmj*, 306, 1737-9.
- ROBERTS, I. (1996) Safely to school? *Lancet*, 347, 1642.
- ROBERTS, I., ALDERSON, P., BUNN, F., CHINNOCK, P., KER, K. & SCHIERHOUT, G. (2004) Colloids versus crystalloids for fluid resuscitation in critically ill patients. *Cochrane Database Syst Rev*, CD000567.
- ROBERTS, I. & ARNOLD, E. (2007) Policy at the crossroads: climate change and injury control. *Inj Prev*, 13, 222-223.
- ROBERTS, I. & BUNN, F. (2002) Egg on their faces. The story of human albumin solution. *Eval Health Prof*, 25, 130-8.
- ROBERTS, I., EDWARDS, P. & MCLELLAND, B. (1999) More on albumin. Use of human albumin in UK fell substantially when systematic review was published. *Bmj*, 318, 1214-5.
- ROBERTS, I. & KWAN, I. (2001) School based driver education for the prevention of traffic crashes. *Cochrane Database Syst Rev*, CD003201.
- ROBERTS, I., WENTZ, R. & EDWARDS, P. (2006) Car manufacturers and global road safety: a word frequency analysis of road safety documents. *Inj Prev*, 12, 320-2.
- ROBERTS, K. A., DIXON-WOODS, M., FITZPATRICK, R., ABRAMS, K. R. & JONES, D. R. (2002) Factors affecting uptake of childhood immunisation: a Bayesian synthesis of qualitative and quantitative evidence. *Lancet*, 360, 1596-9.
- ROGERS, E. M. (1983) *Diffusion of innovation 4th Edition*, The Free Press, New York.
- ROGSTAD, K., THOMAS, A., WILLIAMS, O., FOSTER, G., MUNDAY, P., ROBINSON, A., et al. (2008) United Kingdom National Guideline on the management of sexually transmitted infections and related conditions in children and young people. Clinical Effectiveness Group, British Association for Sexual Health and HIV. <http://www.bashh.org/documents/2028/2028.pdf>.
- ROLLO, J. (1801) A short Account of the Royal Artillery Hospital at Woolwich with some Observations on the Management of Artillery Soldiers, Respecting the Preservation of Health. Addressed to the officers of the regiment and dedicated to the Master-General and Board of Ordinance. . London, Mawman.
- ROTSTEIN, D. & LAUPACIS, A. (2004) Differences between systematic reviews and health technology assessments: a trade-off between the ideals of scientific rigor and the realities of policy making. *Int J Technol Assess Health Care*, 20, 177-83.
- RUTHERFORD, D. (2002) The Amsterdam group. *Globe*, 1, 8-11.
- RYAN, J., PIERCY, J. & JAMES, P. (2004) Assessment of NICE guidance on two surgical procedures. *Lancet*, 363, 1525-6.
- RYCROFT-MALONE, J. (2006) The politics of the evidence based practice movements: Legacies and current challenges. *Journal of Research in Nursing*, 11, 95-108.
- RYCROFT-MALONE, J., HARVEY, G., SEERS, K., KITSON, A., MCCORMACK, B. & TITCHEN, A. (2004a) An exploration of the factors that influence the implementation of evidence into practice. *J Clin Nurs*, 13, 913-24.
- RYCROFT-MALONE, J., SEERS, K., TITCHEN, A., HARVEY, G., KITSON, A. & MCCORMACK, B. (2004b) What counts as evidence in evidence-based practice? *J Adv Nurs*, 47, 81-90.
- SABATIER, P. & JENKINS-SMITH, H. (1993) *Policy Change and Learning: An Advocacy Coalition Approach*, Boulder CO, Westview Press.

- SABATIER, P. A. (1988) An advocacy coalition framework of policy change and the role of policy-oriented learning therein. *Policy Sciences*, 21, 129-168.
- SABATIER, P. A. (2007) *Theories of the policy process*, Westview Pr.
- SACKETT, D. L., ROSENBERG, W. M., GRAY, J. A., HAYNES, R. B. & RICHARDSON, W. S. (1996) Evidence based medicine: what it is and what it isn't. *Bmj*, 312, 71-2.
- SALK, E. D., SCHRIGER, D. L., HUBBELL, K. A. & SCHWARTZ, B. L. (1998) Effect of visual cues, vital signs, and protocols on triage: a prospective randomized crossover trial. *Ann Emerg Med*, 32, 655-64.
- SANDELOWSKI, M. & BARROSO, J. (2007) *Handbook for Synthesizing Qualitative Research*, New York, Springer Publishing Company.
- SANDELOWSKI, M., DOCHERTY, S. & EMDEN, C. (1997) Focus on qualitative methods. Qualitative metasynthesis: issues and techniques. *Res Nurs Health*, 20, 365-71.
- SCHIERHOUT, G. & ROBERTS, I. (1998) Fluid resuscitation with colloid or crystalloid solutions in critically ill patients: a systematic review of randomised trials. *Bmj*, 316, 961-4.
- SCHLAGER, E. (2007) A comparison of frameworks, theories and models of policy processes. IN SABATIER, P. A. (Ed.) *Theories of the policy process 2nd Edition*. Boulder, Westview Press.
- SCHREIBER, R., CROOKS, D. & STEIN, P. (1997) Qualitative Meta-analysis. IN MORSE, J. M. (Ed.) *completing a Qualitative Project: Details and Dialogue*. Thousand Islands, CA, Sage.
- SCHUBERT, A. & BRAUN, T. (1996) Cross-field normalization of scientometric indicators. *Scientometrics*, 36, 311-324.
- SCHUERMAN, J. S., H; MACDONALD, G; FORSLUND, M; DE MOYA, D; BORUCH, R. (2002) The Campbell Collaboration. *Research on Social Work Practice*, 12, 309-317.
- SCHULZ, K. F., CHALMERS, I., HAYES, R. J. & ALTMAN, D. G. (1995) Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *Jama*, 273, 408-12.
- SCHWANDT, T. A. (1996) Farewell to Criteriology. 2, 58-72.
- SEGLIN, P. O. (1997) Why the impact factor of journals should not be used for evaluating research. *BMJ*, 314, 497.
- SELECT COMMITTEE ON TRANSPORT (2002) House of Commons Transport, Local Government and the Regions Committee
Road traffic speed, ninth report of session 2001-2002. Vol 1. London, Stationery Office (HC 557-I.).
- SHAW, R. L., BOOTH, A., SUTTON, A. J., MILLER, T., SMITH, J. A., YOUNG, B., et al. (2004) Finding qualitative research: an evaluation of search strategies. *BMC Med Res Methodol*, 4, 5.
- SHELDON, T. A. (2005) Making evidence synthesis more useful for management and policy-making. *J Health Serv Res Policy*, 10 Suppl 1, 1-5.
- SHELDON, T. A., CULLUM, N., DAWSON, D., LANKSHEAR, A., LOWSON, K., WATT, I., et al. (2004) What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients' notes, and interviews. *Bmj*, 329, 999.
- SHEPHERD, J., HARDEN, A., REES, R., BRUNTON, G., OLIVER, S. & OAKLEY, A. (2001) Young People and Healthy Eating: A Systematic Review of Barriers and Facilitators. Londobn, EPPI-Centre, Social Science Research Unit.
- SHORT, S. (1997) Elective affinities: research and health policy development. . IN GARDNER, H. (Ed.) *Health policy in Australia*. Melbourne, Oxford University Press.

- SIGN (2008) Management of acute upper and lower gastrointestinal bleeding. IN NETWORK, S. I. G. (Ed.).
- SMITH, A. (2008) Benchmarking Google Scholar with the New Zealand PBRF research assessment exercise. *Scientometrics*, 74, 309-316.
- SMITH, R. (2000) The failings of NICE. *Bmj*, 321, 1363-4.
- SMITH, R. (2005) Medical journals are an extension of the marketing arm of pharmaceutical companies. *PLoS Med*, 2, e138.
- SOAR, J., DEAKIN, C. D., NOLAN, J. P., ABBAS, G., ALFONZO, A., HANDLEY, A. J., et al. (2006) Cardiac arrest under special circumstances: Section 7 of the guidelines for resuscitation 2005 of the European Resuscitation Council. *Kreislaufstillstand unter besonderen umständen - Abschnitt 7 der leitlinien zur reanimation 2005 des European Resuscitation Council*, 9, 123-154.
- SOCIAL EXCLUSION UNIT (1999) Teenage Pregnancy report by the Social Exclusion Unit. London: Stationery Office.
- SOLESBURY, W. (2001) Evidence based policy: whence it came and where it's going. *ESRC UK centre for Evidence Based Policy and Practice, Department of Politics, Queen Mary*.
- SOLESBURY., W. (2001) Evidence Based Policy: Whence it Came and Where it's going. IN PRACTICE, E. U. C. F. E. B. P. A. (Ed.), Queen Mary, University of London.
- SONKIN, B., EDWARDS, P., ROBERTS, I. & GREEN, J. (2006) Walking, cycling and transport safety: an analysis of child road deaths. *J R Soc Med*, 99, 402-5.
- SOPER, B. & HANNEY, S. R. (2007) Lessons from the evaluation of the UK's NHS R&D implementation methods programme. *Implement Sci*, 2, 7.
- SPAHN, D. R., CERNY, V., COATS, T. J., DURANTEAU, J., FERNANDEZ-MONDEJAR, E., GORDINI, G., et al. (2007) Management of bleeding following major trauma: A European guideline. *Critical Care*, 11.
- SPAPEN, H., DAMAS, P., D'ORIO, V., SABBE, M., VOGELAERS, D., BYL, B., et al. (1999) Consensus statement on the treatment of septic shock. *Acta Clinica Belgica*, 54, 256-262.
- SPELLER, V., LEARMONTH, A. & HARRISON, D. (1997) The search for evidence of effective health promotion. *Bmj*, 315, 361-3.
- SPENCE, D. (2009) We need ideas based medicine.
- SPENCER, L., RITCHIE, J., LEWIS, J. & DILLON, L. (2003) Quality in Qualitative Evaluation: A Framework for Assessing Research Evidence. London (UK): Government Chief Social Researcher's Office, Cabinet Office. Available from www.gsr.gov.uk/downloads/evaluating_policy/a_quality_framework.pdf.
- SPIEGELHALTER, D. J., MYLES, J. P., JONES, D. R. & ABRAMS, K. R. (2000) Bayesian methods in health technology assessment: a review. *Health Technol Assess*, 4, 1-130.
- SPIELMANS, G. I. & PARRY, P. I. (2010) From Evidence-based Medicine to Marketing-based Medicine: Evidence from Internal Industry Documents. *Journal of Bioethical Inquiry*, DOI 10.1007/s11673-010-9280-8, 1-17.
- STAINSBY, D., MACLENNAN, S. & HAMILTON, P. J. (2000) Management of massive blood loss: A template guideline. *British Journal of Anaesthesia*, 85, 487-491.
- STAINSBY, D., MACLENNAN, S., THOMAS, D., ISAAC, J. & HAMILTON, P. J. (2006) Guidelines on the management of massive blood loss. *British Journal of Haematology*, 135, 634-641.
- STATIONARY OFFICE (1999) Modernising government: presented to parliament by the prime minister and the minister for the Cabinet Office by command of Her Majesty. London: Stationary Office.
- STONE, D. (1989) Upside down prevention. *Health Serv J*, 99, 890-1.

- SUSSMAN, S., VALENTE, T. W., ROHRBACH, L. A., SKARA, S. & ANN PENTZ, M. (2006) Translation in the Health Professions: Converting Science into Action. *Eval Health Prof*, 29, 7-32.
- SUTTON, A. J. & ABRAMS, K. R. (2001) Bayesian methods in meta-analysis and evidence synthesis. *Stat Methods Med Res*, 10, 277-303.
- SWANSON, J., M (2001) The nature of outcomes. IN MORSE, J. M., SWANSON, J. M. & KUZEL, A. J. (Eds.) *The nature of quantitative evidence*. Thousand Oaks, CA, Sage.
- SWOV (2009) Traffic education of children 4-12 years
http://www.swov.nl/rapport/Factsheets/UK/FS_Traffic_education_children.pdf.
- TARNOW-MORDI, W. O. & HEALY, M. J. (1999) Distinguishing between "no evidence of effect" and "evidence of no effect" in randomised controlled trials and other comparisons. *Arch Dis Child*, 80, 210-1.
- TERRY, F. (2000) Transport: beyond predict and provide. IN DAVIES, H. N., S. SMITH, P. (Ed.) *What works? Evidence-based policy and practice in public services*. Bristol, The Policy Press.
- THE ALBUMIN REVIEWERS, BUNN, F., LI WAN PO, A., LI, L., PEARSON, M., ROBERTS, I., et al. (2004) Human albumin solution for resuscitation and volume expansion in critically ill patients. *Cochrane Database of Systematic Reviews*. Chichester, UK, John Wiley & Sons, Ltd.
- THE COCHRANE LIBRARY (2008) 2007 Impact Factor Cochrane Database of Systematic Reviews (CDSR) Frequently Asked Questions. The Cochrane Collaboration.
- THE PAEDIATRIC ACCIDENT AND EMERGENCY RESEARCH GROUP The management of a child with a decreased conscious level. A nationally developed evidence-based guideline for hospital practitioners. The University of Nottingham.
- THOMAS, J. & HARDEN, A. (2008) Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8, 45.
- THOMAS, J., HARDEN, A., OAKLEY, A., OLIVER, S., SUTCLIFFE, K., REES, R., et al. (2004) Integrating qualitative research with trials in systematic reviews. *Bmj*, 328, 1010-2.
- THOMAS, L., CULLUM, N., MCCOLL, E., ROUSSEAU, N., SOUTTER, J. & STEEN, N. (2000) Guidelines in professions allied to medicine. *Cochrane Database Syst Rev*, CD000349.
- THOMPSON, D. C. & RIVARA, F. P. (2000) Pool fencing for preventing drowning in children. *Cochrane Database Syst Rev*, CD001047.
- THOMPSON, D. C., RIVARA, F. P. & THOMPSON, R. (2000) Helmets for preventing head and facial injuries in bicyclists. *Cochrane Database Syst Rev*, CD001855.
- THOMPSON, S. G. & POCOCK, S. J. (1991) Can meta-analyses be trusted? *Lancet*, 338, 1127-30.
- TIMMERMANS, S. & MAUCK, A. (2005) The promises and pitfalls of evidence-based medicine. *Health Aff (Millwood)*, 24, 18-28.
- TOROYAN, T. & PEDEN (2007) Youth and road safety. Geneva, World Health Organization
- TOWNER, E. & ERRINGTON, G. (2004) How can injuries in children and older people be prevented? *Health Evidence Network synthesis report*.
<http://www.euro.who.int/Document/E84938.pdf> accessed March 2009.
Copenhagen, WHO Regional Office for Europe.
- TRANSPORT BILL (1981) Transport bill: compulsory wearing of seat belts. House of Commons Official Report (Hansard) 1981 .
- TRAYNOR, M. (2002) The oil crisis, risk and evidence-based practice. *Nurs Inq*, 9, 162-9.

- TRIVEDI, D., BUNN, F., GRAHAM, M. & WENTZ, R. (2007a) Preventing teenage pregnancy: evidence from the systematic reviews. IN BAKER, P., GURTHRIE, K., HUTCHINSON, C., KAN, R. & WELLINGS, K. (Eds.) *Teenage Pregnancy and Reproductive Health*. London, The Royal College of Obstetricians and Gynaecologists.
- TRIVEDI, D., BUNN, F., GRAHAM, M. & WENTZ, R. (2007b) Update on review of reviews on teenage pregnancy and parenthood Submitted as an Addendum to the first evidence briefing 2003. On behalf of NICE.
- UNICEF (2001) A league table of teenage births in rich nations. Innocenti Report Card No.3 Florence Italy. United Nations Children's Fund.
- UPSHUR, R. E. G. (2001) The status of qualitative research as evidence. IN MORSE, J. M., SWANSON, J. M. & KUZEL, A. J. (Eds.) *The nature of qualitative evidence*. Thousand Oaks, CA, Sage.
- VAN LEEUWEN, T. N., MOED, H. F. & REEDIJK, J. (1999) Critical comments on Institute for Scientific Information impact factors: a sample of inorganic molecular chemistry journals. *Journal of Information Science*, 25, 489-498.
- VANDENBROUCKE, J. P., VON ELM, E., ALTMAN, D. G., GOTZSCHE, P. C., MULROW, C. D., POCOCK, S. J., et al. (2007) Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Med*, 4, e297.
- VERNICK, J. S., LI, G., OGAITIS, S., MACKENZIE, E. J., BAKER, S. P. & GIELEN, A. C. (1999) Effects of high school driver education on motor vehicle crashes, violations, and licensure. *Am J Prev Med*, 16, 40-6.
- WAGENAAR, A. C. (1999) Importance of systematic reviews and meta-analyses for research and practice. *Am J Prev Med*, 16, 9-11.
- WAGG, A. & BUNN, F. (2007) Unassisted pelvic floor exercises for postnatal women: a systematic review. *J Adv Nurs*, 58, 407-17.
- WAKEFIELD, A. J., MURCH, S. H., ANTHONY, A., LINNELL, J., CASSON, D. M., MALIK, M., et al. (1998) Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet*, 351, 637-41.
- WALKER, S., PALMER, S. & SCULPHER, M. (2007) The role of NICE technology appraisal in NHS rationing. *Br Med Bull*, 81-82, 51-64.
- WALT, G. (1994) *Health policy: an introduction to process and power*, London, Zed Books Ltd.
- WALT, G. & GILSON, L. (1994) Reforming the health sector in developing countries: the central role of policy analysis. *Health Policy Plan*, 9, 353-70.
- WALT, G., SHIFFMAN, J., SCHNEIDER, H., MURRAY, S. F., BRUGHA, R. & GILSON, L. (2008) 'Doing' health policy analysis: methodological and conceptual reflections and challenges. *Health Policy Plan*, 23, 308-17.
- WALTER, I. N., SM. & DAVIES, HTO (2003a) Research Impact: A Cross Sector Review', Research Unit for Research Utilisation. University of St Andrews.
- WALTER, I. N., SM. & DAVIES, HTO. (2003b) Developing a Taxonomy of Interventions used to Increase the Impact of Research', Discussion Paper 3, Research Unit for Research Utilisation, . University of St Andrews.
- WALTERS, L. A., WILCZYNSKI, N. L. & HAYNES, R. B. (2006) Developing optimal search strategies for retrieving clinically relevant qualitative studies in EMBASE. *Qual Health Res*, 16, 162-8.
- WARD, H. (2007) One-to-one counselling for STI prevention: not so much whether as how. *Sexually Transmitted Infections*, 83, 1-1.
- WATT, A., CAMERON, A., STURM, L., LATHLEAN, T., BABIDGE, W., BLAMEY, S., et al. (2008) Rapid versus full systematic reviews: validity in clinical practice? *ANZ J Surg*, 78, 1037-40.

- WEBB, M. (2009) Rapid Review of Hospital Element of Unscheduled Care Services in North Wales: Literature Review on Best Practice in Unscheduled Care Services. National Public Health Service for Wales
<http://www.wales.nhs.uk/sites3/Documents/837/UC%20literature%20review.pdf>.
- WEISS, C. (1976) Using research in the policy process: potential and constraints. *Policy Studies Journal*, 4, 224-228.
- WEISS, C. (1977) research for policy's sake: the enlightenment function of social research. *Policy Analysis*, 3, 531-47.
- WEISS, C. (1979) The many meanings of research utilization. *Public Administration Review*, 39, 426-31.
- WEISS, C. H. (1998) Have We Learned Anything New About the Use of Evaluation? *American Journal of Evaluation*, 19, 21-33.
- WENNERBERG, J. E. (1973) "Small area variations in health care delivery". *Science*, 82, 1102-1108.
- WENNERBERG, J. E. (2002) Unwarranted variations in healthcare delivery: implications for academic medical centres. *Bmj*, 325, 961-4.
- WENTZ, R. (2002) Visibility of research: FUTON bias. *Lancet*, 360, 1256.
- WENTZ, R., ROBERTS, I., BUNN, F., EDWARDS, P., KWAN, I. & LEFEBVRE, C. (2001) Identifying controlled evaluation studies of road safety interventions: Searching for needles in a haystack. *Journal of Safety Research*, 32, 267-276.
- WENZEL, V., RUSSO, S., ARNTZ, H. R., BAHR, J., BAUBIN, M. A., BÖTTIGER, B. W., et al. (2006) The new 2005 resuscitation guidelines of the European Resuscitation Council. Comments and supplements. *Die neuen reanimationsleitlinien 2005 des European Resuscitation Council. Kommentar und ergänzungen*, 55, 958-979.
- WHITE, D. G. (2001) Evaluating evidence and making judgements of study quality: loss of evidence and risks to policy and practice decisions. *Critical Public Health*, 11, 3 - 17.
- WHITEHEAD, A. & WHITEHEAD, J. (1991) A general parametric approach to the meta-analysis of randomized clinical trials. *Stat Med*, 10, 1665-77.
- WHITEHEAD, M., PETTICREW, M., GRAHAM, H., MACINTYRE, S. J., BAMBRA, C. & EGAN, M. (2004) Evidence for public health policy on inequalities: 2: assembling the evidence jigsaw. *J Epidemiol Community Health*, 58, 817-21.
- WHO (2001) Alcohol in the European region—consumption, harm and policies. Copenhagen. *WHO Regional Office for Europe*. Copenhagen, World Health Organization.
- WHO (2006) Helmets - A road safety manual for decision makers and practitioners. Geneva, World Health organization.
- WHO (2008) The Global Burden of Disease: 2004 update
http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_AnnexA.pdf. Geneva, World Health Organisation.
- WHO (2009) Global status report on road safety.
http://whqlibdoc.who.int/publications/2009/9789241563840_eng.pdf. Geneva, World Health Organization.
- WILCZYNSKI, N. L., MARKS, S. & HAYNES, R. B. (2007) Search strategies for identifying qualitative studies in CINAHL. *Qual Health Res*, 17, 705-10.
- WILLET, W., KOPLAN, J. & NUGENT, R. (2006) Prevention of chronic disease by means of diet and lifestyle changes *Prevention*. Washington, The International Bank for Reconstruction and Development/ The World Bank.
- WILLIAMS, G. (2007) Should we ditch impact factors? *BMJ*, 334, 568-.

- WILLISON, D. J. & MACLEOD, S. M. (1999) The role of research evidence in pharmaceutical policy making: evidence when necessary but not necessarily evidence. *J Eval Clin Pract*, 5, 243-9.
- WILSON, P., PETTICREW, M. & BOOTH, A. (2009) After the gold rush? A systematic and critical review of general medical podcasts. *J R Soc Med*, 102, 69-74.
- WINGENS, M. (1990) Toward a General Utilization Theory: A Systems Theory Reformulation of the Two-Communities Metaphor. *Science Communication*, 12, 27-42.
- WONG, S. S., WILCZYNSKI, N. L. & HAYNES, R. B. (2004) Developing optimal search strategies for detecting clinically relevant qualitative studies in MEDLINE. *Stud Health Technol Inform*, 107, 311-6.
- WOOD, L., EGGER, M., GLUUD, L. L., SCHULZ, K. F., JUNI, P., ALTMAN, D. G., et al. (2008) Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *Bmj*, 336, 601-5.
- WOODING, S., HANNEY, S., BUXTON, M. & GRANT, J. (2004) The returns from arthritis research: A report prepared by RAND Europe for the Arthritis Research Campaign (ARC) Volume 1 Approach, Analysis and Recommendations. Cambridge, RAND Europe, <http://rand.org/pubs/monograph/MG251/index.html>.
- WOODING, S. N., E. KLAUTZER, L. RUBIN, J. HANNEY, S. GRANT, J. (2007) Policy and practice impacts of research funded by the Economic and Social Research Council. a case study of the Future of Work programme, approach and analysis. Cambridge, RAND Europe.
- WOOLF, S. H. (1997) Shared decision-making: the case for letting patients decide which choice is best. *J Fam Pract*, 45, 205-8.
- WOOLF, S. H., GROL, R., HUTCHINSON, A., ECCLES, M. & GRIMSHAW, J. (1999) Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *Bmj*, 318, 527 - 530.
- YANAGAWA, T., BUNN, F., ROBERTS, I., WENTZ, R. & PIERRO, A. (2000) Nutritional support for head-injured patients. *Cochrane Database Syst Rev*, CD001530.
- ZAHARIADIS, N. (2007) The multiple streams framework: Structure, limitations, prospects. IN SABATIER, P. (Ed.) *Theories of the Policy Process*. Westview Press.

Appendix 1: Full list of authors publications

Bunn F, Kendall S. Does nursing research impact on policy? A case study of health visiting research and UK health policy. *Journal of Research in Nursing* 2011 16:169-191

Bunn F. The Cochrane Injuries Group celebrates the publication of its 100th Review: time to reflect on impact. *Injury Prevention* 2010 16(3)

Pocock M, Trivedi D, Wills W, **Bunn F**, Magnusson J. Parental Perceptions regarding Healthy Behaviours for Preventing Overweight and Obesity in Young Children : A systematic review of qualitative studies. *Obesity Review* 2009 doi: 10.1111/j.1467-789X.2009.00648.x

Brooks F, **Bunn F**, Morgan J. Transition for adolescents with long-term conditions: event to process. *British Journal of Community Nursing* 2009;14(7):301-304

Offredy M, **Bunn F**, Morgan J. Case management in long term conditions: an inconsistent journey? *British Journal of Community Nursing* 2009;14(6):252-257

Wilson P, **Bunn F**, Morgan J. A mapping of the evidence on integrated long term condition services. *British Journal of Community Nursing* 2009;14(5); 202-206

Trivedi D, Brooks F, **Bunn F**, Graham M. Early fatherhood: a mapping of the evidence base relating to pregnancy prevention and parenting support. *Health Education Research Advance Access published online on April 30, 2009 Health Education Research*, doi:10.1093/her/cyp025

Bunn F, Dickinson A, Barnett-Page E, Horton K. A systematic review of older people's perceptions of facilitators and barriers to participation in falls prevention interventions. *Ageing & Society* 2008;28:449-472

Bunn F, Trivedi D, Ashraf S. Colloid solutions for fluid resuscitation (Cochrane Review) [Cochrane Database Syst Rev](#). 2008 Jan 23;(1):CD001319.

Petrie J, **Bunn F**. Parenting programmes for preventing tobacco, alcohol or drugs misuse in children <18: a systematic review. *Health Education Research* 2007;22(2):177-91

Wagg A, **Bunn F**. Unassisted pelvic floor exercises for postnatal women: a systematic review. *Journal of Advanced Nursing* 2007; 58(4):1-11

Brooks F, Kendall S, **Bunn F**, Bindler R, Bruya M. The school nurse as navigator of the school health journey: developing the theory and evidence for policy. *Primary Health care Research & Development* 2007;8:226-234

Bunn F, Cunningham M, Handscomb K. Prophylactic antibiotics to prevent surgical site infection after breast cancer surgery. *The Cochrane Database of Systematic Reviews* 2006, Issue 2. Art. No.: CD005360. DOI: 10.1002/14651858.CD005360.pub2.

Perel P, Yanagawa T, **Bunn F**, Roberts I, Wentz R, Pierro A. Nutritional support for head injured patients. *Cochrane Database of Systematic Reviews* 2006 Oct 18;(4):CD001530

Bunn F, Byrne G, Kendall S. The effects of telephone consultation and triage on health care use and patient satisfaction: a systematic review. *British Journal of General Practice* 2005;55:956-961.

Ker K, Roberts I, Collier T, Renton F, **Bunn F**, Frost C. Post-licence driver education for the prevention of road traffic crashes: a systematic review of randomised controlled trials. *Accident Analysis and Prevention* 2005;37:305-313.

Bunn F, Byrne G, Kendall S. Telephone consultation and triage systems: effects on health care use and patient satisfaction (Cochrane Review): In: *The Cochrane Library*, Issue 2004(2). Chichester, UK: John Wiley & Sons, Ltd

Bunn F, Collier T, Frost C, Ker K, Roberts I. Traffic calming for the prevention of road traffic injuries: systematic review and meta-analysis. *Injury Prevention* 2003;9:200-204

Bunn F, Collier T, Frost C, Ker K, Roberts I. Area-wide traffic calming for preventing traffic related injuries (Cochrane Review): In: *The Cochrane Library* Issue 2 2003, Oxford, Update Software.

Bunn F, Alderson P, Hawkins V. Colloid solutions for fluid resuscitation (Cochrane Review): In: *The Cochrane Library*, Issue 3 2003. Oxford: Update Software

Bunn F, Roberts I, Tasker R, Akpa E. Hypertonic versus isotonic crystalloid for fluid resuscitation in critically ill patients (Cochrane Review). In *the Cochrane Library*, Issue 3 2003. Oxford Update Software

Alderson P, **Bunn F**, Lefebvre C, Li WP, Li L, Roberts I, Schierhout G. Human albumin solution for resuscitation and volume expansion in critically ill patients (Cochrane Review). In *the Cochrane Library*, Issue 3 2003. Oxford Update Software

Duperrex O, Roberts I, **Bunn F**. Safety education of pedestrians for injury prevention (Cochrane Review). In *the Cochrane Library* Issue 3 2003. Oxford Update Software.

Ker K, Roberts I, Collier T, Renton F, **Bunn F**. Post-licence driver education for the prevention of road traffic crashes (Cochrane Review). In *the Cochrane Library* Issue 3 2003. Oxford Update Software.

Kwan I, **Bunn F**, Roberts I; WHO Pre-Hospital Trauma Care Steering Committee. Timing and volume of fluid administration for patients with bleeding (Cochrane Library). In the Cochrane Library Issue 3 2003. Oxford Update Software

Kwan I, **Bunn F**, Roberts I. Spinal immobilisation for trauma patients (Cochrane Review). In the Cochrane Library, Issue 3 2003. Oxford Update Software

Duperrex O, **Bunn F**, Roberts I. Safety education of pedestrians for injury prevention: a systematic review of randomised controlled trials. *BMJ* 2002;324:1129

Kwan I, **Bunn F**, Roberts I, Wentz R. The development of a register of randomised controlled trials in pre-hospital trauma care. *Prehosp Emerg Care* 2002 Jan-Mar;6(1):27-30

Roberts I, **Bunn F**. Egg on their faces. The story of human albumin solution. *Eval Health Prof.* 2002; 25(1):130-8

Roberts I, Evans P, **Bunn F**, Kwan I, Crowhurst E. Is the normalisation of blood pressure in bleeding trauma patients harmful? *Lancet* 2001;357(9253):385-7

Sethi D, Kwan I, Kelly AM, Roberts I, **Bunn F**. Advanced trauma life support training for ambulance crews (Cochrane Review): In: The Cochrane Library, Issue 1 2001: Oxford: Update Software

Wentz R, Roberts I, **Bunn F**, Edwards P, Kwan I, Lefebvre C. Identifying controlled evaluation of road safety interventions: searching for needles in a haystack. *Journal of Safety Research* 2001;32(3):267-276

Dickinson K, **Bunn F**, Wentz R, Edwards P, Roberts I. Size and quality of randomised controlled trials in head injury: review of published studies. *BMJ* 2000; 320:1308-11

Yanagawa T, **Bunn F**, Roberts I, Wentz R, Pierro A. Nutritional support for head injured patients (Cochrane Review). In: The Cochrane Library, Issue 4, 2000. Oxford: Update Software.

Cochrane Injuries Group Albumin Reviewers (Alderson P, **Bunn F**, Lefebvre C, Li Wan Po A, Li L, Roberts I, Schierhout G). Human albumin administration in critically ill patients: systematic review of randomised controlled trials. *BMJ* 1998; 317:235-240

Book chapters

Trivedi D, **Bunn F**, Graham M, Wentz R. Preventing teenage pregnancy: evidence from the systematic reviews. In: *Teenage Pregnancy and Reproductive Health*. Baker P, Guthrie K, Hutchinson C, Kane R and Wellings K (Eds). RCOG Press 2007

Bunn F, Roberts I, DiGiuseppi C. Systematic reviews of Injury Studies. In Injury Control: Research and Program Evaluation Rivara F, Cummings P, Koepsell T, Grossman D, Maier R (Eds). Cambridge University Press 2001.

Bunn F. The needs of families and carers of stroke patients. In: Wolfe C, Rudd T, Beech R, eds. Stroke Services and Research. The Stroke Association, London 1996: 247-59

Bunn F. The role of the primary health care team. In: Wolfe C, Rudd T, Beech R, eds. Stroke Services and Research. The Stroke Association, London 1996: 227-36

Reports

Winter D, Bradshaw S, **Bunn F**, Wellsted D. Counselling and psychotherapy for the prevention of suicide: a systematic review of the evidence. British Association for Counselling & Psychotherapy 2009

Kendall S, Wilson P, Procter S, McNeilly E, **Bunn F**, Brooks F, Gage H. The nursing, midwifery and health visiting contribution to chronic disease management: a whole systems approach. *Report for the National Co-ordinating Centre for NHS Service Delivery and Organisation R&D (NCCSDO). June 2009*

Trivedi D, **Bunn F**, Forbes A, Scott C. Evaluating the nursing, midwifery and health visiting contribution to chronic disease management: An integration of three reviews. *Report for the National Co-ordinating Centre for NHS Service Delivery and Organisation R&D (NCCSDO). July 2009*

Wilson P, Brooks F, **Bunn F**, Offredy M, Morgan J. Long Term Conditions Service Delivery & Provision: National scoping & identification of innovative models fit for purpose in West Essex Primary Care Trust.

Bunn F, Wills WJ, Hughes J. Rapid Review of Brief Interventions to Prevent or Reduce Obesity in Adults and Children. On behalf of East of England Public Health Network. March 2008

Bunn F, McNeilly E, Kendall S, Nijjar A, Procter S, Trivedi D, Wentz R, Wilson P. Evaluating the nursing, midwifery and health visiting contribution to models of chronic disease management: a mapping of the literature. On behalf of the SDO July 2007

Bunn F, Brooks F, Applestone J, et al. Contraceptive advice and provision for the prevention of under 18 conceptions and STIs: a rapid review. On behalf of the National Institute for Health and Clinical Excellence. August 2006

Trivedi D, **Bunn F**, Graham M, Wentz R. Update on review of reviews on teenage pregnancy and parenthood. April 2006. Centre for Research in Primary and Community

Care, University of Hertfordshire. On behalf of the National Institute for Health and Clinical Excellence. August 2006

Bunn F, Dickinson A, Windle K. Interventions for reducing unplanned hospital admissions in older people: a mapping exercise. On behalf of the Innovation Forum Project and Hertfordshire County Council. October 2004.

Bunn F, Kwan I, Roberts I, Wentz R. Effectiveness of pre-hospital care: A report by the Cochrane Injuries Group for the World Health Organisation, January 2001: London

Bunn F, Roberts I, DiGuseppi C. A National Contract on Accidents. In Evidence from systematic reviews of research relevant to implementing the 'wider public health' agenda. NHS Centre for Reviews and Dissemination August 2000

(Available on the web site:

http://www.doh.gov.uk/pub/docs/doh/injury_prevention.pdf)

Bunn F, DiGuseppi C, Roberts I. Health of Londoners Injury Prevention. A rapid review prepared for the NHS Executive London. October 1999.

Appendix 2: Published papers included in submission

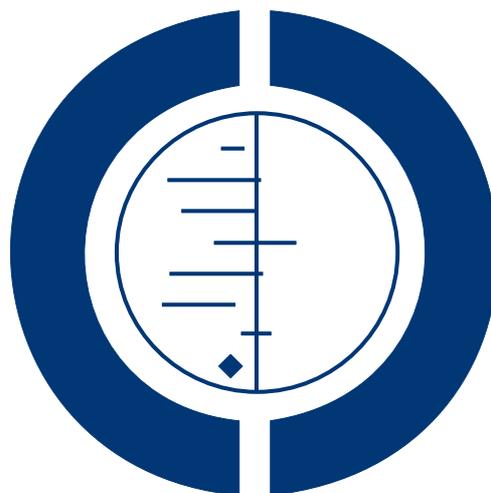
Included paper	Page ¹¹
Bunn F, Trivedi D, Ashraf S. Colloid solutions for fluid resuscitation (Cochrane Review) Cochrane Database Syst Rev. 2008 Jan 23;(1):CD001319.	250
Bunn F, Roberts IG, Tasker R, Trivedi D. Hypertonic versus near isotonic crystalloid for fluid resuscitation in critically ill patients. <i>Cochrane Database of Systematic Reviews</i> 2004, Issue 3. Art. No.: CD002045. DOI: 10.1002/14651858.CD002045.pub2	308
Cochrane Injuries Group Albumin Reviewers (Alderson P, Bunn F , Lefebvre C, Li Wan Po A, Li L, Roberts I, Schierhout G). Human albumin administration in critically ill patients: systematic review of randomised controlled trials. <i>BMJ</i> 1998; 317:235-240	330
Kwan I, Bunn F , Roberts I. Timing and volume of fluid administration for patients with bleeding. <i>Cochrane Database of Systematic Reviews</i> 2003, Issue 3. Art. No.: CD002245. DOI: 10.1002/14651858.CD002245	336
Bunn F, Collier T, Frost C, Ker K, Roberts I. Traffic calming for the prevention of road traffic injuries: systematic review and meta-analysis. <i>Injury Prevention</i> 2003;9:200-204	352
Duperrex O, Bunn F , Roberts I. Safety education of pedestrians for injury prevention: a systematic review of randomised controlled trials. <i>BMJ</i> 2002;324:1129	358
Bunn F, Dickinson A, Barnett-Page E, Horton K. A systematic review of older people's perceptions of facilitators and barriers to participation in falls prevention interventions. <i>Ageing & Society</i> 2008;28:449-472	363
Bunn F, Brooks F, Appleton J, et al. Contraceptive advice and provision for the prevention of under 18 conceptions and STIs: a rapid review. On behalf of the National Institute for Health and Clinical Excellence. August 2006	387
Executive summary included in appendix but full version of report can be found at http://www.nice.org.uk/nicemedia/pdf/STI_University_of_Hertfordshire_Final.pdf	
Bunn F, Byrne G, Kendall S. The effects of telephone consultation and triage on health care use and patient satisfaction: a systematic review. <i>British Journal of General Practice</i> 2005;55:956-961.	403

N.B Colloids versus crystalloids for fluid resuscitation in critically ill patients is not included as although I was an author on an earlier version I am no longer an author on the most recent version of the review. This is because Cochrane Injuries Group policy is to remove authors names from the review if they are no longer involved in updating the review. Previous authors are listed in the acknowledgements section.

¹¹ The page number in this submission is bottom centre

Colloid solutions for fluid resuscitation (Review)

Bunn F, Trivedi D, Ashraf S



**THE COCHRANE
COLLABORATION®**

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2008, Issue 4

<http://www.thecochranelibrary.com>



Colloid solutions for fluid resuscitation (Review)
Copyright © 2008 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
BACKGROUND	2
OBJECTIVES	2
METHODS	2
RESULTS	4
DISCUSSION	6
AUTHORS' CONCLUSIONS	6
ACKNOWLEDGEMENTS	6
REFERENCES	7
CHARACTERISTICS OF STUDIES	12
DATA AND ANALYSES	53
FEEDBACK	54
WHAT'S NEW	54
HISTORY	55
CONTRIBUTIONS OF AUTHORS	55
DECLARATIONS OF INTEREST	55
SOURCES OF SUPPORT	55
INDEX TERMS	55

[Intervention Review]

Colloid solutions for fluid resuscitation

Frances Bunn¹, Daksha Trivedi¹, Syed Ashraf²

¹Centre for Research in Primary and Community Care, University of Hertfordshire, Hatfield, UK. ²Psychiatry, Bedfordshire and Luton Mental Health Social Care Partnership NHS Trust, Bedford, UK

Contact address: Frances Bunn, Centre for Research in Primary and Community Care, University of Hertfordshire, College Lane, Hatfield, Hertfordshire, AL10 9PN, UK. f.bunn@herts.ac.uk.

Editorial group: Cochrane Injuries Group.

Publication status and date: Edited (no change to conclusions), published in Issue 4, 2008.

Review content assessed as up-to-date: 1 October 2007.

Citation: Bunn F, Trivedi D, Ashraf S. Colloid solutions for fluid resuscitation. *Cochrane Database of Systematic Reviews* 2008, Issue 1. Art. No.: CD001319. DOI: 10.1002/14651858.CD001319.pub2.

Copyright © 2008 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Colloids are widely used in the replacement of fluid volume. However doubts remain as to which colloid is best. Different colloids vary in their molecular weight and therefore in the length of time they remain in the circulatory system. Because of this and their other characteristics, they may differ in their safety and efficacy.

Objectives

To compare the effects of different colloid solutions in patients thought to need volume replacement.

Search strategy

We searched the Cochrane Injuries Group specialised register, CENTRAL (2007, Issue 1), MEDLINE (1994 to March 2007), EMBASE (1974 to March 2007), and the National Research Register (2007, issue 1). Bibliographies of trials retrieved were searched, and drug companies manufacturing colloids were contacted for information. The search was last updated in March 2007.

Selection criteria

Randomised and quasi-randomised trials comparing colloid solutions in critically ill and surgical patients thought to need volume replacement. The outcomes measured were death, amount of whole blood transfused, and incidence of adverse reactions.

Data collection and analysis

Two authors independently extracted the data and assessed the quality of the trials.

Main results

Seventy trials, with a total of 4375 participants, met the inclusion criteria. Quality of allocation concealment was judged to be adequate in 24 trials and poor or uncertain in the rest.

Deaths were obtained in 46 trials. For albumin or PPF versus hydroxyethyl starch (HES) 25 trials (n = 1234) reported mortality. The pooled relative risk (RR) was 1.14 (95% CI 0.91 to 1.43). For albumin or PPF versus gelatin, seven trials (n = 636) reported mortality. The RR was 0.97 (95% CI 0.68 to 1.39). For albumin or PPF versus Dextran four trials (n = 360) reported mortality. The RR was 3.75 (95% CI 0.42 to 33.09). For gelatin versus HES 18 trials (n = 1337) reported mortality and RR was 1.00 (95% CI 0.80 to 1.25). RR was not estimable in the gelatin versus dextran and HES versus dextran groups.

Colloid solutions for fluid resuscitation (Review)

Copyright © 2008 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

I

Thirty-seven trials recorded the amount of blood transfused, however quantitative analysis was not possible due to skewness and variable reporting. Nineteen trials recorded adverse reactions, but none occurred.

Authors' conclusions

From this review, there is no evidence that one colloid solution is more effective or safe than any other, although the confidence intervals are wide and do not exclude clinically significant differences between colloids. Larger trials of fluid therapy are needed if clinically significant differences in mortality are to be detected or excluded.

PLAIN LANGUAGE SUMMARY

There is no strong evidence to be certain of the safety of any particular type of colloid solution for replacing blood fluids

When a person is bleeding heavily, the loss of fluid volume in their veins can lead to shock, so they need fluid resuscitation. Colloids and crystalloids are two types of solutions used to replace lost blood fluid (plasma). They include blood and synthetic products. Both colloids and crystalloids appear to be similarly effective at resuscitation. There are different types of colloids and these may have different effects. However, the review of trials found there is not enough evidence to be sure that any particular colloid is safer than any other.

BACKGROUND

Colloids are used as plasma substitutes for short-term replacement of fluid volume, while the cause of the problem is being addressed (for example, stopping bleeding). These solutions can be blood products (human albumin solution, plasma protein fraction [PPF]) or synthetic (modified gelatins, dextrans, etherified starches). Colloid solutions are widely used in fluid resuscitation (Yim 1995) and they have been recommended in a number of resuscitation guidelines and intensive care management algorithms (Armstrong 1994; Vermeulen 1995). Previous systematic reviews have suggested that colloids are no more effective than crystalloids in reducing mortality (CIGAR 2004; Roberts 2004). Despite this, colloid solutions are still widely used as they are thought to remain in the intravascular space for longer than crystalloids and, therefore, be more effective in maintaining osmotic pressure.

It is plausible that colloids may vary in their safety and effectiveness. Different colloids vary in the length of time they remain in the circulatory system. It may be that some low to medium molecular weight colloids (for example, gelatins and albumin) are more likely to leak into the interstitial space (Traylor 1996), whereas some larger molecular weight hydroxyethyl starches are retained for longer (Boldt 1996). In addition it is thought that some colloids may effect coagulation or cause other adverse effects.

The previous review of colloids against crystalloids only allows indirect comparison of the different colloids. This review examines

direct comparisons of the different colloid solutions in randomised trials to complement the earlier reviews on colloids compared to crystalloids (Roberts 2004) and human albumin (CIGAR 2004).

OBJECTIVES

To quantify the relative effects on mortality of different colloid solutions in critically ill and surgical patients requiring volume replacement, by examining direct comparisons of colloid solutions.

METHODS

Criteria for considering studies for this review

Types of studies

Randomised and quasi-randomised (for example, allocation by hospital number or alternation) controlled trials.

Types of participants

Patients clinically assessed as requiring volume replacement or maintenance of colloid osmotic pressure. Administration of fluid

for preoperative haemodilution or volume loading, during plasma exchange, for priming extracorporeal circuits or following paracentesis are excluded.

Types of interventions

The colloid solutions considered are human albumin solutions, plasma protein fraction, modified gelatins, dextran 70, or etherified starch solutions.

Trials of other blood products not used primarily for volume replacement (for example, fresh frozen plasma, pooled serum) were excluded.

The review compares the administration of any regimens of different classes of colloids with each other.

Types of outcome measures

The primary outcome measure is mortality from any cause at the end of the study period.

We also attempted to find data on incidence of adverse reactions, allergies or anaphylactic shock, and the amount of blood (whole blood or red blood cells) transfused in each group. Some of the synthetic colloids may have anticoagulant properties and, therefore, we felt that some measure of blood loss or haemorrhage was important. However, as blood loss is vulnerable to measurement error, we decided to use the amount of blood products transfused as an outcome measure.

Intermediate physiological outcomes were not used for several reasons. These were that they are subject to intra- and inter-observer variation, they have no face value to patients and relatives, and the ones seen as appropriate are not stable over time. Also there would need to exist a strong predictive relationship between the variable and mortality.

Search methods for identification of studies

Electronic searches

We searched the following electronic databases:

- Cochrane Injuries Group trials register (searched 23 March 2007),
- CENTRAL (*The Cochrane Library issue 1, 2007*),
- MEDLINE (1966 to March 2007),
- EMBASE (1974 to March 2007),
- National Research Register (*issue 1, 2007*),
- Zetoc (searched 23 March 2007).

Full search strategies can be found in Appendix 1.

Searching other resources

We searched the bibliographies of the retrieved trials, and contacted drug companies manufacturing colloids for information. We also identified trials by using the searches undertaken for the pre-existing review of colloids versus crystalloids (Roberts 2004), which included BIDS Index to Scientific and Technical Proceedings, drawing on the handsearching of 29 international journals and the proceedings of several international meetings on fluid resuscitation, and checking the reference lists of the trials found. There were no language restrictions in any of the searches.

To identify unpublished trials we searched the register of the Medical Editors' Trial Amnesty and we contacted the UK Medicines Control Agency.

For the first version of the review (published 1999) we also contacted the medical directors of the following companies which all manufacture colloids:

- Alpha Therapeutic UK Limited (Albutein),
- American Critical Care McGraw (Hespan),
- Bayer (Plasbumin),
- Baxter (Gentran),
- Bio Products Laboratory (Zenalb),
- Cambridge Laboratories (Rheomacrodex),
- Centeon Limited (Albuminar),
- CIS UK Ltd,
- CP (Lomodex),
- Common Services Agency,
- Consolidated (Gelofusine),
- DuPont (Hespan),
- Fresenius (eloHAES and HAES-Steril),
- Geistlich Sons Ltd (Hespan and Pentaspan),
- Hoechst (Haemacel),
- Mallinckrodt Medical GMBH (Infoson),
- Nycomed, Oxford Nutrition (Elohes),
- Pharmacia and Upjohn Ltd (Rheomacrodex), and
- Sorin Biomedica Diagnostics Spa.

Data collection and analysis

Selection of studies

One author examined the electronic search results for reports of possibly relevant trials and these reports were then retrieved in full. Two authors applied the selection criteria independently to the trial reports, resolving disagreements by discussion.

Data extraction and management

Two authors independently extracted information on the following:

- method of allocation concealment,
- number of randomised patients,
- type of participants,
- the interventions, and
- outcome data (numbers of deaths, volume of blood transfused, and incidence of adverse or allergic reactions).

The authors were not blinded to the authors or journal when doing this, as the value of this has not been established (Berlin 1997). Results were compared and any differences resolved by discussion. Where there was insufficient information in the published report, we attempted to contact the trial authors for clarification.

Assessment of risk of bias in included studies

Since there is evidence that the quality of allocation concealment particularly affects the results of studies (Schulz 1995), two authors scored this quality on the scale used by Schulz 1995 as shown below, assigning C to poorest quality and A to best quality:

- A = trials deemed to have taken adequate measures to conceal allocation (that is, central randomisation; numbered or coded bottles or containers; drugs prepared by the pharmacy; serially numbered, opaque, sealed envelopes; or other description that contained elements convincing of concealment).
- B = trials in which the authors either did not report an allocation concealment approach at all or reported an approach that did not fall into one of the other categories.
- C = trials in which concealment was inadequate (such as alternation or reference to case record numbers or to dates of birth).

Where the method used to conceal allocation was not clearly reported, the author was contacted, if possible, for clarification. We then compared the scores allocated and resolved differences by discussion.

Data synthesis

The following comparisons were made:

- albumin or PPF versus etherified starch.
- albumin or PPF versus modified gelatin.
- albumin or PPF versus dextran 70.
- modified gelatin versus etherified starch.
- modified gelatin versus dextran 70.
- etherified starch versus dextran 70.

For each trial we calculated the relative risk (RR) of death and 95% confidence interval (CI), such that a RR of more than 1 indicates a higher risk of death in the first group named. We chose RR, as it is more readily applied to the clinical situation.

We examined the groups of trials for statistical evidence of heterogeneity using chi-square and I^2 tests. If there was no obvious

heterogeneity on visual inspection or statistical testing, we calculated pooled RRs and 95% confidence intervals using a fixed-effects model.

We assessed the skewness of continuous data by checking the mean and standard deviation (if available). If the standard deviation is more than twice the mean for data with a finite end point (such as 0 in the case of bleeding), the data are likely to be skewed and it is inappropriate to apply parametric tests (Altman 1996). This is because the mean is unlikely to be a good measure of central tendency. If parametric tests could not be applied, we tabulated the data.

We examined the effect of excluding trials judged to have inadequate (scoring C) allocation concealment in a sensitivity analysis.

RESULTS

Description of studies

See: [Characteristics of included studies](#); [Characteristics of excluded studies](#).

For more detailed descriptions of individual studies, please see the table 'Characteristics of included studies'.

Seventy trials met the inclusion criteria, with a total of 4375 participants. The earliest trial was from 1980 and the most recent from 2006. From the drug companies we contacted, we were sent information by Hoechst, Baxter Health Care Ltd, Fresenius Ltd, CIS UK Ltd, and Rhemocradex. No new trials were identified from the information sent to us.

The trials included the following comparisons:

Albumin or PPF versus starch (n=41 trials with 1839 participants in these groups)

[Arellano 2005](#); [Boldt 1986](#); [Boldt 1993a](#); [Boldt 1995](#); [Boldt 1996a](#); [Boldt 1996b](#); [Boldt 1996c](#); [Boldt 1998](#); [Boldt 2006](#); [Brock 1995](#); [Brutocao 1996](#); [Claes 1992](#); [Diehl 1982](#); [Falk 1988](#); [Fulachier 1994](#); [Gahr 1981](#); [Gallagher 1985](#); [Gold 1990](#); [Hausdorfer 1986](#); [Hiippala 1995](#); [Huskisson 1993](#); [Jones 2004](#); [Kirklin 1984](#); [London 1989](#); [Mastroianni 1994](#); [Moggio 1983](#); [Munoz 1980](#); [Munsch 1988](#); [Niemi 2006](#); [Prien 1990](#); [Rackow 1983](#); [Rackow 1989](#); [Rosencher 1992](#); [Shatney 1983](#); [Veneman 2004](#); [Verheij 2006](#); [Vogt 1994](#); [Vogt 1996](#); [Vogt 1999](#); [von Sommoggy 1990](#); [Woittiez 1997](#).

Albumin or PPF versus dextran (n=6 trials with 410 participants in these groups)

[Hedstrand 1987](#); [Hiippala 1995](#); [Jones 2004](#); [Karanko 1987](#); [Lisander 1996](#); [Tollofsrud 1995](#).

Albumin or PPF versus gelatin (n=12 trials with 1024 participants in these groups)

Boldt 1986; Du Gres 1989; Evans 2003; Huang 2005; Huskisson 1993; Karanko 1987; Niemi 2006; Stockwell 1992; Stoddart 1996; Tollofsrud 1995; Verheij 2006; Wahba 1996.

Starch versus gelatin (n=20 trials with 1545 participants in these groups)

Allison 1999; Asfar 2000; Beards 1994; Berard 1995; Beyer 1997; Boldt 1986; Boldt 2000; Boldt 2001; Carli 2000; Dytowska 1998; Haisch 2001a; Haisch 2001b; Huskisson 1993; Huttner 2000; Molnar 2004; Niemi 2006; Rittoo 2004; Schortgen 2001; Van der Linden 2004; Van der Linden 2005.

Starch versus dextran (n=1 trial with 30 participants in these groups)

Hiippala 1995.

Dextran versus gelatin (n=2 trial with 42 participants in these groups)

Karanko 1987; Tollofsrud 1995.

The trials involved patients with hypovolaemia, sepsis, trauma, and patients who had undergone surgery.

The trials tended to report surrogate outcomes such as hemodynamic variables. Data on death were obtainable from 46 trials. Information on the amount of blood transfused was available in 38 trials. However, the data were reported in a variety of different ways that made combining the data in a meta-analysis unfeasible. Inclusion and exclusion criteria varied, but many of the studies excluded patients with previous adverse reactions to colloids, clotting problems, or renal disease.

Risk of bias in included studies

Using predefined criteria (Schulz 1995) the quality of allocation concealment was judged to be adequate in 24 trials, unclear in 37 trials and inadequate in nine trials. Where the method of allocation concealment was unclear, we attempted to contact all of the trialists and we obtained information from 11 of them. However, due to the lack of reported information on the process of randomisation and allocation concealment, we were unable to properly assess the quality of the majority of the trials.

Thirteen trials mentioned that some form of blinding was used. In nine, some, or all, of the staff giving treatment were blinded, in four those giving post-operative care were blinded, in two the outcome assessors were blinded and in one the statisticians performing the analysis were blinded to treatment group.

Effects of interventions

Mortality

Of the 70 trials identified, 33 reported mortality data. Information on death was obtained from a further 13 trials by contact with the trial authors. We, therefore, had data on death from 46 trials.

Albumin or PPF versus hydroxyethyl starch (HES)

Twenty-five trials (1234 participants) reported mortality data. The pooled RR (relative risk) was 1.14 (95% CI 0.91 to 1.43).

Albumin or PPF versus gelatin

Seven trials (636 participants) reported mortality but only one of those trials had any deaths. The RR was 0.97 (95% CI 0.68 to 1.39).

Albumin or PPF versus dextran

Four trials (360 participants) reported mortality and were included in the meta-analysis. Only one of these (Hedstrand 1987) reported any deaths. The RR was 3.75 (95% CI 0.42 to 33.09).

Gelatin versus HES

Eighteen trials (1337 participants) reported mortality and the pooled RR was 1.00 (95% CI 0.80 to 1.25).

Gelatin versus dextran 70

There were two trials (42 participants) which reported mortality. There were no deaths so the RR was not estimable.

HES versus dextran 70

No trials reported mortality.

Amount of blood transfused

Thirty-eight trials recorded the amount of blood transfused. As the data was reported in various ways, often lacking a measure of variation, and was also skewed we did not attempt a quantitative synthesis. These data can be seen in the 'other data' table.

Adverse events

Nineteen trials reported the incidence of adverse or allergic reactions or anaphylactic shock: all reported that there were no such incidents.

Sensitivity analysis

The effect of excluding trials judged to have inadequate or unclear (scoring B or C) allocation concealment was examined in a subgroup analysis. This made no significant difference to the results (albumin or PPF versus HES pooled RR 1.16 95% CI 0.90 to 1.50; gel versus HES pooled RR 1.07 95% CI 0.78 to 1.46).

DISCUSSION

Despite finding 70 trials we cannot make any conclusions about the relative effectiveness of different colloid solutions. Previous systematic reviews have suggested that colloids are no more effective than crystalloids in reducing mortality (CIGAR 2004; Roberts 2004), but there are too few data available to show in direct comparisons whether any of the colloids are safer or more effective than another. The confidence intervals are wide and do not exclude clinically significant differences between colloids.

Mortality was selected as the main outcome measure in this systematic review for several reasons. In the context of critical illness, death or survival is a clinically relevant outcome that is of immediate importance to patients, and data on death are reported in many of the studies. Furthermore, one might expect that mortality data would be less prone to measurement error or biased reporting than would data on pathophysiological outcomes. The use of a pathophysiological end point as a surrogate for an adverse outcome assumes a direct relationship between the two, an assumption that may sometimes be inappropriate. Finally, when trials collect data on a number of physiological end-points, there is the potential for bias due to the selective publication of end-points showing striking treatment effects.

There was wide variation in the participants, intervention regimens, and the length of follow-up. The length of follow-up is not reported in many of the studies. Where it is reported it ranges from a matter of hours to months, which may explain a lot of the heterogeneity in overall event rates. The effect of these factors was not examined in a sensitivity analysis, as there was felt to be insufficient data to justify examining subgroups.

Many of the trials were small, and some had been done some time ago. Although older trials will not necessarily be of poorer quality, it may be that treatment protocols have subsequently altered making these trials less relevant to current clinical practice.

AUTHORS' CONCLUSIONS

Implications for practice

Previous reviews have failed to show any benefit of colloids over crystalloids for volume replacement (CIGAR 2004; Roberts 2004).

This review does not provide any evidence that one colloid is safer than another, but does not rule out clinically significant differences.

Implications for research

Trials of fluid therapy need to be larger in order to exclude clinically significant differences between colloids in patient relevant outcomes. However, trials should probably first address the question of whether colloids are any more effective than crystalloid solutions.

Use of surrogate outcomes, such as physiological measurements should be discouraged unless there is a strong relationship with outcomes of interest to patients and relatives.

ACKNOWLEDGEMENTS

We wish to acknowledge the contribution of Phil Alderson and Victoria Hawkins who were authors of earlier versions of this review. In addition we acknowledge the help of Ralph Bloch, Olivier Duperrex, Andrew Smith, Peter Smith and Reinhard Wentz, who assisted with translating articles. Also many thanks to the authors who provided us with details of their studies.

We are grateful to the drug companies, Baxter Healthcare Ltd, CIS Ltd, Fresenius, Hoechst, and Pharmalink who responded to our request for information.

REFERENCES

References to studies included in this review

Allison 1999 *{published data only}*

Allison KP, Gosling P, Jones S, Pallister I, Porter K. Randomized trial of hydroxyethyl starch versus gelatine for trauma resuscitation. *Journal of Trauma* 1999;**47**(6):1114–21.

Arellano 2005 *{published data only}*

Arellano R, Gan BS, Salpeter MJ, Yeo E, McCluskey S, Pinto R, et al. A triple-blinded randomized trial comparing the hemostatic effects of large-dose 10% hydroxyethyl starch 264/0.45 versus 5% albumin during major reconstructive surgery. *Anesthesia and Analgesia* 2005;**100**:1846–53.

Asfar 2000 *{published data only}*

Asfar P, Kerani N, Labadie F, Gouello JP, Brenet O, Alquier P. Assessment of hemodynamic and gastric mucosal acidosis with modified fluid versus 6% hydroxyethyl starch: a prospective, randomized study. *Intensive Care Medicine* 2000;**26**(9):1282–7.

Beards 1994 *{published and unpublished data}*

Beards SC, Watt T, Edwards JD, Nightingale P, Farragher EB. Comparison of the hemodynamic and oxygen transport responses to modified fluid gelatin and hetastarch in critically ill patients: a prospective, randomized trial. *Critical Care Medicine* 1994;**22**(4):600–5. [MEDLINE: 1994192356]

Berard 1995 *{published data only}*

Berard JP, Curt I, Piech JJ, Ruiz F. Hydroxyethylamidons versus gelatines: impact on the cost of replacement in an emergency (resuscitation) service [Hydroxyethylamidons versus gelatines: Impact sur le cout du rempiissage dans un service de reanimation]. *Annales Francaises d'Anaesthesia et de Reanimation* 1995;**14**:R335.

Beyer 1997 *{published and unpublished data}*

Beyer R, Harmening U, Rittmeyer O, Zielmann S, Mielck F, Kazmaier S, Kettler D. Use of modified fluid gelatin and hydroxyethyl starch for colloidal volume replacement in major orthopaedic surgery. *British Journal of Anaesthesia* 1997;**78**(1):44–50. [MEDLINE: 1997212347]

Boldt 1986 *{published data only}*

Boldt JV, Von Bormann B, Kling D, Borner U, Mulch J, Hempelmann G. Volume replacement with a new hydroxyethyl starch preparation (3% HES 200/0.5) in heart surgery [Volumenersatz mit einem neuen hydroxyethylstarke – preparat (3% HAS 200/0.5) in der herzchirurgie]. *Infusionstherapie und Klinische Ernährung* 1986;**13**(3):145–151. [MEDLINE: 1986302988]

Boldt 1993a *{published and unpublished data}*

Boldt J, Knothe C, Zickmann B, Andres P, Dapper F, Hempelmann G. Influence of different intravascular volume therapies on platelet function in patients undergoing cardiopulmonary bypass. *Anesthesia and Analgesia* 1993;**76**(6):1185–90.

Boldt 1995 *{published data only}*

Boldt J, Heesen M, Welters I, Padberg W, Martin K, Hempelmann G. Does the type of volume therapy influence endothelial-related coagulation in the critically ill?. *British Journal of Anaesthesia* 1995;**75**(6):740–6. [MEDLINE: 1996246804]

Boldt 1996a *{published data only}*

Boldt J, Heesen M, Muller M, Pabsdorf M, Hempelmann G. The effects of albumin versus hydroxyethyl starch solution on cardiorespiratory and circulatory variables in critically ill patients. *Anesthesia and Analgesia* 1996;**83**(2):254–61. [MEDLINE: 1996302067]

Boldt 1996b *{published data only}*

Boldt J, Heesen M, Padberg W, Martin K, Hempelmann G. The influence of volume therapy and pentoxifylline infusion on circulating adhesion molecules in trauma patients. *Anaesthesia* 1996;**51**(6):529–35. [MEDLINE: 1996296856]

Boldt 1996c *{published data only}*

Boldt J, Mueller M, Menges T, Papsdorf M, Hempelmann G. Influence of different volume therapy regimens on regulators of the circulation in the critically ill. *British Journal of Anaesthesia* 1996;**77**(4):480–7. [MEDLINE: 1997097789]

Boldt 1998 *{published data only}*

Boldt J, Muller M, Mentges D, Papsdorf M, Hempelmann G. Volume therapy in the critically ill: is there a difference?. *Intensive Care Medicine* 1998;**24**(1):28–36. [MEDLINE: 1998163949]

Boldt 2000 *{published data only}*

Boldt J, Suttner S, Kumle B, Huttner I. Cost analysis of different volume replacement strategies in anaesthesia. *Infusionstherapie und Transfusionsmedizin* 2000;**27**(1):38–43.

Boldt 2001 *{published data only}*

Boldt J, Suttner S, Huttner I, Kumle B, Piper S, Krumholz W. Are cost of a crystalloid-based volume replacement regimen lower than of a colloid-based volume replacement strategy?. *Infusionstherapie und transfusionsmedizin* 2001;**28**(3):144–9.

Boldt 2006 *{published data only}*

Boldt J, Schollhorn T, Mayer J, Piper S, Suttner S. The value of an albumin-based intravascular volume replacement strategy in elderly patients undergoing major abdominal surgery. *Anesthesia and Analgesia* 2006;**103**:191–9.

Brock 1995 *{published and unpublished data}*

Brock H, Rapf B, Necek S, Gabriel C, Peterlik C, Polz W. Volume replacement after cardiac surgery. A comparison of small-volume resuscitation and two different colloid solutions [Vergleichende untersuchungen zur postoperativen volumentherapie]. *Anaesthesist* 1995;**44**(7):486–92. [MEDLINE: 1995390424]

Brutocao 1996 *{published and unpublished data}*

Brutocao D, Bratton SL, Thomas JR, Schrader PF, Coles PG, Lynn AM. Comparison of hetastarch with albumin for postoperative volume expansion in children after cardiopulmonary bypass. *Journal of Cardiothoracic and Vascular Anesthesia* 1996;**10**(3):348–51. [MEDLINE: 1996298754]

Carli 2000 *{published data only}*

Carli P, Goldstein P, Lejay M, Facon A, Orliaguette G, Petit P. Prehospital care of hypovolemic trauma patients: 6% hydroxyethyl starch versus gelatin [Remplissage vasculaire prehospitalier en traumatologie: Hesteril 6% versus Plasmion]. *Journal Europeen des Urgences* 2000;**13**(1-2):101–5.

- Claes 1992** *{published data only}*
 Claes Y, Van Hemelrijck J, Van Gerven M, Arnout J, Vermeylen J, Weidler B, et al. Influence of hydroxyethyl starch on coagulation in patients during the perioperative period. *Anesthesia and Analgesia* 1992;**75**(1):24–30. [MEDLINE: 1992312872]
- Diehl 1982** *{published data only}*
 Diehl JT, Lester JL, Cosgrove DM. Clinical comparison of hetastarch and albumin in postoperative cardiac patients. *Annals of Thoracic Surgery* 1982;**34**(6):674–9. [MEDLINE: 1983073643]
- Du Gres 1989** *{published data only}*
 Du Gres B, Gruner MC, Flamens C. A comparison of the hemodynamic effect of Haemaccel and diluted albumin in the immediate postoperative period after heart surgery [Comparaison des effets hemodynamiques de l'Haemaccel et de l'albumine diluee dans la periode postoperatoire immediate apres chirurgie cardiaque]. *Cahiers d'Anesthesiologie* 1989;**37**(5):327–32. [MEDLINE: 1990029584]
- Dytowska 1998** *{published data only}*
 Dytowska B, Karwacki Z, Suchorzewska J, Wujtewicz M. Comparative assessment of 200/0.5 HAES 6% and Gelafundin in the treatment of hypovolaemia in post-coronary bypass patients. *Medical Science Monitor* 1998;**4**(6):1000–3.
- Evans 2003** *{published data only}*
 Evans PA, Heptinstall S, Crowhurst EC, Davies T, Glenn JR, Madira W, et al. Prospective double-blind randomized study of the effects of four intravenous fluids on platelet function and hemostasis in elective hip surgery. *Journal of Thrombosis and Haemostasis* 2003;**1**:2140–8.
- Falk 1988** *{published data only}*
 Falk JL, Rackow EC, Astiz ME, Weil MH. Effects of hetastarch and albumin on coagulation in patients with septic shock. *Journal of Clinical Pharmacology* 1988;**28**(5):412–5. [MEDLINE: 1988273726]
- Fries 2004** *{published data only}*
 Fries D, Streif W, Margreiter J, Klingler A, Kuehbacher G, Schobersberger W, et al. The effects of perioperative administered crystalloids and colloids on concentrations of molecular markers of activated coagulation and fibrinolysis. *Blood Coagulation and Fibrinolysis* 2004;**15**:213–9.
- Fulachier 1994** *{published data only}*
 Fulachier V, Sicard MP, Baille Y, Auffray JP. Effects of fluid expansion using albumin or hydroxyethylstarch on oxygen transport after induction of anesthesia for cardiac surgery. *Journal of Cardiothoracic and Vascular Anesthesia* 1994;**8**(3Suppl2):89.
- Gahr 1981** *{published data only}*
 Gahr R, Bock PR. Effect of hydroxyethyl starch HES 450/0.7 and 5% human albumin on the colloid osmotic pressure and hemodynamic parameters in hypovolemic patients after major abdominal procedures [Wirkung von hydroxyethylstarke HAS 450/0.7 und humanalbumin 5% auf den kolloidosmotischen druck und hamodynamische parameter bei hypovolamischen patienten nach grosseren abdominalen eingriffen]. *Infusionstherapie und Transfusionsmedizin* 1981;**8**(3):147–52. [MEDLINE: 1981262968]
- Gallagher 1985** *{published and unpublished data}*
 Gallagher JD, Moore RA, Kerns D, Jose AB, Botros SB, Flicker S, et al. Effects of colloid or crystalloid administration on pulmonary extravascular water in the postoperative period after coronary artery bypass grafting. *Anesthesia and Analgesia* 1985;**64**(8):753–8.
- Gold 1990** *{published and unpublished data}*
 Gold MS, Russo J, Tissot M, Weinhouse G, Riles T. Comparison of hetastarch to albumin for perioperative bleeding in patients undergoing abdominal aortic aneurysm surgery. *Annals of Surgery* 1990;**211**(4):482–5. [MEDLINE: 1990210743]
- Haisch 2001a** *{published data only}*
 Haisch G, Boldt J, Krebs C, Suttner S, Lehmann A, Isgro F. Influence of a new hydroxyethylstarch preparation (HES 130/0.4) on coagulation in cardiac surgical patients. *Journal of Cardiothoracic and Vascular Anesthesia* 2001;**15**(3):316–21.
- Haisch 2001b** *{published data only}*
 Haisch G, Boldt J, Krebs C, Kumle B, Suttner S, Schulz A. The influence of intravascular volume therapy with a new hydroxyethyl starch preparation (6% HES 130/0.4) on coagulation in patients undergoing major abdominal surgery. *Anesthesia and Analgesia* 2001;**92**(3):565–71. [MEDLINE: 21124037]
- Hausdorfer 1986** *{published data only}*
 Hausdorfer J, Hagemann H, Heine J. Comparison of volume substitutes human albumin 5% and hydroxyethyl starch 6% in paediatric anaesthesia [Vergleich der volumenersatzmittel humanalbumin 5% und hydroxyethylstarke 6% (40.000/0.5) in der kinderanasthesie]. *Anasthesie, Intensivtherapie, Notfallmedizin* 1986;**21**(3):137–42. [MEDLINE: 1986320933]
- Hedstrand 1987** *{published data only}*
 Hedstrand U, Hogman C, Zaren B, Lundkvist B. Postoperative complications after blood replacement with or without plasma. *Acta chirurgica Scandinavica* 1987;**153**(9):501–5.
- Hiippala 1995** *{published data only}*
 Hiippala S, Linko K, Myllyla G, Lalla M, Hekali R, Makelainen A. Replacement of major surgical blood loss by hypo-oncotic or conventional plasma substitutes. *Acta Anaesthesiologica Scandinavica* 1995;**39**(2):228–35. [MEDLINE: 1995313480]
- Huang 2005** *{published data only}*
 Huang Y, Yan B, Yang Z. Clinical study of a formula for delayed rapid fluid resuscitation for patients with burn shock. *Burns* 2004;**31**:617–22.
- Huskisson 1993** *{published data only}*
 Huskisson L, Elliott M, Spitz L. Haemodynamic effects of three colloids following pediatric open heart surgery. *Clinical Intensive Care* 1993;**4**:302.
- Huttner 2000** *{published data only}*
 Huttner I, Boldt J, Haisch G, Suttner S, Kumle B, Schulz H. Influence of different colloids on molecular markers of haemostasis and platelet function in patients undergoing major abdominal surgery. *British Journal of Anaesthesia* 2000;**85**(3):417–23.
- Jones 2004** *{published data only}*
 Jones S, Whitten C, Monk T. Influence of crystalloid and colloid replacement solutions on hemodynamic variables during acute normovolemic hemodilution. *Journal of Clinical Anaesthesia* 2004;**16**:11–7.

- Karanko 1987** *{published and unpublished data}*
Karanko MS. Effects of three colloid solutions on plasma volume and hemodynamics after coronary bypass surgery. *Critical Care Medicine* 1987;**15**(11):1015–21. [MEDLINE: 1988054051]
- Kirklin 1984** *{published data only}*
Kirklin JK, Lell WA, Kouchoukos NT. Hydroxyethyl starch versus albumin for colloid infusion following cardiopulmonary bypass in patients undergoing myocardial revascularization. *Annals of Thoracic Surgery* 1984;**37**(1):40–6. [MEDLINE: 1984103384]
- Lisander 1996** *{published and unpublished data}*
Lisander B, Jacobsson SA, Ivarsson I, Vegfors M, Engdahl O. Giving both enoxaparin and dextran increases the need for transfusion in revision hip arthroplasty. *European Journal of Surgery* 1996;**162**(11):861–6. [MEDLINE: 1997115584]
- London 1989** *{published data only}*
London MJ, Ho JS, Triedman JK, Verrier ED, Levin J, Merrick SH, et al. A randomized clinical trial of 10% pentastarch (low molecular weight hydroxyethyl starch) versus 5% albumin for plasma volume expansion after cardiac operations. *Journal of Thoracic and Cardiovascular Surgery* 1989;**97**(5):785–97. [MEDLINE: 1989218083]
- Mastroianni 1994** *{published data only}*
Mastroianni L, Low HB, Rollman J, Wagle M, Bleske B, Chow MS. A comparison of 10% pentastarch and 5% albumin in patients undergoing open-heart surgery. *Journal of Clinical Pharmacology* 1994;**34**(1):34–40. [MEDLINE: 1994179580]
- Moggio 1983** *{published data only}*
Moggio RA, Rha CC, Somberg ED, Praeger P, Pooley RW, Reed GE. Hemodynamic comparison of albumin and hydroxyethyl starch in postoperative cardiac surgery patients. *Critical Care Medicine* 1983;**11**(12):943–5. [MEDLINE: 1984056648]
- Molnar 2004** *{published data only}*
Molnar Z, Mikor A, Leiner T, Szakmany T. Fluid resuscitation with colloids of different molecular weight in septic shock. *Intensive Care Medicine* 2004;**30**:1356–60.
- Munoz 1980** *{published data only}*
Munoz E, Raciti A, Dove D, Stahl WM, Del Guercio L. Effect of hydroxyethyl starch versus albumin on hemodynamic and respiratory function in patients with shock. *Critical Care Medicine* 1980;**8**(4):255.
- Munsch 1988** *{published data only}*
Munsch CM, MacIntyre E, Machin SJ, Mackie IJ, Treasure T. Hydroxyethyl starch: an alternative to plasma for postoperative volume expansion after cardiac surgery. *British Journal of Surgery* 1988;**75**(7):675–8. [MEDLINE: 1988327292]
- Niemi 2006** *{published data only}*
Niemi T, Suojäranta-Ylinen R, Kukkonen S, Kuitunen A. Gelatin and hydroxyethyl starch, but not albumin, impair hemostasis after cardiac surgery. *Anesthesia and Analgesia* 2006;**102**:998–1006.
- Prien 1990** *{published and unpublished data}*
Prien T, Backhaus N, Pelster F, Pircher W, Bunte H, Lawin P. Effect of intraoperative fluid administration and colloid osmotic pressure on the formation of intestinal edema during gastrointestinal surgery. *Journal of Clinical Anesthesia* 1990;**2**(5):317–23. [MEDLINE: 1991104037]
- Rackow 1983** *{published data only}*
Rackow EC, Falk JL, Fein IA, Siegel JS, Packman MI, Haupt MT, et al. Fluid resuscitation in circulatory shock: A comparison of the cardiorespiratory effects of albumin, hetastarch, and saline solutions in patients with hypovolemic and septic shock. *Critical Care Medicine* 1983;**11**(11):839–50. [MEDLINE: 1984027713]
- Rackow 1989** *{published data only}*
Rackow EC, Mecher C, Astiz ME, Griffel M, Falk JL, Weil MH. Effects of pentastarch and albumin infusion on cardiorespiratory function and coagulation in patients with severe sepsis and systemic hypoperfusion. *Critical Care Medicine* 1989;**17**(5):395–8. [MEDLINE: 1989209912]
- Rittoo 2004** *{published data only}*
Rittoo D, Gosling P, Burnley S, Bonnici C, Millns P, Simms MH, et al. Randomized study comparing the effects of hydroxyethyl starch solution with gelofusine on pulmonary function in patients undergoing abdominal aortic aneurysm surgery. *British Journal of Anaesthesia* 2004;**92**:61–6.
- Rosencher 1992** *{published and unpublished data}*
Rosencher N, Vassilief N, Guignon V, Toulin P, Conseiller C. Comparison of effects of Elohes and albumin on haemostasis in orthopedic surgery [Comparaison des effets de l'Elohes et de l'albumine sur l'hémostase en chirurgie orthopédique]. *Annales francaises d'Anesthésie et de Réanimation* 1992;**11**(5):526–30. [MEDLINE: 1993118965]
- Schortgen 2001** *{published data only}*
Schortgen F, Lacherade J, Bruneel F, Cattaneo I, Hemery F, Lemaire F, et al. Effects of hydroxyethylstarch and gelatin on renal function in severe sepsis: a multicentre randomised study. *Lancet* 2001;**357**(9260):911–6.
- Shatney 1983** *{published data only}*
Shatney CH, Deepika K, Militello PR, Majerus TC, Dawson RB. Efficacy of hetastarch in the resuscitation of patients with multisystem trauma and shock. *Archives of Surgery* 1983;**118**(7):804–9. [MEDLINE: 1983230251]
- Stockwell 1992** *{published data only}*
Stockwell MA, Scott A, Day A, Riley B, Soni N. Colloid solutions in the critically ill. A randomised comparison of albumin and polygeline: 2. Serum albumin concentration and incidences of pulmonary oedema. *Anesthesia* 1992;**47**(1):7–9.
* Stockwell MA, Soni N, Riley B. Colloid solutions in the critically ill. A randomised comparison of albumin and polygeline: 1. Outcome and duration of stay in the intensive care unit. *Anaesthesia* 1992;**47**(1):3–6. [MEDLINE: 1992161241]
- Stoddart 1996** *{published data only}*
Stoddart PA, Rich P, Sury MR. A comparison of 4.5% human albumin solution and haemaccel in neonates undergoing major surgery. *Paediatric Anaesthesia* 1996;**6**(2):103–6. [MEDLINE: 1996243349]
- Tollofsrud 1995** *{published data only}*
Svennevig JL, Tollofsrud S, Kongsgaard U, Noddeland H, Mohr B, Ozer M, Mollnes TE. Complement activation during and after open-heart surgery is only marginally affected by the choice of fluid for volume replacement. *Perfusion* 1996;**11**(4):326–32.
* Tollofsrud S, Svennevig JL, Breivik H, Kongsgaard U, Ozer M, Hysing E, et al. Fluid balance and pulmonary functions during and

- after coronary artery bypass surgery: Ringer's acetate compared with dextran, polygeline, or albumin. *Acta anaesthesiologica Scandinavica* 1995;**39**(5):671–7.
- Van der Linden 2004** *{published data only}*
Van der Linden PJ, De Hert SG, Daper A, Trenchant A, Schmartz D, Defrance P, et al. 3.5% urea linked gelatine is as effective as 6% HES 200/0.5 for volume management in cardiac surgery patients. *Canadian Journal of Anaesthesia* 2004;**51**(3):236–241.
- Van der Linden 2005** *{published data only}*
Van der Linden PJ, De Hert SG, Deraedt D, Cromheecke S, De Decker K, De Paep R, et al. Hydroxyethyl starch 130/0.4 versus modified fluid gelatin for volume expansion in cardiac surgery patients: the effects on perioperative bleeding and transfusion needs. *Anesthesia and Analgesia* 2005;**101**:629–34.
- Veneman 2004** *{published data only}*
Veneman T, Oude Nijhuis J, Woittiez A. Human albumin and starch administration in critically ill patients: a prospective RCT. *Wien Klin Wochenschr* 2004;**116**(9-10):283–5.
- Verheij 2006** *{published data only}*
Verheij J, van Lingen A, Beishuizen A, Christiaans H, de Jong J, Girbes A, et al. Cardiac response is greater for colloid than saline fluid loading after cardiac or vascular surgery. *Intensive Care Medicine* 2006;**32**:1030–8.
Verheij J, van Lingen A, Raijmakers P, Rijnsburger E, Veerman D, et al. Effect of fluid loading with saline or colloids on pulmonary permeability, oedema and lung injury score after cardiac and major vascular surgery. *British Journal of Anaesthesia* 2006;**96**(1):21–30.
- Vogt 1994** *{published data only}*
Vogt N, Bothner U, Georgieff M. Comparison of 5% human albumin and 6% 200/0.5 HES as exclusive colloid components in large surgical interventions [Vergleich von humanalbumin 5% und 6% HES 200/0.5 als ausschliessliche kolloidkomponente bei grossen chirurgischen eingriffen]. *Anesthesiologie, Intensivmedizin, Notfallmedizin, Schmerztherapie* 1994;**29**(3):150–6.
- Vogt 1996** *{published data only}*
Vogt NH, Bothner U, Lerch G, Linder KH, Georgieff M. Large-dose administration of 6% hydroxyethyl starch 200/0.5 for total hip arthroplasty: Plasma homeostasis, hemostasis, and renal function compared to use of 5% human albumin. *Anesthesia and Analgesia* 1996;**83**(2):262–8. [MEDLINE: 1996302068]
- Vogt 1999** *{published data only}*
* Vogt N, Bothner U, Brinkmann A, De Petriconi R, Georgieff M. Peri-operative tolerance to large-dose 6% HES 200/0.5 in major urological procedures compared with 5% human albumin. *Anaesthesia* 1999;**54**(2):121–7.
Vogt N, Bothner U, Lerch G, Georgieff M. [Pharmakokinetik und onkotisches verhalten von hochdosierter hydroxyethylstarke bei operativen eingriffen im vergleich zu humanalbumin 5%]. *Infusion Therapy and Transfusion Medicine* 1998;**25**:212–21.
- von Sommoggy 1990** *{published data only}*
Von Sommoggy S, Fraunhofer J, Jelen-Esselborn S, Stemberger A. Coagulation changes during aortofemoral bifurcation bypass: is volume and plasma substitution possible with hydroxyethyl starch alone? [Gerinnungsveränderungen bei aortofemoralem bifurkationsbypass: ist eine Volumen –und Plasmasubstitution mit hydroxyethylstarke allein möglich?]. *Anaesthesist* 1990;**39**(7):353–60.
- Wahba 1996** *{published and unpublished data}*
Wahba A, Sendtner E, Birnbaum DE. Fluid resuscitation with Haemacel vs. human albumin following coronary artery bypass grafting. *The Thoracic and Cardiovascular Surgeon* 1996;**44**(4):178–83. [MEDLINE: 1997051433]
- Watkins 1990** *{published data only}*
Watkins J, Wild G, Appleyard TN, Hardy G. Complement activation by polystarch and gelatine volume expanders. *Lancet* 1990;**335**(8683):233.
- Woittiez 1997** *{published and unpublished data}*
Hondebrink Y, Jeekel L, Oude Nijhuis J, Woittiez AJJ. Restoration of colloid osmotic pressure in hypoalbuminaemic patients. *Intensive Care Medicine* 1997;**23**(supp 1):S184.
Timmer B, Hondebrink Y, Oude Nijhuis J, Woittiez AJJ. Restoration of colloid osmotic pressure in hypoalbuminaemic patients. *Netherlands Journal of Medicine* 1998;**52**:A42.

References to studies excluded from this review

- Boldt 1993** *{published data only}*
Boldt J, Knothe C, Schindler E, Hammermann H, Dapper F, Hempelmann G. Volume replacement with hydroxyethyl starch solution in children. *British Journal of Anaesthesia* 1993;**70**(6):661–665.
- Boldt 2000b** *{published data only}*
Boldt J, Lehmann A, Rompert R, Haisch G, Isgró F. Volume therapy with a new hydroxyethyl starch solution in cardiac surgical patients before cardiopulmonary bypass. *Journal of Cardiothoracic and Vascular Anaesthesia* 2000;**14**(3):264–8.
- Brehme 1993** *{published data only}*
Brehme S, Keysser G, Turowski A, Schmidt H. Hemorheologic effects of hydroxyethyl starch 200/0.5, dextran 40, oxypolygelatine and full electrolyte sodium over 48 hours [Hemorheologische wirkungen von hydroxyethylstarke 200/0.5, dextran 40, oxypolygelatine und vollelektrolytlosung uber 48 studen]. *Zeitschrift fur die gesamte innere medizin und ihre grenzgebiete* 1993;**48**(10):506–510.
- Bremerich 2000** *{published data only}*
Bremerich DH, Lischke V, Asskali F, Forster H, Behne M. Pharmacodynamics and tolerability of acetyl starch as a new plasma volume expander in patients undergoing elective surgery. *International Journal of Clinical Pharmacology and Therapeutics* 2000;**38**(8):408–414.
- Charlet 1991** *{published data only}*
Charlet P, Zerr C, Robert D, Merville C, Renouf P, Khayat MC. Comparative trials of fluid gelatins on hemostasis in heart surgery in adults [Essais comparatifs des gelatines fluides sur l'hemostase dans la chirurgie cardiaque de l'adulte]. *Cahiers d'Anesthesiologie* 1991;**39**(4):233–238.
- Christ 1997** *{published data only}*
Christ F, Niklas M, Kreimeier U, Lauterjung L, Peter K, Messmer K. Hyperosmotic-hyperoncotic solutions during abdominal aortic aneurysm (AAA) resection. *Acta Anaesthesiologica Scandinavica* 1997;**41**(1):62–70.

- Emery 1992** *{published data only}*
Emery EF, Greenough A, Gamsu HR. Randomised controlled trial of colloid infusions in hypotensive preterm infants. *Archives of Disease in Childhood*, 1992;**67**(10(S)):1185–8.
- Gan 1999** *{published data only}*
Gan TJ, Bennett-Guerrero E, Phillips-Bute B, Wakeling H, Moskowitz DM, Olufolabi Y, et al. Hextend, a physiologically balanced plasma expander for large volume use in major surgery: a randomized phase III clinical trial. *Anesthesia and Analgesia* 1999;**88**(5):992–8.
- Hankeln 1990** *{published data only}*
Hankeln K, Senker R, Beez M. Comparative study of the intraoperative effectiveness of 5% human albumin or 10% hydroxyethyl starch (HAES-steril) on hemodynamics and oxygen transport in 40 patients [Vergleichende Untersuchung zur intraoperativen Wirksamkeit von 5% Humanalbumin oder 10% Hydroxyethylstarch (HAES-steril) auf Hamodynamik und Sauerstofftransport bei 40 Patienten]. *Infusionstherapie* 1990;**17**(3):135–140.
- Harke 1976** *{published data only}*
Harke H, Thoenies R, Margraf I, Momsen W. The influence of different plasma substitutes on blood clotting and platelet function during and after surgery [Der Einfluss verschiedener Plasmaersatzmittel auf Gerinnungssystem und Thrombocytenfunktion während und nach operativen Eingriffen. Vorläufige Ergebnisse einer klinischen Studie]. *Anaesthetist* 1976;**25**(8):366–373.
- Hiippala 1996** *{published data only}*
Hiippala S, Teppo AM. Perioperative volume effect of HES 120/0.7 compared with dextran 70 and Ringer acetate. *Annales cChirurgiae et Gynaecologiae* 1996;**85**(4):333–9.
- Huet 2000** *{published data only}*
Huet RCGG, Simons AW, Baus D, van Rooyen-Butijn WT, Haagenaars JAM, van Oeveren W, Bepperling F. A novel hydroxyethyl starch (Voluven(TM)) for effective perioperative plasma volume substitution in cardiac surgery. *Canadian Journal of Anaesthesia* 2000;**47**(12):1207–15.
- Jones 2004a** *{published data only}*
Jones SB, Whitten CW, Monk TG. Influence of crystalloid and colloid replacement solutions on hemodynamic variables during acute normovolemic hemodilution. *Journal of Clinical Anesthesia* 2004;**16**(1):11–7.
- Jovanovic 1997** *{published data only}*
Jovanovic K, Filipovic N, Romic P, Surbatovic M. Hetastarch in replacement of circulation volume compared to haemaccel and dextran 70 in pre-hospital resuscitation of polytraumatized patients. *Intensive Care Medicine*. 1997; Vol. 23:S184.
- Korttila 1984** *{published data only}*
Korttila K, Grohn P, Gordin A, Sundberg S, Salo H, Nissinen E, et al. Effect of hydroxyethyl starch and dextran on plasma volume and blood hemostasis and coagulation. *Journal of Clinical Pharmacology* 1984;**24**(7):273–82.
- Langeron 2001** *{published data only}*
Langeron ODM, Doelberg M, Ang ET, Bonnet F, Capdevila X, Coriat P. Voluven, a lower substituted novel hydroxyethyl starch (HES 130/0.4) causes fewer effects on coagulation in major orthopedic surgery than HES 200/0.5. *Anesthesia & Analgesia* 2001;**92**(4):855–62.
- Palumbo 2006** *{published data only}*
Palumbo. The effect of HES solution in critically ill patients. *Minerva anesthesiol* 2006;**72**:655–64.
- Puri 1983** *{published data only}*
Puri VK, Howard M, Paidipaty B, Singh S. Resuscitation in hypovolemia and shock: a prospective study of hydroxyethyl starch and albumin. *Critical Care Medicine* 1983;**11**(7):518–23.
- Rauch 2000** *{published data only}*
Rauch S, Sefrin P. Comparison of hydroxyethyl starch solutions derived from potato and corn starch [Vergleich von Hydroxyethylstarchlösungen aus Kartoffel- und Maisstärke]. *Anesthesiologie, Intensivmedizin, Notfallmedizin, Schmerztherapie* 2000;**35**(12):750–5.
- Rehm 2000** *{published data only}*
Rehm M, Orth V, Scheingraber S, Kreimeier U, Brechtelsbauer H, Finsterer U. Acid-base changes caused by 5% albumin versus 6% hydroxyethyl starch solution in patients undergoing acute normovolemic hemodilution: a randomized prospective study. *Anesthesiology* 2000;**93**(5):1174–83.
- Strauss 1985** *{published data only}*
Strauss RG, Stump DC, Henriksen RA, Saunders R. Effects of hydroxyethyl starch and fibrinogen, fibrin clot formation, and fibrinolysis. *Transfusion* 1985;**25**(3):230–4.
- Waxman 1989** *{published data only}*
Waxman K, Holness R, Tominaga G, Chela P, Grimes J. Hemodynamic and oxygen transport effects of pentastarch in burn resuscitation. *Annals of Surgery* 1989;**209**(3):341–5.

References to studies awaiting assessment

- Hopkins 1994** *{unpublished data only}*
Hopkins PM. 6% hydroxyethylstarch with 4% gelatine as perioperative intravenous volume replacement in surgical patients. National Research Register Version 1/1998.
- Romero 1999** *{published data only}*
Romero J, Luna P, Fernandez B, Rojas E, Sarrano X, Alvarez H. The use of HAES-Steril 6% as a plasma expander after cardiopulmonary bypass in aortocoronary surgery [Uso de HAES esteril 6% como expansor plasmático después de la circulación extracorpórea en revascularización coronaria]. *Revista Mexicana de Anestesiología* 1999;**22**:160–7.

Additional references

- Altman 1996**
Altman DG, Bland JM. Detecting skewness from summary information. *BMJ* 1996;**313**:1200.
- Armstrong 1994**
Armstrong RF, Bullen C, Cohen SL, Singer M, Webb AR. *Critical Care Algorithms*. Vol. **Oxford Medical Publications**, Oxford University Press, 1994.
- Berlin 1997**
Berlin JA. Does blinding of readers affect the results of meta-analyses?. *Lancet* 1997;**350**:185–6.

Boldt 1996

Boldt J, Heesen M, Muller M, pabsdorf M, Hempelmann G. The effects of albumin versus hydroxyethyl starch solution on cardiorespiratory and circulatory variables in critically ill patients. *Anesthesia and Analgesia* 1996;**83**:254–61.

CIGAR 2004

The Albumin Reviewers (Alderson P, Bunn F, Li Wan Po A, Li L, Roberts I, Schierhout G). Human albumin solution for resuscitation and volume expansion in critically ill patients. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: CD001208]

Roberts 2004

Roberts I, Alderson P, Bunn F, Chinnock P, Ker K, Schierhout G. Colloids versus crystalloids for fluid resuscitation in critically ill patients. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: CD000567]

Schulz 1995

Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical Evidence of Bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**: 408–12.

Traylor 1996

Traylor RJ, Pearl RG. Crystalloid versus colloid: All colloids are not created equal. *Anesthesia and Analgesia* 1996;**83**:209–12.

Vermeulen 1995

Vermeulen LC Jr, Ratko T A, Erstad BL, Brecher ME, Matuszewski K.A. A paradigm for consensus. The University Hospital Consortium guidelines for the use of albumin, nonprotein colloid, and crystalloid solutions. *Archives of Internal Medicine* 1995;**155** (4):373–9.

Yim 1995

Yim JM, Vermeyken LC, Erstad BL, Matuszewski KA, Burnett DA, Vlasses PH. Albumin and nonprotein colloid solution use in US academic health centers. *Archives of Internal Medicine* 1995;**155** (22):2450–5.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Allison 1999

Methods	Randomised controlled trial. Randomisation was based on date of admission (on even dates patients received HES). Analysis not intention to treat.
Participants	45 patients with blunt trauma who required colloid infusion. Patients were excluded if they were less than 12 years old, did not require admission to the ITU, died within 24 hours, were pregnant or in renal failure. 8 gelatin and 6 HES patients excluded after randomisation.
Interventions	1) HES (200/0.45 Pentaspan)n=24. 2) Gelatin (Gelofusine)n=21. After 24 hours, colloid administration was at the discretion of the clinician.
Outcomes	Death. Glasgow coma score. Volumes of blood and platelets infused. Haematological parameters.
Notes	Data were collected until the patient left the ITU or for a maximum of 5 days. Main outcome of interest was capillary leak.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	No	C - Inadequate

Arellano 2005

Methods	Randomised controlled trial. Study colloids placed in masked container by nurse not involved in other aspects of trial. All participants, health care workers and study personnel blinded to allocation.
Participants	50 adults undergoing surgical ablation of oropharyngeal cancer with free flap reconstruction (mean age 55). Exclusion criteria-ASA physical status classification iii-iv, cardiac insufficiency, pancreatitis, severe hepatic dysfunction, renal dysfunction, anaemia, coagulation abnormalities, ingestion of NSAID or ASA within 10 days of surgery and previous major head and neck surgery with free flap reconstruction.
Interventions	1) 5% HA (n=25). 2) HES 264/0.45 (n=25). CVP was maintained between 7-10 mmHg.

Arellano 2005 (Continued)

Outcomes	Clinical indices of coagulation. Number of units of blood transfused.	
Notes	Follow-up 24 hours. One patient in each group did not complete the study because planned surgical procedure was abandoned.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Asfar 2000

Methods	Randomised controlled trial. Allocation using sequentially numbered sealed opaque envelopes (information obtained on contact with the author).	
Participants	34 septic, hypovolaemic, ventilated and hemodynamically controlled patients. Inclusion criteria: patients aged over 16 years, systolic arterial pressure higher than 90mmHg and hypovolemia defined by PAOP of 12mmHg or less. Patients were excluded if they had an overt hemodynamic, ventilatory or acid base status instability. Sepsis was identified by either positive bacterial blood cultures, bronchoalveolar lavage or clinical evidence of infection.	
Interventions	1) 6% HES (n=16). 2) 4% Modified fluid gelatin (MFG) (n=18).	
Outcomes	Death. Haemodynamic variables.	
Notes	Follow-up was for one hour. Two patients in the HES group were excluded because they experienced haemodynamic instability. The final analysis was made on remaining 16 patients.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Beards 1994

Methods	Randomised controlled trial. Allocation by alternation. (Information on allocation concealment was obtained on contact with the author).	
Participants	28 patients with hypovolaemia, mechanically ventilated for concurrent acute respiratory failure. Patients fulfilled the following inclusion criteria: age >16 years, body weight between 50 and 85kg, mean arterial pressure <80mmHg (or 30mmHg less than previously recorded); pulmonary artery occlusion pressure	

Beards 1994 (Continued)

	<10mmHg with oliguria (i.e urine output <15 ml/hr).	
Interventions	1) Rapid infusion of 500ml modified fluid gelatin (n=15). 2) Rapid infusion of 500ml hetastarch (n=13).	
Outcomes	Death. Haemodynamic variables. Oxygen variables.	
Notes	Follow-up was 30 minutes for haemodynamic variables and until discharge for deaths.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	No	C - Inadequate

Berard 1995

Methods	Randomised controlled trial. A set of 200 tickets (type 1) and another set of 200 tickets (type 2) were mixed in a box. One ticket was drawn at random for each patient. Information on method of randomisation was obtained on contact with the author. Blinding not mentioned.	
Participants	319 patients in a resuscitation service receiving medical (gastrointestinal haemorrhage) and surgical cases. Patients were excluded if they had had a prior allergic reaction.	
Interventions	1) Gelatin (n=153). 2) HES (n=146). The prescribers chose the quantity of colloid, guided by normal practice.	
Outcomes	Death. Amount of colloid and red blood cells given. Cost.	
Notes	20 patients lost to follow up, no explanation given. Follow-up to discharge.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	No	C - Inadequate

Beyer 1997

Methods	Randomised controlled trial. Allocation was by a list of random numbers read by someone not entering patients into the trial (closed list). Information on method of allocation concealment was obtained by contact with the author. No blinding.
Participants	48 patients undergoing major elective hip surgery with an expected blood loss of >1000 ml. Exclusion criteria were haemoglobin concentration < or equal to 11g/dl, heart failure and coronary artery disease, myocardial infarction within the past 6 months, hypertension (>180mmHg systolic), impaired renal function, pregnancy, known hypersensitivity to HES or gelatin, patient taking drugs that may specifically affect blood viscosity, diuresis or clotting.
Interventions	1) 3% modified fluid gelatin (n=22). 2) 6% HES (n=19). Both groups also given Ringer's lactate. Fluids administered according to haemodynamic and clinical parameters.
Outcomes	Death (information on death was obtained by contact with the author). Haemodynamic variables. Packed cell volume, haemoglobin, clotting times. Incidence of allergic reactions.
Notes	Seven patients were lost to follow up but only 5 were accounted for.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Boldt 1986

Methods	Randomised controlled trial, using sealed opaque envelopes. Information on allocation concealment was obtained on contact with the authors. Blinding not mentioned. Loss to follow up not mentioned.
Participants	55 patients undergoing elective aorto-coronary bypass surgery. Exclusion criteria were ejection fraction < 50% and LVEDP >15 mmHg.
Interventions	1) 500ml 20% HA (n=15). 2) 500ml 3% HES (n=13). 3) 500ml 3.5% gelatin (n= 14). A fourth group received no colloid (n=13).
Outcomes	Haemodynamic variables. Incidence of anaphylactic shock. Amount blood transfused.

Boldt 1986 (Continued)

Notes	Follow-up until discharge from intensive care.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Boldt 1993a

Methods	Randomised controlled trial. Allocation by sequentially numbered sealed opaque envelopes (information obtained on contact with author).	
Participants	75 men undergoing elective aortocoronary bypass grafting, who had a pulmonary capillary wedge pressure of less than 5mmHg after induction of anaesthesia.	
Interventions	1) HA 5% (n=15). 2) 6% HES, HMW (n=15). 3) 6% HES, LMW (n=15). 4) Gelatin 3.5% (n=15). 5) No additional volume.	
Outcomes	Death (information obtained on contact with author). Haemodynamic variables.	
Notes	Follow-up 1 day.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Boldt 1995

Methods	Randomised controlled trial. Randomisation was by the use of sequentially numbered sealed opaque envelopes. Information on allocation concealment was obtained on contact with the author. Blinding of outcome assessors not mentioned.	
Participants	30 consecutive trauma patients (injury severity score >15) and 30 consecutive septic patients who underwent major surgery. Exclusions: patients suffering from renal failure requiring haemofiltration, severe liver dysfunction or coagulation abnormalities in their history were excluded as were patients who were receiving aspirin or other cyclooxygenase inhibitors.	
Interventions	1) 10% HES, LMW (n=15 trauma patients and 15 sepsis patients). 2) 20% human albumin (n=15 trauma patients and 15 sepsis patients). Fluid was given to maintain CVP and PCWP between 12 and 16mmHg.	

Boldt 1995 (Continued)

Outcomes	Death. Haemodynamic variables.	
Notes	Follow-up at 5 days. Deaths were reported within the study period and later (time not specified).	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Boldt 1996a

Methods	Randomised controlled trial. Allocation by sequentially numbered sealed opaque envelopes. Outcome assessors blinded to treatment.	
Participants	30 trauma patients and 30 patients suffering from sepsis secondary to major general surgery. Exclusions were patients with renal impairment, liver insufficiency, disseminated intravascular coagulation or septic shock.	
Interventions	1) 10% HES (n=30). 2) 20% HA solution (n=30). All patients also received Ringer's lactate solution. Volume therapy was given to maintain PCWP between 12 and 18mm Hg.	
Outcomes	Death. Haemodynamic variables.	
Notes	Follow-up at 5 days and at discharge from intensive care.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Boldt 1996b

Methods	Randomised controlled trial. Randomisation was by the use of sequentially numbered sealed opaque envelopes. Information on allocation concealment was obtained on contact with the author. The doctors giving the fluid were blinded to the solution but blinding of outcome assessors not mentioned. Loss to follow up not mentioned.	
Participants	45 consecutive trauma patients transferred to the surgical intensive care unit. Inclusion criteria were an injury severity score of >15 points. All patients were haemodynamically stable before being admitted to the study.	

Boldt 1996b (Continued)

Interventions	1) 10% HES (n=15). 2) 20% HA (n=15). 3) unspecified volume therapy regime (n=15). The allocated solution was given to maintain CVP and or PAWP between 12 and 18mmHg.	
Outcomes	Death. Haemodynamic variables. Circulating adhesion molecules.	
Notes	Deaths were reported within the study period and later (left ITU).	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Boldt 1996c

Methods	Randomised controlled trial. Allocation by sequentially numbered sealed opaque envelopes. Outcome variables were collected by an investigator who was blinded to the treatment. Loss to follow up not mentioned.	
Participants	56 patients from the surgical intensive care unit. 28 patients with an injury severity score >15 and 28 patients with sepsis secondary to major surgery. Patients with renal insufficiency, urine output <20ml h, severe liver dysfunction or disseminated intravascular coagulation were excluded.	
Interventions	1) 10% HES, LMW (trauma n=14, sepsis n=14). 2) 20% HA (trauma n=14, sepsis n=14). Fluid was infused to maintain PCWP at 10-15mmHg.	
Outcomes	Death. Haemodynamic variables.	
Notes	Follow-up was 5 days. Deaths were reported within the study period and later (time not specified).	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Boldt 1998

Methods	Randomised controlled trial. Allocation by sequentially numbered sealed opaque envelopes (information on allocation concealment was obtained on contact with the authors). Blinding of outcome assessors not mentioned. Loss to follow up not mentioned.
Participants	150 traumatised patients (injury severity score >15) and 150 postoperative patients with sepsis. Patients suffering from renal failure, severe liver insufficiency, or with major coagulation abnormalities were not included.
Interventions	1) 10% HES, LMW (n=150). 2) 20% HA (n=150). Both for 5 days to maintain the pulmonary wedge pressure between 12 and 15 TORR.
Outcomes	Death. Haemodynamic variables. Organ function. Coagulation.
Notes	Deaths were reported within the study period and after the study period (time not specified).

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Boldt 2000

Methods	Randomised controlled trial. Allocation by sequentially numbered sealed opaque envelopes (information obtained by contacting the author).
Participants	150 patients undergoing major abdominal surgery.
Interventions	1) 6% HES, LMW (n=50). 2) 6% HES, MMW (n=50). 3) 3% modified fluid gelatin (n=50). To keep MAP more than 70 mm Hg and CVP between 10 and 14 mm Hg. Volume was given perioperatively until the morning of the first post-op day. For each hour of surgery 500-800ml of crystalloids was routinely infused.
Outcomes	Death. Haemodynamic variables. Blood loss. Blood transfused. Cost.
Notes	Follow-up for one day post-op. Deaths recorded after study period.

Boldt 2000 (Continued)

<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Boldt 2001

Methods	Randomised controlled trial. Allocation by a "closed envelope system". Volume therapy was done by doctors who did not know the aim of the study.
Participants	75 patients undergoing major abdominal surgery. Volume was administered to keep the CVP between 8 and 12mmHg.
Interventions	1) 6% HES (n=25). 2) 6% HES (n=25). 3) 4% modified fluid gelatin (n=25) All groups also received 500ml of ringers lactate for each hour of surgery.
Outcomes	Death. Haemodynamic variables. Blood loss. Blood units transfused.
Notes	There were no deaths in the study period (until first Follow-up until first day post-op. Deaths until discharge.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Boldt 2006

Methods	Randomised controlled trial. Allocation by closed envelope. Blinding was done the by a pharmaceutical study and the statistician who perform all statistical analyses was also blinded to the grouping.
Participants	50 patients undergoing major abdominal surgery for malignancies (average age 74.5). Exclusion criteria were: cardiac insufficiency, altered liver function, preoperative anaemia or coagulation abnormalities, chronic use of corticosteroids or diuretics.
Interventions	1) 5% HA (n=25). 2) 6% HES (n=25). HA patients received 3960+/-590 ml of HA and 5070+/-1030 ml of RL. HES patients received 3500+/-530 ml of HES and 4500+/-880 ml of RL.

Boldt 2006 (Continued)

Outcomes	Death. Haemodynamic variables.	
Notes	Follow-up at 30 days and also one year for mortality.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Brock 1995

Methods	Randomised controlled trial. Allocation by list of random numbers read by someone entering patients into the trial (open list). Data on allocation concealment was obtained on contact with the authors.	
Participants	21 patients who had undergone cardiac surgery.	
Interventions	1) 10% HES. 200/0.5 in 7.2% saline (n=7). 2) 5% HA (n=7). 3) 6% hydroxyethyl starch in 0.9% saline (n=7).	
Outcomes	Death (data obtained on contact with author). Hemodynamic variables.	
Notes		
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	No	C - Inadequate

Brutocao 1996

Methods	Randomised double-blind controlled trial with pharmacy controlled randomisation. Information on allocation concealment was obtained on contact with the authors.	
Participants	38 children aged 1 year or more who were undergoing surgical repair of a congenital heart disease. Exclusion criteria included aminone therapy, renal disease, coagulopathy or a known bleeding diathesis.	
Interventions	1) 5% albumin (n=18). 2) 6% HES (n=20). Volume expansion was administered as clinically indicated to maintain adequate central venous pressure, perfusion and urine output. The total amount of colloid therapy was determined by care providers blinded to the randomisation.	

Brutocao 1996 (Continued)

Outcomes	Death (information on death was obtained on contact with the authors). Haemodynamic variables. Coagulation variables.	
Notes	Follow-up was until discharge from hospital. 9 children excluded post randomisation because they did not require colloid.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Carli 2000

Methods	Randomised controlled trial. Each centre received instructions from the coordinating Institute on the treatment to give the patient. Not intention to treat analysis.	
Participants	164 trauma patients. Patients were included if their SBP was less than 100mmHg, associated with signs of hypoperfusion.	
Interventions	1) HES (Hesteril 6%) (n= 85). 2) Gelatin (Plasmion) (n=79).	
Outcomes	Glasgow coma score. Haemodynamic variables. Units of blood transfused. Adverse reaction.	
Notes	There were 13 deaths from heart failure but these patients were excluded from the final analysis.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Claes 1992

Methods	Randomised controlled trial. No information given on method of randomisation. Blinding not mentioned. No loss to follow up.	
Participants	20 patients undergoing brain tumor surgery and 20 patients undergoing transabdominal hysterectomy. Exclusion criteria were preexisting coagulopathies; abnormal preoperative coagulation screening tests; intake of drugs affecting haemostasis within 2 week preoperatively as well as liver or kidney dysfunction.	

Claes 1992 (Continued)

Interventions	1000ml of fluid for volume replacement, either as 1) 6% HES (n=19). 2) 5% HA solution in 0.9% NaCl (n=21).	
Outcomes	Haemodynamic variables. Coagulation variables.	
Notes	Follow-up 48 hrs post-op.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Diehl 1982

Methods	Randomised controlled trial. Patients were allocated to groups according to their hospital identification number. Blinding not mentioned. No loss to follow up.	
Participants	60 Patients undergoing coronary artery bypass.	
Interventions	1) 6% HES (n=27). 2) 5% albumin (n=33) for volume expansion during the first 24 hours postoperatively. Neither hetastarch or albumin was used intraoperatively or in the pump prime.	
Outcomes	Death. Coagulation data. Haemodynamic variables.	
Notes	Follow-up 7 days post-op.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	No	C - Inadequate

Du Gres 1989

Methods	Randomised controlled trial. No information given on method of randomisation. Blinding not mentioned. No loss to follow up.	
Participants	30 patients post cardiac surgery. Patients were included if they were haemodynamically stable, were without serious 'rhythm' problems, had a mean arterial pressure less than 90mmHg, a mean pulmonary artery pressure less than 20mmHg and a central venous pressure less than 10mmHg. Patients excluded if they needed blood transfusion, had a hematocrit less than 28% or haemoglobin less than 9g/100ml.	
Interventions	1) 4% HA (n=15). 2) Haemaccel (n=15).	
Outcomes	Haemodynamic parameters.	
Notes	Follow-up 4 hours.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Dytkowska 1998

Methods	Randomised controlled trial. No information given on method of allocation concealment.	
Participants	40 patients post cardiac surgery. Patients were excluded if they had co-existing cardiogenic shock, renal failure with creatine level over 3.0mg or severe clotting disorders.	
Interventions	1) 200/0 HAES 6% (n=20). 2) Gelafundin (n=20). Colloids were administered to patients with diagnosed symptoms of hypovolaemia, during the first 24 hours post-op. Infusion rate was adjusted to patients needs but it did not exceed 1000ml/h.	
Outcomes	Haemodynamic parameters. Biochemical parameters. Adverse reactions.	
Notes	Follow-up 2 hours.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Evans 2003

Methods	Randomised controlled trial. Allocation concealment by use of 'sealed envelopes' (not enough information to be classified as adequate) . Treatment blinded (fluid set up by independent operator & covered with opaque black bag).
Participants	55 Patients undergoing unilateral cemented hip replacement. Exclusion criteria: cardiac insufficiency, renal insufficiency, altered liver function, preoperative anaemia, preoperative coagulation abnormalities and chronic use of corticosteroids and diuretics.
Interventions	1) 4.5% HA (n= 13). 2) 4% Gelosulfine (n=14). 3) Haemacel (n=14). 2L of fluid was infused during the operative period. One other group received normal saline (n=14).
Outcomes	Haemodynamic variables. Total blood loss.
Notes	Follow-up before surgery, at the end of the surgery and 2 hour post-op.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Falk 1988

Methods	Randomised controlled trial. No information given on method of randomisation. Blinding not mentioned. No loss to follow up.
Participants	12 patients with septic shock. Patients were excluded from the study if the pretreatment PAWP was greater than 10mmHg.
Interventions	1) 250ml of 5% albumin (n=6) 2) 250ml of 6% HES (n=6) every 15 minutes until the PAWP was increased to 15mmHg. The test infusion was then continued at 100 mL/hour to maintain PAWP at 15 mm Hg for the next 24 hours.
Outcomes	Haemodynamic variables. Clotting variables.
Notes	Follow-up 24 hours.

Risk of bias

Item	Authors' judgement	Description
------	--------------------	-------------

Falk 1988 (Continued)

Allocation concealment?	Unclear	B - Unclear
-------------------------	---------	-------------

Fries 2004

Methods	Randomised controlled trial. Allocation Concealment unclear. Treatment not blinded.
Participants	60 Patients undergoing primary knee replacement surgery. Exclusion criteria: contraindications for regional anaesthesia and puncture of the radial artery, any known allergies, primary and secondary haemostatic disorder.
Interventions	1) 4% gelofusine (n=20). 2) 6% HES (n=20). A 3rd group received ringers lactate. Before administrating spinal anaesthesia all patient received 500 ml RL. All patient intra operatively.
Outcomes	Haemodynamic variables.
Notes	Follow-up 2 hours post-op.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Fulachier 1994

Methods	Randomised controlled trial. No information given on method of randomisation. Blinding not mentioned. No loss to follow up.
Participants	16 patients undergoing cardiac surgery (8 were undergoing valve replacement and 8 coronary bypass) Patients were excluded if they were over 80, under 18 years of age, had been included in other studies, had received colloids in the month preceding surgery, had coagulation abnormalities or who were undergoing inotropic treatment.
Interventions	1) 500ml of a 4% solution of HA in Ringer's lactate (n=8). 2) 500ml of HES (n=8) until starting cardiopulmonary bypass.
Outcomes	Haemodynamic variables.
Notes	Follow up 30 minutes.

Risk of bias

Fulachier 1994 (Continued)

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Gahr 1981

Methods	Randomised controlled trial. No information given on method of randomisation. No loss to follow up.	
Participants	20 patients with hypovolaemia following abdominal surgery for malignoma.	
Interventions	1) 500ml HES 450/0.7 (n=10). 2) 500ml HA 5% (n=10) during the first 24 hrs after the operation.	
Outcomes	Haemodynamic parameters. Coagulation data.	
Notes	Follow-up 6 hrs.	

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Gallagher 1985

Methods	Randomised controlled trial. Allocation concealment by computerised system - patient details were entered before treatment assignment was revealed (data on allocation given on contact with author).	
Participants	10 patients after coronary artery bypass graft surgery. Exclusions: patients with significant left main coronary artery stenosis, poor left ventricular function or poor pulmonary function.	
Interventions	1) 5% albumin (n=5). 2) 6% HES (n=5).	
Outcomes	Death (data on deaths from author). Haemodynamic data.	
Notes	Follow-up 1 day.	

Risk of bias

Item	Authors' judgement	Description
------	--------------------	-------------

Gallagher 1985 (Continued)

Allocation concealment?	Yes	A - Adequate
-------------------------	-----	--------------

Gold 1990

Methods	Randomised controlled trial. Randomisation was done by alternation. Colloid solution was blinded by covering with foil. Information on allocation concealment was obtained by contact with the author. No loss to follow up.	
Participants	40 Surgical patients undergoing abdominal aortic aneurysm surgery.	
Interventions	1) 1g/kg of albumin 5% solution (n=20). 2) 1g/kg or hetastarch 6% solution (n=20).	
Outcomes	Death (data on death was obtained on contact with the author). Haemodynamic and coagulation variables.	
Notes	Follow-up not specified.	

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	No	C - Inadequate

Haisch 2001a

Methods	Randomised controlled trial. No information given on method of allocation concealment. Patient management by doctors who were blinded to the grouping.	
Participants	42 patients undergoing cardiac surgery. Patients were excluded if they had: an MI within previous 3 months, renal insufficiency, liver insufficiency, non controlled diabetes mellitus, preoperative coagulation abnormalities or patients treated with heparin or cyclooxygenase inhibitors within last 7 days.	
Interventions	1) Gelatin (n=21). 2) HES (n=21).	
Outcomes	Death. Use of blood products.	
Notes	Follow-up until first postoperative day.	

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Haisch 2001b

Methods	Randomised controlled trial using computer generated random numbers. No information given on allocation concealment.
Participants	42 patients undergoing major abdominal surgery for malignancies. Patients were excluded if they had cardiac insufficiency, renal insufficiency, altered liver function, pre-operative anemia, pre-operative coagulation abnormalities or if they had had cyclooxygenase inhibitors.
Interventions	1) HES (n=21). 2) Gelatin (n=21) until the morning of the first post-operative day.
Outcomes	Death. Use of allogenic blood products. Coagulation variables.
Notes	Follow-up until first postoperative day.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Hausdorfer 1986

Methods	Randomised controlled trial. No information given on method of randomisation.
Participants	30 children undergoing major surgery. During about 3 hours of surgery, the patients lost up to 15% of blood volume.
Interventions	1) Human albumin 5% (n=15). 2) HES 6% (n=15) with 14ml/kg body weight each, respectively.
Outcomes	Haemodynamic variables.
Notes	Follow-up 24 hours post-op.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Hedstrand 1987

Methods	Randomised controlled trial. no information given on method of randomisation. Post-op care staff were blinded. No loss to follow up.
Participants	275 patients undergoing major surgery. Patients were excluded if they if they were known to have decreased serum albumin levels or expected to sustain plasma loss, or had pronounced cardiovascular disease.
Interventions	1) PPF (n=142). 2) Dextran (n=133).
Outcomes	Volume transfused. Complication rates. Serum albumin. Deaths.
Notes	Follow-up one month.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Hiippala 1995

Methods	Randomised controlled trial. No information given on method of randomisation. Blinding not mentioned. 3 patients lost to follow up (explanation given).
Participants	60 patients undergoing major abdominal or urological surgery. Patients who had used platelet inhibiting drugs or had a diagnosed haemostatic defect were excluded.
Interventions	1) 3% dextrose (n=15). 2) 4% HES (n=15). 3) 6% HES (n=15). 4) 5% albumin (n=15).
Outcomes	Haemodynamic variables. Clotting variables. Blood loss.
Notes	Follow-up 3 days post-op.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Huang 2005

Methods	Randomised controlled trial. Allocation concealment and blinding was not clear.
Participants	20 patients with burns over 40% of total body surface area admitted 4-8 hour after injury.
Interventions	1) PPF (n=9). 2) Gelofusine (n=11). In a third control group patients did not receive fluid resuscitation.
Outcomes	Haemodynamic variables.
Notes	Follow-up 48 hours. No relevant outcome data.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Huskisson 1993

Methods	Randomised controlled trial. No information given on method of randomisation.
Participants	27 children returning to the intensive care unit following hypothermic open heart surgery.
Interventions	1) Albumin. 2) Gelatin. 3) Hetastarch.
Outcomes	Haemodynamic variables.
Notes	

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Huttner 2000

Methods	Randomised controlled trial. Allocation concealment using 'blind envelopes'. Anaesthetists responsible for patients management were blinded to the grouping.
Participants	60 patients undergoing major abdominal surgery. Patients were excluded if they had any of the following: cardiac insufficiency, renal insufficiency, liver dysfunction, pre-operative anaemia or coagulation abnormalities, or were on cyclooxygenase inhibitors or non steroidal therapy.

Huttner 2000 (Continued)

Interventions	1) 4% Gelatin (n=20). 2) 6% LMW HES (n=20). 3) 6% MMW HES (n=20).	
Outcomes	Haemodynamic variables. Clotting variables. Death.	
Notes	Follow up until first day post-op.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Jones 2004

Methods	Randomized controlled trial. Surgeons blinded to the fluid administered although the anaesthetist was aware of the fluid administered to a given patient. Allocation Concealment-unclear.	
Participants	40 adults scheduled to undergo radical retropubic prostatectomy. Exclusion criteria: coagulation disorder, platelet count <100,000/mm ³ , preoperative Hb <12 gm/dL, if anticoagulant therapy within 10 days of the surgery, aspirin or NSAID use less than 10 days before surgery or if they had documented allergy to any of the IV fluids used in the protocol.	
Interventions	1) 5% HA (n=10). 2) 6% dextran 70 (n=10). 3) 6% HES (n =10). A 4th group received ringers lactate. Hemodilution was done with the target of 9 gm/dL. All patient underwent moderate hemodilution to a target of Hb 9 gm/dL.	
Outcomes	Haemodynamic variables. Blood loss and units transfused.	
Notes	Follow-up was for 3 days.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Karanko 1987

Methods	Randomised controlled trial. Patients were randomised in blocks of four. Paper was put into a hat and taken out by an independent person. Information on method of randomisation was obtained on contact with the author. Blinding not mentioned. No loss to follow up.	
Participants	48 patients who had undergone coronary bypass surgery 20 hrs earlier.	
Interventions	1) 4% PPF (n=15). 2) 6% dextran-70 (n=10). 3) 5.5% Oxypolygelatin (n=12). A 4th group (not randomly selected) acted as a control (n=11).	
Outcomes	Death (data on death was obtained on contact with the author). Hemodynamic variables.	
Notes	Follow-up 28 hrs.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	No	C - Inadequate

Kirklın 1984

Methods	Randomised controlled trial. No information given on method of randomisation. Blinding not mentioned. No loss to follow up.	
Participants	30 patients undergoing coronary artery operations. Patients were excluded if they had undergone previous cardiac operations, if they had severe coagulopathies, anemia or chronic renal failure.	
Interventions	1) 6% HES (n=15). 2) 5% albumin (n=15). Both fluids infused over 24 hours to maintain left atrial pressure between 6 and 12mmHg and cardiac index greater than 2.0L/min/m ² .	
Outcomes	Death. Haemodynamic and coagulation variables. Adverse reactions.	
Notes	Follow-up until discharge from intensive care. 34 patients were originally included in the trial but data from 4 of them was not included in the final analysis.	
<i>Risk of bias</i>		

Kirklin 1984 (Continued)

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Lisander 1996

Methods	Randomised controlled trial. Randomisation using sequentially numbered sealed opaque envelopes. Information on allocation concealment was obtained on contact with the author. No loss to follow up. Blinding not mentioned.
Participants	40 patients undergoing revision hip arthroplasty.
Interventions	1) albumin 40g/L (n=20). 2) dextran 70 60g/L (n=20). Patients all received enoxaparin 40mg daily.
Outcomes	Death (data obtained from contact with author). External blood loss. Red cell balance. Packed cell volume.
Notes	Follow-up until discharge from hospital.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

London 1989

Methods	Randomised controlled trial. No information given on method of randomisation. Blinding not mentioned. No loss to follow up.
Participants	93 male cardiac surgical patients. Patients were excluded from the study if they had a significant coagulopathy or were anaemic (haematocrit value <30%).
Interventions	1) 10% pentastarch in 0.9% saline (n=50), 2) 5% HA in 0.9% saline (n=44), to provide volume expansion during the first 24 hours after cardiac operations.
Outcomes	Haemodynamic variables. Coagulation variables. Death. Length of stay.

London 1989 (Continued)

Notes	One patient was treated twice with an 8-month interval. Follow up until discharge from hospital.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Mastroianni 1994

Methods	Randomised controlled trial. No information given on method of randomisation. Blinding not mentioned.	
Participants	34 patients undergoing open heart surgery were enrolled.	
Interventions	1) 10% pentastarch. (n=12). 2) 5% albumin (n=17).	
Outcomes	Death. Haemodynamics variables. Clotting variables. Pulmonary oedema.	
Notes	Follow-up 7 days. Four patients in the pentastarch group, and one patient in the albumin group were excluded after randomisation.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Moggio 1983

Methods	Patients were randomised according to the last digit of their hospital identification numbers. No loss to follow up. Blinding not mentioned.	
Participants	47 postoperative open heart surgery patients. Operations performed included coronary revascularisation, valve operations, and combined coronary and valve procedures. Patients with pre existing hepatic or renal disease were not eligible for the study.	
Interventions	1) 5% albumin in 0.9% NaCl (n=23). 2) 6% HES in 0.9% NaCl (n=24).	

Moggio 1983 (Continued)

Outcomes	Haemodynamic variables. Clotting variables.	
Notes	Follow-up not specified.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	No	C - Inadequate

Molnar 2004

Methods	Randomised controlled trial. Allocation concealment and blinding unclear.	
Participants	30 hypovolemic patients with Intra thoracic blood volume index (ITBVI)<850 in septic shock with acute lung injury. Exclusion criteria: CVS failure (NYHA class iv), chronic respiratory failure (chronic hypoxia, hypercapnia) CRF requiring renal replacement therapy, chronic liver failure or those with diabetes mellitus or with known aortic aneurysm.	
Interventions	1) 6% HES (n= 15). 2) 4% GEL (n=15). 250ml/15min boluses (max 1000 ml) were given until the end point ITBVI>900 ml/m2.	
Outcomes	Death. Haemodynamic variables.	
Notes	Follow-up 60 min after the end point was reached. Follow up for deaths was not clear.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Munoz 1980

Methods	Randomised controlled trial. No information given on method of allocation. Blinding not mentioned. No mention of loss to follow up.	
Participants	14 patients with shock due to haemorrhage or sepsis.	

Munoz 1980 (Continued)

Interventions	Patients received either. 1) HES (hespan). 2) 5% albumin. Number in each group not reported.	
Outcomes	Haemodynamic variables.	
Notes	Follow-up 4 hours post infusion.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Munsch 1988

Methods	Randomised controlled trial. No information given on method of randomisation. Blinding not mentioned. No loss to follow up.	
Participants	40 consecutive patients undergoing elective coronary artery bypass graft surgery.	
Interventions	1) HES 6% (n=20) or 2) PPF (n=20) as their postoperative volume expander.	
Outcomes	Haemodynamic variables. Clotting variables. Death. Adverse reactions.	
Notes	Follow-up 7 days post-op.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Niemi 2006

Methods	Randomised controlled trial. Allocation by closed envelope (not enough information provided to classify as adequate). Blinding not clear.
Participants	45 patients post cardiac surgery. Exclusion criteria: pre operative coagulation disorders, renal or hepatic failure or taking medication with coumarin anticoagulants, heparin, and/or salicylic acids within the previous 5 days.
Interventions	1) 4% HA (n=15). 2) 4% gelatine (n=15). 3) 6% HES (n=15).
Outcomes	Death (data on death obtained on contact with the author). Clotting variables. Blood transfused.
Notes	Follow-up 1 day post-op. 54 patients gave consent but 9 later excluded.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Prien 1990

Methods	Randomised controlled trial. Method of allocation concealment unknown. Blinding not mentioned. Loss to follow up not mentioned.
Participants	18 patients undergoing modified Whipple's operation (hemipancreato-duodenectomy). Patients were eligible for the study if there was an absence of major organ dysfunction and serum protein, sodium, glucose, blood urea nitrogen, haematocrit, aPTT and PT times, and platelet times were within normal limits. Specific exclusion criteria included compensated myocardial insufficiency, chronic hypertension, chronic obstructive airways disease and insulin-dependent diabetes mellitus.
Interventions	1) 10% HES (n=6). 2) 20% HA (n=6). A third group were given ringer's lactate (n=6) All given as a volume replacement solution, which was given to maintain central venous pressure at the pre-operative level.
Outcomes	Death (data on death was obtained on contact with the author). Haemodynamic variables. Clotting variables.

Prien 1990 (Continued)

Notes	Follow-up unspecified. Study was intra-operative.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Rackow 1983

Methods	Randomised controlled trial, method of allocation concealment not mentioned. Blinding not mentioned.	
Participants	18 patients with hypovolemic and septic shock. Patients were excluded if they were less than 18 yrs of age, considered to be in a terminal state, or had a significant coagulopathy.	
Interventions	1) albumin (n=9). 2) HES (n=9). Patients received 250ml of the treatment fluid every 15 minutes as a fluid challenge. The fluid challenge ended when the WP equalled 15 mmHg. Thereafter the treatment fluid was given in sufficient quantities to maintain the WP at 15 mmHg for the next 24h, at which point the study was completed.	
Outcomes	Death. Haemodynamic variables. Respiratory variables.	
Notes	Deaths given for study period and for length of hospital stay. Survival until discharge was used for the mortality data for this review.	

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Rackow 1989

Methods	Randomised controlled trial. Method of allocation concealment was not recorded. No loss to follow up. Blinding not mentioned.	
Participants	20 patients with severe sepsis and systemic hypoperfusion. Patients were excluded from the study if they were <21 yrs of age, pregnant, considered to be terminal, or they manifested spontaneous bleeding.	

Rackow 1989 (Continued)

Interventions	1) 5% albumin (n=10). 2) 10% hydroxyethyl starch (pentastarch)(n=10). Each group received 250ml of the treatment fluid every 15 mins until either the WP was > or equal to 15mm Hg or a maximum volume of 2000ml of study colloid was infused.	
Outcomes	Death. Haemodynamic variables. Clotting variables. Allergic reactions.	
Notes	Follow-up unspecified.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Rittoo 2004

Methods	Randomised controlled trial. Allocation by sealed envelopes (not enough information provided to classify as adequate). Blinding-not clear.	
Participants	40 patients undergoing abdominal aortic aneurysm surgery. Exclusion criteria: ejection fraction of <40% with poor pulmonary function with micro albuminuria and a creatinine concentration of >150 micro ml per litre.	
Interventions	1) 6% HES (n=20). 2) 4% Gelosulfine (n=20). All patients received crystalloid. Colloid infused to maintain stable heart rate, CVP 8-10 cm H2O and steady mean arterial pressure and urine output of >40ml/hr.	
Outcomes	Lung function. Adverse events.	
Notes	Follow-up 24 hours.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Rosencher 1992

Methods	Randomised controlled trial. Randomisation done using sequentially numbered, sealed, opaque envelopes. Information on allocation concealment was obtained on contact with the author. No mention of blinding. Loss to follow up not mentioned.	
Participants	32 patients undergoing total hip replacement.	
Interventions	1) 4% albumin (n=16). 2) LMW HES (n=16).	
Outcomes	Death (data obtained on contact with author). Bleeding. Clotting variables.	
Notes	Follow-up for 5 days post-op.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Schortgen 2001

Methods	Randomised controlled trial. Allocation was by sealed opaque envelopes serially numbered and used in sequence.	
Participants	129 patients with severe sepsis or septic shock over 18 years of age. Patients were excluded if they were pregnant, had a history of allergy to HES or gelatin, had severe acute or chronic renal dysfunction or previous administration of HES or mannitol.	
Interventions	1) 6% hydroxyethyl starch (n=65). 2) 3% fluid-modified gelatin (n= 64).	
Outcomes	Death (data obtained on contact with author). Length of stay in ITU. Acute renal failure.	
Notes	Follow-up while in ITU.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Shatney 1983

Methods	Controlled clinical trial. Patients were assigned to groups in an alternating fashion. No loss to follow up. No mention of blinding.	
Participants	32 patients with multisystem trauma and/or haemorrhagic shock. Patients with cardiac arrest on hospital admission or during the first half hour after admission were excluded from the study.	
Interventions	1) PPF 5% solution (n=16), 2) Hetastarch 6% (n=16). Study patients continued to receive the assigned colloid solution for the first 8 days whenever colloid was thought necessary.	
Outcomes	Hepatic, pulmonary and renal function. Clotting variables. Volume of fluids infused. Deaths.	
Notes	Follow-up 8 days.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	No	C - Inadequate

Stockwell 1992

Methods	Randomised controlled trial. No information given on method of randomisation. No loss to follow up. Blinding not mentioned.	
Participants	475 patients admitted to the intensive care unit. Patients were excluded from the study if they were under 18 years or if admitted for cardiac monitoring or cardiac thrombolytic therapy.	
Interventions	1) 4.5% albumin (n=226). 2) A synthetic colloid polygeline (Haemaccel)(n=249) for intravenous volume replacement.	
Outcomes	Death. Length of stay in ITU. Incidence of renal failure. Pulmonary oedema.	
Notes	Follow-up until discharge from ITU.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description

Stockwell 1992 (Continued)

Allocation concealment?	Unclear	B - Unclear
-------------------------	---------	-------------

Stoddart 1996

Methods	Randomised blinded trial. Sequentially numbered sealed opaque envelopes were used. Information on allocation concealment was obtained on contact with the author. Anaesthetist unaware of intervention. No loss to follow up.
Participants	30 neonates undergoing major surgery. They were excluded if the body weight was less than 2kg or more than 5kg, the preoperative haemoglobin was less than 14g/dl, they had previously received blood or colloid, or they had suspected major cardiac, renal, metabolic or chromosomal abnormalities. Neonates were withdrawn from the study if either blood or more than 40ml/kg of colloid was required either during or within the first 24hr after surgery.
Interventions	1) HA 4.5% (n=15). 2) Haemaccel (n=15).
Outcomes	Haemodynamic variables. Plasma albumin. Haemoglobin.
Notes	Follow-up 24 hrs post op.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Tollofsrud 1995

Methods	Randomised controlled trial. Allocation concealment was by the use of sequentially numbered sealed opaque envelopes. Information on allocation concealment was obtained on contact with the authors. No loss to follow up. Blinding not mentioned.
Participants	30 patients undergoing elective coronary artery bypass surgery. Patients with left ventricular ejection fraction less than 40%, valvular heart disease, ventricular aneurysm, arrhythmia, diabetes mellitus, renal failure or lung disease were excluded.
Interventions	1) Polygeline (Haemaccel)(n=10). 2) Dextran 70 (n=10). 3) Albumin 40 (n=10). A 4th group received ringers lactate (n=10).

Tollofsrud 1995 (Continued)

Outcomes	Death. Haemodynamic variables. Respiratory data. Cost of fluid regimens.	
Notes	Follow-up 48 hours during and after surgery.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Van der Linden 2004

Methods	Randomised controlled trial. Patients were randomly allocated by opening an envelope (not enough information provided to classify as adequate). Blinding not clear.	
Participants	110 patients (average age 63) undergoing cardiac surgery under cardiopulmonary bypass (elective coronary artery or single valve surgery). Exclusion criteria: undergoing combined cardiac surgery or redo operations, history of allergic reactions to starches or gelatins, significant liver or renal dysfunction.	
Interventions	1) 6% HES (n=55). 2) 3.5% urea-lined gelatine (n=55). If additional colloid required 4.5% HA given.	
Outcomes	Death. Hemodynamic variables. Blood transfused.	
Notes	Follow-up 18 hours after surgery.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Van der Linden 2005

Methods	Randomised controlled trial. Allocation concealment and blinding unclear.	
---------	--	--

Van der Linden 2005 (Continued)

Participants	132 patients with a preoperative left ventricular ejection fraction more than 35% undergoing elective primary cardiac surgery.	
Interventions	1) 6% HES 130/0.4 [48.9+/-17.2 ml/kg] (n=64). 2) 3% GEL [48.9+/-14.6 ml/kg] (n=68).	
Outcomes	Hemodynamic variables. Blood loss.	
Notes	Follow-up until 5 days post-op.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Veneman 2004

Methods	Randomised controlled trial. Allocation by sealed envelopes kept outside of hospital.	
Participants	61 critically ill hypoalbuminemic patients (serum concentration < 20g/l).	
Interventions	1) Albumin (n=15). 2) HES 10% 500ml (n=15). 3) HES 10% 1000ml (n=15). A fourth group received saline.	
Outcomes	Death. Hemodynamic variables. Adverse events (from author).	
Notes	Follow-up 72 hours post-op, mortality 30 days.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Verheij 2006

Methods	Randomised controlled trial. Hospital pharmacy assigned patients via sealed envelope method.	
Participants	67 patients undergoing either vascular (n=28) or cardiac surgery (n=40). Exclusion criteria: age > 79 years and known anaphylactoid reactions to colloids.	

Verheij 2006 (Continued)

Interventions	1) 4% Gelatine (n=16). 2) 6% HES (n=18). 3) 5% HA (n=18). A 4th group received normal saline.	
Outcomes	Death. Haemodynamic variables.	
Notes	Follow-up not clear.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Vogt 1994

Methods	Randomised controlled trial. No information given on method of randomisation.	
Participants	40 patients undergoing major surgery. Exclusion criteria included anaemia, renal, liver and coagulation disorders.	
Interventions	1) 5% HA (n=20). 2) 6% hydroxyethyl starch (n=20).	
Outcomes	Haemodynamic variables. Coagulation. Haematological parameters. Blood loss and blood intake.	
Notes		
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Vogt 1996

Methods	The patients were divided into two groups using random numbers. Blinding not mentioned. No loss to follow up.	
---------	---	--

Vogt 1996 (Continued)

Participants	41 patients undergoing total hip arthroplasty during the perioperative period. Exclusion criteria were weight <60 kg, age <18 yrs, ASA grade>III, haematocrit <34% or >44%, history of coagulopathies or a Quick's prothrombin test of < 75%, partial thromboplastin time (PTT) > 45s, platelet count < 100,000/mm ³ , impaired liver function and renal failure.	
Interventions	1) 6% HES (n=20). 2) 5% HA (n=21).	
Outcomes	Haemodynamic and clotting variables.	
Notes	Follow-up 6 hrs post-op.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Vogt 1999

Methods	Randomised controlled trial. No information given on method of randomisation.	
Participants	50 patients undergoing radical prostatectomy or cystectomy with bladder replacement. Exclusion criteria were: weight less than 60kg, age less than 21 years, ASA 1 or 2, haemoglobin less than 12g/dl, history of clotting disorders, liver function disorders, advanced renal insufficiency or hypoproteinaemia.	
Interventions	1) 5% HA. 2) 6% HES 200/0.5.	
Outcomes	Haemodynamic variables. Blood loss.	
Notes	Follow-up for 3 days.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

von Sommoggy 1990

Methods	Randomised controlled trial. No information given on method of randomisation. No loss to follow up.	
Participants	24 patients undergoing infrarenal aortofemoral bifurcation grafting.	
Interventions	1) FFP and 5% HA (n=13). 2) HES 200 10% and HES 450 6% (n=11).	
Outcomes	Haemodynamic variables. Clotting variables. Influence on organ function.	
Notes	Follow-up 6 hours post-op.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Wahba 1996

Methods	Randomised controlled trial. Computerised system was used for randomisation. Data on method of allocation concealment was obtained on contact with the author. Blinding not mentioned. Loss to follow up not mentioned.	
Participants	20 patients who had had coronary artery bypass grafting. Patients with abnormal left-ventricular function as judged from cine-angiography were excluded as were patients on anticoagulants less than 10 days before the operation.	
Interventions	1) 5% albumin (n=10). 2) Haemaccel (n=10).	
Outcomes	Death (data on death was obtained on contact with the author). Haemodynamic variables.	
Notes	Follow-up was 2 weeks.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Watkins 1990

Methods	Randomised controlled trial. No information given on method of randomisation.	
Participants	12 patients undergoing major surgery.	
Interventions	1) LMW polystarch or 2) Polygelatine (Haemaccel) for postoperative volume replacement.	
Outcomes	Death. Adverse reactions.	
Notes	Follow-up was for 24 hours after the infusion.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Woittiez 1997

Methods	Randomised controlled trial. Allocation concealment by sequentially numbered sealed opaque envelopes.	
Participants	60 patients who had developed hypoalbuminaemia (<20g/l) after major surgery. 2 patients died after randomisation and before treatment started. These were excluded from the analysis.	
Interventions	1) albumin 20% (300 ml/24h) (n=15). 2) HES 10% (500ml/24h) for 3 days (n=27). Aim was to restore colloid osmotic pressure. A 3rd group received saline (n=16).	
Outcomes	Death (data on death obtained on contact with the author). Changes in fluid balance, serum albumin, COP and clinical signs of oedema were followed daily.	
Notes	Follow-up unspecified.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

ALI = Acute lung injury

CVS = cardiovascular system

COP = Colloid osmotic pressure

CVP = central venous pressure

CRF = chronic renal failure
 EVLW = Extra vascular lung water
 HA = Human albumin
 HES = hydroxyethyl starch
 HMW = high molecular weight
 ITBVI = Intra thoracic blood volume index
 LIS = Lung Injury Scale
 LMW = Low molecular weight
 LVEDP = left ventricular end diastolic pressure
 MMW = medium molecular weight
 MAP = mean arterial pressure
 PAWP = pulmonary artery wedge pressure
 PAOP = pulmonary artery occlusion pressure
 PCWP = pulmonary capillary wedge pressure
 PPF = plasma protein fraction

Characteristics of excluded studies *[ordered by study ID]*

Boldt 1993	Pre-bypass volume loading.
Boldt 2000b	Compares two starches with each other.
Brehme 1993	Haemodilution.
Bremerich 2000	Compares two different starches (acetyl starch with hydroxyethyl starch).
Charlet 1991	Study compared two different gelatins with each other and not with other colloids.
Christ 1997	Non-randomised trial.
Emery 1992	Trial compares 20% and 4.5% albumin with each other and not with other colloids.
Gan 1999	Compares Hextend (a plasma volume expander based upon 6% hetastarch) with 6% hetastarch in saline (HES).
Hankeln 1990	Haemodilution.
Harke 1976	Unable to find out if a randomised controlled trial. Methodology unclear.
Hiiippala 1996	Patients were expected to have minimal blood loss.
Huet 2000	Compares two starches with each other.
Jones 2004a	Haemodilution.
Jovanovic 1997	Does not mention if study was randomised. Unable to contact author for further information.
Korttila 1984	Healthy volunteers and cross over trial.

(Continued)

Langeron 2001	Compares two starches with each other.
Palumbo 2006	Authors do not report the number of patients randomised to each group.
Puri 1983	There is no mention of a method of randomisation. Just says "Twenty-five patients studied in each group were well matched".
Rauch 2000	Compares two starches with each other.
Rehm 2000	Haemodilution.
Strauss 1985	Healthy volunteers.
Waxman 1989	Cross-over study.

DATA AND ANALYSES

Comparison 1. Albumin or PPF versus HES

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Death	25	1234	Risk Ratio (M-H, Fixed, 95% CI)	1.14 [0.91, 1.43]
2 Blood/red cells transfused (skewed or inadequate data)			Other data	No numeric data

Comparison 2. Albumin or PPF versus Gelatin

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Death	7	636	Risk Ratio (M-H, Fixed, 95% CI)	0.97 [0.68, 1.39]
2 Blood/red cells transfused (skewed or inadequate data)			Other data	No numeric data

Comparison 3. Albumin or PPF versus Dextran

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Death	4	360	Risk Ratio (M-H, Fixed, 95% CI)	3.75 [0.42, 33.09]
2 Blood/red cells transfused (skewed or inadequate data)			Other data	No numeric data

Comparison 4. Modified Gelatin versus HES

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Death	18	1337	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.80, 1.25]
2 Blood/red cells transfused (skewed or inadequate data)			Other data	No numeric data

Comparison 5. Modified Gelatin versus Dextran

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Death	2	42	Risk Ratio (M-H, Fixed, 95% CI)	Not estimable
2 Blood/red cells transfused (skewed or inadequate data)			Other data	No numeric data

Comparison 6. HES versus Dextran

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Death	0	0	Risk Ratio (M-H, Fixed, 95% CI)	Not estimable
2 Blood/red cells transfused (skewed or inadequate data)			Other data	No numeric data

FEEDBACK

Colloid solutions for fluid resuscitation

Summary

1. Please explain, in the 'what's new' section, in what respects this update differs from the previous version.
2. The drug companies listed in the acknowledgments are not in alphabetic order: please do so or explain the reason for the order shown (e.g. in order of helpfulness).
3. Fresenius is misspelt.
4. In the references to included trials, please use an asterisk to identify those trials which are the main publication where there are more than one article referring to a trial.

Reply

1. The review has been marked as an update by mistake. As of September 1999 no substantial updates have been made.
2. The drug companies have been re-ordered alphabetically.
3. The spelling of Fresenius is corrected.
4. The primary reference has been marked with an asterisk.

Contributors

Comment by Andrew Herxheimer

Response by Frances Bunn

WHAT'S NEW

Last assessed as up-to-date: 1 October 2007.

11 July 2008	Amended	Converted to new review format.
--------------	---------	---------------------------------

HISTORY

Protocol first published: Issue 4, 1998

Review first published: Issue 2, 1999

2 October 2007	New search has been performed	The search for the review was updated in March 2007 and thirteen new studies were added to the review.
----------------	-------------------------------	--

CONTRIBUTIONS OF AUTHORS

FB screened citations for eligibility, obtained references, contacted authors, extracted data, entered data and wrote up the review. DT and SA obtained references, screened citations for eligibility and extracted data. PA and VH contributed to earlier versions of the review.

DECLARATIONS OF INTEREST

None known.

SOURCES OF SUPPORT

Internal sources

- University of Hertfordshire, UK.

External sources

- NHS Research and Development Programme, UK.

INDEX TERMS

Medical Subject Headings (MeSH)

*Fluid Therapy; Blood Proteins [*therapeutic use]; Colloids [therapeutic use]; Dextrans [*therapeutic use]; Plasma Substitutes [*therapeutic use]; Randomized Controlled Trials as Topic; Rehydration Solutions [*therapeutic use]

MeSH check words

Humans

Hypertonic versus near isotonic crystalloid for fluid resuscitation in critically ill patients (Review)

Bunn F, Roberts IG, Tasker R, Trivedi D



**THE COCHRANE
COLLABORATION®**

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2008, Issue 4

<http://www.thecochranelibrary.com>



Hypertonic versus near isotonic crystalloid for fluid resuscitation in critically ill patients (Review)
Copyright © 2008 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
BACKGROUND	2
OBJECTIVES	3
METHODS	3
RESULTS	4
DISCUSSION	4
AUTHORS' CONCLUSIONS	5
ACKNOWLEDGEMENTS	5
REFERENCES	5
CHARACTERISTICS OF STUDIES	7
DATA AND ANALYSES	18
FEEDBACK	18
WHAT'S NEW	19
HISTORY	19
CONTRIBUTIONS OF AUTHORS	19
DECLARATIONS OF INTEREST	19
SOURCES OF SUPPORT	19
INDEX TERMS	20

[Intervention Review]

Hypertonic versus near isotonic crystalloid for fluid resuscitation in critically ill patients

Frances Bunn¹, Ian G Roberts², Robert Tasker³, Daksha Trivedi¹

¹Centre for Research in Primary and Community Care, University of Hertfordshire, Hatfield, UK. ²Cochrane Injuries Group, London School of Hygiene & Tropical Medicine, London, UK. ³University of Cambridge School of Clinical Medicine, Department of Paediatrics, Cambridge, UK

Contact address: Frances Bunn, Centre for Research in Primary and Community Care, University of Hertfordshire, College Lane, Hatfield, Hertfordshire, AL10 9PN, UK. f.bunn@herts.ac.uk.

Editorial group: Cochrane Injuries Group.

Publication status and date: Edited (no change to conclusions), published in Issue 4, 2008.

Review content assessed as up-to-date: 14 October 2007.

Citation: Bunn F, Roberts IG, Tasker R, Trivedi D. Hypertonic versus near isotonic crystalloid for fluid resuscitation in critically ill patients. *Cochrane Database of Systematic Reviews* 2004, Issue 3. Art. No.: CD002045. DOI: 10.1002/14651858.CD002045.pub2.

Copyright © 2008 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Hypertonic solutions are considered to have a greater ability to expand blood volume and thus elevate blood pressure and can be administered as a small volume infusion over a short time period. On the other hand, the use of hypertonic solutions for volume replacement may also have important disadvantages.

Objectives

To determine whether hypertonic crystalloid decreases mortality in patients with hypovolaemia.

Search strategy

We searched the Cochrane Injuries Group's specialised register, MEDLINE, EMBASE, The Cochrane Library, issue 3, 2007, The National Research Register issue 3, 2007 and the British Library's Electronic Table of Contents ZETOC. We also checked reference lists of all articles identified. The searches were last updated in October 2007

Selection criteria

Randomised trials comparing hypertonic to isotonic and near isotonic crystalloid in patients with trauma or burns or who were undergoing surgery.

Data collection and analysis

Two authors independently extracted the data and assessed the quality of the trials.

Main results

Fourteen trials with a total of 956 participants are included in the meta-analysis. The pooled relative risk (RR) for death in trauma patients was 0.84 (95% confidence interval [CI] 0.69 to 1.04); in patients with burns 1.49 (95% CI 0.56 to 3.95); and in patients undergoing surgery 0.51 (95% CI 0.09 to 2.73). In the one trial that gave data on disability using the Glasgow outcome scale, the relative risk for a poor outcome was 1.00 (95% CI 0.82 to 1.22).

Hypertonic versus near isotonic crystalloid for fluid resuscitation in critically ill patients (Review)
Copyright © 2008 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

1

Authors' conclusions

This review does not give us enough data to be able to say whether hypertonic crystalloid is better than isotonic and near isotonic crystalloid for the resuscitation of patients with trauma or burns, or those undergoing surgery. However, the confidence intervals are wide and do not exclude clinically significant differences. Further trials which clearly state the type and amount of fluid used and that are large enough to detect a clinically important difference are needed.

PLAIN LANGUAGE SUMMARY

More evidence needed as to the best concentration of crystalloid to use in resuscitation fluids

Fluid resuscitation is usually given when a patient has lost a lot of blood, but there is continuing uncertainty as to the best sort of fluid to use. Some of the fluids used contain substances classified as "crystalloids", but should the concentration of crystalloids in the fluid be about the same as their concentration in human blood ("isotonic") or higher ("hypertonic")? It is commonly believed that hypertonic crystalloid is the more effective at increasing blood volume but that there could be some disadvantages to using it. This review has assessed the evidence from studies that compared the use of the two types of fluid with patients who had been injured or burned, or were having surgery. Not enough evidence is available, however, to decide which crystalloid concentration is best. More research is needed.

BACKGROUND

Fluid resuscitation is a mainstay of the medical management of haemorrhagic hypovolaemia. However, there is continuing uncertainty about the most appropriate fluid (Krausz 1995). Isotonic crystalloid solutions are often used to replace blood loss until a blood transfusion can be administered, but the wish to administer large volumes (advanced trauma life support [ATLS] guidelines suggest two litres of isotonic crystalloid), particularly in the pre-hospital phase when there may be problems with venous access, has stimulated the development of alternative approaches. One such approach is the use of hypertonic saline. Hypertonic solutions are considered to have a greater ability to expand blood volume and thus elevate blood pressure, and can be administered as a small volume infusion over a short time period (Krausz 1995). Infusion of hypertonic saline is believed to act by causing an osmotic shift of fluid from the intracellular and interstitial spaces to the extracellular compartment. The resulting auto-transfusion of fluid increases blood pressure and circulating volume. The use of hypertonic solutions has the potential to provide rapid volume resuscitation but with less interstitial oedema than with isotonic saline solutions (Shackford 1983).

It has also been suggested that hypertonic solutions may be the fluid of choice in hypovolaemic patients with head injuries (Khanna 2000; Peterson 2000; Walsh 1991). Cerebral perfusion pressure (CPP) depends on both intracranial pressure (ICP) and

mean arterial blood pressure. (CPP = mean arterial blood pressure - mean ICP.) Patients in hypovolaemic shock who have head injuries may require rapid blood pressure elevation to maintain cerebral perfusion pressure, but excessive fluid and salt administration may result in brain swelling with an increase in intracranial pressure. Hypertonic solutions, however, are believed to reduce intracranial pressure by establishing an osmotic gradient across the blood brain barrier that draws water from the brain tissue into the vascular space (Fisher 1992). Hypertonic solutions, therefore, have the potential to restore blood pressure rapidly, but without increasing intracranial pressure. Hypertonic solutions are also thought to be beneficial in preventing the 'water logging' effect when there is interstitial lung injury, for example as occurs both in elective surgery and in trauma.

On the other hand, the use of hypertonic solutions for volume replacement may also have important disadvantages. In situations where haemorrhage is on-going, hypertonic solutions may result in continued bleeding from injured vessels. A potential problem in head injuries is that, in patients with a disrupted blood brain barrier, excess sodium may leak into brain tissue drawing water with it, thus worsening cerebral oedema. At present, there are no clinical ways to assess the integrity of the blood brain barrier. Furthermore, not only could the integrity of the blood brain barrier vary among patients with head injury, but it might also vary in

different parts of the brain in a single patient. The possibility that hypertonic fluids may worsen outcome following head injury cannot therefore be dismissed (Krausz 1995, Shenkin 1976).

OBJECTIVES

To determine whether hypertonic crystalloid decreases mortality in patients with hypovolaemia with and without head injuries, we conducted a systematic review of randomised controlled trials.

METHODS

Criteria for considering studies for this review

Types of studies

Randomised controlled trials. Crossover trials were excluded.

Types of participants

Patients with trauma, burns or surgery. Trials in both the pre-hospital and hospital setting were included.

Types of interventions

Trials that compare hypertonic to isotonic and near isotonic crystalloid. Trials of hypertonic crystalloid with an add-on colloid (e.g. hypertonic saline and dextran) are not included. This comparison has been dealt with in a previous systematic review by the Cochrane Injuries Group (Perel 2007).

Types of outcome measures

The principal outcome measure is mortality from all causes and disability assessed at the end of the follow-up period scheduled for each trial. Disability was assessed using the Glasgow outcome scale (Jennett 1975) which includes the following categories: death, persistent vegetative state, severely disabled, moderately disabled and good recovery. For the purpose of this review, the scale was dichotomised with death, persistent vegetative state and severe disability denoting a poor outcome, and moderate disability and good recovery denoting a good outcome.

Intermediate physiological outcomes were not used for several reasons. Such outcomes are subject to intra and inter-observer variation, they have no face value to patients and relatives, and those seen as appropriate are not stable over time. Also, there would need to exist a strong predictive relationship between the variable and mortality.

Search methods for identification of studies

The search was last updated in October 2007.

Electronic searches

We searched the following electronic databases;

- Cochrane Injuries Group's specialised register
- CENTRAL
- MEDLINE
- EMBASE
- National Research Register
- Zetoc.

The original search strategies were based on the terms listed below. The full search strategies for the most recent search update are listed in Appendix 1.

1. Saline solutions hypertonic (MeSH)
2. Isotonic solutions (MeSH)
3. Hypertonic or isotonic or hyperosmotic or hyperoncotic
4. Hypotensive resuscitation
5. #1 or #2 or #3 or #4
6. random*
7. double-blind-procedure (MeSH)
8. #6 or #7
9. #5 and #8

Data collection and analysis

Selection of studies

One reviewer (FB) examined the electronic search results for reports of possibly relevant trials and these reports were retrieved in full. Two reviewers applied the selection criteria independently to the trial reports, resolving disagreements by discussion.

Data extraction and management

Two reviewers independently extracted information on the following: study quality, number of randomised patients, type of participants and the interventions. The outcome data sought were number of deaths and disability. The reviewers were not blinded to the authors or journal when doing this, as evidence for the value of this is far from conclusive (Berlin 1997). Results were compared and any differences resolved by discussion.

For each trial the relative risk of death and 95% confidence interval were calculated. The relative risk was chosen as it is more readily applied to the clinical situation.

The groups of trials were examined for statistical evidence of heterogeneity using a chi squared test. As there was no obvious heterogeneity on visual inspection or statistical testing, pooled relative risks (RR) and 95% confidence intervals (CIs) were calculated using a fixed effects model.

Assessment of risk of bias in included studies

Since there is evidence that the quality of allocation concealment particularly affects the results of studies (Schulz 1995), two reviewers scored this quality on the scale used by Schulz (Schulz 1995) as shown below, assigning C to poorest quality and A to best quality: A = trials deemed to have taken adequate measures to conceal allocation (i.e. central randomisation; numbered or coded bottles or containers; drugs prepared by the pharmacy; serially numbered, opaque, sealed envelopes; or other description that contained elements convincing of concealment).

B = trials in which the authors either did not report an allocation concealment approach at all or reported an approach that did not fall into one of the other categories.

C = trials in which concealment was inadequate (such as alternation or reference to case record numbers or to dates of birth)

In addition, data was extracted on blinding and loss to follow-up. Where the method used to conceal allocation was not clearly reported, the author was contacted, if possible, for clarification. We then compared the scores allocated and resolved differences by discussion.

RESULTS

Description of studies

See: [Characteristics of included studies](#); [Characteristics of excluded studies](#).

Eighteen randomised controlled trials were identified by the searches. However, four (Gunn 1989; McGough 1990; Younes 1988a; Younes 1988b) did not provide data on the outcomes specified in the review. Details of these studies are also reported in the table of included studies for completeness.

In the 14 trials reported in the meta-analysis, patients with burns were included in three (n=72) (Bortolani 1996; Caldwell 1979; Jelenko 1978), patients undergoing surgery in five (n= 230) (Croft 1992; Cross 1989; Jarvela 2002; Shackford 1983; Shackford 1987) and trauma patients in six (n= 654) (Cooper 2004; Simma 1998; Vassar 1990; Vassar 1993a; Vassar 1993b; Younes 1992).

Eleven trials compared hypertonic saline versus Ringer's lactate (Bortolani 1996; Caldwell 1979; Cooper 2004; Croft 1992; Jelenko 1978; Shackford 1983; Shackford 1987; Simma 1998; Vassar 1990; Vassar 1993a; Vassar 1993b), and the rest compared hypertonic saline with normal saline.

For more detailed descriptions of individual studies, please see the table of included studies. No additional studies were identified for this latest update.

Risk of bias in included studies

Allocation concealment was judged to be adequate in five trials (Cooper 2004; Simma 1998; Vassar 1990; Vassar 1993a; Vassar 1993b), inadequate in three (Caldwell 1979; Shackford 1983; Shackford 1987), and unclear in the rest. Five trials reported the use of identical bags or containers for the fluids, so that staff were blinded to the identify of the solutions (Cooper 2004; Cross 1989; Vassar 1990; Vassar 1993a; Vassar 1993b).

Effects of interventions

Death was reported either in the paper, or the information was obtained by contacting the researcher, in 14 studies. Data on death were not obtained for four trials (Gunn 1989; McGough 1990; Younes 1988a; Younes 1988b). Data on disability was obtained from the author of one trial (Cooper 2004).

Due to the clinical heterogeneity of the different patient groups it was felt to be inappropriate to pool them; therefore, only the results for the subgroups are given. The pooled relative risk for death in trauma patients was 0.84 (95% CI 0.69 to 1.04), for patients with burns 1.49 (95% CI 0.56 to 3.95) and for patients undergoing surgery 0.51 (95% CI 0.09 to 2.73). Only one trial gave data on disability (Cooper 2004) and the relative risk for a poor outcome was 1.00 (95% CI 0.82 to 1.22).

DISCUSSION

This review does not give us enough data to be able to say whether hypertonic crystalloid is better than isotonic crystalloid for the resuscitation of patients with trauma or burns, or those undergoing surgery. However, the confidence intervals are wide and do not exclude clinically significant differences between hypertonic and isotonic crystalloid. A previous review (Perel 2007) found there was a trend towards a favourable effect on mortality for colloids in hypertonic crystalloid compared to isotonic crystalloids. However, those results are compatible with the play of chance.

We chose not to pool the results of the burns, surgery and trauma patients, as we felt these groups were too clinically heterogeneous. Bleeding and fluid management in patients undergoing elective surgery would tend to be more controlled and, therefore, different to that in trauma patients.

Most of the trials are small and quality was judged to be adequate in only five of them. There was variation in the type of participants, and length of follow-up, and little standardisation in terms of fluid regimes. Also some of the trials were old. Although older trials will not necessarily be of poorer quality, it may be that treatment protocols have subsequently altered, making these trials less relevant to current clinical practice. Indeed in the 1970s and 1980s there were few protocols on fluid resuscitation in the critically ill.

Since the late 1980s, there have been more clear guidelines and standardisation of fluid resuscitation regimes, although many areas of contention still exist.

Mortality was selected as the main outcome measure in this systematic review for several reasons. In the context of critical illness, death or survival is a clinically relevant outcome that is of immediate importance to patients, and data on death are reported in many of the studies. Furthermore, one might expect that mortality data would be less prone to measurement error or biased reporting than would data on pathophysiological outcomes. The use of a pathophysiological end-point as a surrogate for an adverse outcome assumes a direct relationship between the two, an assumption that may sometimes be inappropriate. Finally, when trials collect data on a number of physiological end-points, there is the potential for bias, due to the selective publication of end-points showing striking treatment effects.

Hypertonic solutions have been proposed as the fluid of choice in patients with head injuries (Walsh 1991), as they may maintain cerebral perfusion pressure without causing brain swelling with an increase in intracranial pressure. However, we found only one small trial (Simma 1998) among people with head injuries.

AUTHORS' CONCLUSIONS

Implications for practice

This review does not provide any evidence that hypertonic crystalloid is better than isotonic crystalloid, but it does not rule out clinically important differences.

Implications for research

Further trials are needed comparing hypertonic to isotonic crystalloid. These trials should be multi-centre prospective randomised controlled trials, that are large enough to detect a clinically important difference. Clinically relevant outcomes such as mortality should be used and trials should specify the type and amount of fluid used.

ACKNOWLEDGEMENTS

Thanks to Reinhard Wentz for help with the searches and to Phil Alderson for overseeing the editorial process. Thanks also to E Akpa for assistance during the review process.

REFERENCES

References to studies included in this review

Bortolani 1996 *{published data only}*

Bortolani A, Governa M, Barisoni D. Fluid replacement in burned patients. *Acta Chirurgiae Plasticae* 1996;**38**(4):132–6.

Caldwell 1979 *{published data only}*

Caldwell FT, Bowser BH. Critical Evaluation of Hypertonic Solutions to Resuscitate Severely Burned Children. *Annals of Surgery* 1979;**18**(5):546–52.

Cooper 2004 *{published data only}*

Cooper J, Myles P, McDermott F, Murray L, Laidlaw J, Cooper G, et al. Prehospital hypertonic saline resuscitation of patients with hypotension and severe traumatic brain injury. *Journal of the American Medical Association* 2004;**291**(11):1350–7.

Croft 1992 *{published data only}*

Croft D, Dion YM, Dumont M, Langlois D. Cardiac compliance and effects of hypertonic saline. *Canadian Association of General Surgeons* 1992;**35**(2):139–44.

Cross 1989 *{published data only}*

Cross JS, Gruber DP, Burchard KW, Singh AK, Moran JM, Gann DS. Hypertonic saline fluid therapy following surgery: a prospective study. *Journal of Trauma* 1989;**29**(6):817–26.

Gunn 1989 *{published data only}*

Gunn ML, Hansbrough JF, Davis JW, Furst SR, Field TO. Prospective, randomized trial of hypertonic sodium lactate versus Ringer's solution for burn shock resuscitation. *Journal of Trauma* 1989;**29**(9):1261–7.

Jarvela 2002 *{published data only}*

Jarvela K, Kaukinen S. Hypertonic saline (7.5%) after coronary artery bypass grafting. *European Journal of Anaesthesiology* 2001;**18**(2):100–7.

Jarvela K, Kaukinen S. Hypertonic saline (7.5%) decreases perioperative weight gain following cardiac surgery. *Journal of Cardiothoracic and Vascular Anesthesia* 2002;**16**(1):43–6.

* Jarvela K, Koskinen M, Kaukinen S, Koobi T. Effects of hypertonic saline (7.5%) on extracellular fluid volumes compared with normal saline (0.9%) and 6% hydroxyethyl starch after aortocoronary bypass graft surgery. *Journal of Cardiothoracic and Vascular Anesthesia* 2001;**15**(2):210–15.

Jelenko 1978 *{published data only}*

Jelenko C, Wheeler ML, Calloway D, Divilio T, Bucklen KR, Holdredge TD. Shock and resuscitation II: volume repletion with minimal edema using the "HALFD" method. *JACEP* 1978;**7**(9):326–33.

McGough 1990 *{published data only}*

McGough EK, Kirby RR. Hypertonic saline for intraoperative fluid therapy. *Critical Care Medicine* 1990;**18**(4):5193.

Shackford 1983 *{published data only}*

Shackford SR, Sise MJ, Fridlund PH, Rowley WR, Peters RM, Virgilio RW, Brimm JE. Hypertonic sodium lactate versus lactated ringers solution for intravenous fluid therapy in operations on the abdominal aorta. *Surgery* 1983;**94**(1):41–51.

Shackford 1987 *{published data only}*

Shackford SR, Fortlage DA, Peters RM, Hollingsworth-Fridlund,

Sise MJ. Serum osmolar and electrolyte changes associated with large infusions of hypertonic sodium lactate for intravascular volume expansion of patients undergoing aortic reconstruction. *Surgery, Gynecology & Obstetrics* 1987;**164**(2):127–136.

Simma 1998 {published data only}

Simma B, Burga R, Falk M, Sacher P, Fanconi S. A prospective, randomized, and controlled study of fluid management in children with severe head injury: lactated ringer's solution versus hypertonic saline. *Critical Care Medicine* 1998;**26**(7):1265–70.

Simma B, Burger R, Falk M, Sacher P, Torresani T, Fanconi S. The release of antidiuretic hormone is appropriate in response to hypovolemia and/or sodium administration in children with severe head injury: A trial of lactated ringer's solution versus hypertonic saline. *Anesthesia and Analgesia* 2001;**92**(3):641–5.

Vassar 1990 {published data only}

Vassar MJ, Perry CA, Holcroft JW. Analysis of potential risks associated with 7.5% sodium chloride resuscitation of traumatic shock. *Archives of Surgery* 1990;**125**(10):1309–15.

Vassar 1993a {published data only}

Vassar MJ, Perry CA, Holcroft JW. Prehospital resuscitation of hypotensive trauma patients with 7.5% NaCl versus 7.5% NaCl with added dextran: a controlled trial. *The Journal of Trauma* 1993;**34**(5):622–633.

Vassar 1993b {published data only}

Vassar MJ, Fischer RP, O'Brien PE, Bachulis BL, Chambers JA, Hoyt DB, Holcroft JW. A multicenter trial for resuscitation of injured patients with 7.5% sodium chloride. *Arch Surg* 1993;**128**:1003–1013.

Younes 1988a {published data only}

Younes RN, Aun F, Birolini D, Kawahara NT, Takeuti M, Casale L, Accioly C, Szajn bok I, Brito N. Initial treatment of hypovolemic patients with 7.5% solution of NaCl [O Tratamento inicial de pacientes hipovolemicos: Emprego da solucao hipertonica de NaCl a 7.5%]. *Rev. Hosp. Clin. Fac. Med. S. Paulo* 1988;**43**(3):138–141.

Younes 1988b {published data only}

Younes R, Bechara M, Langer B, Aun F, Birolini D, Kuznieck S, Guimaraes P. Hypertonic saline 7.5% in the prevention of hypotension in post abdominal aorta unclamping [Emprego da solucao hipertonica de NaCl a 7.5% na prevencao da hipotensao pos-desclampamento da aorta abdominal]. *Rev Ass Med Brasil* 1988;**34**(5):150–154.

Younes 1992 {published data only}

Younes RN, Aun F, Accioly CQ, Casale LPL, Szajn bok I, Birolini. Hypertonic solutions in the treatment of hypovolemic shock: A prospective, randomized study in patients admitted to the emergency room. *Surgery* 1992;**111**(4):380–5.

References to studies excluded from this review

Fisher 1992 {published data only}

Fisher B, Thomas D, Peterson B. Hypertonic saline lowers raised intracranial pressure in children after head trauma. *Journal of Neurosurgical Anesthesiology* 1992;**4**(1):4–10.

Holcroft 1987 {published data only}

Holcroft J, Vassar M, Turner J, Derlet R, Kramer G. 3% NaCl and 7.5% NaCl/Dextran 70 in the resuscitation of severely injured patients. *Annals of Surgery* 1987;**206**(3):279–87.

Shackford 1998 {published data only}

Shackford SR, Bourguignon PR, Wald SL, Rogers FB, Osler TM, Clark DE. Hypertonic saline resuscitation of patients with head injury: a prospective, randomized clinical trial. *Journal of Trauma* 1998;**44**(1):50–8.

Shao 2005 {published data only}

Shao Y, Zhang Y, Peng K, Quan Z. Effects of 7.5% hypertonic saline on fluid balance after radical surgery for gastrointestinal carcinoma. *World J Gastroenterol* 2005;**11**(11):1577–1581.

Additional references

Berlin 1997

Berlin JA for the University of Pennsylvania Meta-analysis Blinding Study Group. Does blinding of readers affect the results of meta-analyses?. *Lancet* 1997;**350**(9072):185–6.

Jennett 1975

Jennett B, Bond M. Assessment of outcome after severe brain damage. A practical scale. *Lancet* 1975;**1**(7905):480–4.

Khanna 2000

Khanna S, Davis D, Peterson B, Fisher B, Tung H, O'Quigley J, Deutsch R. Use of hypertonic saline in the treatment of severe refractory posttraumatic intracranial hypertension in pediatric traumatic brain injury. *Critical Care Medicine* 2000;**28**(4):1144–51.

Krausz 1995

Krausz MM. Controversies in shock research: hypotonic resuscitation - pros and cons. *Shock* 1995;**3**(1):69–72.

Perel 2007

Perel P, Roberts I. Colloids versus crystalloids for fluid resuscitation in critically ill patients (Cochrane Review). *Cochrane Database of Systematic Reviews* 2007, Issue 4. [Art. No.: CD000567. DOI: 10.1002/14651858.CD000567.pub3]

Peterson 2000

Peterson B, Khanna S, Fisher B, Marshall L. Prolonged hypernatremia controls elevated intracranial pressure in head-injured pediatric patients. *Critical Care Medicine* 2000;**28**(4):1136–43.

Schulz 1995

Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**(5):408–12.

Shenkin 1976

Shenkin HA, Bezier HS, Bouzarth WF. Restricted fluid intake: rational management of the neurosurgical patient. *Journal of Neurosurgery* 1976;**45**(4):432–6.

Walsh 1991

Walsh JC, Zhuang J, Shackford SR. A comparison of hypertonic to isotonic fluid in the resuscitation of brain injury and hemorrhagic shock. *Journal of Surgical Research* 1991;**50**(3):284–92.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Bortolani 1996

Methods	Randomised controlled trial. Method of allocation concealment is not described. No mention of blinding.
Participants	40 patients, with burns over 30% of the body surface area admitted within 4 hours of trauma. Country: Italy
Interventions	1. Hypertonic lactated saline (n=20) 2. Ringer's lactated saline (n=20)
Outcomes	Haemodynamic variables. Death. Complications.
Notes	Length of follow-up not clear.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Caldwell 1979

Methods	Treatments were alternated. No mention of blinding.
Participants	20 children with thermal burns covering 30% or more of the body surface area. Country: USA
Interventions	1. Hypertonic lactated Ringer's (n=17). 2. Lactated Ringer's (n=20). i.v. treatment discontinued after 48 hrs.
Outcomes	Death. Haemodynamic variables.
Notes	Length of follow-up not clear. No loss to follow-up reported.

Risk of bias

Item	Authors' judgement	Description
------	--------------------	-------------

Caldwell 1979 (Continued)

Allocation concealment?	No	C - Inadequate
-------------------------	----	----------------

Cooper 2004

Methods	Double blind randomised controlled trial. Identical bags of sequentially numbered, computer randomised fluid were packed in groups of 4 in each ambulance.
Participants	229 patients with traumatic brain injury. Country: Australia.
Interventions	1. 250 ml of Hypertonic saline (7.5%). 2. 250 ml of ringers lactate.
Outcomes	Death. Disability (glasgow outcome scale). Cognitive score. Functional independence score.
Notes	Follow up 6 months. One patient in control group declined to participate and two in intervention were lost to follow-up.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Croft 1992

Methods	Randomised controlled trial. Patients were divided consecutively, with a random allocation chart. Method of allocation concealment was not described. No mention of blinding.
Participants	28 patients undergoing major intra-abdominal surgery. Patients with an abnormality of a cardiac valve, liver failure, pacemaker, shock, septicemia or presence of myocardial ischemia less than 24 hours before the study were excluded. Country: Canada.
Interventions	1. Hypertonic saline (n=13). 2. Isotonic Ringer's lactate (n=15). Preoperatively RL or HS were infused at a rate sufficient to maintain a PAWP and a CVP within 3 mm Hg of the initial value.
Outcomes	Haemodynamic variables. Death.

Croft 1992 (Continued)

Notes	Follow up 72 hours.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Cross 1989

Methods	Randomised double-blind study. Method of allocation concealment not described. Doctors and nurses directly involved in pt care did not know identity of solutions.	
Participants	20 post-op coronary artery bypass patients. Patients with history of significant arrhythmias, congestive heart failure, renal, hepatic, or pulmonary failure were excluded. Country: USA	
Interventions	<ol style="list-style-type: none"> 1. Hypertonic saline (n=11). 2. Normal saline (n=9). for 24 hr period following arrival at ITU. Study solutions were initially infused at 100 ml/hr, subsequent rates were adjusted according to the clinical status and were infused to maintain hemodynamic stability.	
Outcomes	Death. Hemodynamic variables.	
Notes	Follow-up 24 hours.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Gunn 1989

Methods	Randomised controlled trial. Method of allocation concealment not described.	
Participants	51 adult patients who sustained at least 20% body surface area burns and who were admitted within 12 hours of injury.	
Interventions	<ol style="list-style-type: none"> 1. Hypertonic sodium lactate 2. Ringer's lactate. Intravenous fluid was administered to maintain the urine output at a target rate of 0.5-1.0 cc/kg/hour, and maintain a minimal or zero base deficit in serial blood gas analyses.	

Gunn 1989 (Continued)

Outcomes	Hemodynamic variables. Enteral intake.	
Notes	Follow-up was for 72 hours. High drop-out rate due to need for surgery, excision and grafting. These patients were not followed up after surgery.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	D - Not used

Jarvela 2002

Methods	Randomised controlled trial. Patients were randomly allocated according to a list of random digits to 2 groups. Blinding not specified.	
Participants	72 patients undergoing elective coronary artery bypass graft surgery. Patients were excluded if they had a left ventricular ejection fraction less than 0.4, a serum creatinine more than 130 umol/L, hepatic or renal disease, or continuous medication with diuretics. Country: Finland.	
Interventions	<ol style="list-style-type: none"> 1. Hypertonic saline (7.5%) (n=36). 2. Normal saline (0.9%) (n=36). Both groups received 4ml/kg during 30 minutes, when volume loading was needed during the postoperative warming period in ICU. The infusion was stopped if systemic arterial pressure exceeded 170 mmHg.	
Outcomes	Hemodynamic variables. MAP. Cardiac index.	
Notes	Follow-up first post-op morning.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Jelenko 1978

Methods	Randomised controlled trial, method of allocation concealment not described. Blinding not mentioned. No loss to follow up.	
Participants	12 patients with burns covering more than 20% of body surface. Country: USA.	
Interventions	1. Hypertonic saline (240MeQ Na+, 120 MeQ Chloride, 120 MeQ lactate) (n=5). 2. Ringer's lactate (n=7). Allocated fluid was used, guided by haemodynamic variables, to the end of resuscitation.	
Outcomes	Death. Hemodynamic variables.	
Notes	Follow-up to the end of resuscitation.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

McGough 1990

Methods	Randomised controlled trial. Method of allocation concealment not described.	
Participants	50 patients undergoing total hip arthroplasty, hysterectomy, or radical prostatectomy.	
Interventions	1. Hypertonic saline at 4 ml/kg/hr. 2. Ringers lactate at 8 ml/kg/hr.	
Outcomes	Complications.	
Notes		
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Shackford 1983

Methods	Patients were assigned by random number to one of two groups. Allocation was done by list of random numbers read by someone entering the patient into the trial (open list). No mention of blinding.
Participants	58 patients undergoing aortic reconstruction. Country: USA.
Interventions	1. Group one received a hypertonic solution (HSL) (n=30). 2. Group two received ringers lactate (n=28). Fluid was given to maintain the cardiac filling pressure within 3 torr of the preoperative level and the cardiac output at or above the preoperative level. All pts received 5% dextrose in 0.25N saline as a maintenance solution.
Outcomes	Hemodynamic variables. Death.
Notes	Follow-up three days.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	No	C - Inadequate

Shackford 1987

Methods	Patients were assigned by random number to one of two groups. Allocation was done by list of random numbers read by someone entering the patient into the trial (open list). No mention of blinding.
Participants	52 patients undergoing aortic reconstruction. Country: USA
Interventions	1. Hypertonic lactated saline (n=26). 2. Ringer's lactate (n=26). During and immediately after the operation fluid was given to maintain the CO equal to preoperative levels and the cardiac filling pressures within 3 torr of the preoperative value. Post-op all of the patients received 5% dextrose in normal saline as a maintenance solution, this was continued until the first day post-op. During this same period, additional fluid (either HSL or RL) was given to maintain cardiac filling pressures and CO at pre-op levels.
Outcomes	Hemodynamic variables. Serum compositional changes.
Notes	Follow-up three days.

Risk of bias

Shackford 1987 (Continued)

Item	Authors' judgement	Description
Allocation concealment?	No	C - Inadequate

Simma 1998

Methods	Randomized controlled trial. Randomization was done by an independent investigator. Staff were not blinded to the type of fluid.	
Participants	32 head-injured children under the age of 16 with Glasgow coma scores of <8. The patients entered the study at the time when ICP was first measured. Country: Switzerland.	
Interventions	<ol style="list-style-type: none"> 1. Hypertonic saline (n=15). 2. Ringer's lactate (n=17). Over 72 hours.	
Outcomes	Hemodynamic variables. Complications. Length of hospital stay. Death.	
Notes	Follow-up until discharge from hospital.	

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Vassar 1990

Methods	Randomised controlled trial, allocation concealment unclear. Double blind study (solutions prepared in identical containers). No loss to follow up.	
Participants	59 injured patients were entered into the trial. Participants were emergency department admissions with trauma and a systolic blood pressure below 80mm Hg and were 18 years or older. Pregnant women and people with preexisting cardiac, hepatic or renal disease were excluded. Country: USA	
Interventions	<ol style="list-style-type: none"> 1. 7.5% saline (n=32). 2. Ringer's lactate (n=27). Allocated fluids were given as the initial resuscitation fluid in the emergency department.	
Outcomes	Haemodynamic variables. Death.	

Vassar 1990 (Continued)

Notes	Follow-up until discharge from hospital.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Vassar 1993a

Methods	Randomised controlled double blind trial. Allocation concealed by random sequence of identical containers. 36 people excluded post randomisation as deemed not to have met eligibility criteria. No loss to follow up.	
Participants	169 pre-hospital trauma patients, who were undergoing ambulance transport to an emergency centre, had systolic blood pressure 90 mmHg or less, and were 18 years or older. Exclusions: asystolic, undergoing CPR, lack sinus complex on ECG, more than 2 hours after trauma, pregnant, preexisting seizures, bleeding disorder, hepatic, cardiac or renal disease. Country: USA	
Interventions	<ol style="list-style-type: none"> 1. 7.5% saline (n=85). 2. 0.9% saline (n=84). Participants received 250mL of the allocated fluid in the pre-hospital setting. Additional isotonic crystalloids were given as needed.	
Outcomes	Deaths reported. Haemodynamic variables. Trauma scores and neurological outcome scores.	
Notes	Follow-up until discharge from hospital.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Vassar 1993b

Methods	Randomised controlled trial, allocation concealed by sequential use of coded identical containers. Only the manufacturer could know the treatment assignment. Double blind study. 39/233 patients excluded as deemed not to meet eligibility criteria, unclear from which groups.	
---------	---	--

Vassar 1993b (Continued)

Participants	95 pre-hospital trauma patients undergoing helicopter transport to an emergency centre, had a systolic blood pressure of 100mmHg or less and were 18 years or older. Exclusions: asystolic, undergoing CPR, lack sinus complex on ECG, more than 2 hours after trauma, pregnant, preexisting seizures, bleeding disorder, hepatic, cardiac or renal disease. Country: USA	
Interventions	1. 7.5% saline. (n=50) 2. Ringer's lactate. (n=45) Participants received 250mL of the allocated fluid in the pre-hospital setting. Additional isotonic crystalloids were given as needed.	
Outcomes	Deaths reported. Haemodynamic variables. Trauma scores and neurological outcome scores.	
Notes	Follow-up until discharge from hospital.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Younes 1988a

Methods	Random assignment. Method of allocation concealment not mentioned.	
Participants	33 patients admitted to the emergency ward in hypovolemic shock (mean arterial pressure < 60 mmHg)	
Interventions	1. Hypertonic 7.5% saline (n=18). 2. Isotonic NaCl (n=15). Both fluids received at infusion rate of 10ml/minute, over 15 minutes. No other fluid was given after the infusion unless MAP fell below 80mmHg, until typed-crossmatched blood was available. Patients were excluded from the study as soon as given fluid or blood.	
Outcomes	MAP	
Notes	Length of follow-up not recorded.	
Risk of bias		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Younes 1988b

Methods	Random assignment. Method of allocation concealment not mentioned.	
Participants	31 patients admitted for abdominal aorta reconstructive surgery.	
Interventions	<ol style="list-style-type: none"> 1. Hypertonic 7.5% NaCl, (n=18) 2. Isotonic 0.9% saline (n=13) Both groups received fluid as the volume of 4ml/kg of body weight, infused during 15 minutes. The infusion was started 2 minutes before the release of the aortic clamp.	
Outcomes	MAP. Haemodynamic variables.	
Notes	Length of follow up not recorded.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Younes 1992

Methods	Randomised 'in a double blind fashion'. Blinding by use of similar bottles. No loss to follow up.	
Participants	70 emergency department admissions, who had a systolic blood pressure of less than 80mm Hg and were 19 years and older. Exclusions: pregnant, preexisting cardiac or metabolic disease.	
Interventions	<ol style="list-style-type: none"> 1. 7.5% saline (n=35). 2. 0.9% saline (n=35). Allocated fluid was for initial bolus of 250mL, followed by isotonic crystalloids as needed.	
Outcomes	Deaths reported. Fluid balance.	
Notes	Follow-up until discharge from hospital.	
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

MAP = mean arterial pressure.

ICU = Intensive care unit.

PCWP = Pulmonary capillary wedge pressure.

Characteristics of excluded studies *[ordered by study ID]*

Fisher 1992	Cross-over study.
Holcroft 1987	The study was not randomised. Fluid was administered depending on the attending surgeon.
Shackford 1998	Study compared hypertonic fluid versus hypotonic.
Shao 2005	The patients were 'assigned' and not randomised.

DATA AND ANALYSES

Comparison 1. Hypertonic versus isotonic crystalloid

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Death	14		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.1 Trauma	6	651	Risk Ratio (M-H, Fixed, 95% CI)	0.84 [0.69, 1.04]
1.2 Burns	3	89	Risk Ratio (M-H, Fixed, 95% CI)	1.49 [0.56, 3.95]
1.3 Surgery	5	230	Risk Ratio (M-H, Fixed, 95% CI)	0.51 [0.09, 2.73]
2 Poor outcome (GOS)	1	226	Risk Ratio (M-H, Fixed, 95% CI)	1.0 [0.82, 1.22]
2.1 Trauma	1	226	Risk Ratio (M-H, Fixed, 95% CI)	1.0 [0.82, 1.22]

FEEDBACK

Lactated Ringer's not isotonic

Summary

Lactated Ringer's solution is not a truly isotonic fluid. In one report (Tommasino C, Moore S, Todd MM. Crit Care Med 1988;16 p867) the measured osmolality was stated to be approximately 254 mosm/l while the calculated osmolality was 273 mosm/l.

The treatment of traumatized patients will include the infusion of many liters of crystalloids during the first hours. In comparison, the 250 ml of lactated Ringer's or saline used at intervention in the studies concerned probably does not matter very much. The problem addressed by the review, rather than one of "isotonic versus hypotonic", may be more precisely formulated as something like "early supplementation or not" of hypertonic fluid to the continued use of many liters of a weakly hypotonic fluid.

The hypothesis that 250 ml of hypertonic fluid is beneficial, may easily lead to the idea that many litres of a hypotonic fluid is detrimental. Or is the hypertonic fluid of benefit only when it is added to adjust for the hypotonic one? Will a test with really isotonic crystalloid do better than the hypotonic one and show the supplementation with hypertonic fluid not only to be unnecessary, but even harmful?

The reports often conceal the true nature of the fluids used behind designations like "conventional isotonic solutions" or "standard of care", and the amount of fluid given after arrival in hospital may not be stated.

Implications for research is that studies with the continued use of truly isotonic solutions have to be done to decide whether hypertonic or weakly hypotonic solutions are beneficial or detrimental. The nature and amount of fluids used in future studies should be clearly stated.

I certify that I have no affiliations with or involvement in any organisation or entity with a direct financial interest in the subject matter of my criticisms.

Reply

We agree that lactated Ringer's is not a truly isotonic fluid and have, therefore, changed the title of the review to reflect this. The title is now 'Hypertonic versus near isotonic crystalloid for fluid resuscitation in critically ill patients'.

We also agree that the nature and amount of fluids used in future studies should be clearly stated, and have included a statement to this effect in the conclusions.

Contributors

Comment by Per Størset (anesthesiologist), December 2002.

Reply from Frances Bunn, May 2004.

WHAT'S NEW

Last assessed as up-to-date: 14 October 2007.

10 July 2008	Amended	Converted to new review format.
--------------	---------	---------------------------------

HISTORY

Protocol first published: Issue 2, 2000

Review first published: Issue 4, 2000

20 February 2008	New search has been performed	This review was originally published in the Cochrane Library in 2000. It has subsequently been updated in 2001, 2004, and, most recently, 2008. No new studies were added in the most recent update.
------------------	-------------------------------	--

CONTRIBUTIONS OF AUTHORS

FB screened citations for eligibility, obtained references, contacted authors, extracted data, entered data and wrote up the review. IR helped to write the review. RT commented on the protocol and review. DT screened citations for eligibility.

DECLARATIONS OF INTEREST

None known.

SOURCES OF SUPPORT

Internal sources

- University of Hertfordshire, UK.

External sources

- NHS Research and Development, UK.

INDEX TERMS

Medical Subject Headings (MeSH)

*Plasma Substitutes; Critical Illness; Hypertonic Solutions [*therapeutic use]; Hypovolemia [mortality; *therapy]; Isotonic Solutions [*therapeutic use]; Randomized Controlled Trials as Topic; Rehydration Solutions [*therapeutic use]

MeSH check words

Humans

Human albumin administration in critically ill patients: systematic review of randomised controlled trials

Cochrane Injuries Group Albumin Reviewers

Abstract

Objective: To quantify effect on mortality of administering human albumin or plasma protein fraction during management of critically ill patients. **Design:** Systematic review of randomised controlled trials comparing administration of albumin or plasma protein fraction with no administration or with administration of crystalloid solution in critically ill patients with hypovolaemia, burns, or hypoalbuminaemia.

Subjects: 30 randomised controlled trials including 1419 randomised patients.

Main outcome measure: Mortality from all causes at end of follow up for each trial.

Results: For each patient category the risk of death in the albumin treated group was higher than in the comparison group. For hypovolaemia the relative risk of death after albumin administration was 1.46 (95% confidence interval 0.97 to 2.22), for burns the relative risk was 2.40 (1.11 to 5.19), and for hypoalbuminaemia it was 1.69 (1.07 to 2.67). Pooled relative risk of death with albumin administration was 1.68 (1.26 to 2.23). Pooled difference in the risk of death with albumin was 6% (95% confidence interval 3% to 9%) with a fixed effects model. These data suggest that for every 17 critically ill patients treated with albumin there is one additional death.

Conclusions: There is no evidence that albumin administration reduces mortality in critically ill patients with hypovolaemia, burns, or hypoalbuminaemia and a strong suggestion that it may increase mortality. These data suggest that use of human albumin in critically ill patients should be urgently reviewed and that it should not be used outside the context of rigorously conducted, randomised controlled trials.

Introduction

In patients with acute and chronic illness serum albumin concentration is inversely related to risk of death. A systematic review of cohort studies meeting specified criteria estimated that for each 2.5 g/l decrement in serum albumin concentration the risk of death increases by between 24% and 56%.¹ The association persists after adjustment for other known risk factors and pre-existing illness, and some commentators have suggested the possibility of the albumin molecule hav-

ing a direct protective effect.¹ Partly as a result of the association between serum albumin and mortality, human albumin solutions are now used in the management of a diverse range of medical and surgical problems. Licensed indications for human albumin solution are the emergency treatment of shock and other conditions in which restoration of blood volume is urgent, the acute management of burns, and clinical situations associated with hypoproteinaemia.²

Compared with other colloidal solutions and with crystalloid solutions, human albumin solutions are expensive.³ Volume for volume, human albumin solution is twice as expensive as hydroxyethyl starch and over 30 times more expensive than crystalloid solutions such as sodium chloride or Ringer's lactate. Because of the high cost and limited availability of human albumin, it is imperative that its use should be restricted to the indications for which it has been shown to be effective. To quantify the effect on mortality of human albumin solution in the management of critically ill patients with hypovolaemia from injury or surgery, burns, and hypoproteinaemia, we conducted a systematic review of randomised controlled trials.

Methods

Identification of trials

Our aim was to identify all relevant randomised controlled trials that were available for review by March 1998. A randomised controlled trial was defined as a trial in which the subjects followed were assigned prospectively to one of two (or more) interventions by random allocation or some quasi-random method of allocation. This definition was agreed at an international meeting held in Oxford in November 1992 in association with the official opening of the UK Cochrane Centre. We sought to identify all randomised controlled trials of administration of human albumin or plasma protein fraction (supplemental albumin or plasma protein fraction compared with no albumin or plasma protein fraction or with a crystalloid solution) in critically ill patients with hypovolaemia from trauma or surgery, with burns, or with hypoalbuminaemia. Studies that compared different levels of albumin supplementation were also included.

Editorial
by Ofringa and
Letters p 277

Cochrane Injuries
Group, Department
of Epidemiology
and Public Health,
Institute of Child
Health, London
WC1N 1EH

Ian Roberts,
*director, child health
monitoring unit*

Ian.Roberts@ich.ucl.
ac.uk

BMJ 1998;317:235-40

Trials were identified by computerised searches of the Cochrane Controlled Trials Register, Medline, Embase, and BIDS Index to Scientific and Technical Proceedings (search strategies are available from IR); by hand searching 29 international journals and the proceedings of several international meetings on fluid resuscitation; by checking the reference lists of all included trials; and by contacting the authors of identified trials and asking them about any other published or unpublished trials that may have been conducted. There were no language restrictions. To identify unpublished trials we searched the register of the Medical Editors' Trial Amnesty,⁴ and contacted the

Medical Directors of Bio Products Laboratory (Zenalb), Centeon (Albuminar), and Alpha Therapeutic UK (Albutein).

Outcome measures and data extraction

The outcome measure was mortality from all causes at the end of the follow up period scheduled for each trial. For all trials we collected data on the type of participants, details about the interventions, the quality of concealment of allocation, and mortality at the end of follow up. We rated quality of allocation concealment using the method proposed by Schulz et al.⁵ We sought mortality data in simple categorical form, and we did

Summary of randomised trials comparing albumin with no albumin or crystalloid that met criteria for inclusion

Trial	Critical illness	No of patients	Intervention	Control	Length of follow up	Total No of deaths	Allocation concealment*
Hypovolaemia							
Skillman et al ³¹	Surgery	16	25% concentrated salt-poor albumin 1 g/kg and 5% albumin in saline	Ringer's lactate with 5% dextrose	1 day	Not known	2
Shah et al ²⁷	Trauma	20	5% salt-poor albumin in Ringer's lactate	Ringer's lactate	Unspecified	5	3
Lowe et al ²⁰	Trauma	171	50 g albumin/200 ml Ringer's lactate	Ringer's lactate	5 days	6	3
Boutros et al ⁹	Surgery	24	Albumin in 5% dextrose	5% dextrose in lactated Ringer's (n=9) 5% dextrose in 0.45% NaCl (n=8)	4 days	2	2
Virgilio et al ³³	Surgery	29	5% albumin in Ringer's lactate	Ringer's lactate	2½ weeks	2	2
Lucas et al ²¹	Trauma	52	150 g salt-poor albumin during operation, 150 g/day for 5 days postoperatively	No albumin	To positive fluid balance or oral intake	7	1
Zetterstrom et al ³⁷	Surgery	30	20% albumin 100 ml at end of operation, 200 ml on day of operation, 100 ml/day for next 3 days	No albumin	Unspecified	1	3
Zetterstrom ³⁸	Surgery	18	5% albumin to keep pulmonary arterial occlusion pressure equal to preoperative level	Balanced electrolyte solution of Ringer's type to keep pulmonary arterial pressure equal to preoperative level	Unspecified	2	3
Grundman et al ¹⁷	Surgery	17	Human albumin and crystalloid	Crystalloid only	5 days	1	2
Rackow et al ³⁰	Trauma and sepsis	17	5% albumin	0.9% NaCl	To discharge	12	2
Gallagher et al ¹²	Surgery	10	5% albumin	Ringer's lactate	1 day	0	3
Nielsen et al ²³	Surgery	26	80 g albumin in units of 100 ml 20% albumin on day of operation, 20 g daily for next 3 days	No albumin	4 days	0	2
Prien et al ²⁶	Surgery	12	20% albumin to maintain central venous pressure at preoperative level	Ringer's lactate	Unspecified	0	2
Boldt et al ⁸	Surgery	30	5% albumin	No albumin	1 day	0	3
McNulty et al ²²	Surgery	28	5% albumin	Isotonic crystalloid	Unspecified	Not known	2
Woods et al ³⁶	Surgery	69	Albumin supplementation	No supplementation	To discharge	1	1
Pockaj et al ²⁵	Vascular leak syndrome	107	5% albumin in 0.9% NaCl	0.9% NaCl	Unspecified	0	2
Tølløfsrud et al ³²	Surgery	20	4% albumin when fluid required	Ringer's acetate	48 hours	1	3
So et al ²⁸	Hypotensive preterm infant	63	5% albumin 10 ml/kg over 30 minutes	0.9% NaCl 10 ml/kg over 30 minutes	To discharge	12	3
Woittiez et al ³⁴	Surgery	31	20% albumin	0.9% NaCl	Unspecified	12	3
Burns							
Jelenko et al ¹⁸	Burns	14	Hypertonic crystalloid with 12.5 g/l albumin	Ringer's lactate	5 days	3	2
Goodwin et al ¹⁴	Burns	79	2.5% albumin in Ringer's lactate	Ringer's lactate	To discharge	14	2
Greenhalgh et al ¹⁵	Burns	70	25% albumin to maintain serum levels between 2.5 and 3.5 g/dl	No albumin unless levels dropped below 1.5 g/dl	To discharge	10	3
Hypoproteinaemia							
Bland et al ⁷	Hypoproteinaemia	27	25% albumin 8 ml/kg	5% glucose 8 ml/kg	Unspecified	5	2
Nilsson et al ²⁴	Hypoalbuminaemia (postoperative)	59	20-25 g albumin/day for 3 days starting day after operation	No supplemental albumin	To discharge	1	3
Brown et al ¹⁰	Hypoalbuminaemia	67	TPN with added albumin	No supplemental albumin	To discharge	10	1
Foley et al ¹¹	Hypoalbuminaemia	40	TPN with added albumin (25-50 g/day 25% albumin)	No supplemental albumin	To discharge	13	1
Kanarek et al ¹⁹	Hypoalbuminaemia	24	TPN with added albumin	No supplemental albumin	Unspecified	5	3
Wojtylak et al ³⁵	Hypoalbuminaemia	30	TPN with added albumin	No supplemental albumin	5 days	0	1
Greenough et al ¹⁶	Hypoalbuminaemic sick preterm infants	40	20% salt-poor albumin 5 ml/kg with maintenance fluids	5 ml/kg maintenance fluid placebo	24 hours after infusion	10	3
Golub et al ¹³	Hypoalbuminaemia	219	37.5 g/day albumin until serum albumin >3.0 g/dl	No supplemental albumin	To discharge	18	3
Rubin et al ²⁹	Hypoalbuminaemia	36	TPN with added albumin	No supplemental albumin	To discharge	3	3

TPN=Total parenteral nutrition. *Allocation concealment: 1=inadequate, 2=unclear, 3=adequate.

not extract data on time to death. If a report did not include the numbers of deaths in each group, we sought these data from the authors. Two reviewers independently extracted the data, and any disagreements were resolved by discussion.

Data analysis and statistical methods

We used the Mantel-Haenszel method to calculate relative risks, risk differences, and 95% confidence intervals for death for each trial on an intention to treat basis using RevMan (Review Manager) statistical software. When there are no events in one group the software adds 0.5 to each cell of the 2×2 table. We tested heterogeneity between trials using χ^2 tests, with $P \leq 0.05$ indicating significant heterogeneity. As long as statistical heterogeneity did not exist, we used a fixed effects model to calculate summary relative risks and 95% confidence intervals.

To examine the extent to which the results of the meta-analyses may have been biased as a result of the selective inclusion of randomised trials with positive findings (publication and other selection bias), we prepared a funnel plot and used the regression approach to assessing funnel plot asymmetry proposed by Egger et al.⁶ We used the log odds ratio in the funnel plot because this is the measure that is used in the regression test of funnel plot asymmetry as described by Egger et al. Using simple unweighted linear regression, we regressed the standard normal deviate (defined as the log odds ratio divided by its standard error) against the estimate's precision (defined as the inverse of the standard error). The larger the deviation of the intercept of the regression line from zero, the greater the asymmetry and the more likely it is that the meta-analysis will yield biased estimates of effect. As suggested by Egger et al, we considered $P < 0.1$ to indicate significant asymmetry.

Results

We identified a total of 32 randomised controlled trials that met the study's inclusion criteria.⁷⁻³⁸ The table shows details of these trials. Mortality data were available either from the published report or on contact with the authors in 30 of these trials. The two trials for which mortality data could not be obtained included a total of 42 randomised patients, comprising 3% of the total number of randomised patients in all trials meeting our inclusion criteria.²²⁻³¹ One of the trials was an unpublished trial registered in the Medical Editors' Trial Amnesty, and we obtained further details, including data on mortality, directly from the trialist. In six trials there were no deaths in either the intervention or comparison groups.⁸⁻¹²⁻²³⁻²⁵⁻²⁶⁻³⁵

The trial by Lucas et al was reported in five publications.²¹⁻³⁹⁻⁴² An early report gave the mortality data for 52 randomised patients, 27 allocated to receive albumin and 25 allocated to receive no albumin.²¹ Subsequent publications indicated that recruitment to the trial continued until 94 patients were randomised. Mortality data for all the 94 patients were not published, nor were they available on contact with the author. Consequently, we present the outcome data for the 52 patients.

Of the 24 trials in which one or more deaths occurred in either the intervention or control groups,

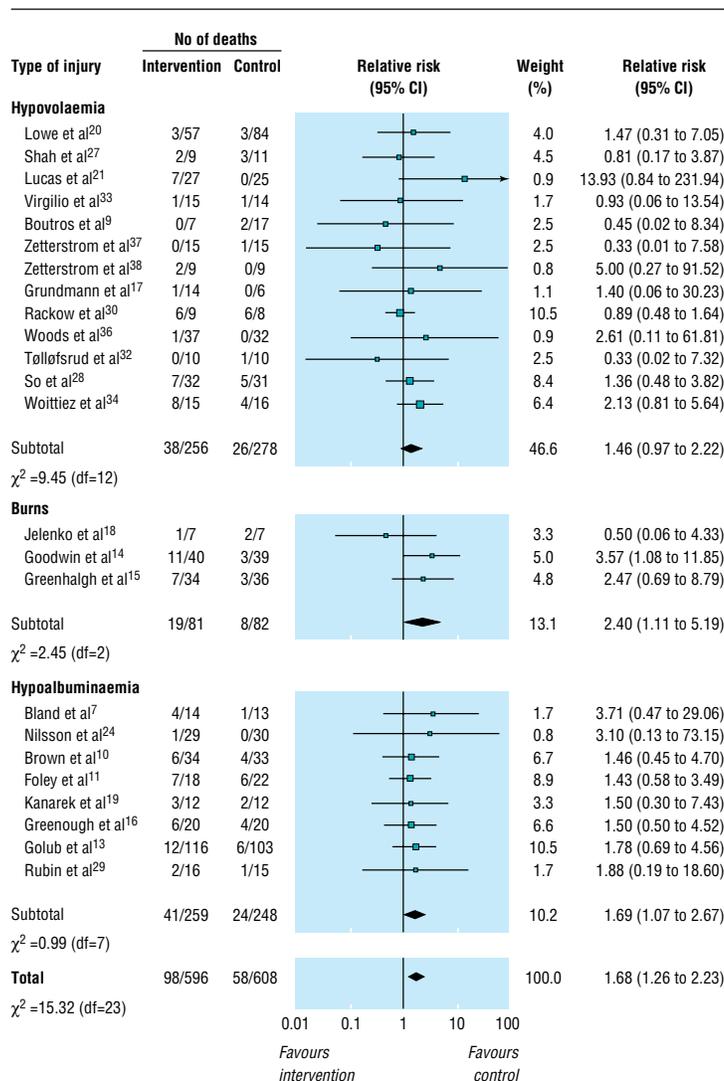


Fig 1 Fixed effects model of relative risks (95% confidence interval) of death associated with intervention (fluid resuscitation with albumin or plasma protein fraction) compared with control (no albumin or plasma protein fraction or resuscitation with a crystalloid solution) in critically ill patients

13 included a method of allocation concealment that would be expected to reduce the risk of foreknowledge of treatment allocation (pharmacy controlled randomisation or serially numbered sealed opaque envelopes). In seven trials this was unclear, and in four trials concealment was inadequate (table).

In each of the patient categories the risk of death in the albumin treated group was higher than in the comparison group (fig 1). For hypovolaemia the relative risk of death after albumin administration was 1.46 (95% confidence interval 0.97 to 2.22), for burns the relative risk was 2.40 (1.11 to 5.19), and for hypoalbuminaemia the it was 1.69 (1.07 to 2.67). There was no significant heterogeneity either between or within the groups of trials, or overall ($\chi^2 = 15.32$, $df = 23$, $P > 0.2$). The pooled relative risk of death with albumin administration was 1.68 (1.26 to 2.23).

There was no significant heterogeneity in the risk difference for mortality ($\chi^2 = 36.69$, $df = 29$, $P > 0.1$). The pooled difference in the risk of death with albumin was 6% (95% confidence interval 3% to 9%).

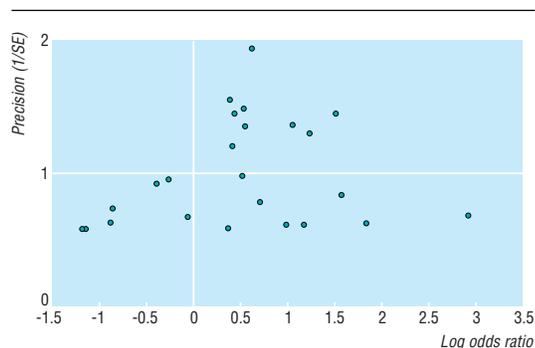


Fig 2 Funnel plot for the 24 trials in which deaths occurred and that were used in systematic review

Figure 2 shows a funnel plot for the 24 trials in which deaths occurred. There was no clear evidence of asymmetry, and the regression approach to funnel plot asymmetry yielded an intercept of -0.39 and $P=0.33$, indicating no statistical evidence of selection bias.

We repeated the analyses for the 13 trials with deaths in which allocation concealment was adequate.^{13 15 16 19 20 24 27-29 32 34 37 38} For hypovolaemia the relative risk of death with albumin administration was 1.39 (0.80 to 2.40), for burns the relative risk was 2.47 (0.69 to 8.79), and for hypoalbuminaemia it was 1.71 (0.92 to 3.18). There was no substantial heterogeneity between the trials in the various categories ($\chi^2=4.42$, $df=12$, $P>0.2$), and the pooled relative risk of death with albumin administration was 1.61 (1.09 to 2.38). Thus, restricting the analyses to the adequately concealed trials had almost no effect on the relative risks in each group or overall.

Discussion

We found no evidence that albumin reduced mortality and a strong suggestion that it might increase the risk of death in patients with hypovolaemia, burns, or hypoproteinaemia. Overall, the risk of death in patients treated with albumin was 6% (95% confidence interval 3% to 9%) higher than in patients not given albumin.

Limitations of study

Mortality was selected as the outcome measure in this systematic review for several reasons. In the context of critical illness, death or survival is a clinically relevant outcome that is of immediate importance to patients, and data on death are reported in nearly all studies. Furthermore, one might expect that mortality data would be less prone to measurement error or biased reporting than would data on pathophysiological outcomes. The use of a pathophysiological end point as a surrogate for an adverse outcome assumes a direct relationship between the two, an assumption that may sometimes be inappropriate. Finally, when trials collect data on a number of physiological end points, there is the potential for bias due to the selective publication of end points showing striking treatment effects. Because we obtained mortality data for all but two of the included trials, the likelihood of bias due to selective publication of trial outcomes is minimal. We examined mortality from all causes because the attribution of cause of death in critically ill patients, many of whom

may have multiorgan failure, can be problematic and may be prone to bias. Length of follow up was not specified in many of the trials, but when these data were available, follow up was for the first week or until hospital discharge.

Although publication bias is a potent threat to the validity of systematic reviews, it is unlikely to have had an important impact in this study. There was no evidence of funnel plot asymmetry on visual inspection, and there was no statistical evidence of asymmetry from linear regression analysis.

In some of the trials included in this review allocation concealment was inadequate or unclear. As a result, it is possible that more severely ill patients were preferentially allocated to albumin treated groups, which could account for the increased mortality in these groups. Nevertheless, when we repeated the analyses for only those trials in which the method of allocation concealment would be expected to reduce the risk of foreknowledge of allocation, the point estimates were almost identical.

Implications of results

To what extent are the results of this review of 30 relatively small randomised trials of albumin administration generalisable to clinical practice? We believe that this is a matter for judgment by the responsible clinician faced with an individual patient.⁴³ However, the advantage of an overview such as ours is that, since it includes many studies, the results are based on a wide range of patients. Because the results were consistent across the studies, they might reasonably be taken to apply to this wide variety of patients.⁴⁵ Moreover, the evidence that we have brought together is, as far as we can ensure, the totality of the available randomised evidence for the use of albumin in hypovolaemia, burns, and hypoalbuminaemia, the indications for which albumin is currently licensed.

Is there a plausible mechanism by which human albumin might increase mortality? Albumin is used in hypovolaemia and hypoalbuminaemia because it is believed to be effective in replacing volume and supporting colloid oncotic pressure.⁴⁴ However, albumin is also believed to have anticoagulant properties, inhibiting platelet aggregation and enhancing the inhibition of factor Xa by antithrombin III.⁴⁴ Such anticoagulant activity might be detrimental in critically ill patients, particularly those with haemorrhagic hypovolaemia. Furthermore, albumin has been shown to distribute across the capillary membrane, a process that is accelerated in critically ill patients.⁴⁵ It has been suggested that increased leakage of albumin into the extravascular spaces might reduce the oncotic pressure difference across the capillary wall, making oedema more likely.⁴⁵

Conclusions

Because this review was based on relatively small trials in which there were only a small number of deaths the results must be interpreted with caution. Nevertheless, we believe that a reasonable conclusion from these results is that the use of human albumin in the management of critically ill patients should be reviewed. A strong argument could be made that human albumin should not be used outside the context of a properly concealed and otherwise rigorously conducted randomised controlled trial with

Key messages

- Human albumin solution has been used in the treatment of critically ill patients for over 50 years
- Currently, the licensed indications for use of albumin are emergency treatment of shock, acute management of burns, and clinical situations associated with hypoproteinaemia
- Our systematic review of randomised controlled trials showed that, for each of these patient categories, the risk of death in the albumin treated group was higher than in the comparison group
- The pooled relative risk of death with albumin was 1.68 (95% confidence interval 1.26 to 2.23) and the pooled difference in the risk of death was 6% (3% to 9%) or six additional deaths for every 100 patients treated
- We consider that use of human albumin solution in critically ill patients should be urgently reviewed

mortality as the end point. Until such data become available, there is also a case for a review of the licensed indications for albumin use.

This review will also be published in the *Cochrane Library*, where it will be regularly updated to take account of new data and comments on this version.

We thank the Intensive Care National Audit and Research Centre in London for help with identifying trials for this review and for their extensive hand searching. We thank A J Woittiez for providing unpublished trial data from the trial that was registered in the Medical Editors' Trial Amnesty. We thank Elizabeth Bryant, information officer at Centeon, and Martin O'Fobve, at Bio Products, for searching their databases for albumin trials. We thank Anne Greenough for re-examining individual patient records in order to provide data on mortality. We thank Iain Chalmers, Jos Kleijnen, Richard Peto, Dave Signorini, and David Yates for their comments on the manuscript.

Contributors (listed alphabetically): Phil Alderson (UK Cochrane Centre) searched The Cochrane Controlled Trials Register for relevant trials, extracted the data from the trials, and commented on the paper. Frances Bunn (Institute of Child Health) searched the Cochrane Injuries Group Specialised Register for relevant trials, obtained copies of relevant papers, wrote to authors for further information on allocation concealment, and commented on the paper. Carol Lefebvre (UK Cochrane Centre) designed the search strategies for The Cochrane Controlled Trials Register and Embase, and searched these two databases for relevant trials. Leah Li (Institute of Child Health) did the funnel plot and the regression test of funnel plot asymmetry. Alain Li Wan Po (Centre for Evidence-Based Pharmacotherapy, University of Nottingham) helped to write the paper. Ian Roberts (Institute of Child Health) designed the protocol, extracted data from the trials, contacted authors for unpublished data, and wrote the paper. Gillian Schierhout proposed the study hypothesis and conducted preliminary searches of Medline, Embase, and BIDS Index to Scientific and Technical Proceedings.

Funding: The infrastructure of the Cochrane Injuries Group is supported by the NHS Research and Development Programme.

Conflict of interest: None.

- 1 Goldwasser P, Feldman J. Association of serum albumin and mortality risk. *J Clin Epidemiol* 1997;50:693-703.
- 2 *ABPI compendium of data sheets and summaries of product characteristics 1998-99*. London: Association of the British Pharmaceutical Industries, 1998.
- 3 McClelland DB. Human albumin solutions. *BMJ* 1990;300:35-7.
- 4 Medical Editors' Trial Amnesty, The Cochrane Controlled Trials Register. In: *The Cochrane Library*. Oxford: Update Software, 1998. Updated quarterly.

- 5 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.
- 6 Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analyses detected by a simple graphical test. *BMJ* 1997;315:629-34.
- 7 Bland RD, Clarke TL, Harden LB. Rapid infusion of sodium bicarbonate and albumin into high-risk premature infants soon after birth: a controlled, prospective trial. *Am J Obstet Gynecol* 1976;124:263-7.
- 8 Boldt J, Knothe C, Zickmann B, Andres P, Dapper F, Hempelmann G. Influence of different intravascular volume therapies on platelet function in patients undergoing cardiopulmonary bypass. *Anesth Analg* 1993;76:1185-90.
- 9 Boutros AR, Ruess R, Olson L, Hoyt JL, Baker WH. Comparison of hemodynamic, pulmonary, and renal effects of use of three types of fluids after major surgical procedures on the abdominal aorta. *Crit Care Med* 1979;7:9-13.
- 10 Brown RO, Bradley JE, Bekemeyer WB, Luther RW. Effect of albumin supplementation during parenteral nutrition on hospital morbidity. *Crit Care Med* 1988;16:1177-82.
- 11 Foley EF, Borlase BC, Dzik WH, Bistran BR, Benotti PN. Albumin supplementation in the critically ill: a prospective randomised trial. *Arch Surg* 1990;125:739-42.
- 12 Gallagher JD, Moore RA, Kerns D, Jose AB, Botros SB, Flicker S, et al. Effects of colloid or crystalloid administration on pulmonary extravascular water in the postoperative period after coronary artery bypass grafting. *Anesth Analg* 1985;64:753-8.
- 13 Golub R, Sorrento JJ Jr, Cantu R Jr, Nierman DM, Moideen A, Stein HD. Efficacy of albumin supplementation in the surgical intensive care unit: a prospective, randomized study. *Crit Care Med* 1994;22:613-9.
- 14 Goodwin CW, Dorethy J, Lam V, Pruitt BA Jr. Randomized trial of efficacy of crystalloid and colloid resuscitation on hemodynamic response and lung water following thermal injury. *Ann Surg* 1983;197:520-31.
- 15 Greenhalgh DG, Housinger TA, Kagan RJ, Kiemann M, James L, Novak S, et al. Maintenance of serum albumin levels in paediatric burn patients: a prospective, randomized trial [with discussion]. *J Trauma* 1995;39:67-73. (Discussion 1995;39:73-4.)
- 16 Greenough A, Emery E, Hird MF, Gamsu HR. Randomised controlled trial of albumin infusion in ill pre-term infants. *Eur J Pediatr* 1993;152:157-9.
- 17 Grundmann R, Meyer H. The significance of colloid osmotic pressure measurement after crystalloid and colloid infusions. *Intensive Care Med* 1982;8:179-86.
- 18 Jelenko C 3rd, Williams JB, Wheeler ML, Callaway BD, Fackler VK, Albers CA, et al. Studies in shock and resuscitation. I: Use of a hypertonic, albumin-containing, fluid demand regimen (HALFD) in resuscitation. *Crit Care Med* 1979;7:157-67.
- 19 Kanarek KS, Williams PR, Blair C. Concurrent administration of albumin with total parenteral nutrition in sick new-born infants. *J Parenter Enteral Nutr* 1992;16:49-53.
- 20 Lowe RJ, Moss GS, Jilek J, Levine HD. Crystalloid vs colloid in the aetiology of pulmonary failure after trauma: a randomized trial in man. *Surgery* 1977;1:676-83.
- 21 Lucas CE, Weaver D, Higgins RF, Ledgerwood AM, Johnson SD, Bouwman DL. Effects of albumin versus non-albumin resuscitation on plasma volume and renal excretory function. *J Trauma* 1978;18:565-70.
- 22 McNulty SE, Sharkey SJ, Asam B, Lee JH. Evaluation of STAT-CRIT hematocrit determination in comparison to coulter and centrifuge: the effects of isotonic hemodilution and albumin administration. *Anesth Analg* 1993;76:830-4.
- 23 Nielsen OM, Engell HC. Extra-cellular fluid volume and distribution in relation to changes in plasma colloid osmotic pressure after major surgery. A randomised study. *Acta Chir Scand* 1985;151:221-5.
- 24 Nilsson E, Lamke O, Liljedahl SO, Elfstrom K. Is albumin therapy worthwhile in surgery for colorectal cancer? *Acta Chir Scand* 1980;146:619-22.
- 25 Pockaj BA, Yang JC, Lotze MT, Lange JR, Spencer WF, Steinberg S, et al. A prospective randomized trial evaluating colloid versus crystalloid resuscitation in the treatment of the vascular leak syndrome associated with interleukin-2 therapy. *J Immunother* 1994;15:22-8.
- 26 Prien T, Backhaus N, Pelster F, Pircher W, Bunte H, Lawin P. Effect of intraoperative fluid administration and colloid osmotic pressure on the formation of intestinal edema during gastrointestinal surgery. *J Clin Anesth* 1990;2:317-23.
- 27 Shah DM, Browner BD, Dutton RE, Newell JC, Powers SR. Cardiac output and pulmonary wedge pressure. *Arch Surg* 1977;112:1161-4.
- 28 So KW, Fok TF, Ng PC, Wong WW, Cheung KL. Randomised controlled trial of colloid or crystalloid in hypotensive pre-term infants. *Arch Dis Child* 1997;76:F43-6.
- 29 Rubin H, Carlson S, DeMeo M, Ganger D, Craig R. Randomized, double-blind study of intravenous human albumin in hypoalbuminaemic patients receiving total parenteral nutrition. *Crit Care Med* 1997;25:249-52.
- 30 Rackow EC, Falk JL, Fein IA, Siegel JS, Packman MI, Haupt MT, et al. Fluid resuscitation in circulatory shock: a comparison of the cardiorespiratory effects of albumin, hetastarch, and saline solutions in patients with hypovolemic and septic shock. *Crit Care Med* 1983;11:839-50.
- 31 Skillman JJ, Restall DS, Salzman EW. Randomized trial of albumin vs. electrolyte solutions during abdominal aortic operations. *Surgery* 1975;78:291-303.
- 32 Tølløfsrud S, Svennevig JL, Breivik H, Kongsgaard U, Øzer M, Hysing E, et al. Fluid balance and pulmonary functions during and after coronary artery bypass surgery: Ringer's acetate compared with dextran, polygeline, or albumin. *Acta Anaesthesiol Scand* 1995;39:671-7.
- 33 Virgilio RW, Rice CL, Smith DE, James OR, Zarins CK, Hobelmann CF, et al. Crystalloid vs. colloid resuscitation: is one better? A randomized clinical study. *Surgery* 1979;85:129-39.
- 34 Woittiez AJ. Restoration of colloid osmotic pressure in post operative intensive care patients. A randomised placebo controlled trial with albumin.

- men 20% and hydroxy-ethyl starch. In: Medical Editors' Trial Amnesty, The Cochrane Controlled Trials Register. In: *The Cochrane Library*. Issue 2. Oxford: Update Software, 1998. Updated quarterly.
- 35 Wojtyciak SL, Brown RO, Roberson D, Powers DA, Kudsk KA. Effect of hypoalbuminaemia and parenteral nutrition on free water excretion and electrolyte-free water resorption. *Crit Care Med* 1992;20:164-9.
- 36 Woods MS, Kelley H. Oncotic pressure, albumin and ileus: the effect of albumin replacement on postoperative ileus. *Am Surg* 1993;59:758-63.
- 37 Zetterstrom H, Hedstrand U. Albumin treatment following major surgery. I: Effects on plasma oncotic pressure, renal function and peripheral oedema. *Acta Anaesth Scand* 1981;25:125-32.
- 38 Zetterstrom H. Albumin treatment following major surgery. II: Effects on postoperative lung function and circulatory adaptation. *Acta Anaesth Scand* 1981;25:133-41.
- 39 Weaver DW, Ledgerwood AM, Lucas CE, Higgins R, Bouwman DL, Johnson SD. Pulmonary effects of albumin resuscitation for severe hypovolaemic shock. *Arch Surg* 1978;113:387-92.
- 40 Johnson SD, Lucas CE, Gerrick SJ, Ledgerwood AM, Higgins RF. Altered coagulation after albumin supplements for treatment of oligoemic shock. *Arch Surg* 1979;114:379-83.
- 41 Lucas CE, Bouwman DL, Ledgerwood AM, Higgins R. Differential serum protein changes following supplemental albumin resuscitation for hypovolaemic shock. *J Trauma* 1980;20:47-51.
- 42 Clift DR, Lucas CE, Ledgerwood AM, Sardesai V, Kithier K, Grabow D. The effect of albumin resuscitation for shock on the immune response to tetanus toxoid. *J Surg Res* 1982;32:449-52.
- 43 Oxman A, Cook D, Guyatt GH for the Evidence-based Medicine Working Group. User's guides to the medical literature. VI: How to use an overview. *JAMA* 1994;272:1367-71.
- 44 Soni N. Wonderful albumin? *BMJ* 1995;310:887-8.
- 45 Fleck A, Raines G, Hawker F, Trotter J, Wallace P, Ledingham I, et al. Increased vascular permeability: a major cause of hypoalbuminaemia in disease and injury. *Lancet* 1985;i:781-4.

(Accepted 1 July 1998)

Science commentary

Why albumin may not work

Starling's principle is often represented as the leakage of fluids from the arterial end of capillaries, where the hydrostatic pressure is greater than the oncotic pressure (derived from the plasma proteins), and the reabsorption of fluid into the venous end, where the oncotic pressure exceeds the hydrostatic pressure. A small excess of fluid in the interstitial space—when filtration from the capillaries is greater than reabsorption—is dealt with by lymphatic drainage from the interstitial space. The rationale for giving albumin solutions rather than crystalloid solutions in cases of hypovolaemic shock is that fluid reabsorption from the interstitial space is enhanced, and fluid therefore remains in the vascular system for longer.

But in recent years the assumed reabsorption of fluid at the venous end of capillaries has been challenged. There is now good evidence to show that, except in the gut and the renal circulation, there is no sustained reabsorption of fluid at the venous end of capillaries. Instead, there is a small constant level of filtration from the capillaries, restrained by the osmotic pressure of the plasma proteins. In some rare circumstances—for example, in hypovolaemic shock—there is a transient reabsorption of fluid, but this lasts for only a few minutes and it amounts to an "internal transfusion" of about 500 ml of fluid over 15 minutes.

The production of life threatening pulmonary oedema begins when the loss of protein and fluid from

the blood vessels exceeds the volume of fluid that can be drained from the interstitial space by the lymphatics. In some disease states or when tissue is damaged, as in severe burns, the capillary walls become very much more permeable under the influence of direct cellular damage and from inflammatory mediators. The filtration of fluids, together with proteins, out into the interstitial space is greatly increased and cannot be matched by lymphatic drainage. The filtration rate may be further increased by a fall in the hydrostatic pressure in the interstitial space as a result of tissue damage, so that even more fluid is sucked out of the capillaries.

Conventionally, colloids such as albumin are administered to these patients in an attempt to maintain their intravascular volume, but because of the increased permeability of the vessels, the albumin solution becomes much less effective in maintaining plasma volume than in healthy individuals who have normal vessel permeability. Thus the rationale for administering albumin solutions becomes much less clear. In disease states such as the nephrotic syndrome, for example, there is new evidence to show that protein is lost not only from the renal circulation owing to greater permeability of the renal vessels, but also from the rest of the systemic circulation. This being the case, it is difficult to see how the administration of albumin could ever replace the deficit without causing further problems.

Abi Berger—*Science editor, BMJ*

A memorable patient "I got no counselling"

Examining war pensioners can provide an opportunity to listen, unstressed by the constraints imposed by active disease or the length of the appointment. Occasionally, you are exposed to tales of immense courage or distress recounted with characteristic British understatement.

The gentle former bank messenger described how his warship was ordered alongside a burning merchant ship which was packed full of ammunition. The inevitable happened and the pensioner found himself floating in the water. He was taken ashore to a hospital and after four weeks of convalescence his bed was required and he was sent back to his ship on "light duties."

What had these "light" duties consisted of? "Well by then," he recounted, "our ship had been beached and we had to go below decks to bring out the bodies and sew them into canvas hammocks. When the padre found out what we were doing it was

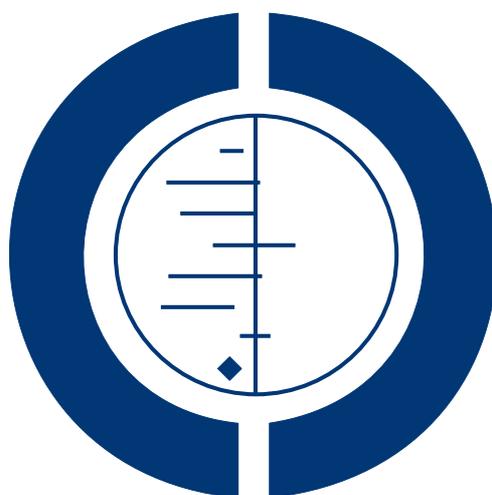
stopped, but, you know doc, I got no counselling," he added with a wry smile.

Close to tears, he described his visit to bereaved parents whose only son he had taught to wash and iron his own clothes. Amazingly, my patient had no subsequent experience of flashbacks or nightmares. But what he did have was a strong feeling of the shared experience of working with fellow survivors and their relatives to lay to rest shipmates with whom he had sailed and fought. The existence of a common enemy allowed comfort to be obtained from even this gruesome task, spared from the modern distraction of searching through a sequence of events for someone to blame and the possibility of eventual financial compensation.

Jim Ford, *senior medical officer, Department of Health, Leeds*

Timing and volume of fluid administration for patients with bleeding (Review)

Kwan I, Bunn F, Roberts I



**THE COCHRANE
COLLABORATION®**

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2009, Issue 1

<http://www.thecochranelibrary.com>



Timing and volume of fluid administration for patients with bleeding (Review)
Copyright © 2009 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
BACKGROUND	2
OBJECTIVES	3
METHODS	3
RESULTS	5
DISCUSSION	6
AUTHORS' CONCLUSIONS	7
ACKNOWLEDGEMENTS	7
REFERENCES	7
CHARACTERISTICS OF STUDIES	8
DATA AND ANALYSES	13
WHAT'S NEW	13
HISTORY	13
CONTRIBUTIONS OF AUTHORS	13
DECLARATIONS OF INTEREST	13
SOURCES OF SUPPORT	13
INDEX TERMS	14

[Intervention Review]

Timing and volume of fluid administration for patients with bleeding

Irene Kwan¹, Frances Bunn², Ian Roberts³

¹National Collaborating Centre For Women's and Children's Health, Royal College of Obstetricians & Gynaecologists, London, UK.

²Centre for Research in Primary and Community Care, University of Hertfordshire, Hatfield, UK. ³Cochrane Injuries Group, London School of Hygiene & Tropical Medicine, London, UK

Contact address: Irene Kwan, National Collaborating Centre For Women's and Children's Health, Royal College of Obstetricians & Gynaecologists, 2-16 Goodge Street, London, W1T2QA, UK. ikwan@ncc-wch.org.uk.

Editorial group: Cochrane Injuries Group.

Publication status and date: New search for studies and content updated (no change to conclusions), published in Issue 1, 2009.

Review content assessed as up-to-date: 21 October 2008.

Citation: Kwan I, Bunn F, Roberts I. Timing and volume of fluid administration for patients with bleeding. *Cochrane Database of Systematic Reviews* 2003, Issue 3. Art. No.: CD002245. DOI: 10.1002/14651858.CD002245.

Copyright © 2009 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Treatment of haemorrhagic shock involves maintaining blood pressure and tissue perfusion until bleeding is controlled. Different resuscitation strategies have been used to maintain the blood pressure in trauma patients until bleeding is controlled. However, while maintaining blood pressure may prevent shock, it may worsen bleeding.

Objectives

To assess the effects of early versus delayed, and larger versus smaller volume of fluid administration in trauma patients with bleeding.

Search strategy

We searched the CENTRAL (*The Cochrane Library* 2008, Issue 4), the Cochrane Injuries Group's Specialised Register (searched October 2008), MEDLINE (to October 2008), EMBASE (to October 2008), the National Research Register (in Current controlled trials.gov; searched October 2008) and the Science Citation Index (to October 2008). We checked reference lists of identified articles and contacted authors and experts in the field.

Selection criteria

Randomised trials of the timing and volume of intravenous fluid administration in trauma patients with bleeding. Trials in which different types of intravenous fluid were compared were excluded.

Data collection and analysis

Two authors independently extracted data and assessed trial quality.

Main results

We did not combine the results quantitatively because the interventions and patient populations were so diverse.

Early versus delayed fluid administration

Three trials reported mortality and two coagulation data.

Timing and volume of fluid administration for patients with bleeding (Review)

Copyright © 2009 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

1

In the first trial (n=598) relative risk (RR) for death with early fluid administration was 1.26 (95% confidence interval of 1.00–1.58). The weighted mean differences (WMD) for prothrombin time and partial thromboplastin time were 2.7 (95% CI 0.9–4.5) and 4.3 (95% CI 1.74–6.9) seconds respectively.

In the second trial (n=50) RR for death with early blood transfusion was 5.4 (95% CI 0.3–107.1). The WMD for partial thromboplastin time was 7.0 (95% CI 6.0–8.0) seconds. In the third trial (n=1309) RR for death with early fluid administration was 1.06 (95% CI 0.77–1.47).

Larger versus smaller volume of fluid administration

Three trials reported mortality and one coagulation data.

In the first trial (n=36) RR for death with a larger volume of fluid resuscitation was 0.80 (95% CI 0.28–22.29). Prothrombin time and partial thromboplastin time were 14.8 and 47.3 seconds in those who received a larger volume of fluid, as compared to 13.9 and 35.1 seconds in the comparison group.

In the second trial (n=110) RR for death with a high systolic blood pressure resuscitation target (100mmHg) maintained with a larger volume of fluid, as compared to low systolic blood pressure resuscitation target (70mmHg) maintained with a smaller volume of fluid was 1.00 (95% CI 0.26–3.81). In the third trial (n=25) there were no deaths.

Authors' conclusions

We found no evidence from randomised controlled trials for or against early or larger volume of intravenous fluid administration in uncontrolled haemorrhage. There is continuing uncertainty about the best fluid administration strategy in bleeding trauma patients. Further randomised controlled trials are needed to establish the most effective fluid resuscitation strategy.

PLAIN LANGUAGE SUMMARY

No evidence from trials to support or not to support the use of early or larger volume intravenous fluid in uncontrolled bleeding

About one third of injury deaths are due to shock from blood loss. Preventing shock in people with uncontrolled bleeding is, therefore, very important and is generally done by giving fluids intravenously. The aim is to maintain blood pressure and reduce tissue damage. The review of trials found that there is uncertainty about the best time to give fluid and what volume of fluid should be given. While increasing fluids will maintain blood pressure, it may also worsen bleeding by diluting clotting factors in the blood. More research is needed.

BACKGROUND

In 1990 approximately five million people died worldwide as a result of injury (Murray 1996). For people younger than 35 years, injury is now the leading cause of death. Nevertheless, the global epidemic of injury is only beginning. It is estimated that by 2020, deaths from injury will have increased from 5.1 million to 8.4 million (Murray 1997). About one third of injury deaths are due to haemorrhagic shock (Deakin 1994). Acute blood loss following injury leads to a reduction in tissue perfusion and tissue oxygen delivery that, if prolonged, causes lactic acidosis and organ fail-

ure. Treatment of haemorrhagic shock involves maintaining blood pressure and tissue perfusion until the bleeding is controlled. Over the past 50 years, a number of resuscitation strategies have been used to maintain the blood pressure in trauma patients until bleeding is controlled. The evidence for the effectiveness of these approaches has been the subject of a number of systematic reviews by the Cochrane Injuries Group and by others.

Pre-hospital use of medical anti-shock trousers

Medical anti-shock trousers (MAST) were first used in the Vietnam War to stabilise patients with haemorrhagic shock during transportation. After the war, MAST became widely used in the care of bleeding trauma patients. MAST increases blood pressure by compressing the blood vessels in the legs, thus increasing systemic vascular resistance, and by shunting blood from the lower body to the brain, heart and lungs. It was hoped that by increasing venous return to the heart, MAST would maintain blood flow to vital organs until definitive care was given. Nevertheless, a systematic review of randomised controlled trials of MAST use in pre-hospital trauma care provided no evidence that MAST increases survival, and a suggestion that it may increase the risk of death. The pooled relative risk of death with MAST was 1.13 (95% CI 0.97–1.32) (Dickinson 1999).

Paramedic ambulance crews

In high-income countries, an increasing number of ambulance crews include a paramedic trained in advance life support. Paramedics receive extra training in intubation, intravenous cannulation, and the administration of intravenous fluids. Only a small proportion of paramedic-attended trauma patients require intubation (1%), but a larger proportion (18%) receive intravenous fluids (Nicholl 1998). Because of the strong conviction amongst the public and medical profession that paramedic intervention is beneficial, it has been difficult to conduct randomised controlled trials comparing paramedic and non-paramedic trauma care. However, a review and meta-analysis of four cohort studies gave a significantly increased ($p=0.03$) risk of death in paramedic attended patients (RR=1.26) (Nicholl 1998). Because of the potential for confounding by injury severity, the validity of inferences from cohort studies must be questioned. Nevertheless, the results are consistent with the hypothesis that efforts by paramedics to raise the blood pressure of bleeding trauma patients may be counterproductive.

Colloid fluid resuscitation

Intravenous fluid administration, with colloid or crystalloid solutions, is the mainstay of the non-surgical management of bleeding trauma patients. Colloids are better than crystalloid solutions in expanding the circulation, because they are retained within the blood vessels to a greater extent. Crystalloid solutions rapidly leak out of the blood vessels into the interstitial spaces. After a colloid infusion, the increase in the circulating volume is about the same as the volume of colloid infused, whereas only about one quarter of the volume of a crystalloid infusion remains in the blood vessels (Weil 1999). Although colloids are effective in expanding the circulation there is no evidence that this improves outcome in critically ill patients (Alderson 2000).

The systematic reviews of medical anti-shock trousers, paramedic resuscitation and colloid administration call into question the benefits of raising the blood pressure in bleeding trauma patients. But by what mechanisms could fluid resuscitation adversely affect outcome? Stern, using a swine model of near-fatal haemorrhage, found that attempts to restore blood pressure with crystalloid resulted in increased haemorrhage volume and markedly higher mortality (Stern 1993). It was postulated that the increased pulse pressure from crystalloid resuscitation might cause the mechanical disruption of blood clots and worsen bleeding. It has also been proposed that fluid administration might also worsen bleeding by diluting clotting factors. In view of the concerns raised by the previous systematic reviews and by the results of animal models of haemorrhagic hypovolaemia, we have conducted a systematic review of the effect on mortality of early versus delayed fluid resuscitation, and of larger versus smaller fluid volumes.

OBJECTIVES

To examine the effect on mortality and coagulation times of two intravenous fluid administration strategies in the management of haemorrhagic hypovolaemia: early compared to delayed administration and larger compared to smaller volume of fluid administered.

METHODS

Criteria for considering studies for this review

Types of studies

All unconfounded randomised and quasi-randomised controlled trials of the timing or volume of intravenous fluid administration in haemorrhagic hypovolaemia.

Types of participants

Patients of all ages with haemorrhagic hypovolaemia of traumatic or non-traumatic origin. Because the physiological response to bleeding and to fluid resuscitation is likely to be similar among patients with internal bleeding (e.g. bleeding peptic ulcer) and those with external bleeding (e.g. penetrating trauma), both types of participants were included.

Types of interventions

Intravenous fluids including crystalloid solutions, colloids, plasma and blood. Trials in which the timing or volume of fluid administration is confounded by the type of intravenous fluid given – for example, a trial comparing the administration of 1000ml of colloid with 500ml blood – were excluded.

Types of outcome measures

Mortality from all causes at the end of the follow-up period scheduled for each trial. We sought mortality data in simple categorical form, and we did not extract data on time to death. If a report did not include the numbers of deaths in each group, we sought these data from the authors. We also sought data on prothrombin time and partial thromboplastin time during fluid administration.

Search methods for identification of studies

Searches were not restricted language, date or publication status.

Electronic searches

We searched:

- CENTRAL (*The Cochrane Library* 2008, Issue 4),
- MEDLINE (1966 to October 2008),
- National Research Register (October 2008),
- EMBASE (1980 to October 2008),
- Web of Science; Science Citation Index (to October 2008),
- Cochrane Injuries Group Trials Register (searched October 2008),
- Current Controlled Trials (Searched October 2008).

The search strategy can be found in Appendix 1.

Searching other resources

We checked the reference lists of all included studies and contacted authors and experts in the field. The Science Citation Index was checked for eligible papers that cited two of the trials (Bickell 1994, Blair 1986) included in this review.

Data collection and analysis

Selection of studies

One reviewer (IK) examined the electronic search results for reports of possibly relevant trials and these reports were then retrieved in full. The first reviewer (IK) also contacted experts in the field for unpublished and ongoing trials. A second reviewer (FB) examined 10% of the electronic search results to check for

agreement on eligibility criteria. Two reviewers (FB, IK) applied the selection criteria independently to the trial reports, resolving disagreements by discussion with a third (IR).

Data extraction and management

Two reviewers (IK, FB) independently extracted information on the following: method of allocation concealment, number of randomised patients, type of participants and the interventions, loss to follow-up and length of follow-up. The outcome data sought were numbers of deaths, prothrombin time and partial thromboplastin time. The reviewers were not blinded to the authors or journal when doing this. Results were compared and any differences resolved by discussion.

Where there was insufficient information in the published report we attempted to contact the authors for clarification.

Assessment of risk of bias in included studies

Since there is evidence that the quality of allocation concealment particularly affects the results of studies, two reviewers (IK, FB) scored this quality on the scale used by Higgins 2008 as shown below, assigning 'No' to poorest quality and 'Yes' to best quality:

- Yes = trials deemed to have taken adequate measures to conceal allocation (i.e. central randomisation; serially numbered, opaque, sealed envelopes; or other description that contained elements convincing of concealment)
- Unclear = trials in which the authors either did not report an allocation concealment approach at all or reported an approach that did not fall into one of the other categories
- No = trials in which concealment was inadequate (such as alternation or reference to case record numbers or to dates of birth).

Where the method used to conceal allocation was not clearly reported, the author(s) were contacted, if possible, for clarification. We then compared the scores allocated and resolved differences by discussion.

Data synthesis

The following comparisons were made:

- early versus delayed intravenous fluids administration
- larger versus smaller volume of intravenous fluids administration.

The relative risk of death and 95% confidence interval (95% CI) was calculated, such that a relative risk of more than 1 indicated a higher risk of death in the first group named. The relative risk was chosen as it is more readily applied to the clinical situation. The weighted mean difference was calculated for coagulation times. Because of differences in the types of patients and in the nature of the trial interventions we did not pool the data in our analysis.

RESULTS

Description of studies

See: [Characteristics of included studies](#).

Our original search strategy found 4,487 potentially eligible reports of which six unpublished trials met the inclusion criteria. A further 655 abstracts were retrieved from a search carried out in August 2007, and another search in October 2008 retrieved 261 abstracts. All search results were scanned by two individuals for potentially relevant studies. No new trials that met the inclusion criteria were indentified from these search results.

A. Early versus delayed intravenous fluids administration

[Bickell 1994](#)

This trial compared early versus delayed administration of Ringer's acetate solution, an isotonic crystalloid, in patients with penetrating torso injuries during the prehospital phase. Participants were adults over 16 years of age, with gunshot or stab wounds to the torso, and who had a systolic blood pressure of <90mmHg. Participants with head injury, a Revised Trauma Score of zero or minor injuries were excluded. During the trial, 22 patients (8%) in the delayed resuscitation group were inadvertently given fluid prior to surgery in violation of the protocol. Follow-up was until hospital discharge.

[Blair 1986](#)

This trial compared early versus delayed blood transfusion in patients with acute gastrointestinal haemorrhage during the first 24 hours after admission. Patients with oesophageal varices were excluded because of abnormal coagulation related to liver diseases. Follow-up was until hospital discharge.

[Turner 2000](#)

This trial compared early versus no/delayed fluid administration in trauma patients. Participants were all trauma patients with moderate to severe injuries, over the age of 16 years of age. Patients who were pregnant or without vital signs were excluded. Fluids given were crystalloids. Protocol compliance was poor with 31% of patients in the early fluid group receiving fluids and 80% of the delayed/no fluid group not given fluids. Follow-up was for six months.

B. Larger versus smaller volume of intravenous fluids administration

[Dunham 1991](#)

This trial compared fluid resuscitation using the rapid infusion system and conventional fluid administration method in trauma patients during the first 24 hours of admission. Participants were between 14 and 60 years of age and had a systolic blood pressure of <90mmHg. Patients with a Glasgow Coma Score of <5, cardiac

arrest, quadriplegia and myocardial infarct on admission were excluded. Fluids given included red blood cells, platelets, fresh frozen plasma and crystalloids. Follow-up was until hospital discharge.

[Dutton 2002](#)

This trial compared the maintenance of target systolic blood pressures of 70 and 100mmHg respectively with fluid restriction (Plasma, Plasmalyte-A and red blood cells in the first 24 hours) in patients with blunt and penetrating trauma injuries. All participants suffered haemorrhagic shock with a systolic blood pressure (SBP) of <90mmHg. Patients with head or spinal cord injury were excluded. Length of follow-up period was until death or hospital discharge.

[Fortune 1987](#)

This trial compared the maintenance of haematocrit at 30% and 40% respectively with blood transfusion in patients following acute injuries and haemorrhage during the first 72 hours of admission. All participants had sustained Class III/IV haemorrhage with a systolic blood pressure of <90mmHg, heart rate > 100 beats per minute. Follow-up was for three days.

Risk of bias in included studies

A. Early versus delayed intravenous fluids administration

[Bickell 1994](#)

Randomisation was by alternate day allocation which allowed foreknowledge of treatment allocation. Data were analysed as randomised, on an intention-to-treat basis. Blinding of outcome assessment was not stated. There was no loss to follow-up.

[Blair 1986](#)

Contact with the author of this trial established the adequacy of the randomisation method used. Allocation was by opening sealed envelopes at the time of patient presentation. Blinding of outcome assessment was not stated. Data were analysed as randomised, on an intention-to-treat basis. There was no loss to follow-up.

[Turner 2000](#)

Paramedics rather than trauma patients were randomised, using computer-generated random numbers, stratified by base stations. The paramedics crossed over to alternate fluid protocol halfway through the trial and they were not blinded. Data were analysed as randomised, on an intention-to-treat basis. There was no blinding in outcome assessment.

B. Larger versus smaller volume of intravenous fluids administration

[Dunham 1991](#)

Method of randomisation and allocation concealment was unclear. Blinding of outcome assessment was not stated. Data on eight

patients who died during the first 12 hours were excluded from the analysis except for the outcome of death.

Dutton 2002

Randomisation was by selecting the next numbered envelope from a supply maintained in the Trauma Resuscitation Unit. The envelopes were made up in batches of 20 (10 to each group), thoroughly mixed, and then numbered for selection. Allocation was blinded to all Unit personnel prior to enrolment. Only the patients were blinded to the allocation in this trial after randomisation. Data were analysed as randomised, on an intention-to-treat basis. Blinding of outcome assessment was not stated. There was no loss to follow-up.

Fortune 1987

Contact with the co-author of this trial established the adequacy of the randomisation method used. Sequences of random allocations were generated by a statistician not involved with the study, in sets of sealed opaque envelopes, differentiated by sex and age groups. Both patients and physicians had no prior knowledge of which arm the patient would be assigned to. Blinding of outcome assessment was not stated. There was no loss to follow-up.

The characteristics of each trial are listed in [Characteristics of included studies](#).

Effects of interventions

A. Early versus delayed fluid administration

One trial (Bickell 1994) reported mortality and coagulation time on a total of 598 hypotensive trauma patients with penetrating torso injuries. Mortality was 116/309 (38%) in the early and 86/289 (30%) in the delayed administration group. The relative risk for death with early fluid administration was 1.26 (95% CI 1.00–1.58). Prothrombin time and partial thromboplastin time were 14.1 and 31.8 seconds in the early, as compared to 11.4 and 27.5 seconds in the delayed administration group. The weighted mean difference (WMD) for prothrombin time and partial thromboplastin time was 2.7 (95% CI 0.90–4.5) and 4.3 seconds (95% CI 1.7–6.9) respectively.

One trial (Blair 1986) reported mortality and coagulation time on a total of 50 hypotensive patients with acute upper gastrointestinal haemorrhage. Mortality was 2/24 (8%) in the early as compared to 0/26 (0%) in the delayed transfused group. The relative risk for death with early blood transfusion was 5.4 (95% CI 0.3–107.1). Activated partial thromboplastin time was 48 in the early, as compared to 41 seconds in the delayed administration group. The WMD for partial thromboplastin time was 7.0 seconds (95% CI 6.0–8.0).

In one trial (Turner 2000) on a total of 1309 trauma patients, mortality was 73/699 (10.4%) in the early as compared to 60/610 (9.8%) in the delayed/no fluid administration group. The relative

risk for death with early fluid administration was 1.06 (95% CI 0.77–1.47). There were no data on coagulation times.

B. Larger versus smaller volume of fluid administration

One trial (Dunham 1991) reported mortality and coagulation time on a total of 36 hypotensive trauma patients. Mortality was 5/20 (25%) in the group who received a larger volume of fluids administered conventionally, as compared to 5/16 (31%) in the group who received a smaller volume of fluids administered using the Rapid Infusion System. The relative risk for death is 0.80 (95% CI 0.28–2.29). Prothrombin time and partial thromboplastin time were 14.8 and 47.3 seconds in the group who received a larger volume of fluid as compared to 13.9 and 35.1 seconds in the comparison group.

In one trial (Dutton 2002) on a total of 110 hypotensive patients with blunt and penetrating injuries, mortality was 4/55 (7.3%) in the group administered a larger volume and 4/55 (7.3%) in the group administered a smaller volume (1000ml less than in the intervention group). The relative risk for death is 1.00 (95% CI 0.26–3.81). There were no data on coagulation times.

In one trial (Fortune 1987) on a total of 25 hypotensive patients with acute injury and haemorrhage, there were no data on mortality in both the groups administered with larger or smaller volume of blood. Contact with the co-author established that there were no deaths in either group. There were no data on coagulation times.

DISCUSSION

This review found insufficient evidence for or against the use of early or larger volume fluid resuscitation in the treatment of uncontrolled haemorrhage. While vigorous fluid resuscitation may be life-saving in some patients, results from clinical trials are inconclusive.

Every year, tens of thousands of patients receive intravenous fluids in the management of bleeding. The Advanced Trauma Life Support (ATLS) protocol of the American College of Surgeons recommends the liberal use of isotonic crystalloid to correct hypotension in bleeding trauma patients. Nevertheless, we could find no reliable evidence to support or not to support this recommendation. While we cannot exclude the possibility that we overlooked a large high-quality randomised controlled trial showing that early or larger volume fluid resuscitation is beneficial, we believe that this is unlikely. To identify eligible trials we screened over 4,000 potentially relevant reports, we searched the reference lists of included trials, and contacted authors and experts in the field.

Six published trials were reviewed. Due to their heterogeneity, in terms of types of patients and types of fluids used, we did not attempt to perform a meta-analysis of the studies.

Death was chosen as the primary end-point in this review for two reasons. First, death is a clinically relevant outcome that matters to patients. Second, death is not prone to measurement error and to reporting bias, as are pathophysiological end points. Mortality data were available for all six included trials, three on the effect of early fluid resuscitation (Bickell 1994; Blair 1986; Turner 2000) and three on the effect of larger volume fluid resuscitation (Dunham 1991, Dutton 2002, Fortune 1987). Three trials examined the effect of fluid administration on coagulation. Clotting times were significantly elevated in the immediate resuscitation groups (Bickell 1994, Blair 1986) and the group who received a larger volume (Dunham 1991). Method of randomisation was inadequate in two trials (Bickell 1994, Turner 2000) and unclear in another (Dunham 1991). Allocation concealment was inadequate in two trials (Bickell 1994; Turner 2000). Because inadequate randomisation and poorly concealed allocation can bias the results of randomised controlled trials, and because this bias can be large and can operate in either direction, the impact of early or larger volume fluid resuscitation on mortality remains difficult to estimate.

Interpretation of results also needs to be cautious due to the heterogeneous nature of traumatic injuries encountered in these trials. Haemorrhagic shock can be caused by a variety of underlying anatomic injuries. Some of these injuries, such as posterior pelvic fractures, may be more amenable to hypotensive management maintained by smaller volume of fluid, than liver injuries where haemostasis can be difficult to achieve (Dutton 2002).

The use of medical anti-shock trousers, early and larger volume fluid administration and colloid resuscitation are based on the idea that raising the blood pressure in bleeding trauma patients will

maintain tissue perfusion and so prevent haemorrhagic shock and its consequences. However, while maintaining blood pressure may prevent shock, it may worsen bleeding. In view of the lack of evidence for or against the effectiveness of currently recommended resuscitation protocols and the potential for harm, the balance of risks and benefits of contemporary resuscitation practice warrants careful consideration. Further randomised controlled trials are required to identify the most effective strategies for the fluid management of bleeding trauma patients.

AUTHORS' CONCLUSIONS

Implications for practice

We found no evidence for or against the use of early or larger volume intravenous fluid administration in uncontrolled haemorrhage. There is uncertainty about the effectiveness of fluid resuscitation in patients with bleeding following trauma.

Implications for research

Large, well concealed, randomised controlled trials are urgently needed to establish the optimal fluid resuscitation strategy in haemorrhaging trauma patients, with a focus on specific types of injuries likely to benefit from the appropriate resuscitation strategy in terms of timing and volume of fluids given.

ACKNOWLEDGEMENTS

We thank the trialists and their co-authors for supplying additional information. We thank Drs. D. Mohan, C. Mock, R. Norton, M. Varghese of the WHO Pre-Hospital Trauma Care Steering Committee for their comments and advice on the review, and Mr. Reinhard Wentz for his help with the searching.

REFERENCES

References to studies included in this review

Bickell 1994 *{published data only}*

* Bickell WH, Wall MJ, Pepe PE, Martin RR, Ginger VF, Allen MK, Mattox KL. Immediate versus delayed fluid resuscitation for hypotensive patients with penetrating torso injuries. *New England Journal of Medicine* 1994;**331**(17):1105–9.
Martin RR, Bickell WH, Pepe PE, Burch JM, Mattox KL. Prospective evaluation of preoperative fluid resuscitation in hypotensive patients with penetrating truncal injury: a preliminary report. *Journal of Trauma* 1992;**33**(3):354–62.

Blair 1986 *{published data only}*

Blair SD, Janvirin SB, McCollum CN, Greenhalgh RM. Effect of early blood transfusion on gastrointestinal haemorrhage. *British*

Journal of Surgery 1986;**73**(10):783–5.

Dunham 1991 *{published data only}*

Dunham CM, Belzberg H, Lyles R, Weireter L, Skurdal D, Sullivan G, Esposito T, Namini M. The rapid infusion system: a superior method for the resuscitation of hypovolemic patients. *Resuscitation* 1991;**21**(2-3):207–27.

Dutton 2002 *{published data only}*

Dutton RP, MacKenzie CF, Scalea TM. Hypotensive resuscitation during active haemorrhage: impact on in-hospital mortality. *Journal of Trauma* 2002;**52**:1141–6.

Fortune 1987 *{published data only}*

Fortune JB, Feustel PJ, Saifi J, Stratton HH, Newell JC, Shah DM. Influence of hematocrit on cardiopulmonary function after acute

hemorrhage. *Journal of Trauma* 1987;**27**(3):243–9.

Turner 2000 {*published data only*}

Turner J, Nicholl J, Webber L, Cox H, Dixon S, Yates D. A randomised controlled trial of pre-hospital intravenous fluid replacement therapy in serious trauma. *Health Technology Assessment* 2000;**4**(31):1–47.

Additional references

Alderson 2000

Alderson P, Schierhout G, Roberts I, Bunn F. Colloids versus crystalloids for fluid resuscitation in critically ill patients (Cochrane Review). *Cochrane Database of Systematic Reviews* 2000, Issue 2. [DOI: 10.1002/14651858.CD000567.pub3]

Deakin 1994

Deakin CD, Hicks IR. AB or ABC: pre-hospital fluid management in major trauma. *Journal of Accident and Emergency* 1994;**11**(3): 154–7.

Dickinson 1999

Dickinson K, Roberts I. Medical anti-shock trousers (pneumatic anti-shock garments) for circulatory support in patients with trauma. *Cochrane Database of Systematic Reviews* 1999, Issue 4. [DOI: 10.1002/14651858.CD001856]

Higgins 2008

Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions*. Vol. **5.0.0** [updated February 2008],

Available from www.cochrane-handbook.org: The Cochrane Collaboration, 2008.

Murray 1996

Murray CJL, Lopez AD. *Global Statistics: a compendium of incidence, prevalence and mortality estimates for over 200 conditions*. Boston: Harvard School of Public Health, Harvard University Press, 1996.

Murray 1997

Murray CLJ, Lopez AD. Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet* 1997;**349**(9064):1498–1504.

Nicholl 1998

Nicholl J, Hughes S, Dixon S, Turner J, Yates D. The costs and benefits of paramedic skills in pre-hospital trauma care. *Health Technology Assessment* 1998; Vol. 2, issue 17.

Stern 1993

Stern SA, Dronen SC, Birrer P, Wang X. Effect of blood pressure on hemorrhage volume and survival in a near-fatal hemorrhage model incorporating a vascular surgery. *Annals of Emergency Medicine* 1993;**22**(2):155–63.

Weil 1999

Weil MH. Crystalloids, colloids and fluid compartments. *Critical Care Medicine* 1999;**27**(1):3.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Bickell 1994

Methods	Quasi-randomised controlled trial. (Allocation by alternation - odd and even numbered days of the month.)
Participants	598 trauma patients >16 years of age with penetrating injuries and hypotension. Mean age = 31 years. Exclusion: pregnancy, Revised Trauma Score = 0, minor injuries not requiring surgery.
Interventions	1) 870ml of Ringer's solution pre-hospital (n=309). 2) 92ml of Ringer's solution with IV cannulation pre-hospital (n=289).
Outcomes	<ul style="list-style-type: none"> • Haemodynamic variables, • amount of fluids given, • intraoperative blood loss, • post-op • complications, • process of care, • death.
Notes	Patients in both groups were treated with a standard paramedical protocol as appropriate until after IV cannulation. 22/289(8%) in delayed fluids group were inadvertently given fluids in violation of the protocol. Results were analysed as randomised on an intention-to-treat basis.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	No	

Blair 1986

Methods	Randomised controlled trial. (Allocation by opening sealed envelopes.)
Participants	50 patients with acute severe upper gastrointestinal haemorrhage. Mean age = 62 Exclusion: patients with oesophageal varices due to abnormal coagulation.
Interventions	1) \geq 2 units of blood in first 24 hr (n=24). 2) No blood transfusion during first 24 hr (n=26).

Blair 1986 (Continued)

Outcomes	<ul style="list-style-type: none"> • Coagulation times, • Haematocrit, • re-bleeding rate, • volume of blood given, • death.
Notes	5/26 patients in the no-blood group received blood in the first 24 hours when their Hgb < 8g/dl. Results were analysed as randomised on an intention-to-treat basis.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	

Dunham 1991

Methods	Randomised controlled trial. (Allocation unclear.)
Participants	36 trauma patients >14 <60 years of age with hypotension. Mean age = 35. Exclusion: Glasgow Coma Score <5, cardiac arrest, quadriplegia, myocardial infarct.
Interventions	<ol style="list-style-type: none"> 1. 23,661ml of IV fluids (red blood cells, fresh frozen plasma, platelets and Plasmalyte-A) in first 24 hours via conventional fluid administration (CFA) (n=20). 2. 20,224 ml of IV fluids (red blood cells, fresh frozen plasma, platelets and Plasmalyte-A, given via the RIS (Rapid Infusion System). (n=16).
Outcomes	<ul style="list-style-type: none"> • Blood loss, • temperature, • Haematocrit, • coagulation times, • serum Lactate, • base excess, • ionised calcium, • costs, • death.
Notes	Data from 3/20 patients in CFA group and 5/16 patients in RIS group who died in the first 12 hours were excluded from subsequent analyses except for death.

Dutton 2002

Methods	Randomised controlled trial. (Allocation by drawing the next numbered envelope from a batch of 20, thoroughly mixed but sequentially numbered envelopes.)
Participants	110 trauma patients >16, <55 years of age with blunt and penetrating injuries and in shock. Mean age = 31 Exclusion: pregnancy, no pulse, head or spinal injury, known end-organ ischaemic disease.
Interventions	1. Bolus of fluids (plasma, Plasmalyte-A and packed red blood cells) to maintain systolic blood pressure of >100mmHg (n=55). 2. 1000ml less of fluids (plasma, Plasmalyte-A and packed red blood cells) to maintain a lower blood pressure of 70mmHg (n=55).
Outcomes	<ul style="list-style-type: none"> • Duration of bleeding, • average ISS, • death.
Notes	All patients maintained at a haematocrit of at least 25%. Results analysed as randomised on an intention-to-treat basis.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	

Fortune 1987

Methods	Randomised controlled trial. (Random allocations generated by a statistician blinded to the study, in sets of sealed opaque envelopes.)
Participants	25 patients with acute injury and haemorrhage, hypotensive, urine output < 20ml/hr. Mean age = 46. Exclusion: history of myocardial infarction in previous year as a higher haematocrit could be harmful.
Interventions	1. ≥ 5 units of blood to maintain Haematocrit at 40% (n=13). 2. < 5 units of blood transfusion to maintain Haematocrit at 30% (n=12).
Outcomes	<ul style="list-style-type: none"> • Cardiopulmonary status, • death not reported, but later obtained from the author.
Notes	Study was designed to test the hypothesis that sufficient oxygen can be provided at lower haematocrit of 30%. Allocation concealment was later reported to be adequate when co-author was contacted.

Risk of bias

Item	Authors' judgement	Description
------	--------------------	-------------

Fortune 1987 (Continued)

Allocation concealment?	Yes
-------------------------	-----

Turner 2000

Methods	Randomised controlled trial (of paramedics using computer-generated random numbers, stratified by base stations).
Participants	1309 trauma patients > 16 years of age. Exclusion: pregnancy, no vital signs.
Interventions	1. \geq 1 unit of fluids of Hartmann's solution and Haemacell pre-hospital (n=699) 2. Delayed/no fluids pre-hospital (n=610).
Outcomes	<ul style="list-style-type: none"> • Post-op complications, • process of care, • costs, • death.
Notes	Protocol compliance was poor with 31% of the fluid group receiving fluids and 80% of the no fluid group not given fluids.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	No	

DATA AND ANALYSES

This review has no analyses.

WHAT'S NEW

Last assessed as up-to-date: 21 October 2008.

6 November 2008	New search has been performed	Search updated. No new trials were found for inclusion in the review.
-----------------	-------------------------------	---

HISTORY

Protocol first published: Issue 3, 2000

Review first published: Issue 1, 2001

20 August 2008	Amended	Converted to new review format.
12 April 2003	New search has been performed	Data have now become available from a trial (Dutton 2002) which was ongoing in 2000. These data did not affect the results or the overall conclusion of the review. Types of injuries may be an important consideration in the research design of future trials in fluids resuscitation.

CONTRIBUTIONS OF AUTHORS

IK screened citations, extracted data, contacted authors, entered data into RevMan and helped to write the review. IR developed the protocol, and helped to write the review. FB screened citations, extracted data, and commented on the review.

DECLARATIONS OF INTEREST

None known.

SOURCES OF SUPPORT

Internal sources

- Institute of Child Health, University of London, UK.

External sources

- Global Programme on Evidence for Health Policy (GPE), World Health Organisation, Switzerland.

INDEX TERMS

Medical Subject Headings (MeSH)

Hemorrhage [*therapy]; Infusions, Intravenous; Plasma Substitutes [*administration & dosage]; Randomized Controlled Trials as Topic; Time Factors; Wounds and Injuries [blood; *complications]

MeSH check words

Humans



Traffic calming for the prevention of road traffic injuries: systematic review and meta-analysis

F Bunn, T Collier, C Frost, K Ker, I Roberts and R Wentz

Inj. Prev. 2003;9;200-204
doi:10.1136/ip.9.3.200

Updated information and services can be found at:
<http://ip.bmjournals.com/cgi/content/full/9/3/200>

These include:

References

4 online articles that cite this article can be accessed at:
<http://ip.bmjournals.com/cgi/content/full/9/3/200#otherarticles>

Rapid responses

You can respond to this article at:
<http://ip.bmjournals.com/cgi/eletter-submit/9/3/200>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

Topic collections

Articles on similar topics can be found in the following collections

[Systematic reviews \(incl meta-analyses\): examples](#) (340 articles)
[Other Public Health](#) (2612 articles)

Notes

To order reprints of this article go to:
<http://www.bmjournals.com/cgi/reprintform>

To subscribe to *Injury Prevention* go to:
<http://www.bmjournals.com/subscriptions/> 352

SYSTEMATIC REVIEW

Traffic calming for the prevention of road traffic injuries: systematic review and meta-analysis

F Bunn, T Collier, C Frost, K Ker, I Roberts, R Wentz

Injury Prevention 2003;9:200–204**Objective:** To assess whether area-wide traffic calming schemes can reduce road crash related deaths and injuries.**Design:** Systematic review and meta-analysis.**Data sources:** Cochrane Injuries Group Specialised Register, Cochrane Central Register of Controlled Trials, Medline, EMBASE, Sociological Abstracts Science (and social science) citation index, National Technical Information service, Psychlit, Transport Research Information Service, International Road Research Documentation, and Transdoc, and web sites of road safety organisation were searched; experts were contacted, conference proceedings were handsearched, and relevant reference lists were checked.**Inclusion criteria:** Randomised controlled trials, and controlled before/after studies of area-wide traffic calming schemes designed to discourage and slow down through traffic on residential roads.**Methods:** Data were collected on road user deaths, injuries, and traffic crashes. For each study rate ratios were calculated, the ratio of event rates before and after intervention in the traffic calmed area divided by the corresponding ratio of event rates in the control area, which were pooled to give an overall estimate using a random effects model.**Findings:** Sixteen controlled before/after studies met our inclusion criteria. Eight studies reported the number of road user deaths: pooled rate ratio 0.63 (95% confidence interval (CI) 0.14 to 2.59). Sixteen studies reported the number of injuries (fatal and non-fatal): pooled rate ratio 0.89 (95% CI 0.80 to 1.00). All studies were in high income countries.**Conclusion:** Area-wide traffic calming in towns and cities has the potential to reduce road traffic injuries. However, further rigorous evaluations of this intervention are needed, especially in low and middle income countries.

See end of article for authors' affiliations

Correspondence to:
Frances Bunn, Centre for
Research in Primary and
Community Care,
University of Hertfordshire,
College Lane, Hatfield,
Herts AL10 9AB, UK;
F.bunn@herts.ac.uk

The worldwide epidemic of road traffic injuries is only just beginning. At present, over a million people die each year and some 10 million people sustain permanent disabilities in road traffic crashes. For people under 44 years, road traffic crashes are a leading cause of death and disablement, second only to HIV and AIDS.¹ Many developing countries are still at comparatively low levels of motorisation and the incidence of road traffic injuries in these countries is likely to increase. It is estimated that by 2020 road traffic crashes will have moved from ninth to third in the world disease burden ranking, as measured in disability adjusted life years.²

Most of the road deaths in developing countries involve vulnerable road users such as pedestrians and cyclists. In Ethiopia, pedestrian injuries account for 84% of all road traffic fatalities compared with 32% in Britain and 15% in the USA.³ In the heavily motorised countries drivers and passengers account for the majority of road deaths but pedestrians account for a large proportion of road deaths involving children. The identification of effective strategies for the prevention of road traffic injuries is of global health importance.

In urban areas, road traffic crashes are scattered widely, and in such situations localised interventions for high risk sites are not appropriate. In high income countries area-wide traffic calming schemes, including the treatment of both main roads and residential roads, have been proposed as a strategy for reducing such scattered crashes. Traffic calming has been defined as the combination of mainly physical measures that reduce the negative effects of motor vehicle use, alter driver behaviour, and improve conditions for non-motorised street users.⁴ It has been estimated that area-wide traffic calming schemes can reduce the number of road traffic injuries by

about 15%.⁵ However, this estimate was based on a review that included uncontrolled before/after studies in which the effect of traffic calming could be confounded by other factors that influence road traffic injury rates. In particular, in high income countries there is evidence that pedestrian injury rates have fallen because of a reduction in walking.⁶ In this case, the inclusion of uncontrolled studies could exaggerate the apparent effect of traffic calming. We conducted a systematic review of controlled studies to assess the effect of area-wide traffic calming on road user deaths, injuries (fatal and non-fatal), and numbers of road traffic crashes.

METHODS

Inclusion criteria

We included randomised controlled trials and controlled before/after studies of area-wide traffic calming schemes. Eligible schemes included those that involved a number of specific changes to the road layout, road hierarchy or road environment, for example road narrowing, road closures, creation of one way streets, changes at junctions, mini-roundabouts, road surface treatment, or speed humps. We excluded studies describing the enforcement of legal interventions, financial incentives or disincentives, and interventions investigating alteration to road signage or traffic lights alone, or studies solely describing interventions to separate different road users (cycle lanes, bus lanes, pedestrian walkways). The outcomes of interest were all road user deaths, all road user injuries (fatal and non-fatal), and the number of traffic crashes.

Box 1: Strategy for identification of studies**Search strategy for electronic databases; searches run in 2000**

- Terms describing the intervention, outcomes, and study methodology were combined.
- A: the intervention—area traffic control* or TRAFFIC RESTRAINT* or traffic calming or traffic engineering or road design or road layout or roundabout* or humps or bumps or traffic distribution or traffic redistribution or traffic flow or crosswalk* or speed cushion* or chicane* or road narrowing or refuges or road hierarchy or traffic hierarchy or four way* stop* or access only or sheltered parking or left turn lane* or wooner* or junction layout or road layout or lateral clearance.
- B: the outcome—accident* or injur* or fatalit* or death or safety.
- C: the study methodology—evaluation or assess* or stud* or evaluation or assess* or (controlled near2 stud*) or comparison or comparative or intervention near2 stud* or controls.

Web sites searched; searches conducted in 2001

- AAA Foundation for Traffic Safety (USA): www.aaafoundation.org
- ARRB, Australian Road Research Board: www.arrb.org.au
- Australian Transport Safety Bureau: www.atsb.gov.au
- CROW, Information and Technology Centres for Transport and Infrastructure (Netherlands): www.crow.nl
- Danish Council for Road Safety Research: www.trm.dk/eng/veje/rft
- Danish Transport Research Institute: www.dtf.dk
- DVR, Deutscher Verkehrssicherheitsrat Road Safety Institute (Germany): www.dvr.de/
- FINNRA, Finnish National Road administration: www.tieh.fi
- INRETS, Institut National de Recherche sur les Transports et leur Sécurité (France): www.inrets.fr
- ITE, Institute of Transportation Engineers (USA): www.ite.org
- LET, Laboratoire d'économie des transports (France): www.lsh-lyon.cnrs.fr
- NHTSA, National Highway Traffic Safety Administration (USA): www.nhtsa.dot.gov
- Swedish National Roads Administration: www.vv.se/for_lang/english/
- SWOV, Institute for Road Safety Research (Netherlands): www.swov.nl
- TOI, Institute of Transport Economics (Norway): www.toi.no
- TC, Transport Canada: www.tc.gov
- TRB, Transportation Research Board: www.nas.edu/trb/
- TRL, Transport Research Laboratory (UK): www.trl.co.uk
- US Department of Transport, Federal Highway Administration (USA): www.fhwa.dot.gov
- VTI, Swedish National Road and Transport Research Institute: www.vti.se
- VTT, Finland www.vtt.fi/indexe.htm

Conference proceedings handsearched

- Australian Road Research Board (ARRB). Proceedings of the 12th ARRB conference; Hobart, Tasmania 27–31 August 1984.
- Australian Road Research Board (ARRB). Proceedings of the 15th ARRB conference; Darwin 26–31 August 1990.
- Institution of Professional Engineers New Zealand (IPENZ). Annual conference, Christchurch February 1992 volumes 1 and 2.
- Institution of Professional Engineers New Zealand (IPENZ). Proceedings of the technical session of the group at the annual conference of IPENZ; Auckland 8–12 February 1982.
- Institute of Transportation Engineers (ITE). Proceedings of the 45th to 71st ITE annual meeting, 1975–2001.
- Institute of Transportation Engineers (ITE). Transportation and traffic theory 9th international symposium; Netherlands 1984.
- Institute of Transportation Engineers (ITE). Residential street design and traffic control 1989.
- Israel Institute of Technology. International conference on pedestrian safety; Haifa 20–23 December 1976.
- Lander Publishing Ltd. The third national traffic calming conference; London 18 October 1996.
- Organisation for Economic Co-operation and Development (OECD). Proceedings of the symposium on techniques of improving urban conditions by restraint of road traffic; 25–27 October 1971.
- PTRC Transport, Highways and Planning Summer 13th–18th Annual Meetings, 1985–90.
- Swedish Road and Traffic Research Institute. Proceedings of the conference on strategic highway research program and traffic safety on two continents; Gothenburg, Sweden 18–20 September 1991.
- Swedish Road and Traffic Research Institute. Proceedings of the conference on strategic highway research program and traffic safety on two continents; Gothenburg, Sweden 27–29 September 1989.
- Swedish Road and Traffic Research Institute. Proceedings of the conference on road safety and traffic environment in Europe; Gothenburg, Sweden 26–28 September 1990.
- The Technion Israel Institute of Technology. The second international conference on new ways for improved road safety and quality of life; Tel-Aviv Hilton Hotel, Israel 7–10 October 1991.
- Transportation Research Institute. International conference on new ways and means for improved safety; Tel Aviv, Israel 20–23 February 1989.
- Transport Research Laboratory. Safety 91 Papers on vehicle safety, traffic safety and road user safety research; TRL Laboratory, Berks 1–2 May 1991.

Identification of studies

We searched the following electronic databases; Cochrane Injuries Group Specialized Register, Cochrane Central Register of Controlled Trials, Medline, EMBASE, Sociological Abstracts Science (and Social Science) Citation Index, National Technical Information Service, Psychlit, Transport Research Information Service, International Road Research Documentation, and TRANSDOC (the last three combined in the TRANSPORT database). One reviewer examined titles, abstracts, and keywords of citations, as given on electronic databases, for eligibility. Where possible the full text of all of potentially relevant citations was obtained. We also searched the web sites of road safety organisations, contacted experts, hand searched

conference proceedings, and checked reference lists of relevant papers. There were no language restrictions. Further details of the search strategy can be seen in box 1.

Data extraction and analysis

One reviewer decided whether studies met the inclusion criteria, and this was checked by a second reviewer. Using a data collection form two reviewers independently extracted data on road user deaths, injuries (fatal and non-fatal), traffic crashes, characteristics of the intervention and control area, and types of measures implemented. To assess study quality we collected information on how the intervention and control areas were matched, duration of the before and after periods,

Table 1 Table of included studies

Study ID	Methods	Participating areas	Interventions
Charlottenburg ¹⁶ (Germany 1977–84)	CBA 2 years before data 2 years after data	(I) Residential area with small businesses. Area of about 60 hectares with 15000 inhabitants (C) Similar area in the same city	Different levels of road surface, road narrowing, chicanes, staggered lanes, speed restrictions
GST Borgentreich ¹⁰⁻¹² (Germany 1983–90)	CBA 3 years before data 3 years after data	(I) Whole town centre: mixture of residential, commercial, and farm properties (C) Similar area in different town	Road narrowing, redesigning major roads, traffic free zones, speed restrictions
GST Buxtehude ¹⁰⁻¹² (Germany 1981–87)	CBA 2 years before data 2 years after data	(I) Mixture of shopping and residential areas. Area of about 268 hectares population of about 11000 (C) Similar area in the same city	Road narrowing, speed restrictions, and a wide range of traffic restraint measures
GST Esslingen ¹⁰⁻¹² (Germany 1983–90)	CBA 2 years before data 2 years after data	(I) Mixture of residential, industrial, and commercial properties (C) Similar area in another town	Reconstruction of major roads, speed restrictions, and renewal of residential roads
GST Ingolstadt ¹⁰⁻¹² (Germany 1982–90)	CBA 2 years before data 2 years after data	(I) Most of the old part of the town, 5500 inhabitants (C) Similar area in another town	A wide range of traffic restraint measures
GST Mainz ¹⁰⁻¹² (Germany 1983–90)	CBA 2 years before data 2 years after data	(I) Rural suburb of 200 hectares with 11000 inhabitants (C) Similar area in the same city	Reconstruction of public spaces including road narrowing and narrowing of road entrances
GST Moabit ¹⁰⁻¹² (Germany 1982–88)	CBA 2 years before data 2 years after data	(I) Residential area of about 120 hectares near the city centre (C) Similar area in the same city	Rebuilding of major traffic roads, increasing level of vegetation in streets
Rijswijk/Eindhoven ¹⁵ (Netherlands 1972–86)	CBA 6 years before data 5 years after data	(I) Road districts in Rijswijk and Eindhoven (C) Residential zones bordering on main traffic arteries within the boundaries of Rijswijk and Eindhoven	Road humps, road closures and narrowing, raised cross roads. Public spaces reclassified
Swindon ¹³ (UK 1975–81)	CBA 2 years before data 3 years after data	(I) 2.8 km section of an all purpose road in Swindon (C) 3 routes of similar layout and function	Roundabouts, pedestrian crossings, changes to intersections
Sydney-Canterbury ¹⁴ (Australia 1981–87)	CBA 3 years before data 2.5 years after data	(I) Predominantly residential area in city (C) Similar area in the same city	Speed humps, roundabouts, slow points, speed limits
Sydney-Willoughby ¹⁴ (Australia 1980–87)	CBA 2 years after data 2 years before data	(I) Predominantly residential area in city (C) Similar area in the same city	Entry thresholds, slow points, speed humps, T-intersection treatments, roundabouts, and road closures
USP Bradford ¹⁷ (UK 1981–88)	CBA 5 years before data 2 years after data	(I) Mainly residential area, population approximately 33000 (C) Similar area in same city	Junction redesign, closure of through roads, and installation of central refuges
USP Bristol ¹⁸ (UK 1981–88)	CBA 5 years before data 2 years after data	(I) Mainly residential area of approximately 10 square km, population was approximately 32000 in about 12000 households (C) Similar area in same city	Junction redesign, mini-roundabouts, right turn bans, improvement of pedestrian crossings, improved road signs and markings, road closures
USP Nelson ¹⁹ (UK 1980–87)	CBA 5 years before data 2 years after	(I) An area of 7 square km, population of approximately 30000 people (C) Similar area in same city	Junction redesign, road closures, and mini-roundabouts
USP Reading ²⁰ (UK 1979–86)	CBA 5 years before data 2 years after data	(I) Approximately 8 square km, with a population of about 36000 people (C) Similar area in same city	Road closures, right turn bans, mini-roundabouts
USP Sheffield ²¹ (UK 1979–87)	CBA 5 years before data 5 years after data	(I) Mostly residential area covering approximately 9 square km, population approximately 50000 (C) Similar area in same city	Road closures, traffic islands, central refuges, turning restrictions

CBA, controlled before after study; I, intervention area; C, control area; GST, German six towns project; USP, UK Urban Safety Project.

and, because of the potential for contamination, we also noted the proximity of the intervention and control areas.

For each study we calculated a rate ratio: the ratio of event rates before and after intervention in the traffic calmed area divided by the corresponding ratio of event rates in the control area. This gives the reduction in the incident rate in the intervention area compared to that in the control area. For example, a rate ratio of 0.8 corresponds to a 20% reduction in events compared with that predicted from the rates in the control area. For the calculation of 95% confidence intervals, standard errors of the logarithms of the rate ratios were constructed assuming that the number of events in each area in each period followed a Poisson distribution,⁷ provided there was at least one event in each period. For studies with no events in one or more periods exact confidence intervals were calculated where the rate ratio was defined. Rate ratios were combined on a logarithmic scale using a random effects meta-analysis model. The assumption of random effects means that the effect estimates and confidence intervals allow for variation in study specific rate ratios over and above that due to variability within studies.⁸ In this meta-analysis such additional variability reflects both underlying heterogeneity in rate

ratios across studies and any variability arising through overdispersion⁹ if the assumption that events follow Poisson distributions is violated.

For studies with no events in one or more periods 1/2 was added to all counts in the pooled analysis. In the analysis of road user deaths, where the majority of studies had no events in at least one period, no test of heterogeneity was carried out, and a pooled estimate of the rate ratio was obtained from the column totals. Analyses were carried out in Stata version 7.0 (Stata corporation, College Station, Texas 77845, USA).

RESULTS

The searches identified 12 986 published and unpublished reports which were screened for eligibility. We obtained the full text of 586 reports and of these 12 reports, describing 16 controlled before/after studies, met our inclusion criteria (see table 1).¹⁰⁻²¹ We found no randomised controlled trials. Seven studies were done in Germany,¹⁰⁻¹² six in the UK,¹³ two in Australia,¹⁴ and one in the Netherlands¹⁵; all were done in the 1970s and 1980s. In most studies attempts had been made to match the intervention and control sites. However, in three

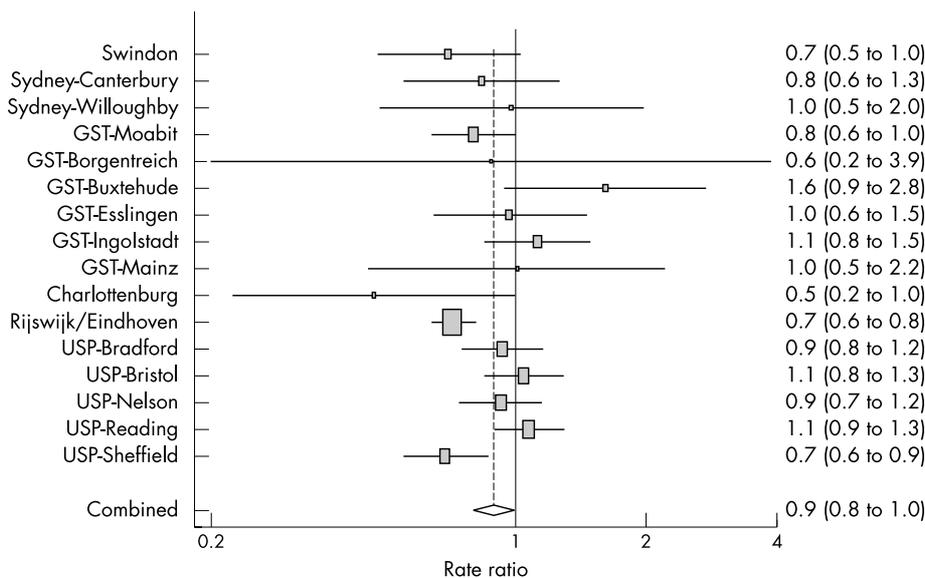


Figure 1 Number of road traffic injuries (fatal and non-fatal).

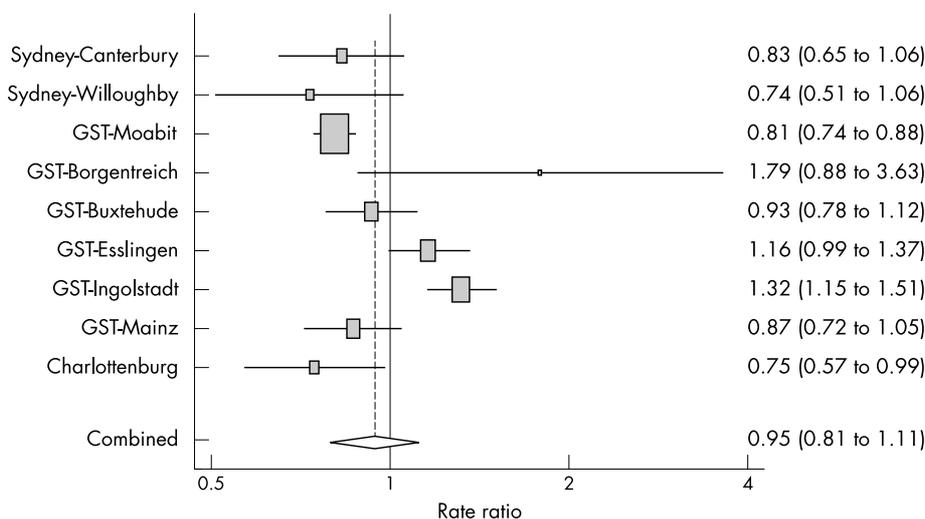


Figure 2 Number of road traffic crashes.

differences in the land use characteristics or type of district are reported,^{14, 15} and in one the control area was much larger than the intervention area.¹⁵ Outcome data was collected from police or local authority records in all studies.

Eight studies reported the number of road user deaths.¹⁰⁻¹⁴ The pooled rate ratio was 0.63 (95% confidence interval (CI) 0.14 to 2.59). This result should be interpreted with caution since many of the studies include at least one period in which no road user deaths were observed. Sixteen studies reported the number of road traffic injuries (fatal and non-fatal).¹⁰⁻²¹ The pooled rate ratio was 0.89 (95% CI 0.80 to 1.00) (fig 1), with statistically significant heterogeneity between the studies ($p = 0.05$). Nine studies reported the total number of road traffic crashes.^{10, 19, 20} The pooled rate ratio was 0.95 (95% CI 0.81 to 1.11) (fig 2), again with statistically significant heterogeneity between the studies ($p = 0.001$). Thirteen trials reported the number of pedestrian crashes.^{10, 14, 17-21} The pooled rate ratio was 1.00 (95% CI 0.84 to 1.18). There was no significant heterogeneity ($p = 0.21$).

DISCUSSION

This systematic review of controlled before/after studies shows that area-wide traffic calming has the potential to prevent road traffic injuries. Although the effect of traffic calming on road user deaths is in the same direction as for injuries (fatal and non-fatal), because the number of road user deaths in the

included studies is low the estimated rate ratio is imprecise. Indeed, the imprecision in the rate ratio may be understated by the confidence interval because the way that the confidence interval was calculated ignores the likely heterogeneity between studies. Although we found no reliable evidence that traffic calming reduces the number of road traffic crashes, because traffic calming may reduce vehicle speeds,²² this is not inconsistent with a reduction in the occurrence of injury. Our estimates of the effectiveness of traffic calming provide a basis for future cost effectiveness analyses that would be important in informing decisions about resource allocation.

Several methodological issues may have a bearing on the validity of these results. Publication and other selection biases are a potential threat to validity in all systematic reviews, but this is a particular problem in road safety where a large proportion of the available research is published in the grey literature. In this review only two of the included studies were published in journals. There are also problems identifying published controlled studies in the road safety databases.²³ Search strategies for identifying controlled studies in medical databases can achieve high sensitivity because terms describing the study methodology are included among the indexing (descriptor) terms. Road safety databases, however, have a very limited range of indexing terms describing the study methodology. Despite our considerable efforts to identify all eligible studies, published and unpublished, irrespective of

Key points

- Injuries as a result of road traffic crashes are a global problem and are likely to increase.
- A previous meta-analysis, including uncontrolled before after studies, found area-wide traffic calming can reduce road traffic injuries by about 15%.
- This systematic review, of 16 controlled before after studies, found an 11% reduction in road traffic injuries (fatal and non-fatal).
- Traffic calming has the potential to prevent road traffic injuries but further rigorous evaluations, particularly in low and middle income countries, is required.

language of publication, we cannot exclude the possibility that some studies were missed resulting in reduced precision and the potential for bias.

Although we found no randomised controlled trials of traffic calming schemes, the inclusion of studies with well matched intervention and control areas, with adequate before and after periods, may avoid the problem of confounding by changes in the background rate of injury. All but one¹⁵ of the included studies had attempted to match the intervention and control areas and all had collected at least two years before and two years after data, with a number collecting up to five years before or after data.

Because there was significant heterogeneity between the studies reporting the number of road traffic injuries and crashes, these results should be interpreted with caution. The observed heterogeneity may be due to differences in study design, in the types of traffic calming schemes involved, or in the way outcomes were defined and data collected.

The included studies were all conducted in the 1970s and 1980s, and, apart from two Australian studies, were all done in Europe. As a result it may make it more difficult to generalise from this systematic review and make inferences about the effectiveness of present day area-wide traffic calming schemes. In addition road traffic crashes are a major cause of death and injury in low and middle income countries where most of the casualties are pedestrians, cyclists, and riders of motorised two wheelers. Although traffic calming appears to be a promising intervention for preventing road traffic injuries because none of the included studies were conducted in low and middle income countries further rigorous evaluation is required in these settings.

ACKNOWLEDGEMENTS

We thank the Medical Research Council for funding to undertake this review.

.....

Authors' affiliations

F Bunn, University of Hertfordshire, Centre for Research in Primary and Community Care

T Collier, C Frost, London School of Hygiene and Tropical Medicine, Medical Statistics Unit

K Ker, I Roberts, London School of Hygiene and Tropical Medicine, Public Health Intervention Research Unit

R Wentz, Imperial College Library and Information Service, Chelsea and Westminster Campus

REFERENCES

- 1 **Krug E**, ed. *Injury. A leading cause of the global burden of disease*. Geneva: World Health Organisation, WHO/HSC/PVI/99.11.
- 2 **Murray CJL**, Lopez AD. Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet* 1997;**349**:1498–504.
- 3 **Barss P**. *Injury prevention: an international perspective—epidemiology, surveillance and policy*. New York: Oxford University Press, 1998.
- 4 **Lockwood I**. ITE Traffic calming definition. *Institute of Transportation Engineers Journal* July 1997: 22.
- 5 **Elvik R**. Area-wide urban traffic calming schemes: a meta-analysis of safety effects. *Accid Anal Prev* 2001;**33**:327–36.
- 6 **Roberts I**. Why have child pedestrian death rates fallen? *BMJ* 1993;**306**:1737–9.
- 7 **Sutton AJ**, Jones DR, Abrams KR, et al. *Methods for meta-analysis in medical research*. Chichester: Wiley and Sons, 2000.
- 8 **DerSimonian R**, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;**7**:177–86.
- 9 **"Student"**. An explanation of deviations from Poisson's law in practice. *Biometrika* 1919;**66**:585–90.
- 10 **Blanke H**. *Geschwindigkeitsverhalten und Verkehrssicherheit bei flachenhafter Verkehrsberuhigung*. Bochum: Ruhr-Universität, 1993.
- 11 **Brilon W**. Area-wide traffic calming measures and their effects on traffic safety in residential area. *Proceedings of the 14th ARRB Conference*. Part 2: 199–205.
- 12 **Brilon W**, Blanke H. Extensive traffic calming: results of the accident analyses in 6 model towns. *1993 Compendium of technical papers*. Washington, DC: Institute of Transportation Engineers, 1993:119–23.
- 13 **Dalby E**. Applications of low-cost road accident countermeasures according to an area-wide strategy. *Traffic Engineering & Control* 1981;**22**:567–74.
- 14 **Fairlie R**, Taylor M. Evaluating the safety benefits of local area traffic management. *Proceedings of the 15th ARRB conference*. Vol 7. Darwin, Australia, 1990: 141–66.
- 15 **Janssen S**. Final results of accident studies in the Dutch demonstration projects of the 1970s. *Traffic Engineering & Control* 1991;**32**:292–6.
- 16 **Kahrman B**. Area-wide traffic restraint measures: analysis of accidents in Berlin-Charlottenburg. *2nd International Conference on Road Safety*. Groningen, Netherlands, 1988: 424–34.
- 17 **Walker R**, McFetridge M. *Urban Safety Project: the Bradford scheme*. Crowthorne, Berks: Transport and Road Research Laboratory, 1989 (contractor report 190).
- 18 **Ward H**, Norrie J, Allsop R, et al. *Urban Safety Project: the Bristol scheme*. Crowthorne, Berks: Transport and Road Research Laboratory, 1989 (contractor report 192).
- 19 **Walker R**, Gardner G, McFetridge M. *Urban Safety Project: the Nelson scheme*. Crowthorne, Berks: Transport and Road Research Laboratory, 1989 (contractor report 191).
- 20 **Ward H**, Norrie J, Sang A. *Urban Safety Scheme: the Reading project*. Crowthorne, Berks: Transport and Road Research Laboratory, 1989 (contractor report 138).
- 21 **Ward H**, Norrie J, Sang A, et al. *Urban Safety Project: the Sheffield scheme*. Crowthorne, Berks: Transport and Road Research Laboratory, 1989 (contractor report 134).
- 22 **Fairlie RB**. *Evaluation of the safety benefits of local area traffic management*. (MengSc thesis.) Department of Civil Engineering, Monash University, 1989.
- 23 **Wentz R**, Roberts I, Bunn F, et al. Identifying controlled evaluation studies of road safety interventions searching for needles in a haystack. *J Safety Res* 2001;**32**:267–76.

Safety education of pedestrians for injury prevention: a systematic review of randomised controlled trials

Olivier Duperrex, Frances Bunn, Ian Roberts

Abstract

Objectives To quantify the effectiveness of safety education of pedestrians.

Design Systematic review of randomised controlled trials of safety education programmes for pedestrians of all ages.

Main outcome measures Effect of safety education on pedestrians' injuries, behaviour, attitude, and knowledge and on pedestrian-motor vehicle collisions. Quality of trials: methods of randomisation; and numbers lost to follow up

Results We identified 15 randomised controlled trials of safety education programmes for pedestrians.

Fourteen trials targeted children, and one targeted institutionalised adults. None assessed the effect of safety education on the occurrence of pedestrian injury, but six trials assessed its effect on behaviour. The effect of pedestrian education on behaviour varied considerably across studies and outcomes.

Conclusions Pedestrian safety education can change observed road crossing behaviour, but whether this reduces the risk of pedestrian injury in road traffic crashes is unknown. There is a lack of good evidence of effectiveness of safety education for adult pedestrians, specially elderly people. None of the trials was conducted in low or middle income countries.

Introduction

Each year about one million people die and about 10 million are seriously injured on the world's roads.¹ The World Health Organization has indicated that, for people aged 3-35 years, road traffic crashes are now the leading cause of death and disablement. The global economic burden of road traffic crashes is estimated at \$500bn (£300bn, €500bn).² Most of the casualties are in low and middle income countries, and most are vulnerable road users: pedestrians, cyclists, and riders of two wheeled motor vehicles. Children as pedestrians are particularly vulnerable, and pedestrian injuries account for most of the 280 000 childhood road deaths each year.^{1,3,4} Elderly pedestrians constitute another particularly vulnerable group.⁴

In the prevention of pedestrian injuries, educational measures to teach pedestrians how to cope with the traffic environment are considered to be an essential component of any strategy, and pedestrian education has been recommended in high, middle, and

low income countries.² Because the resources available for road safety are limited, a key question for road safety policy concerns the relative effectiveness of different prevention strategies. The aim of this systematic review of randomised controlled trials was to quantify the effectiveness of safety education programmes for pedestrians in improving their knowledge, attitudes, and behaviour and, most importantly, in preventing pedestrian-motor vehicle collisions.

Methods

Identification of trials

We aimed to identify all randomised controlled trials of road safety education programmes for pedestrians of all ages. We also included community based interventions such as media awareness campaigns and parental education programmes. We excluded studies where safety education of pedestrians was confounded by another intervention and studies that tried to modify the behaviour of drivers towards pedestrians.

We identified trials by computerised searches of the Cochrane Injuries Group specialised register, Cochrane Controlled Trials Register, Transport, Medline, Embase, ERIC, PsychLit, Spectr, and the World Health Organization's database on the internet; by checking the reference lists of relevant reviews, books, and articles; by contacting authors of relevant papers; by use of the citation analysis facility of the Science Citation Index and Social Sciences Citation Index; and by contacting relevant professionals, organisations, and voluntary agencies. No methodological filters were used, and we made no language restrictions and repeated searches with key words translated into French, German, Italian, Spanish, Dutch, and Danish.

Outcome measures and data extraction

Two reviewers independently extracted data on pedestrians' injuries, behaviour, attitude, and knowledge; pedestrian-motor vehicle collisions; methods of randomisation; and numbers lost to follow up. We assessed trial quality using the method proposed by Schulz.⁵ Disagreements were resolved by discussion with a third reviewer. When the method used to conceal allocation of intervention was not clearly reported we contacted the study author, if possible, for clarification.

Institut de Médecine Sociale et Préventive, CMU, CH-1211 Geneva 4, Switzerland

Olivier Duperrex
research fellow in community paediatrics and public health

Public Health Intervention Research Unit, London School of Hygiene and Tropical Medicine, UK

Frances Bunn
review group coordinator of Cochrane Injuries Group

Ian Roberts
professor of epidemiology and public health

Correspondence to: O Duperrex
Olivier.Duperrex@imsp.unige.ch

bmj.com 2002;324:1129

Data analysis and statistical methods

Wherever possible we performed an intention to treat analysis. Meta-analysis was not considered appropriate because of the differences across studies in the types of interventions and the types of outcomes. We calculated effect estimates with RevMan version 4.1 and report these as relative risks (95% confidence intervals) for dichotomous outcomes (relative probability of presenting the measured outcome in trained pedestrians compared with non-trained ones) and as standardised

mean difference (95% CI) for continuous outcomes. If the variance for the change score was not presented and could not be obtained from the authors, we ascribed a value using a correlation factor between pretest and post-test scores of $r=0.50$.^{6,7} We report the post-test data or the change between pretest and post-test when available, grouped by age categories and by type of outcomes. Outcomes are expressed as “positive” expected behaviour, attitude, or knowledge, so that a relative risk of >1 and a standardised mean

Table 1 Characteristics of included randomised controlled trials of safety education of pedestrians

Study and country	Participants	Allocation concealment*	Interventions	Outcomes	Loss to follow up
Ampofo Boateng et al (1993), ¹⁰ UK	26 children aged 5	C	Direct education: 1—Trained using a tabletop model of traffic environment 2—Trained in a real traffic environment 3—No training	Child's perception about safest place to cross road in real traffic situation (attitude)	37.5% for intervention group
Bouck (1992), ¹¹ UK	40 children aged 8-11	C	Indirect education provided by teachers 1—Trained in classroom and in semi-real environment 2—No training	Child's knowledge	20% for both groups
Cross et al (1988), ¹² Australia	138 children aged 7-8	C	Direct education 1—Trained in classroom during teaching unit on speed 2—No training	Child's response and documentation of verbal explanation in a play situation (tabletop model) (attitude)	Not stated
Downing et al (1981), ¹³ UK	1560 children aged 3 and their parents	A	Indirect education provided by parents 1—Road safety booklet after an interview 2—Interview but no booklet 3—Road safety booklet with a letter 4—No intervention	Child's knowledge	44% overall
Limboung et al (1981), ¹⁴ Germany	658 parents volunteered to teach their children aged 3-6	C	Indirect education provided by parents 1—Behavioural road safety training by parent with psychologist's supervision 2—Behavioural road safety training by parent without psychologist's supervision 3—Parents shown a film and given a booklet on road safety problems in childhood 4—No training	Child's behaviours in real traffic situations with and without distraction	15% overall
Luria et al (2000), ¹⁵ USA	246 children aged 5	C	Direct education 1—Trained with Safety City programme 2—No training	Child's knowledge	26% for both groups
Matson (1980), ¹⁶ USA	30 “mentally retarded” institutionalised adults aged 21-55	B	Direct education 1—Individual training in classroom using tabletop model 2—Independence training in a semi-real traffic situation 3—Training in how to cook and to make the bed	Steps performed correctly on a set of target behaviours	Not stated
Miller et al (1982), ¹⁷ USA	550 children (2nd grade)	A	Indirect education provided by teachers 1—Beltman programme 2—Beltman programme with booster course at 4 months 3—Normal safety teaching	Child's safety knowledge and behaviour	6% for knowledge test and 65% and 77% for reported behaviour
Nishioka et al (1991), ¹⁸ Japan	79 children aged 4-5	A	Direct education 1—Caution advising how to behave safely (“A motorcycle is running. If you come around here, stop and look at the right and left side, as it is dangerous”) 2—Simple caution (“A motorcycle is running. Be careful as it is dangerous”) 3—No caution (“A motorcycle is running”)	Child's behaviour	10%
Renaud et al (1989), ¹⁹ Canada	136 children aged 5	C	Direct education 1—Simulation game, targeted attitude 2—Simulation game, targeted behaviour 3—Simulation game, targeted attitude and behaviour 4—No simulation game	Child's behaviour, attitude, and knowledge	None
Singh (1979), ²⁰ UK	4024 children aged 5-13	B	Indirect education provided by teachers 1—Traffic education materials used by class teachers 2—No road safety education	Child's knowledge	7 classes in intervention, none in control group
Thomson et al (1992), ²¹ UK	30 children aged 5	C	Direct education 1—Trained in a real traffic environment 2—Trained using tabletop model of traffic environment 3—No training	Child's perception about safest place to cross road in real traffic situation (attitude)	None
Thomson et al (1997), ²² UK	201 children aged 5: 104 in year 1, 97 in year 2	C	Indirect education provided by 10 parent volunteers 1—Trained in a real traffic environment 2—No training	Child's behaviour when crossing between parked cars, and when crossing near junction. Child's perception about safest place to cross road in real traffic situation (attitude)	None
Thomson et al (1998), ²³ UK	60 children aged 5	C	Direct education 1—Trained using tabletop model of traffic environment and real traffic environment 2—No training	Child's perception about safest place to cross road in real traffic situation (attitude)	None

*Score of quality on scale used by Schulz et al⁶ assigning A to best quality and C to poorest quality: A—trials deemed to have taken adequate measures to conceal allocation (that is, central randomisation, numbered or coded bottles or containers, drugs prepared by the pharmacy, serially numbered, opaque, sealed envelopes, or other description that contained elements convincing of concealment); B—trials in which authors either did not report allocation concealment approach or reported an approach that did not fall into one of the other categories; C—trials in which concealment was inadequate (such as alternation or reference to case record numbers or to dates of birth).

difference of >0 represent a beneficial effect of the intervention programme.

In the included studies, training was provided either directly to the target population (direct education) or by training “intermediate” educators such as parents or teachers (indirect education). The way safety education is provided and the age of the target group are potential effect modifiers, but we did not explore their influence because we did not perform a meta-analysis.

For cluster randomised trials, we calculated an “effective sample size” if the intra-cluster coefficient was available.⁸ We excluded studies in which there were less than five randomised clusters because, in order to analyse at the individual level, one would have to assume that there is no clustering of individual responses within the community, which is almost always untenable.⁹

Results

We identified 13 899 studies, of which 674 (5%) were potentially relevant based on the title or abstract of the report. After a full text review, we identified 15 trials that met our inclusion criteria,^{10–23} two of which are reported in the same document.²² Table 1 shows the basic characteristics of these trials.

The methodological quality of the included trials was generally poor. The method of allocation concealment was adequate in only three trials,^{13 17 18} outcome assessment was blinded in eight,^{10 11 14–16 21–23} and in most of the studies large numbers of participants were lost to follow up. The participants were children in 14 of the studies and institutionalised adults in one.¹⁶ No trial focused on the other vulnerable pedestrian group,

elderly people. All trials were conducted in high income countries. Eight studies involved the direct education of study participants,^{10 12 15 16 18 19 21 23} and seven involved the use of parents^{13 14 22} or teachers^{11 17 20} as educators. Outcomes were measured before and after the intervention in 12 studies^{10 12–17 20–23} and only after the intervention in three studies.^{11 18 19} None of the trials assessed the effect of safety education on the occurrence of pedestrian injury, but five assessed the effect on observed behaviour,^{14 16 18 19 22} one assessed reported behaviour,¹⁷ six assessed attitude,^{10 12 19 21–23} and five assessed knowledge.^{11 13 15 17 20}

Each research group used different tools to measure outcomes, and the delay for the post-test measurement varied from less than one month to eight months. Six trials measured the effect of safety education at different times after the intervention.^{14 17 21–23} The effect of the intervention was lower in the later follow up period for 18 of the 24 behavioural outcomes,^{14 22} for two of the four attitude outcomes,^{21–23} and for the two knowledge outcomes.¹⁷

In some studies, the post-test conditions varied and influenced the results. For example, Limbourg and Gerber¹⁴ reported that 5–6 year old children given safety education were, at five months after intervention, more likely to stop and look at the line of vision when crossing roads than controls (relative probability 1.79 (95% confidence interval 1.18 to 2.72) for children without distraction). However, when the children were distracted by racing with another child the relative probability increased to 2.80 (1.39 to 5.64).

Table 2 shows the most pertinent outcomes and only the longest period to post-test measurements. (More detailed results are available in the *Cochrane Library*.) Overall, the effect of safety education on

Table 2 Selected outcomes of randomised controlled trials of safety education of pedestrians

Population	Injuries, deaths, collisions	Behaviour	Attitude	Knowledge
Children and adolescents:				
<5 years old	No RCT found	Trained children more likely to stop and look at line of vision than controls (RR 1.71 (95% CI 0.62 to 4.70)) ¹⁴	No RCT found	Trained children knew slightly more often that they had to “walk or stay on pavement” than controls (RR 1.05 (0.79 to 1.39)) ^{13*†}
5–9 years old	No RCT found	Trained children more likely to stop and look at line of vision than controls (RR 1.79 (1.18 to 2.72)) ¹⁴ Trained children more likely to stop at line of vision when crossing between parked cars than controls (RR 1.73 (1.39 to 2.14)) ^{22*} Trained children more likely to “always cross in crosswalks” according to their parents than controls (RR 1.63 (0.89 to 3.00)) ^{17*} Trained children more likely to exhibit “safe behaviour” than controls (RR 2.13 (1.01 to 4.47)) ^{18*} Trained children had better “post-test transfer score” than controls (SMD 0.83 (0.31 to 1.35)) ¹⁹	Trained children had greater proportion of routes categorised as “safe” at post-test than controls (SMD 1.28 (0.30 to 2.26)) ¹⁰ Change between pretest and post-test in proportion of routes categorised as “safe” greater in trained children than controls: (SMD 0.80 (–0.12 to 1.72)) ^{21‡} (SMD 0.17 (–0.21 to 0.55)) ^{22‡} (SMD 0.92 (0.39 to 1.46)) ^{23‡} Trained children had better “post-test attitude score” than controls (SMD 0.85 (0.35 to 1.35)) ¹⁹ Trained children more likely to apply “concept of speed” than controls (RR 1.27 (1.07 to 1.50)) ¹²	Change between pretest and post-test scores of “crossing the street” test slightly greater in trained children than controls (SMD 0.16 (–0.13 to 0.45)) ¹⁵ Change between pretest and post-test in score of “traffic safety knowledge” test was greater in trained children than controls (SMD 0.81 (0.60 to 1.02)) ^{17‡} Change between pretest and post-test scores of “cognitive” test greater in trained children than controls (5–7 year olds, SMD 0.47 (0.36 to 0.57); 7–9 year olds, 0.96 (0.85 to 1.08)) ^{20‡}
10–14 years old	No RCT found	No RCT found	No RCT found	Change between pretest and post-test scores of “cognitive” test greater in trained children than controls (SMD 0.57 (0.46 to 0.68)) ^{20‡} Trained children had better post-test score of “conspicuity, mass, speed and control” test than controls (SMD 2.39 (1.46 to 3.33)) ¹¹
15–20 years old	No RCT found	No RCT found	No RCT found	No RCT found
Adults	No RCT found	Trained institutionalised adults had higher “post-test mean proportion of steps correctly performed” than controls (RR 5.17 (3.48 to 7.67)) ^{16*}	No RCT found	No RCT found
Elderly people	No RCT found	No RCT found	No RCT found	No RCT found

RCT=randomised controlled trial (only most relevant outcomes are reported here with longest period to post-test measurements). RR=relative risk. SMD=standardised mean difference
*Intervention groups pooled. †Control groups pooled. ‡Variance of change between pretest and post-test measurements ascribed.

pedestrian behaviour varied considerably. The relative probability of trained pedestrians behaving correctly compared with controls ranged between 1.63 and 2.13 for the selected outcomes in table 2 but varied overall between 0.49 (control group performed better than trained group) and 9.29 for all the studies and outcomes (data not shown). Safety education improved pedestrians' attitude and intentions (with standardised mean differences ranging from 0.17 to 1.28) and their knowledge about road safety when outcomes were measured before and after intervention (standardised mean differences from 0.16 to 2.39), but for dichotomous outcomes the range of effect was wide (relative probability ranging from 0.72 to 1.66) (data not shown).

Discussion

Despite a thorough search in several databases in many languages and by contact with various interested parties, we could not identify good evidence of effectiveness of safety education for adult pedestrians and only limited evidence for child pedestrians. None of the included trials assessed the effect of safety education on the occurrence of pedestrian injury, but six trials assessed the effect on observed behaviour. Some of these trials showed evidence of behavioural change after safety education, but for various reasons it is difficult to predict what effect this might have on pedestrian injury risk.

Firstly, we cannot be sure that the observed behaviour is causally related to the occurrence of pedestrian injury. For example, Nishioka et al¹⁸ considered that slowing down or stopping before crossing a road to be the safe response. However, even if this behavioural change, observed in a simulated traffic environment, was repeated in a real traffic situation it is difficult to estimate what effect it would have on injury risk. Once a child has established that a road is clear, it may be safer to run across before another vehicle approaches because it reduces the time of exposure to risk. Secondly, assuming that the measured behaviours are causally related to risk of pedestrian injury, we have no reliable information about the size of this effect, and so we cannot predict how much a given behavioural change will reduce the risk of injury. Finally, there is uncertainty about the extent to which the observed behavioural changes persist over time, although the apparent declines may have been due to chance alone.

Limitations of review

Certain methodological issues could have an important bearing on the validity of our findings. In particular, publication and other selection biases may have resulted in the over-representation of studies showing promising intervention effects. This is especially likely in the context of road safety, where a large proportion of the available research information is published in the grey literature of road safety research organisations. Most of the statistical methods that can be used to assess the possibility of publication bias require the use of meta-analysis and so could not be used in this systematic review.

Although we made considerable efforts to identify all eligible trials, published and unpublished irrespective of language, we cannot exclude the possibility of

What is already known on this topic

Road traffic crashes are a leading cause of death and disablement, and pedestrians are particularly vulnerable road users

Several organisations strongly recommend road safety education

As resources are limited, a key question concerns the relative effectiveness of different prevention strategies, including road safety education of pedestrians

What this study adds

This systematic review showed safety education for pedestrians could improve children's knowledge and change their observed road crossing behaviour

However, effects on pedestrian injury were unknown

There is a lack of good evidence of effectiveness of safety education for adult pedestrians, especially elderly people, and in low and middle income countries

selection bias. The validity of the inferences from any systematic review depends on the quality of the included studies, and in this case many of the studies were of poor quality. It has been shown that inadequate allocation concealment, lack of blinding of outcome assessment, and large losses to follow up can result in the overestimation of intervention effects in randomised controlled trials,⁵ and many of these methodological weaknesses were present in the included trials.

Several included studies were conducted more than 10 years ago, and so their relevance to the current situation is open to question. Walking habits and the pedestrian environment have dramatically changed during the past two decades. All the included trials compared groups that were in the same surroundings, allowing the effect of the intervention to be isolated. Another limitation of this study is that we could not identify any randomised controlled trial conducted in low and middle income countries.

Implications of results

The Global Road Safety Partnership strongly recommends road safety education of children worldwide.²⁴ Our review indicates that there is no reliable evidence supporting the effectiveness of pedestrian education for preventing injuries in children and inconsistent evidence that it might improve their behaviour, attitudes, and knowledge. While the value of safety education of pedestrians remains in doubt, environmental modification and the enforcement of appropriate speed limits may be more effective strategies to protect children from road traffic.

Conclusions

Pedestrian safety education can improve children's knowledge of the road crossing task and can change observed road crossing behaviour, but whether this reduces the risk of pedestrian-motor vehicle collision is

unknown. No trial focused on the other vulnerable road users, elderly pedestrians. None of the trials was conducted in low and middle income countries.

Large scale, randomised controlled trials with injury outcomes (or end points that are likely to predict injury outcomes, such as near misses) are needed to establish the effectiveness of safety education of pedestrians. Although some existing trials showed evidence of behavioural change after safety education, these changes cannot be assumed to decrease pedestrian injury risk.

We thank Reinhard Wentz and Irene Kwan for help with database searching and obtaining papers; Angela Huertas, Maaïke Kruseman, Valdo Pezzoli, and Finn Johnsen for help with translation; Marjan Loep from the Dutch Cochrane Centre for help with the Dutch titles; Toshihiko Yanagawa for help with translation and contacting Japanese experts; and Kathryn Kilburn for proof reading. This review is also published in the *Cochrane Library* where it will be regularly updated to take account of new data and comments on this version.

Contributors: OD designed the protocol, searched databases, screened records, extracted data, contacted authors, and wrote the review. FB helped design the protocol, extract data, and write the review. IR helped design the protocol and write the review. OD is guarantor for the paper.

Funding: Institut de Médecine Sociale et Préventive, Geneva, Switzerland, and the Medical Research Council.

Competing interests: None declared.

- 1 Murray CJL, Lopez AD. *Global health statistics: a compendium of incidence, prevalence and mortality estimates for over 200 conditions*. Geneva: WHO, 1996.
- 2 World Bank Group. Road safety. www.worldbank.org/html/fpd/transport/roads/safety.htm (accessed 24 Nov 2001).
- 3 Rivara FP. Child pedestrian injuries in the United States. Current status of the problem, potential interventions, and future research needs. *Am J Dis Child* 1990;144:692-6.
- 4 Barss P. *Injury prevention: an international perspective epidemiology, surveillance and policy*. Oxford: Oxford University Press, 1998.
- 5 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.

- 6 Mulrow CD, Oxman AD, eds. *Cochrane Collaboration handbook*. Oxford: Update Software, 1997, Issue 4.
- 7 Follmann D, Elliott P, Suh I, Cutler J. Variance imputation for overviews of clinical trials with continuous response. *J Clin Epidemiol* 1992;45:769-73.
- 8 Donner A, Klar N. Confidence interval construction for effect measures arising from cluster randomization trials. *J Clin Epidemiol* 1993;46:123-31.
- 9 Donner A, Klar N. *Design and analysis of cluster randomization trials in health research*. London: Arnold, 2000.
- 10 Ampofo Boateng K, Thomson JA, Grieve R, Pitcairn T, Lee DN, Demetre JD. A developmental and training study of children's ability to find safe routes to cross the road. *Br J Dev Psychol* 1993;11:31-45.
- 11 Bouck LH. Development of a British road safety education support materials curriculum [thesis]. College Station, TX: Texas A&M University, 1992.
- 12 Cross RT, Pitkethly A. Speed, education and children as pedestrians: a cognitive change approach to a potentially dangerous naive concept. *Int J Sci Educ* 1988;10:531-40.
- 13 Downing CS, Murray G, Durow C. Trials of a road safety booklet for a pre-school traffic club. *TRRL Laboratory Report* 1981;LR 992:1-27.
- 14 Limbourg M, Gerber D. A parent training program for the road safety education of preschool children. *Accid Anal Prev* 1981;13:255-67.
- 15 Luria JW, Smith GA, Chapman JL. An evaluation of a safety education program for kindergarten and elementary school children. *Arch Pediatr Adolesc Med* 2000;154:227-31.
- 16 Matson JL. A controlled group study of pedestrian-skill training for the mentally retarded. *Behav Res Ther* 1980;18(2):99-106.
- 17 Miller DA, Davis L. Evaluation of Beltman traffic safety program for children. *J Traffic Safety Educ* 1982;30:13-4.
- 18 Nishioka N, Ieda S, Watanabe M, Takahashi H, Yamakawa M, Okajima Y, et al. An experimental study on the safety behavior of children in a dashing-out situation: effects of verbal instructions and traffic conditions on safety behavior. *IATSS Res* 1991;15:39-45.
- 19 Renaud L, Suissa S. Evaluation of the efficacy of simulation games in traffic safety education of kindergarten children. *Am J Public Health* 1989;79:307-9.
- 20 Singh A. Children and traffic. *Traffic Educ* 1979;4:8-12.
- 21 Thomson JA, Ampofo Boateng K, Pitcairn TK, Grieve R, Lee DN, Demetre JD. Behavioural group training of children to find safe routes to cross the road. *Br J Educ Psychol* 1992;62:173-83.
- 22 Thomson JA, Whelan KM. A community approach to road safety education using practical training methods. *Road Safety Res Rep* 1997;3:1-49.
- 23 Thomson JA, Ampofo Boateng K, Lee DN, Grieve R, Pitcairn TK, Demetre JD. The effectiveness of parents in promoting the development of road crossing skills in young children. *Br J Educ Psychol* 1998;68:475-91.
- 24 Global Road Safety Partnership. www.grsproadsafety.org/ (accessed 24 Nov 2001).

(Accepted 13 March 2002)

A systematic review of older people's perceptions of facilitators and barriers to participation in falls-prevention interventions

FRANCES BUNN*, ANGELA DICKINSON*, ELAINE BARNETT-PAGE*, ELIZABETH MCINNES† and KHIM HORTON**

ABSTRACT

The prevention of falls is currently high on the health policy agenda in the United Kingdom, which has led to the establishment of many falls-prevention services. If these are to be effective, however, the acceptability of services to older people needs to be considered. This paper reports a systematic review of studies of older people's perceptions of these interventions. The papers for review were identified by searching electronic databases, checking reference lists, and contacting experts. Two authors independently screened the studies and extracted data on the factors relating to participation in, or adherence to, falls-prevention strategies. Twenty-four studies were identified, of which 12 were qualitative. Only one study specifically examined interventions that promote participation in falls-prevention programmes; the others explored older people's attitudes and views. The factors that facilitated participation included social support, low intensity exercise, greater education, involvement in decision-making, and a perception of the programmes as relevant and life-enhancing. Barriers to participation included fatalism, denial and under-estimation of the risk of falling, poor self-efficacy, no previous history of exercise, fear of falling, poor health and functional ability, low health expectations and the stigma associated with programmes that targeted older people.

KEY WORDS – health attitudes, health-related behaviour, adherence, older people, falls prevention, systematic review.

Background

Falls are the leading cause of serious accidental injury (resulting in admission to hospital for four or more days) amongst people aged 65 or more years in the United Kingdom (Cryer 2001). Hip fractures are

* Centre for Research in Primary and Community Care, University of Hertfordshire, Hatfield, Hertfordshire, UK.

† Royal College of Nursing Institute, Oxford, UK.

** Division of Health and Social Care, University of Surrey, Guildford, Surrey, UK.

an especially grave complication of falls in older adults, and result in more hospital admissions than any other type of injury (Jensen *et al.* 1982), which during 2000 cost the National Health Service (NHS) in England around £1.7 billion (Easterbrook *et al.* 2001). There is a 10–20 per cent reduction in expected survival in the first year following a hip fracture (Cummings *et al.* 1985; Magaziner *et al.* 1989; Lu-Yao *et al.* 1994), and roughly one-half of survivors never recover normal function (Magaziner *et al.* 1989). Falling can also have serious psycho-social consequences, such as increased anxiety and depression, and can raise the fear of falling and reduce activity (Chandler *et al.* 1996; Lachman *et al.* 1998; Yardley and Smith 2002). The prevention and management of falls in older people is a key target of the Department of Health (DOH 2001), and national guidelines on the topic have recently been produced by the UK National Institute for Health and Clinical Excellence (NICE 2004).

Older people are at particular risk of falls and fall-related injuries. Physiological changes with age, such as osteoporosis, postural instability, gait disturbances, diminished muscle strength, poor vision and cognitive impairment, as well as multiple medications, are all risk factors for falling. Environmental hazards, such as steps, stairs, beds, baths, showers, lighting, loose rugs, and the absence of grab rails and banisters, have also been identified as contributing to falls (Parker, Twemlow and Pryor 1996; Lilley, Arie and Chilvers 1995; Cryer 2001). Apart from the injury sustained in the fall, there are other potential consequences, such as loss of mobility, increased dependency and disability, hypothermia, pressure-related injuries and infections (DOH 2001).

In recent years, much attention in health-promotion research and in health care has focused on falls and falls prevention among older adults. In particular, several trials and systematic reviews have investigated the effectiveness of various falls-prevention strategies, and shown that effective interventions include multi-disciplinary, multi-factorial risk-factor screening and intervention programmes, muscle strength and balance training, individually-tailored home exercise programmes, home modification, *Tai Chi* programmes, medication review, and the follow up of patients who have fallen (Cryer 2001; Easterbrook *et al.* 2001; Gillespie *et al.* 2003; Parker, Gillespie and Gillespie 2005). Previous reviews have concentrated on quantitative evaluations of effectiveness, however, and have neglected the patients' views about the acceptability of the programmes. As a result, service or guideline developers have little information by which to improve acceptability or adherence.

Aims and objectives

The aim was to undertake a systematic review of the research evidence on the barriers and facilitators which influence older people's participation in, and adherence to, falls-prevention programmes and interventions, and to identify the measures that promote acceptance. The exercise was expected to identify examples of good practice. Five review questions were to be asked of each published report:

1. What influences whether older people participate in falls-prevention programmes?
2. What factors prevent older people from taking part in falls-prevention programmes?
3. What do older people perceive to be the benefits of falls-prevention programmes?
4. What interventions are effective in promoting participation in falls-prevention programmes?
5. What are the key components of successful interventions for promoting participation in falls-prevention programmes?

The search strategy and coding

The studies of interest were those that evaluated interventions to promote adherence to, or participation in, a falls-prevention programme or strategy, and that identified the factors that influenced whether older people participated and were compliant. Further details of the inclusion criteria are provided in Table 1. We searched for all potentially relevant literature, both published and unpublished, with no date restrictions, and included relevant evidence regardless of country of origin. In order to find all potential studies we used a broad, topic-oriented approach. Methodological search filters were not used, as many non-randomised studies are not key-worded by study design (Peersman *et al.* 1998). The search terms were both free-text and 'medical subject heading' (MeSH) terms and they were combined with the appropriate Boolean operators. Details of the search terms and the databases that were searched are given in Table 2. In addition, lists of references in the selected papers were checked for otherwise unfound contributions. To identify unpublished or grey literature, we contacted field researchers and experts, including the guideline development group of NICE. The searches were conducted in January 2005.

TABLE 1. *Inclusion criteria for the systematic review*

Attribute	Intervention studies	Other studies of facilitators and barriers
Type of study	Randomised-control trials, controlled trials, controlled before/after studies	Any type, including non-intervention and qualitative. Excluding editorials/policy documents/single case studies.
Participants	Older people (mainly aged 65+ years)	Older people (mainly aged 65+ years)
Intervention	Main focus: intervention to promote participation in falls-prevention programmes	Examining barriers and facilitators to participation in programmes/interventions.
Setting	All settings (including A&E/care home/hospital)	All settings (including A&E/care home/hospital)
Outcomes	Rates of compliance/adherence with programme/strategy Predictors of compliance/adherence Participants' views and experiences of programme/strategy	Measures of/self reports of barriers/benefits of participation. Participant views/experiences of falls prevention-strategies.

Note: A&E Accident and emergency hospital departments (= casualty wards).

TABLE 2. *The search strategy*

Search terms and sequence

#1. (fall or falls or falling or faller* or fallen or slip or trips or tripped). #2. accidental falls.
 #3. femoral neck fractures. #4. (#1 or #2 or #3).
 #5. (old or older or senior* or elder* or aged or geriatric* or middleage*) #6. (#4 and #5)
 #7. (impact* or psycholog* or psychosocial or emotion or experience* or subjective* or status
 or perception* or consequence* or sequelae or effect* or meaning* or rating* or view*)
 #8. (#6 and #7)

Abstract databases searched

AMED (Allied and Complementary Medicine Database) 1985–2005
 CINAHL (Cumulative Index to Nursing and Allied Health Literature) 1982–2005
 CCTR (Cochrane Controlled Trials Register) Cochrane Library Issue 1, 2005
 DARE (Database of Reviews of Effectiveness); Cochrane Library Issue 1, 2005
 HTA (Health Technology Assessment Database) 1988–2005
 HMIC (Health Management and Information Consortium) 1983–2005. Includes the database of
 The King's Fund (London), the Department of Health Library and Information Service, and
 The Nuffield Institute for Health (Leeds).
 MEDLINE 1966–2005
 UK National Research Register (Issue 1 2005).
 NHS Economic Evaluations Database 1968–2005
 PsychInfo 1872–2005
 SIGLE (System for Information on Grey Literature) 1976–2005
 ZETOC (British Library Electronic Table of Contents) 1993–2005

All citations identified by the above searches were downloaded into an *Endnote* database. Two authors independently screened the titles and abstracts against the inclusion criteria and extracted data from the full papers

TABLE 3. Core principles of quality assessment for main study designs

Study type	Scoring strategy and criteria
Randomised-controlled trials	Quality scoring: Allocation to treatment groups concealed Study blinded, if possible All randomised participants included in the analysis (intention to treat) Withdrawals/drop-outs, reasons given for each group
Cross-sectional studies/surveys	Quality scoring: Selected subjects are representative (all eligible or a random sample) 80 per cent or more agreed to participate Exposure/outcome status ascertained in a standardised way
Qualitative studies	Assessed on seven criteria, scored as 'yes', 'no', 'partly' or 'unclear': Scope and purpose, <i>e.g.</i> clearly stated question, clear outline of theoretical framework Design, <i>e.g.</i> discussion of why particular approach/methods chosen Sample, <i>e.g.</i> adequate description of sample used and how sample identified and recruited Data collection, <i>e.g.</i> systematic documentation of tools/guides/researcher role, recording methods explicit Analysis, <i>e.g.</i> documentation of analytic tools/methods used, evidence of rigorous/systematic analysis Reliability and validity, <i>e.g.</i> presentation of original data, how categories/concepts/themes developed and were they checked by more than one author, interpretation, how theories developed, triangulation with other sources Generalisability, <i>e.g.</i> sufficient evidence for generalisability or limits made clear by author(s)

onto a specially designed form. Disagreements were resolved by discussion. Two reviewers independently assessed the quality of the selected studies using design assessment checklists. The quality criteria were informed by several sources (Higgins 2006; Thomas *et al.* 2003; Spencer *et al.* 2003) and are similar to established tools (Mays and Pope 1995; Giacomini and Cook 2000). The core quality-assessment principles are summarised in Table 3.

The papers were categorised by study design using the following categories: randomised-controlled trial, controlled trial, before/after study (with or without control), cohort study (with or without concurrent controls), case control, survey, process evaluation and qualitative study. Process evaluations were categorised in terms of the intervention's implementation, its acceptability, and the explanations given about why an intervention was successful or unsuccessful. Non-intervention studies (cohort, case-control and cross-sectional survey designs) were differentiated by whether they aimed to identify or analyse the factors that influence adherence with falls and fracture prevention, and whether they

sought older people's views about such programmes. Data were also extracted on the type, location and duration of the intervention, the characteristics of the participants and providers, the country, the main aims of the study, and the outcome measures.

The reviewed studies

The electronic searches yielded 6,191 records from all data bases, including duplicates. Of those, 134 appeared potentially relevant and a hard copy was obtained for screening. After full text review, 24 studies met the inclusion criteria. One study (Yardley and Todd 2005) was in two parts using both quantitative and qualitative methods. The remaining studies were excluded because they did not meet the inclusion criteria or because the quality was poor. Details of the studies' aims, settings and methods, of the interventions (where appropriate), and of the participants are summarised in Table 4 (further details are available on request from the authors).

Seven studies took place in the UK (Table 4, rows 1, 2, 5, 7, 8/9, 19, 20); seven in the USA (6, 10, 11, 14, 16–18); five in Australia (3, 4, 12, 23, 24); and five in Canada (13, 15, 21, 22, 25). Thirteen studies focused on people living in the community (1, 3–5, 8/9, 10–15, 19, 20); one on a combination of community dwellers and nursing/residential home residents (6); and three on people living in a continuing-care retirement village (16–18). Two studies were conducted in hospital (2, 7). Of the studies in English-speaking countries, only a few examined the health-promotion needs of non-English speaking groups (5, 21, 25). In the majority of the studies, the participants were aged 60 or more years, but in one (13) they were aged 55 or more years. The participants were variously those at high risk of falling or the 'healthy and active'. Sample sizes ranged from eight to 89 in the qualitative studies, and from 19 to 1,500 in the quantitative studies.

Only one study (18) evaluated an intervention (more exercise) explicitly to promote adherence to, or participation in, a falls-prevention programme or strategy. Two studies were process evaluations. One (1) looked at the implementation of and adherence to a nurse-led falls-prevention programme, and the other (13) examined the acceptability to older people of line-dancing and *Tai Chi* classes. The other studies all identified factors involved in older people participating or complying with falls-prevention programmes. Five of these examined older people's general views on falls and falls prevention (2, 4, 5, 6, 8), two examined specific adherence factors, such as self-efficacy (3, 12), and the rest aimed to identify the factors that

influence adherence to falls and fractures prevention interventions. The following sections of the paper present the findings about three groups of studies concerned respectively with falls and falls prevention in general ($N=8$), exercise interventions (11), and home modifications or assistive devices, *e.g.* canes (5).

The methodologies and quality of the studies

The *qualitative* studies employed methodologies that ranged from phenomenology to discourse analysis, although some did not make the method clear (Tables 5 and 6). In general, the studies had clearly defined aims, and gave adequate descriptions of the sampling and data collection, but there was limited evidence of sample validation, triangulation or assessments of generalisability. All examined people's views or knowledge of falls prevention or strategies to reduce falls, in most cases using semi-structured interviews and focus groups, and with the findings reported as themes and categories. Five examined perceptions, motivations and barriers to physical activity (Gavin and Myers 2003; Grossman and Stewart 2003; Resnick and Spellbring 2000; Sharon *et al.* 1997; Stead *et al.* 1997), and two these same reactions to home modification or assistive devices (Aminzadeh and Edwards 1998; Clemson, Cusick and Fozzard 1999), and the remainder examined more general reactions.

The *quantitative* studies used various methods to measure or review predictors of increased exercise adherence, behaviour change, falls history, fear of falling, ability and confidence, self-efficacy, participation rates, and activity levels. Overall the quality was low to fair. Some had small sample sizes (*e.g.* Hinman 1998; Resnick 2002) and some reached conclusions and recommendations that were questionably supported by the data. For example, although Cheal and Clemson (2001) was treated as a qualitative study, it also had a quantitative component, but the uncontrolled 'before and after' design and very small sample were inadequate to assess the efficacy of the intervention. In both the randomised-controlled trials, the allocation to the samples was not fully described and neither reported a sample size calculation or an intention-to-treat analysis. Only Resnick (2002) gave details of numbers lost to follow-up.

Of the surveys, two stated that they used a random sample (Bruce, Devine and Prince 2002; Edwards *et al.* 2003), and three a convenience sample (Aminzadeh and Edwards 2000; Hinman 1998; Yardley and Smith 2002). Only three studies gave the response rates (Bruce, Devine and Prince 2002; Edwards *et al.* 2003; Yardley and Smith 2002), none of

TABLE 4. Summary characteristics of included studies

Study	Authors	Aim	Method/intervention	Participants	Setting
A. Falls prevention: general					
1	Allen (1999)	Explored recruitment to and attendance at falls-prevention programme.	Part of RCT of a nurse-led falls-prevention programme (included medication review and exercise).	Recently fallen aged 65+, N=202	Community, UK
2	Ballinger (2000)	Explored perspectives on falls and falling among those with hip fracture.	Qualitative study with: semi-structured interviews/discourse analysis.	Age: 65+ (M=81) N=8 (7 f, 1 m).	Hospital, UK
3	Cheal (2001)	Evaluation of <i>Steady As You Go</i> programme (enhance self-efficacy)	Before/after study. Qualitative semi-structured interviews.	Age: 65+, N=8	Community, Australia
4	Commonwealth DHAC (2000)	Explored information needs/perceptions of falls and their prevention.	Qualitative study. 7 discussion groups, 10 in-depth interviews. Thematic analysis.	Age: 65+, and carers. N=59	Rural/metro community dwellers, Australia
5	Health Education Board for Scotland (2003)	Explored constructions of the risks of falling.	Qualitative study, 2 phases. Second phase used to validate and explore data from first phase. Individual and group interviews.	Age: 60+. Phase 1 N=39 (14 m, 25 f), Phase 2 N=50 (40 f, 10 m)	Community (rural and urban), UK
6	Hinman (1998)	Described beliefs held by older adults regarding stability, cause and control of their falls.	Survey. 8 brief closed questions (no mention of validation). Convenience sample, no information on response rate.	Age: 64-91 (M=79). N=25 (15 f, 10 m)	Community and residential care, USA
7	Simpson (2003)	Examined the precautions older people are prepared to take to prevent falls.	Qualitative, semi-structured interviews. Recorded by hand-taken notes.	Age: 65+ (M=83). N=32 (26 f, 6 m).	Hospital, UK
8	Yardley (2005) Part 1	Identified negative aspects of falls prevention communications and improving messages.	Qualitative study including focus groups and 21 individual interviews.	Age: 61-94. N=66 (41 f, 24 m)	Community, UK

9	Yardley (2005) Part 2	To determine which beliefs and feelings have greatest impact on intentions to undertake falls-prevention activities. Influence of different messages.	Questionnaires/ structured interviews on attitudes to balance training (BT), beliefs/feelings about falling. Randomised to one of four groups: (a) leaflet on positive benefits of BT, (b) leaflet describing need to do BT, (c) both leaflets; (d) no leaflet.	Age: 60–95 (M = 74.7), N = 715, 73% f	Community, UK
B. Falls prevention: exercise					
10	Boyette (1997)	Assess initiation to and adherence with a strength-training programme.	Follow-up subset for 6-months of a previous 4-month strength-training and flexibility intervention. Initiation defined: completing programme/attending 75% of sessions. Adherence defined: continuing exercises 6 months after intervention.	Age: M = 71.3, N = 46 (33 f, 13 m). Healthy older adults.	Community, USA
11	Sharon (1997)	To understand older adults' attitudes and concerns about a strength-training intervention and to identify factors that determined their adherence.	Qualitative study, 3 focus groups (1 10 f, 1 7 m, 1 of 7 m/f). Participants drawn from Boyette (1997). Excluded if depressed, cognitive impairment or cardiovascular disease.	Aged: 61–82, N = 23	Community, USA
12	Bruce (2002)	Examined whether fear of falling reduced recreational physical activity levels in healthy older women.	Cross-sectional analysis of data from RCT of calcium supplements to prevent fractures. Measured fear of falling and physical activity (validated scales). 27% response rate.	Age: 70–85 (M = 75.2). Ambulatory f, N = 1500	Community, Australia
13	Gavin (2003)	Explored participation and adherence to <i>Tai Chi</i> and line-dancing classes.	Process evaluation with qualitative and survey elements. People enrolling for <i>Tai Chi</i> /line-dancing classes. 75% reported good health.	Age: 55+, N = 328. Mostly f (85%), white (97%)	Community, Canada
14	Grossman (2003)	Explored perceptions, motivations and barriers to physical activity	In-depth interviews using open-ended questions. Details of data analysis not given.	Age: 75+ (M = 80), N = 33. Sedentary/ inactive.	Community, USA

TABLE 4 (*Cont.*)

Study	Authors	Aim	Method/intervention	Participants	Setting
15	Rejeski (1997)	Examined predictors of exercise compliance.	RCT. Compared health education, aerobic exercise and resistance exercise on self-reported disability. Knee osteoarthritis and self-reported difficulties with ADLs.	Age: M = 68.6. N = 439 (308 f, 131 m)	Unclear, Canada
16	Resnick (2000)	Explored factors influencing adherence to an exercise programme.	Qualitative study. Open-ended interviews with members of a walking group.	Age: M = 81. N = 23 (91 % f).	Continuing care retirement village, USA
17	Resnick (2001)	Tested model of exercise behaviour, explored factors influencing adherence to an exercise programme.	Survey. Assessed self-efficacy/motivation, fear of falling and health status using validated scores. Association between above factors and exercise adherence (session attendance).	Age: 65+ (M = 85). N = 201 (77 % f). No cognitive impairment.	Continuing care retirement village, USA
18	Resnick (2002)	Assessed effect of intervention on self-efficacy, exercise activity, falls and fall-related injuries.	RCT, WALC intervention (walk, address pain, fear and fatigue, learn about exercise and verbal encouragement). Control: routine care, assessment and treatment when necessary. Outcomes: exercise self-efficacy, health status, exercise behaviour and activity.	Age: M = 88. N = 20 (3 lost to follow up). Sedentary women.	Community care retirement community, USA
19	Stead (1997)	Investigated factors influencing participation in physical activity.	Qualitative study. Focus groups.	Age: 65-75+. N = 6 groups of 6-8 (3 groups aged 65-74, 3 aged 75+)	Community, UK
20	Yardley (2002)	Identified commonly-feared consequences of falling and whether these lead to activity avoidance.	Survey (random sub-sample of large RCT of Vit D). 75 % response rate. Measured fear of falling at baseline and at six months (validated scales).	Age: 75+ (M = 81). N = 224 (106 m, 118 f), Healthy people	Community, UK

C. Falls prevention: home modification/assistive devices

21	Aminzadeh (2000)	Explored factors associated with cane use.	Questionnaires. Convenience sample. Used cognitive mediator instrument – designed for this study (not clear if validated).	Age: 65+ (M=77). N=106. 48% used canes	Community, Canada
22	Aminzadeh (1998)	Views on the use of assistive devices.	Qualitative study. 4 focus groups. Taped and transcribed categories established.	Age: M=72.2. N=30 (Italian and British Canadian) 70% f	Community, Canada
23	Clemson (1999)	Examined compliance with home-hazard advice.	Qualitative study. Semi-structured in-depth ethnographic interviews	Age: 65+, N=9 (received home safety advice)	Community, Australia
24	Cumming (2001)	Examined adherence/predictors of adherence, to home modification recommendations by an OT.	Compliance study (part of RCT of OT falls-prevention intervention). 12 month follow-up.	Age: 65+ (M=76.4). N=178, 56% f	Community (recruited at hospital), Australia.
25	Edwards (2003)	Identified predictors of bathroom safety-device use (<i>e.g.</i> grab-bar/rails).	Descriptive comparative study, face-to-face interviews. Measured grab-bar use and falls history (validated scales) and used logistic regression to identify predictors of grab-bar use. 62.5% response rate.	Age: 60+ (M=73.9), 76% f. N=550 (French/English-speaking, no cognitive impairment) 32% fell previous year	Community, Canada

Notes: DHAC: Department of Health and Aged Care, Commonwealth of Australia. f: female. m: male. M=average or arithmetic mean. RCT: randomised controlled trial. Vit: Vitamin. WALC: Women against lung cancer. Full citations of the included studies are in the List of References.

TABLE 5. *Methodological approaches used in included studies*

Methodology	Studies
Qualitative (n = 12)	Aminzadeh 1998, Ballinger 2000, Cheal 2001, Clemson 1999, Commonwealth of Australia 2000, Grossman 2003, Health Education Board for Scotland 2003, Resnick 2000, Sharon 1997, Simpson 2003, Stead 1997, Yardley and Todd 2005 (<i>n.b.</i> Cheal 2001, Resnick 2000, Yardley 2005 had quantitative components).
Randomised controlled trials (n = 2)	Rejeski 1997, Resnick 2002.
Cross sectional/survey (n = 6)	Aminzadeh 2000, Bruce 2002, Edwards 2003, Hinman 1998, Resnick 2001, Yardley 2002.
Evaluation of compliance (on intervention group of an RCT/controlled trial) (n = 2)	Boyette 1997, Cumming 2001.
Process evaluation (n = 2)	Allen 1999, Gavin 2000 (included a qualitative element).

TABLE 6. *Quality assessment of the qualitative studies*

Study	Scope/ purpose	Design	Sample	Data collection	Analysis	Reliability/ validity	Generalisability/ transferability	Aggregate score ³
Aminzadeh 1998	+	~	+	+	+	+	+	6.5
Ballinger 2000	+	+	~	+	+	~	+	6.0
Butler 1998	~	+	+	+	+	-	?	4.5
Cameron 1994	+	+	+	+	~	~	~	5.5
Cheal 2001	+	-	+	+	~	~	~	4.5
Clemson 1999	+	~	+	+	+	+	?	5.5
C. of Australia 2000 ¹	~	+	+	~	-	-	~	3.5
Gavin 2003	~	~	+	~	-	~	~	3.5
Grossman 2003	~	-	+	~	~	~	+	4.0
HEBS 2003 ²	+	+	+	+	+	+	+	7.0
Resnick 2000	+	+	+	+	+	+	+	7.0
Sharon 1997	+	-	~	+	~	+	~	4.5
Simpson 2003	+	-	+	+	+	+	~	5.5
Stead 1997	+	+	~	+	~	~	~	5.0
Yardley 2005	+	+	+	~	-	-	+	4.5

Notes + yes, - no, ~ partly, ? not clear. 1. Health Education Board for Scotland. 2. Department of Health and Age Care, Commonwealth of Australia. 3. Aggregate of scores for each attribute (1 for 'yes', 0.5 for 'partly').

which were over 80 per cent. Four reported that they used validated outcome measures (Bruce, Devine and Prince 2002; Edwards *et al.* 2003; Resnick 2001; Yardley and Smith 2002).

Reactions to three types of intervention

General studies of falls and falls prevention

Four of the eight studies of older people's perspectives on falls and falling in this category used qualitative methods (Ballinger and Payne 2000; Commonwealth of Australia 2000; Health Education Board for Scotland 2003; Simpson, Darwin and Marsh 2003) and one used quantitative methods (Hinman 1998). One study (Cheal and Clemson 2001) used quantitative and qualitative methods to examine falls and self-efficacy, and one (Yardley and Todd 2005) used mixed methods to study issues around the communication of the risk of falling and prevention strategies. One study (Allen and Simpson 1999) was a process evaluation that explored recruitment to, and attendance at, a falls-prevention programme.

It emerged that the term 'falls prevention' is unfamiliar to many older people (Commonwealth of Australia 2000) and that some are unaware of the benefits of falls-prevention interventions (Simpson, Darwin and Marsh 2003). Although many people accepted that environmental and personal changes might prevent falling, they tended to advocate change for others rather than themselves (Health Education Board for Scotland 2003; Hinman 1998). Other issues around falls and falls prevention that were noted included stigma, denial, identity, the attribution of falls to external factors, and low health expectations.

Studies of exercise interventions

Eleven studies looked at factors that affect participation in, and adherence to, physical activity routines. Some examined exercise in general whereas others focused on specific exercise programmes such as strength training (Boyette, Sharon and Brandon 1997; Sharon *et al.* 1997), walking (Resnick and Spellbring 2000), and *T'ai Chi* and line dancing (Gavin and Myers 2003). One study examined whether fear of falling deterred physical activity (Bruce, Devine and Prince 2002), and a randomised controlled trial assessed the effect of an intervention to promote exercise (Resnick 2002). Many of the studies did not treat exercise as a falls-prevention intervention but as a challenge that promotes general health; those that did showed that many people are unaware of the benefits of exercise in preventing falls (Simpson, Darwin and Marsh 2003; Yardley and Todd 2005).

The factors shown to increase participation in exercise programmes were high exercise self-efficacy, past exercise history, good general health, and unimpaired functional abilities (Rejeski *et al.* 1997; Resnick and Spellbring 2000; Resnick 2001). Among the programme characteristics shown to improve adherence were: frequent bouts of activity of moderate

duration (Rejeski *et al.* 1997), accessibility, transport, convenience, having a social dimension, strong leadership (Boyette, Sharon and Brandon 1997; Gavin and Myers 2003; Sharon *et al.* 1997), and exercise tailored to individual needs and capabilities (Gavin and Myers 2003). In addition, an intervention to help people learn about exercise and overcome barriers appeared to increase overall activity and exercise at six months, but the sample size was small and the follow-up was limited (Resnick 2002). The factors shown to be associated with the avoidance of exercise included greater age (Yardley and Smith 2002), fear of falling (Bruce, Devine and Prince 2002), fear of exertion (Grossman and Stewart 2003), and discomfort such as pain or shortness of breath (Simpson, Darwin and Marsh 2003; Resnick and Spellbring 2000).

Studies of home modifications and assistive devices

Two studies explored factors in adherence to prescribed home modifications (Clemson, Cusick and Fozzard 1999; Cumming *et al.* 2001) and one examined the use of bathroom-safety devices such as grab-bars (Edwards *et al.* 2003). Three of the general studies (reviewed above) also looked at home modifications (Commonwealth of Australia 2000; Hinman 1998; Simpson, Darwin and Marsh 2003). The other two studies (Aminzadeh and Edwards 1998, 2000) examined the factors that affect the use of assistive devices such as canes. Shared findings were that many see 'home-health checks' as intrusive and unnecessary, that people dislike changes to their home, and that many have a perception of low risk (Clemson, Cusick and Fozzard 1999; Simpson, Darwin and Marsh 2003). The tendency to reject home-safety advice may be related to older people's wish to maintain their independence and control over their lives and homes (Clemson, Cusick and Fozzard 1999). Barriers to the use of canes and walking aids included stigma, embarrassment, fear of dependence, and denial of the need. Home modifications and walking aids were more acceptable than eyesight and footwear checks, medication reviews, and balance and exercise programmes (Commonwealth of Australia 2000).

Cross-cutting themes

Barriers and facilitators

The most common barriers and facilitators relating to participation and adherence to falls-prevention programmes that were identified by two reviewers in the quantitative and qualitative studies are shown in Table 7. Among the barriers to participation and concordance, one that emerged

TABLE 7. *Facilitators and barriers for older people participating in falls-prevention interventions*

Facilitators	Barriers
A. General	
Information that falls can be preventable	Fatalism/attributing falls to external causes/lack of knowledge about effectiveness of falls prevention
Communicating life-enhancing aspects of strategies, <i>e.g.</i> maintaining independence and control	Perception that physical deterioration inevitable with age
Accessible, appealing information format, from a variety of sources and in different languages	Lack of relevant information in appropriate formats/language
Choice of interventions for different people and lifestyles	Provision of 'one size fits all' advice. Advice seen as common sense/patronising
High self-efficacy	Low self-efficacy. Fear of loss of independence/risk-taking ability
Personalised modifications	No perception of need for help (no previous falls)
Emphasis on social aspects of interventions	Provoking fear of falling by using scare tactics Social stigma: association with old age/frailty Differing agenda of older people and health professionals
B. Exercise	
Previous exercise 'habit'	No previous exercise 'habit'
Making exercise fun/enjoyable/sociable	Physical discomfort/unpleasant sensations associated with exercise
Good leadership/facilitation	Underlying beliefs about personality type (<i>e.g.</i> too lazy, no willpower)
Motivation/information about physical and psychological benefits of exercise	Self perception: too old to exercise
Programmes tailored to needs or lifestyle	Poor knowledge of suitable exercises
Convenient scheduling/reasonable pricing/good access and transport	Commitment and high cost. Poor access/awareness
C. Home modifications/assistive devices	
Facilitate feeling of ownership of interventions, shared decision-making	Dislike of interventions seen as intrusive/didactic
Referral from health-care professional (especially doctor)	Stigma of devices associated with old age

strongly was fatalistic attitudes. Falls were often attributed to chance or bad luck and therefore not regarded as preventable. Several of the studies present similar quotations from their respondents that demonstrate this attitude, for example, 'I don't know how you can be told how to prevent falling. You don't do it on purpose ... it just happens' (Yardley and Todd

2005: 16). This idea extended to people's rational for not exercising. Some respondents were reported as fatalistically believing that they were basically a lazy person or lacked willpower and therefore could not exercise (Resnick and Spellbring 2000).

A widespread finding was the attribution of falls to external causes, and a complementary resistance to admitting 'intrinsic' risk factors, such as poor eyesight or dizziness (Commonwealth of Australia 2000). Instead, falls were often attributed to external causes, such as the incompetence of others or inevitable physical deterioration, over which the individual had no control (Allen and Simpson 1999; Ballinger and Payne 2000). One of many participants' quotations that illustrated this well is found in the Health Education Board for Scotland (2003: 17) study: 'You couldn't have prevented it; it was just the corner of the pavement or something. You know something that anybody of any age could do'. In contrast, Hinman (1998) reported that 68 per cent of the participants related their falls to intrinsic rather than environmental factors, but this study had a small sample and used a brief questionnaire with closed questions that may have been leading.

Some people felt falls-prevention interventions were not appropriate for them because they had low health expectations and saw physical decline as an inevitable consequence of ageing. A characteristic expression was, 'I did a lot of exercise in years gone by ... but in recent years, no[t so]. At my age I guess I don't believe it makes a difference' (Resnick and Spellbring 2000: 39). The opposite views were also reported. Evidently many older people reject the idea that they need falls-prevention advice or help because they see themselves as fit, healthy and able to manage (Allen and Simpson 1999). They may wish to distance themselves from the identity of an 'older' person, and see falls prevention as more relevant for 'older' or 'frailer' people (Yardley and Todd 2005: 13). They may, therefore, be alienated by information that is explicitly targeted at older people or that encourages participation by stereotyping older people (Aminzadeh and Edwards 1998; Ballinger and Payne 2000; Health Education Board for Scotland 2003; Stead *et al.* 1997; Yardley and Todd 2005), and dislike advice that they see as common sense and find patronising (Yardley and Todd 2005: 14). Although some fit and active older people may understandably deny the impact of ageing, some of those at high risk and of those with a history of falls also saw falls-prevention advice as relevant to others not themselves (Yardley and Todd 2005).

Related to the issue of identity is the perception of stigma. People often felt stigmatised by interventions aimed at 'older people', for example canes, walking frames and grab-bars (Aminzadeh and Edwards 2000: 300). Pride also played a part in people being reluctant to accept advice or

help (Yardley and Todd 2005: 14). Activities that allowed older people to defy stereotypes (*e.g.* intense exercise programmes) may facilitate participation (Sharon *et al.* 1997). A contradiction was found, however, between people not wanting to be seen as 'old' and distancing themselves from other 'old' people, and the fact that many valued programmes or interventions that involved contact with people of a similar age and outlook (Stead *et al.* 1997; Sharon *et al.* 1997).

Independence and risk negotiation

The studies confirm that many older people see independence as very important. Some disliked interventions that they saw as didactic or intrusive (Simpson, Darwin and Marsh 2003). They wanted to maintain their independence and make their own decisions and risk assessments (Clemson, Cusick and Fozzard 1999). A Scottish participant put this well: 'I think if you've been independent all your life, it's an embarrassment to be dependent' (Health Education Board for Scotland 2003: 27). A fear of falling can lead to a loss of confidence, adversely affect participation in daily activities, and be a barrier to participation in interventions such as exercise (Bruce, Devine and Prince 2002; Cheal and Clemson 2001; Grossman and Stewart 2003; Yardley and Smith 2002). Commonly feared consequences of falling included physical damage, loss of independence, damage to identity, and embarrassment and stigma associated with falling, particularly in a public place (Health Education Board for Scotland 2003; Yardley and Smith 2002). Fear of falling could be made worse by hazard-reduction advice that was frightening and oppressive (Yardley and Todd 2005).

Social interaction and support

For some participants who disliked group activities, the social aspect of falls-prevention interventions was a barrier (Allen and Simpson 1999), but it seems the majority of people preferred interventions with a strong social and recreational component (Allen and Simpson 1999; Boyette, Sharon and Brandon 1997; Stead *et al.* 1997; Sharon *et al.* 1997). Social support was important at several levels. Family and friends had a role in encouraging participation in, and adherence to, falls-prevention programmes (Cameron and Quine 1994; Grossman and Stewart 2003; Sharon *et al.* 1997). In addition, the desire to remain healthy and active so as not to be a 'burden' and to keep up with family and grandchildren was an incentive for some to exercise (Grossman and Stewart 2003). Support from programme leaders was also important with strong leadership a facilitator to exercise adherence (Boyette, Sharon and Brandon 1997; Sharon *et al.* 1997).

Previous experience

Previous experience of falls-prevention programmes was an important factor in acceptance of and participation in the interventions, and similarly participation in and adherence to exercise was more likely among those with a history of taking exercise (Resnick and Spellbring 2000; Resnick 2001; Stead *et al.* 1997). One study found that a 'habit' of exercise was the strongest predictor of future exercise behaviour (Rejeski *et al.* 1997). There was mixed evidence about whether having fallen previously might affect attitudes towards falls prevention. Some studies found that those who had fallen before were more likely to be receptive to falls-prevention interventions (Commonwealth of Australia 2000; Edwards *et al.* 2003), but this was not corroborated by another study (Cumming *et al.* 2001).

The role of the healthcare professional

Health-care professionals, particularly physicians, emerged as important social referents for older people (Aminzadeh and Edwards 2000; Commonwealth of Australia 2000; Grossman and Stewart 2003), although one study (Stead *et al.* 1997) found that they were not perceived as a credible source of information or advice on exercise. Home visits were sometimes seen as an intrusion but this reaction depended on the perceived authority of the person making the visits and recommendations: (Simpson, Darwin and Marsh 2003). Therapists and patients may not share the same agenda and perspectives about falls (Ballinger and Payne 2000); that is, concordance is not present, and professionals need to take into account older people's views and understand and empathise with their risk-taking behaviour (Clemson, Cusick and Fozzard 1999, Simpson, Darwin and Marsh 2003).

Discussion

The systematic review found 24 studies that examined some aspect of older people's attitudes towards falls prevention. Twelve used qualitative methods and the rest quantitative designs. The majority were exploratory studies and collected older people's views on falls and falls prevention. The studies identified a number of factors that affected participation in falls-prevention interventions and programmes. These included denial, fatalism, self-efficacy, past exercise habits, a fear of falling, general health and functional ability, health expectations, under-estimation of personal risk of falling, stigma, embarrassment, and the inconvenience of some assistive devices.

Of particular interest were those aspects of falls-prevention programmes that improved participation and adherence, and the studies provided evidence that social support and interaction, low intensity exercise (*e.g.* walking), education, and the perception that a programme was relevant and beneficial had these effects. Social support was very important in reinforcing engagement with falls-prevention interventions, both at the individual level (*i.e.* from health-care professionals, family, friends) and at the societal level (*i.e.* wider cultural norms that support the idea of older people remaining active).

Many of the themes identified by the review, *e.g.* identity, stigma, independence, denial of the ageing process and health expectations, concern the ways in which older people view themselves and believe they are seen by others. The social identity of older people is sometimes stigmatised, incorporating references to disability, disenfranchisement and other negative attributes. The way dependency is emphasised as a concomitant of old age is culturally constructed and historically located (Chater 1999; McCormack 2003). Discrimination towards older people has been described throughout the British National Health Service and social-care services (Grimley Evans 1997; Department of Health 2001), and is pervasive in society, which adds to older people's disempowerment (Bytheway 1995; Tones 1998). Those working with older people to prevent falls therefore need to be aware of, and to challenge, the factors that sustain their marginalised position (Ryles 1999).

One of the aims of the review was to assess the effectiveness of interventions used to promote the acceptance of falls-prevention strategies and to identify examples of good practice, but only one very small study that evaluated the promotion of adherence to a falls-prevention intervention was found (Resnick 2002). The majority of the reviewed studies were exploratory. The facilitators of participation were identified more often by inference or by the investigators' 'subjective synthesis' than by a statistically significant effect from a controlled study. Many of the studies examined only beliefs and attitudes, not actual behaviour. The presented evidence allows us to speculate about the key factors in successful interventions but these need further evaluation.

Several paradoxes that are challenges for those designing falls-prevention programmes emerged from the review. On the one hand, some people reject interventions that stereotype them as 'old' and wishing to avoid contact with other older people, but on the other hand, many people valued interventions that involved contact with people of a similar age and outlook. Another evident challenge is how health practitioners make people aware of their potential risk of falling without causing distress or denial of the problem.

Limitations and strengths of the review

A number of the review's methodological features could influence the validity of its results. Publication and other selection biases threaten the validity of all systematic reviews, but this is a particular problem when searching for studies that used non-randomised designs. These are more difficult to identify than randomised-controlled trials, because of the diversity of designs, the absence of standardised terminology, and the limitations of key-wording or cataloguing (Peersman *et al.* 1998). Despite our efforts to identify all eligible published and unpublished studies, we cannot exclude the possibility that some were missed. The included studies use several different methodologies. Although methods for conducting syntheses and meta-analyses of trials are well established (Egger, Davey-Smith and Altman 2001; Green and Higgins 2005), several approaches have been proposed for reviewing qualitative and non-randomised studies and they are still being developed (Campbell 2003; Dixon-Woods, Fitzpatrick and Roberts 2001; Harden *et al.* 2004).

There are also issues around quality assessment, both in terms of which quality criteria should be used and how that information should be applied to the review findings (Dixon-Woods Fitzpatrick and Roberts 2001; Sandelowski, Docherty and Emden 1997). In addition, there is no consensus on whether studies should be weighted by design and quality. Therefore, although we critically appraised the qualitative studies in the review, we did not use this information to exclude studies or weight the results. More work is needed on the tools and procedures for quality assessments of qualitative studies in systematic reviews. Despite these issues, a narrative synthesis provides a valuable overview of evidence from disparate studies, and usefully identifies the contributions of such studies to the evidence base. In particular, it collates the evidence about patient perspectives on falls prevention, which have been neglected in the development and implementation of the interventions.

The generalisability of the review findings must be considered. The studies in this review sampled older people from both community and extended-care settings. Although a few attempted to assess the efficacy of an intervention or reported associations with very few respondents, the majority had acceptable sample sizes. In addition, although some of the studies acknowledged the problems of generalisation, there was impressive consistency in the prominent themes. On the other hand, few participants were from non-English speaking backgrounds, there were few data on gender or social-class differences, and those with cognitive impairments were usually excluded.

Conclusions

Much research has been done on which interventions are effective in preventing falls and there have been systematic reviews (Cryer 2001; Easterbrook *et al.* 2001; Gillespie *et al.* 2003; Parker, Gillespie and Gillespie 2005), but the little evidence of the factors that influence participation and long-term adherence has not previously been collated. Gender and ethnicity may affect attitudes towards and participation in falls-prevention strategies but there is no research on these factors (Horton 2002). Currently, the health-care and other professionals that are developing or providing falls-prevention services have little knowledge of either older people's views or the barriers to their participation that older people perceive. Further research that raises understanding of the factors that influence older people's ambivalence towards falls-prevention interventions and that promote their continuing involvement will improve service design and effectiveness.

Acknowledgements

This work was as part of a larger project, 'Facilitators and barriers to older people accepting and complying with interventions to reduce falling and fractures', that was funded by the Department of Health for England through the *Accidental Injury Prevention Programme*. The views expressed in the publication are those of the authors and not necessarily those of the Department of Health.

References

- Allen, A. and Simpson, J. M. 1999. A primary care based falls prevention programme. *Physiotherapy Theory and Practice*, **15**, 2, 121–33.
- Aminzadeh, F. and Edwards, N. 1998. Exploring seniors' views on the use of assistive devices in falls prevention. *Public Health Nursing*, **15**, 4, 297–304.
- Aminzadeh, F. and Edwards, N. 2000. Factors associated with cane use among community dwelling older adults. *Public Health Nursing*, **17**, 6, 474–83.
- Ballinger, C. and Payne, S. 2000. Falling from grace or into expert hands? Alternative accounts about falling in older people. *British Journal of Occupational Therapy*, **63**, 12, 573–9.
- Boyette, L. W., Sharon, B. F. and Brandon, L. J. 1997. Exercise adherence for a strength training program in older adults. *Journal of Nutrition, Health and Aging*, **1**, 2, 93–7.
- Bruce, D. G., Devine, A. and Prince, R. L. 2002. Recreational physical activity levels in healthy older women: the importance of fear of falling. *Journal of the American Geriatric Society*, **50**, 1, 84–9.
- Butler, M., Coggan, C. and Norton, R. 1998. A qualitative investigation into the receptivity to hip protective underwear among staff and residents of residential institutions. *New Zealand Journal of Medicine*, **111**, 1075, 383–5.
- Bytheway, B. 1995. *Ageism*. Open University Press, Buckingham.

- Cameron, I. D. and Quine, S. 1994. External hip protectors: likely non-compliance among high risk elderly people living in the community. *Archives of Gerontology and Geriatrics*, **19**, 3, 273–81.
- Campbell, R., Pound, P., Pope, C., Britten, N., Pill, R. and Morgan, M. 2003. Evaluating meta-ethnography: a synthesis of qualitative research on lay experiences of diabetes and diabetes care. *Social Science and Medicine*, **56**, 40, 671–84.
- Chandler, J. M., Duncan, P. W., Sanders, L. and Sudenski, S. 1996. The fear of falling syndrome: relationship to falls, physical performance and activities of daily living in frail older persons. *Topics in Geriatric Rehabilitation*, **11**, 1, 55–63.
- Chater, K. 1999. Risk and representation: older people and non-compliance. *Nursing Inquiry*, **6**, 2, 132–8.
- Cheal, B. and Clemson, L. 2001. Older people enhancing self-efficacy in fall-risk situations. *Australian Occupational Therapy Journal*, **48**, 2, 80–91.
- Clemson, L., Cusick, A. and Fozzard, C. 1999. Managing risk and exerting control: determining follow through with falls prevention. *Disability and Rehabilitation*, **21**, 12, 531–41.
- Commonwealth of Australia 2000. *National Falls Prevention for Older People Initiative 'Step Out with Confidence': A Study into the Information Needs and Perceptions of Older Australians Concerning Falls and Their Prevention*. Department of Health and Aged Care, Commonwealth of Australia, Canberra.
- Cryer, C. 2001. *What Works to Prevent Accidental Injury Among Older People?* Health Development Agency, London.
- Cumming, R. G., Thomas, M., Szonyi, G., Frampton, G., Salkeld, G. and Clemson, L. 2001. Adherence to occupational therapist recommendations for home modifications for falls prevention. *American Journal of Occupational Therapy*, **55**, 6, 641–8.
- Cummings, S. R., Kelsey, J., Nevitt, M. and O'Dowd, K. 1985. Epidemiology of osteoporosis and osteoporotic fractures. *Epidemiological Reviews*, **7**, 178–208.
- Department of Health 2001. *National Service Framework for Older People*. Department of Health, London.
- Dixon-Woods, M., Fitzpatrick, R. and Roberts, K. 2001. Including qualitative research in systematic reviews: opportunities and problems. *Journal of Evaluation in Clinical Practice*, **7**, 2, 125–33.
- Easterbrook, L., Horton, K., Arber, S. and Davidson, K. 2001. *International Review of Falls in Older People*. Report for the Health Development Agency and Department of Trade and Industry (DTI), DTI, London.
- Edwards, N., Lockett, D., Aminzadeh, F. and Nair, R. 2003. Predictors of bath grab-bar use among community-living older adults. *Canadian Journal on Aging*, **22**, 2, 217–27.
- Egger, G., Davey-Smith, G. and Altman, D. (eds) 2001. *Systematic Reviews in Health Care: Meta-analysis in Context*. BMJ Books, London.
- Gavin, T. S. and Myers, A. M. 2003. Characteristics, enrolment, attendance, and dropout patterns of older adults in beginner *Tai-chi* and line-dancing programs. *Journal of Aging and Physical Activity*, **11**, 1, 123–41.
- Giacomini, M. and Cook, D. 2000. User's guide to the medical literature: XXIII. Qualitative research in health care. B. What are the results and how do they help me care for my patients? Evidence Based Medicine Working Group. *Journal of American Medical Association*, **284**, 4, 478–82.
- Gillespie, L. D., Gillespie, W. J., Robertson, M. C., Lamb, S. E., Cumming, R. G. and Rowe, B. H. Interventions for preventing falls in elderly people. *Cochrane Database of Systematic Reviews* 2003, Issue 4. Art. No.: CD000340. DOI: 10.1002/14651858.CD000340.
- Green, S. and Higgins, J. (eds) 2005. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 4.2.5 (updated May 2005), Available online at <http://www.cochrane.org/resources/handbook/hbook.htm> [Accessed June 2005].

- Grimley Evans, J. 1997. The rationing debate. Rationing healthcare by age: the case against. *British Medical Journal*, **314**, 822.
- Grossman, M. D. and Stewart, A. L. 2003. 'You aren't going to get better by just sitting around': physical activity perceptions, motivations, and barriers in adults 75 years of age or older. *American Journal of Geriatric Cardiology*, **12**, 1, 33-7.
- Harden, A., Garcia, J., Oliver, S., Rees, R., Shepherd, J., Brunton, G. and Oakley, A. 2004. Applying systematic review methods to studies of people's views: an example from public health research. *Journal of Epidemiology and Community Health*, **58**, 9, 794-800.
- Health Education Board for Scotland 2003. *The Construction of the Risks of Falling in Older People: Lay and Professional Perspectives*. Health Education Board for Scotland, Edinburgh.
- Higgins, J. P. T. and Green, S. (eds) 2006. *Cochrane Handbook for Systematic Reviews of Interventions 4.2.6*. Online Cochrane Library. Available online at <http://www.cochrane.org/resources/handbook/hbook.htm> [Accessed 6 October 2006].
- Hinman, M. R. 1998. Causal attributions of falls in older adults. *Physical and Occupational Therapy in Geriatrics*, **15**, 3, 71-84.
- Horton, K. 2002. Gender and Falls: Older People and Key Family Members' Perspectives. Unpublished PhD thesis, Department of Sociology, University of Surrey, Guildford, Surrey.
- Jensen, G. F., Christiansen, C., Boesen, J., Hegedus, V. and Transbol, I. 1982. Epidemiology of postmenopausal spinal and long bone fractures: a unifying approach to postmenopausal osteoporosis. *Clinical Orthopaedics*, **166**, 75-81.
- Lachman, M. E., Howland, J., Tennstedt, S., Jette, A., Assmann, S. and Peterson, E. W. 1998. Fear of falling and activity restriction: the survey of activities of fear of falling in the elderly (SAFE). *Journal of Gerontology: Psychological Sciences*, **53B**, 1, 43-50.
- Lilley, J., Arie, T. and Chilvers, C. 1995. Accidents involving older people: a review of the literature. *Age and Ageing*, **24**, 4, 346-65.
- Lu-Yao, G. L., Baron, J. A., Barrett, J. A. and Fisher, E. S. 1994. Treatment and survival among elderly Americans with hip fractures: a population-based study. *American Journal of Public Health*, **84**, 8, 1287-91.
- Magaziner, J., Simonsick, E. M., Kashner, T. M., Hebel, J. R. and Kenzora, J. E. 1989. Survival experience of aged hip-fracture patients. *American Journal of Public Health*, **79**, 3, 274-8.
- Mays, N. and Pope, C. 1995. Qualitative research: rigour and qualitative research. *British Medical Journal*, **311**, 109-12.
- McCormack, B. 2003. A conceptual framework for person-centred practice with older people. *International Journal of Nursing Practice*, **9**, 3, 202-9.
- National Institute for Health and Clinical Excellence (NICE) 2004. *Falls: The Assessment and Prevention of Falls in Older People: Clinical Guideline*. NICE, London.
- Parker, M. J., Twemlow, T. R. and Pryor, G. A. 1996. Environmental hazards and hip fractures. *Age and Ageing*, **25**, 4, 322-5.
- Parker, M. J., Gillespie, W. J. and Gillespie, L. D. 2005. *Hip Protectors for Preventing Hip Fractures in Older People*. Cochrane Database of Systematic Reviews 2005, Issue 3. Available online at <http://www.cochrane.org/reviews/en/ab001255.html> [Accessed 26 September 2007].
- Peersman, G., Harden, A., Oliver, S. and Oakley, A. I. 1998. *Reviews of Effectiveness in Health Promotion*. Report for the Department of Health, EPPI-Centre, London.
- Rejeski, W. J., Brawley, L. R., Ettinger, W., Morgan, T. and Thompson, C. 1997. Compliance to exercise therapy in older participants with knee osteoarthritis: implications for treating disability. *Medicine and Science in Sports and Exercise*, **29**, 8, 977-85.
- Resnick, B. 2001. A prediction model of aerobic exercise in older adults living in a continuing-care retirement community. *Journal of Aging and Health*, **13**, 2, 287-310.

- Resnick, B. 2002. Testing the effect of the WALC intervention on exercise adherence in older adults. *Journal of Gerontological Nursing*, **28**, 1, 40–9.
- Resnick, B. and Spellbring, A. M. 2000. Understanding what motivates older adults to exercise. *Journal of Gerontological Nursing*, **26**, 3, 34–42.
- Ryles, S. M. 1999. A concept analysis for empowerment: its relationship to mental health nursing. *Journal of Advanced Nursing*, **29**, 3, 600–7.
- Sandelowski, M., Docherty, S. and Emden, C. 1997. Qualitative metasynthesis: issues and techniques. *Research in Nursing and Health*, **20**, 4, 365–71.
- Sharon, B. H., Hennessy, C. H., Brandon, J. and Boyette, L. 1997. Older adults' experiences of a strength training program. *Journal of Nutrition, Health and Aging*, **1**, 2, 103–8.
- Simpson, J. M., Darwin, C. and Marsh, N. 2003. What are older people prepared to do to avoid falling? A qualitative study in London. *British Journal of Community Nursing*, **8**, 4, 152, 154–9.
- Spencer, L., Ritchie, J., Lewis, J. and Dillon, L. 2003. *Quality in Qualitative Evaluation: A Framework for Assessing Research Evidence*. Government Chief Social Researcher's Office, Cabinet Office, London. Available online at http://www.policyhub.gov.uk/docs/a_quality_framework.pdf [Accessed 26 September 2007].
- Stead, M., Wimbush, E., Eadie, D. and Teer, P. 1997. A qualitative study of older people's perceptions of ageing and exercise: the implications for health promotion. *Health Education Journal*, **56**, 1, 3–16.
- Thomas, J., Sutcliffe, K., Harden, A., Oakley, A., Oliver, S., Rees, R., Brunton, V. and Kavanagh, J. 2003. *Children and Healthy Eating: A Systematic Review of Barriers and Facilitators*. Social Science Research Unit, EPPI Centre, Institute of Education, University of London, London.
- Tones, K. 1998. Empowerment for health: the challenge. In Kendall, S. (ed.), *Health and Empowerment*. Arnold, London, 185–204.
- Yardley, L. and Smith, H. 2002. A prospective study of the relationship between feared consequences of falling and avoidance of activity in community-living older people. *The Gerontologist*, **42**, 1, 17–23.
- Yardley, L. and Todd, C. 2005. *Encouraging Positive Attitudes to Falls Prevention in Later Life*. Help the Aged, London.

Accepted 28 September 2007

Address for correspondence:

Frances Bunn, Senior Research Fellow, Centre for Research in Primary and Community Care, University of Hertfordshire, College Lane Campus, Hatfield AL10 9AB, UK.

E-mail: f.bunn@herts.ac.uk

Review 1
**Contraceptive advice and provision for the prevention of
under 18 conceptions and STIs: a rapid review**

F Bunn, F Brooks, J Appleton,
M Mead, J Magnusson, S Norton,
D Trivedi, R Wentz, W Wills

August 2006

**Centre for Research in Primary and Community Care
University of Hertfordshire**

On behalf of the National Institute for Health and Clinical Excellence

GLOSSARY OF TERMS

Adolescents

This term includes people aged 12-18

Young people

This term includes people aged 18-25

MSM

Men who have sex with men is a broad term and includes gay and bisexual men and those who have sex with men but do not identify themselves as either gay or bisexual.

Relative risk (RR)

This is the ratio of risk in two groups. In an intervention study it is the ratio of the risk in the intervention group to the risk in the control group. A risk ratio of one indicates no difference between the groups. For a desirable outcome (e.g number of STIs) a risk ratio that is greater than one indicates that the intervention was effective in reducing the risk of that outcome.

Odds Ratio (OR)

The ratio of the odds of an event in one group to the odds of an event in another group. An odds ratio of one indicates no difference between comparison groups. For a desirable outcome (e.g number of STIs) a odds ratio that is greater than one indicates that the intervention was effective in reducing the risk of that outcome.

95% Confidence Interval

This is a measure of the precision of an estimated value. For example, the confidence interval of an odds ratio tells us the boundaries within which we can be 95% certain the true value for the population falls. Moreover, if we collected 100 samples, we are saying 95 of these would give rise to an odds ratio within the boundaries of the confidence interval. Wide intervals indicate lower precision and narrow intervals greater precision.

Effect Size

This is a measure of the magnitude of the differences between two variables, also known as a treatment effect. For example, an odds ratio or a relative risk represents the size of the difference in two possible outcomes. There are many different methods of calculating effect size dependant on the properties of the data, whether it is continuous or discrete and the manner in which it is distributed.

EXECUTIVE SUMMARY

Background

Over the last ten years there has been a large increase in sexually transmitted infections (STIs) and the UK continues to have the highest rate of teenage pregnancy in Western Europe. Government policy has set targets to decrease under 18 conceptions and improve sexual health. The National Institute for Health and Clinical Excellence (NICE) has been asked by the Department of Health to develop public health intervention guidance to reduce the rate of sexually transmitted diseases (STIs), including HIV, and under 18 conceptions. The rapid review presented here is intended to assist with this guidance development by assessing the effectiveness of one to one interventions, with special reference to vulnerable and high-risk groups.

Objectives

The objectives of the review were to:

- To review the evidence of effectiveness of one to one interventions for the prevention of STI (including HIV)
- To review the evidence of effectiveness of one to one interventions for the prevention of conceptions in the under 18's

Methods

Selection criteria

We included randomised controlled trials and controlled before/after studies of one to one interventions to prevent under 18 conceptions and STIs. This included one to one interventions to provide information or education, advice, therapy, promotion of contraception or condom use, and activities to increase self confidence, self-esteem and to develop skills. Our primary outcomes were conceptions and STIs (including HIV). In addition we looked at secondary outcomes such as condom use, knowledge, number of sexual partners and general sexual risk behaviours. We also included qualitative studies that looked at the process of the interventions (e.g. how and why they do/do not work) and/or those focusing specifically on the user perspective of potential barriers and facilitators.

Data sources

We searched the following electronic databases: AMED, CINAHL, Cochrane Database of systematic Reviews, Cochrane Central Register of Controlled Trials, DARE, EMBASE, HMIC, HTA, IBSS, Psychinfo, PubMed and SIGLE from 1990-November 2005; and we handsearched reference lists from included studies.

Data extraction and quality assessment

Two reviewers independently screened electronic records, extracted data and assessed study quality using specially designed forms. Study quality was assessed using the NICE quality assessment checklists and each study was assigned a quality rating of ++ (best quality), + and – (poorest quality).

Data synthesis

Owing to the wide scope of the research question, and the heterogeneity in interventions, participants, follow up, and outcomes, an overall meta-analysis was not considered to be appropriate. If data were available we calculated relative risks (RR) with 95% confidence intervals. Data are presented in tables with an indication of whether the intervention had a positive effect (+), a negative effect (-) or no statistically significant effect (0). In addition, where possible, forest plots, without a pooled summary statistic, are presented to give a visual representation of the data. Results are presented in two sections, one for the prevention of STIs and one for the prevention of under 18 conceptions.

Main results

We found 62 studies that met our inclusion criteria. Of these, 56 were quantitative and six were qualitative studies.

Additional studies included.

Prevention of STIs

In the initial rapid review we included a systematic review of psychosocial interventions to reduce sexual risk behaviours among drug users (Van Empelen 2003). However, this was a relatively poor quality systematic review that provided little data from the original studies. Therefore, in this update we have included the original studies instead of the systematic review. This is an additional three studies (Gibson 1999, Kotranski 1998, O'Neill 1996).

Prevention of under 18 conceptions

This update of the review includes an additional five studies (Nor 2003, Olds 1997, Olds 2002, Olds 2004, Shlay 2003). These were initially excluded as they did not focus solely on under 18's but included older women as well. However, as literature in this area was scarce and as all the studies included at least 40% of under 20's we have included them in this update. Of these additional studies one looked at contraceptive care in an STI clinic (Shlay 2003) and the rest evaluated home visiting programmes for pregnant women or mothers.

The effectiveness of one to one interventions for the reduction of STIs (including HIV)

Forty-four studies evaluated one to one interventions for the prevention of STIs/HIV infection. Of those, 43 were RCTs, and one was an uncontrolled before/after study. Six studies were graded as having a low risk of bias (++), 11 as having a medium risk of bias (+) and the rest as having a high risk of bias (-). Five qualitative studies that explored barriers and facilitators to the effectiveness of one to one interventions were also included; of these three were graded as having a low risk of bias (++) and two as a medium risk of bias (+).

The quantitative studies included a variety of populations, settings, providers and types of intervention. Many of the populations included were groups at particular risk for STIs/HIV infection. For example: adolescents, MSM, black and minority groups, people with a history of a previous STI, drug users, prisoners and people with HIV. All interventions included safer sex counselling and education of some sort. In 27 studies this was based on some form of theoretical model, and 25 reported that they involved skills development. This included areas such as the development of social skills, self-esteem, self-efficacy or negotiation skills for condom use. Of the qualitative studies one was concerned with the prevention of STI, three with the prevention of HIV and one with the promotion of sexual health.

The main results are presented below:

STIs (including HIV)

(Forest Plot Figures 1, 2 & 3. Effect of one to one interventions on STIs)

Evidence Statement 1.1

In summary the evidence on the effectiveness of one to one interventions for the prevention of STIs is mixed but on balance marginally supports the interventions. There is evidence from Project RESPECT a large (++) US study (Kamb 1998) that both a two session and a four session one to one counselling intervention can reduce STIs in the long and very long term in heterosexuals, and from one (+) study that STIs in men can be reduced in the long term after one 90 minute session (Kalichman). However, the effect appears to decrease over time, with one study finding a reduction in effect after six months (Kamb 1998).

Evidence Statement 1.2

In addition EXPLORE a large (++) US study of ten session one to one counselling for MSM found a 15.7% reduction in HIV infection but this was not statistically significant (EXPLORE 2004). The other studies found no statistically significant effect on STIs but may have been underpowered for this outcome.

Evidence Statement 1.3

Interventions with adolescents appeared to be particularly effective. A subgroup analysis of Project RESPECT (Bolu 2004) found a significant reduction in sexually transmitted infections with both the four and two session interventions versus a didactic control. Although this was the only study to show a statistically significant difference the general trend in this group of studies was towards a reduction in STIs.

Condom use

(Forest Plot Figures 4&5 Effect of one to one interventions on consistent condom use.)

Evidence Statement 1.4

Twenty-five studies reported condom use, of which only eight showed a statistically significant increase in condom use in the intervention group compared to the control. However, overall there is weak evidence (that is it is mixed or conflicting but on balance marginally supports) that one to one STI/HIV prevention interventions can increase short and long-term condom use compared to control. Project RESPECT, a large good quality (++) US study found an increase in condom use in both the four and two session counselling intervention groups compared to a didactic control (Kamb 1998). However,

several studies found the effect of an intervention appears to decrease, or disappear over time. Greater uniformity is needed in the way in which condom use is measured in studies.

Unprotected sex

(Forest Plot Figures 6&7 Effect of one to one interventions on unprotected sex.)

Evidence Statement 1.5

Fifteen studies reported unprotected sex. Only six studies found a statistically significant difference between intervention and control and in general the evidence is conflicting on whether or not one to one STI/HIV interventions reduce unprotected sex. However, EXPLORE a large high quality (++) US RCT found that there was a 13.9% reduction in unprotected sex at very long term follow up after a 10 session + boosters HIV prevention counselling intervention (EXPLORE 2004). At present there seems to be support for multi-session interventions but conflicting evidence on shorter interventions.

Number of sexual partners/initiation of intercourse

Evidence Statement 1.6

Ten studies reported number of partners, initiation of intercourse, or abstinence as an outcome. No high quality studies reported this outcome; three were graded as (+) and seven as (-). Only two studies, one (+) and one (-) found a statistically significant effect (Downs 2004, Metzler 2000) and in one the effect was not maintained after 6 months (Downs 2004). In summary there is weak evidence that one-to-one interventions for the prevention of STIs/HIV are ineffective in reducing the number of sexual partners or in promoting abstinence. However, it should be noted that the interventions included in this review appeared to be designed to promote safer sexual behaviour rather than abstinence.

Risk taking behaviour/perception of risk

Evidence Statement 1.7

Seven studies measured overall risk taking behaviour (e.g. sexual risk taking scores). One (+) study set in a UK STI clinic found a significant effect on risk perception (James 1998). The remaining six (-) RCTs did not find any significant effect on risk taking behaviour or risk perception (Baker 1994, O'Neill 1996, Deas 2000, Ashworth 1994, Proude 2004, Gibson 1999, O'Neill 1996). However, three of the studies involved HIV prevention for drug users where much of the focus was on safer injecting and drug use behaviour rather than safer sexual behaviour (Baker 1994, Gibson 1999, O'Neil 1996). In

summary, there is little evidence that one to one interventions can reduce risk taking behaviour or perception of risk but the quality of studies is poor.

How does the content of the intervention (what?) influence effectiveness?

(Forest Plot Figures 9 & 10 Studies with face to face counselling effect on STIs.)

Evidence Statement 1.8

Nineteen studies compared a theory based/ skills training intervention with a more didactic control. Of those ten measured STIs (Boekeloo 1999, Boyer 1997, El-Bassel 2003, Kalichman 2005, Kamb 1998, Maher 2003, Metzler 2000, Orr 1996, Scholes 2003, Shrier 2001). In general the effects on STIs were mixed. However, Project RESPECT (Kamb 1998) a large (++) US study found that two and four session theory based interventions are more likely to be effective than a didactic control. These interventions were, however, both longer than the control. Further large scale evaluations of theory based interventions are needed to establish which components of interventions are the most effective.

Qualitative studies supported the idea of skills based interventions and found participants wanted practical and psychological strategies to increase self-efficacy for contraception and condom and safe sex negotiation (Choi 2004, Seal 2005).

Does the way that the intervention is carried out e.g. Type/mode of communication, influence effectiveness?

Evidence Statement 1.9

There was a range of types of one to one communication used. The majority of studies evaluated face to face communication between a health care professional, trained counsellor, or health educator and an individual client. Other types of communication evaluated in a few studies included computer assisted interventions, leaflets, personal diaries, and video. Three poor quality studies (-) compared a face to face intervention with a video intervention and found no statistically significant differences (Ashworth 1994, DeLamater 2000, Robert 1990), and one (+) study found no difference between face to face and telephone counselling (Rotheram-Borus 2004). Therefore, there is insufficient evidence to say whether or not face-to-face delivery is superior to other methods of delivery such as telephone, computer assisted or video based interventions. However, the majority of effective interventions involved face-to-face communication.

Does the effectiveness depend on the job title/position or other factors such as age, gender, sexuality, ethnicity, of the deliverer (leader)? What are the significant features of an effective deliverer (leader)?

Evidence Statement 1.10

Evidence from Project RESPECT a large (++) US study, which found a decrease in STIs and an increase in safe sexual behaviour, suggests that clinic staff do not need extensive experience of counselling to deliver a one to one counselling intervention, but that enthusiasm and motivation are key (Kamb 1998). In a large HIV prevention trial, which reduced HIV and unsafe sex, counsellors had 40 hours of training. Both of these studies highlight the importance of training and quality control (Kamb 1998, EXPLORE 2004). Although qualitative studies reported the importance of peers we found only one evaluation of a one to one peer led intervention. Further research is needed to evaluate different types of leaders for one to one interventions, in particular evaluating the effect of peer-led programmes.

Setting (where?). Does the site/setting of delivery of the intervention influence effectiveness?

Evidence Statement 1.11

The majority of interventions were delivered in a clinic setting of some sort, for example STI/GUM clinics, family planning clinics, primary care clinics and HIV clinics. None of the studies compared one setting with another so there is insufficient evidence to say whether the site/setting of delivery of one to one interventions influences effectiveness. However, the authors of Project RESPECT, a (++) trial which showed a counselling intervention to be effective in reducing STIs and increasing condom use, suggest that STI clinics may be appropriate places to deliver interventions as it is possible those seeking treatment for a STI may be particularly amenable to behaviour change (Kamb 1998).

Does the intensity (or length) of the intervention influence effectiveness/duration of effect?

Evidence Statement 1.12

Evidence is mixed on whether the intensity or length of one to one interventions for the prevention of STIs influences effectiveness. A (++) 10 session HIV prevention intervention for MSM found a significant reduction in unprotected sex and a reduction in HIV (EXPLORE 2004). However, longer interventions may not necessarily be better than shorter ones. A (++) study (Kamb 1998) found that both a brief two session and an enhanced four session intervention were effective in reducing STIs and increasing condom use, although the four session intervention was marginally more effective than

the two session intervention. Two studies evaluated the addition of booster sessions to an intervention. Both, Project RESPECT 2 a (++) study (Metcalf 2005) and a (-) study (Patterson 2003), found no evidence that a counselling intervention with additional booster sessions was more effective, in reducing STIs, than a counselling intervention without booster sessions.

Does the effectiveness vary with age, gender, sexuality, socio-economic status, ethnicity?

Evidence Statement 1.13

Age

A subgroup analysis of Project RESPECT (Bolu 2004) found a significant reduction in sexually transmitted infections in adolescents with both the four and two session interventions versus a didactic control. The intervention was more effective with adolescents than with other age groups. Although this was the only study with adolescents to show a statistically significant difference the general trend in this group of studies was towards a reduction in STIs.

Ethnicity

Evidence Statement 1.14

In 15 studies all or the majority of participants were black, and in the majority of the rest the populations were multiethnic. One important exception is a (++) HIV prevention study which found a 10 session counselling intervention reduced HIV and unsafe sex in MSM (EXPLORE 2004). The majority of participants in this study were white and they reported difficulty in recruiting and retaining black and Hispanic participants. In subgroup analyses of Project RESPECT (Bolu 2004) they found that a four session intervention was more effective than a two session intervention for white participants but that conversly the two session intervention was more effective than the four session intervention for black participants.

Sexuality

Evidence Statement 1.15

Project RESPECT a large (++) US study of a STI prevention intervention included heterosexuals only. They found significant reductions in STIs and an increase in condom use after a four and two session counselling intervention (Kamb 1998). EXPLORE a large high quality (++) US RCT with MSM found a non significant reduction in HIV and a 13.9% reduction in unprotected sex at very long term follow up after a 10 session + boosters HIV prevention counselling intervention (EXPLORE 2004).

THE EFFECTIVENESS OF ONE TO ONE INTERVENTIONS FOR PREVENTING UNDER 18 CONCEPTIONS?

The effectiveness of one to one interventions for the prevention of under 18 conceptions

(Forest Plot Figure 11. Effect of one to one interventions on pregnancies (includes repeat pregnancies)

We found only twelve studies that evaluated the effectiveness of one to one interventions to prevent conceptions in the under 18s. On the quality assessment score three out of eleven RCTs scored (++), three (+), and five (-), and a controlled study scored (-). In addition, we included three qualitative studies that looked at barriers and facilitators to the prevention of under 18 conceptions. Of these two were graded as (++), and one was graded as (+).

Two studies evaluated the advanced provision of emergency contraception, six looked at health care programmes for pregnant women/ mothers, two looked at contraceptive care and advice in clinics, and two looked at sexual/reproductive health education.

Of the qualitative studies two included information on the prevention of pregnancy in teenagers and one looked at sexual health promotion.

Pregnancy

Eleven studies reported data on pregnancy or repeat pregnancies.

Advanced Emergency contraception

Evidence Statement 1.16

Of the two studies (++) of advanced provision of emergency contraception, one (Gold 2004) found a trend towards a reduction in pregnancies but this was not statistically significant (Gold 2004), and the other found a non significant reduction in the pharmacy access group but not advanced provision group (Harper 2005).

Support for pregnant women/mothers

(Forest Plot Figure 12. Effect of one to one home visiting or support for pregnant women/mothers on repeat pregnancies).

Evidence Statement 1.17

Six studies evaluated interventions to support pregnant women or mothers. Although only two of the studies focused solely on adolescents (O'Sullivan 1992, Quinlivan 2003) all included at least 40% of adolescents and focused on disadvantaged, low-income women. There is good evidence that multi-session support and home visiting for disadvantaged low-income pregnant women or mothers can prevent repeat pregnancies with two (+) (Olds 2002, Olds 2004) and one (-) (O'Sullivan 1992) studies showing a significant reduction in repeat pregnancies in the intervention group compared to control. In addition one (-) study (Olds 1997) found a reduction in repeat pregnancies in poor unmarried women, although not in the sample as a whole.

Clinic based contraception care

(Forest Plot Figure 13: Effect of one to one interventions on contraception use).

Evidence Statement 1.18

One (-) RCT and one (2+) non randomised controlled study evaluated contraception advice and support in a clinic based setting (Shlay 2003, Winter 1991). One (Winter 1991) found a significant reduction in pregnancies and the other (Shlay 2003) showed a trend towards a reduction in the intervention group compared to control but this was not significant.

In summary although only four studies showed a statistically significant reduction in pregnancy (O'Sullivan 1992, Olds 2002, Olds 2004, Winter 1991) the general trend was towards a reduction. Therefore, there appears to be evidence that one to one interventions with adolescents can reduce pregnancies. Multi-session nurse home visiting appears particularly effective, especially with low-income disadvantaged women (Olds 1997, Olds 2002, Olds 2004). However, more research, is needed in this area with a focus on the under 18s and studies powered to detect a change in pregnancies.

Contraception use

Evidence Statement 1.19

Seven studies reported contraception use. This was measured in various different ways, including oral contraception, emergency contraception and condom use. Four studies showed a statistically significant effect on contraception use. Two increased oral contraceptive use. These were a (++) RCT (Quinlivan 2003) and a (+) RCT (Danielson 1990) that found one to one interventions with teenagers can improve contraception use

in the long term. Of the two (++) studies of advanced provision of emergency contraception one found an increase in the use of EC (Harper 2005) and one an increase in condom use (Gold 2004). In the other studies the general trend was towards an increase in contraception use although one (-) study found the effect on contraception use was no longer significant at 12 months (Winter 1991). Therefore, there is some evidence that one to one interventions with under 18s can increase contraception use. However, further research in this area is needed.

How does the content of the intervention influence effectiveness?

Evidence Statement 1.20

There are few studies evaluating interventions to prevent under 18 conceptions and in general there is insufficient evidence to say whether or not the content of one to one interventions influences effectiveness. However, there is good evidence that multi-session support and home visiting for disadvantaged low-income pregnant women or mothers can prevent repeat pregnancies with two (+) (Olds 2002, Olds 2004) and two (-) (Olds 1997, O'Sullivan 1992) studies showing a significant reduction in repeat pregnancies in the intervention group compared to control.

Does the way the intervention is carried out e.g type/mode of communication, influence effectiveness?

Evidence Statement 1.21

There is insufficient evidence to say whether or not the type/mode of communication of one to one interventions to prevent under 18 conceptions influence effectiveness.

Does the effectiveness depend on the job title/position or other factors such as age, gender, sexuality, ethnicity, of the deliverer (leader)? What are the significant features of an effective deliverer (leader)?

Evidence Statement 1.22

In general there is insufficient evidence to say whether or not the type of leader influences the effectiveness of one to one interventions for preventing under 18 conceptions. However, one (+) US study of home visiting for mothers (Olds 2002) found that nurses were more effective than paraprofessionals in reducing repeat pregnancies.

Setting (where?). Does the site/setting of delivery of the intervention influence effectiveness?

Evidence Statement 1.23

Most intervention were delivered in clinics or via home visiting. There is good evidence that multi-session support and home visiting for disadvantaged low-income pregnant women or mothers can prevent repeat pregnancies with two (+) (Olds 2002, Olds 2004) and one (-) (O'Sullivan 1992) studies showing a significant reduction in repeat pregnancies in the intervention group compared to control, and one (++) study an increase in reliable contraception use (Quinlivan 2003). In addition one (-) study (Olds 1997) found a reduction in repeat pregnancies in poor unmarried women, although not in the sample as a whole.

Does the intensity (or length) of the intervention influence effectiveness/duration of effect?

Evidence Statement 1.24

There is insufficient evidence that the length of clinic based one to one interventions, for the prevention of under 18 conceptions, influences the effectiveness/duration of effect. There is good evidence from one (++) study (Quinlivan 2003) two (+) (Olds 2002, Olds 2004) and two (-) studies (Olds 1997, O'Sullivan 1992) that multi-session one to one interventions may increase effective contraception use and prevent repeat pregnancies.

Does the effectiveness vary with age, gender, sexuality, socio-economic status, ethnicity?

Gender

Evidence Statement 1.25

In summary, there is insufficient evidence to say whether or not gender influences the effectiveness of one to one interventions to prevent under 18 conceptions. Most studies included in the review were aimed at females and there would appear to be a need for further research that evaluates interventions that include, or are specifically targeted at, males.

Socio-economic status

Evidence Statement 1.26

There is good evidence from one (++) study (Quinlivan 2003) two (+) (Olds 2002, Olds 2004) and two (-) studies (Olds 1997, O'Sullivan 1992) that multi-session home visiting or support can be effective in increasing effective contraception use and preventing pregnancies in low-income disadvantaged women.

Conclusions

There is evidence that one to one interventions can reduce STIs and may increase condom use and prevent unsafe sexual behaviours. However, effectiveness decreases over time. A brief US STI prevention intervention, Project RESPECT, delivered in the context of routine health services with existing staff has been shown to be effective (Kamb 1998) in reducing STIs and increasing condom use. Components of Project RESPECT included:

- Client centred intervention tailored to individual's personal risk
- Behavioural goal setting and risk reduction strategies
- Standardised training and structured protocols for clinic staff
- Quality control through observation and feedback

For MSM a multi-session intervention was shown to be more effective than the brief Project RESPECT model (EXPLORE 2004). However, this involved over 10 sessions.

One to one interventions can also improve contraception use and prevent pregnancies in the under 18's. Multi-session interventions involving home visiting appear to be particularly effective in preventing repeat pregnancies in high-risk groups.

Limitations of the review

- Lack of research, in particular there was little UK based research
- In many US studies treatment as usual or control groups received interventions which are more structured and detailed than usual care currently provided in GUM clinics in the UK which makes generalisability to the UK difficult
- Lack of objective primary outcome measures such as incidence of STIs/HIV and conceptions.
- A number of poor quality and underpowered studies

Barriers to implementation

- The provision of resources for multi-session interventions
- Recruitment and retention of participants, particularly for multi-session interventions
- Difficulty generalising current research to a UK setting

Recommendations for future research

There were a number of gaps in the evidence base identified by this review, in particular for the prevention of under 18 conceptions. Overall the effectiveness of many STI and under 18 conception prevention programmes remains in doubt. For this reason further high-quality large scale research is needed with evaluation an integral part of programmes. Areas for future research identified by the review include the following:

Prevention of STIs (including HIV)

- Evaluations aimed specifically at vulnerable groups – e.g young people in or leaving care, young people from some ethnic backgrounds, sex workers, refugees and asylum seekers
- Evaluations of interventions in the UK as most of the included studies were from the USA
- Replication, and evaluation, in the UK of successful US interventions (e.g. Project RESPECT) to evaluate applicability in the UK setting
- Studies large enough to detect a reduction in STIs/HIV infections
- Evaluations of peer-led interventions

Prevention of under 18 conceptions

- Evaluations aimed specifically at vulnerable groups – e.g. young people in or leaving care, young people from some ethnic backgrounds, refugees and asylum seekers
- Evaluations of interventions in the UK as most of the included studies were from the USA
- Studies large enough to detect a reduction in conceptions
- The development and evaluation of one to one interventions in different settings (e.g school based, clinic based)

The effects of telephone consultation and triage on healthcare use and patient satisfaction: a systematic review

Frances Bunn, Geraldine Byrne and Sally Kendall

ABSTRACT

Background

In recent years there has been a growth in the use of the telephone consultation for healthcare problems. This has developed, in part, as a response to increased demand for GP and accident and emergency department care.

Aim

To assess the effects of telephone consultation and triage on safety, service use, and patient satisfaction.

Design of study

We looked at randomised controlled trials, controlled studies, controlled before/after studies, and interrupted time series of telephone consultation or triage in a general healthcare setting.

Setting

All healthcare settings were included but the majority of studies were in primary care.

Method

We searched the Cochrane Central Register of Controlled Trials, EPOC specialised register, PubMed, EMBASE, CINAHL, SIGLE, and the National Research Register and checked reference lists of identified studies and review articles. Two reviewers independently screened studies for inclusion, extracted data, and assessed study quality.

Results

Nine studies met our inclusion criteria: five randomised controlled trials; one controlled trial; and three interrupted time series. Six studies compared telephone consultation with normal care; four by a doctor, one by a nurse, and one by a clinic clerk. Three of five studies found a significant decrease in visits to GPs but two found an increase in return consultations. In general at least 50% (range = 25.5–72.2%) of calls were handled by telephone consultation alone. Of seven studies reporting accident and emergency department visits, six showed no difference between the groups and one — of nurse telephone consultation — found an increase. Two studies reported deaths and found no difference between nurse telephone consultation and normal care.

Conclusions

Although telephone consultation appears to have the potential to reduce GP workload, questions remain about its effect on service use. Further rigorous evaluation is needed with emphasis on service use, safety, cost, and patient satisfaction.

Keywords

consultation; hotlines; review, systematic; telephone; triage.

INTRODUCTION

In recent years there has been an increase in the use of telephone consultation and triage (the process where calls from people with a healthcare problem are received, assessed, and managed by giving advice or by referral to a more appropriate service).¹ One impetus for the development of telephone consultation has been to reduce the workload of GPs and accident and emergency (A&E) departments. A&E attendances in the UK have increased,² as has demand for the service of GPs, although it has been estimated that more than half of out-of-hours calls can be handled by telephone advice alone.^{3–5}

Although some telephone consultation is done by doctors,³ much is now done by qualified nurses using computer-based clinical decision support systems. This reflects changes in the role of the nurse in recent years and the move towards nurses undertaking some tasks previously carried out by doctors. One of the largest telephone consultation systems in operation is NHS Direct; this is a 24-hour nurse-led telephone advice system, based in England, that aims to help callers self-manage problems and reduce unnecessary demands on other NHS services.^{6,7}

F Bunn, BSc(Hons), MSc, research fellow in evidence-based practice; S Kendall, BSc(Hons), PhD, RGN, RHV, professor of nursing, Centre for Research in Primary and Community Care; G Byrne, BA(Hons), PhD, RNT/PGCE, principal lecturer, School of Nursing and Midwifery, University of Hertfordshire, Hatfield.

Address for correspondence

Frances Bunn, Centre for Research in Primary and Community Care, University of Hertfordshire, College Lane, Hatfield AL10 9AB. E-mail: f.bunn@herts.ac.uk

Submitted: 10 November 2004; Editor's response: 4 February 2005; final acceptance: 22 April 2005.

©British Journal of General Practice 2005; 55: 956–961.

To date relatively little information exists on whether telephone consultation reduces pressure on other services. In Denmark, demand for home visits fell by 28% after the introduction of telephone consultation by doctors.³ In the UK there was a small decrease in the use of GP cooperatives, although no significant decrease in the use of A&E departments or ambulance services after the introduction of NHS Direct.⁸

Caller satisfaction with NHS Direct has been found to be high.^{9,10} However, it has been argued that older people, minority ethnic groups, and other disadvantaged groups underuse the service, and that it may in fact have increased, not decreased, the workload of other healthcare services.¹¹ Concerns about telephone consultation include the quality and safety of advice given;^{12–17} although other research has found it safe and effective.^{18–20} In an attempt to clarify the situation we conducted a systematic review of telephone consultation and triage services to assess their effect on safety, satisfaction, and service usage.

METHOD

Inclusion criteria

We included randomised controlled trials, controlled trials, controlled before/after studies and interrupted time series of telephone consultation or triage. This included telephone consultation, by any healthcare worker, compared with a face-to-face consultation or normal care (not including telephone consultation), or telephone consultation by one type of healthcare worker versus another (for example, nurse-led versus doctor-led telephone consultation). Disease-specific phone lines were excluded. The outcomes of interest were: mortality; adverse events; service use; calls handled by telephone alone; patient satisfaction; and cost.

Box 1. Search terms and strategy.

- ▶ 1. Triage (MeSH) all fields
- ▶ 2. Helpline* (free text)
- ▶ 3. Hotlines (MeSH) all fields
- ▶ 4. Family practice/organisation and administration (MeSH)
- ▶ 5. Emergency medicine/organisation and administration (MeSH)
- ▶ 6. #1 OR #2 OR #3 OR #4 OR #5
- ▶ 7. Telephone* (free text)
- ▶ 8. #6 AND #7
- ▶ 9. Telephone consultation
- ▶ 10. Telephone triage
- ▶ 11. NHS direct
- ▶ #8 OR #9 OR #10 OR #11

Identification of studies

We searched for published and unpublished studies using the following databases: Cochrane Central Register of Controlled Trials (Cochrane Library Issue 1 2003), specialised register of the Cochrane Effective Practice and Organisation of Care Group (EPOC) (March 2003), PubMed (1966–February 2003), EMBASE (February 2003), CINAHL (1983–February 2003), SIGLE (System for Information on Grey Literature) (1980–February 2003), and the National Research Register (Issue 2 2003). For details of the search terms used see Box 1. We checked reference lists of identified studies and review articles and contacted experts in the field. There were no language restrictions.

Data extraction and analysis

Two reviewers independently screened titles and abstracts of citations identified by the electronic search, applied the selection criteria to potentially relevant papers, and extracted data from included studies using a standardised checklist. We extracted information on participants, outcomes, and the intervention, which included the comparison, setting, service provider, use of algorithms or computer-based clinical decision support systems, and hours covered. We assessed methodological quality using the criteria of the Cochrane EPOC Group.²¹

Due to heterogeneity in study design, interventions, outcomes, and participating health professionals we did not pool studies in a meta-analysis. Instead a narrative and tabular summary of findings is presented and where possible we have reported post-intervention differences and 95%

How this fits in

In recent years there has been a growth in the use of telephone consultation for the management of healthcare problems. Previous research suggests that more than half of calls can be handled by telephone advice alone, but there have been doubts about safety, and the quality of advice given. In this review we found that telephone consultation can reduce GP visits and that, in general, 50% of calls can be handled by telephone. There was no apparent increase in adverse effects or use of other services and patients were satisfied. However, there was limited data on patient satisfaction and adverse effects.

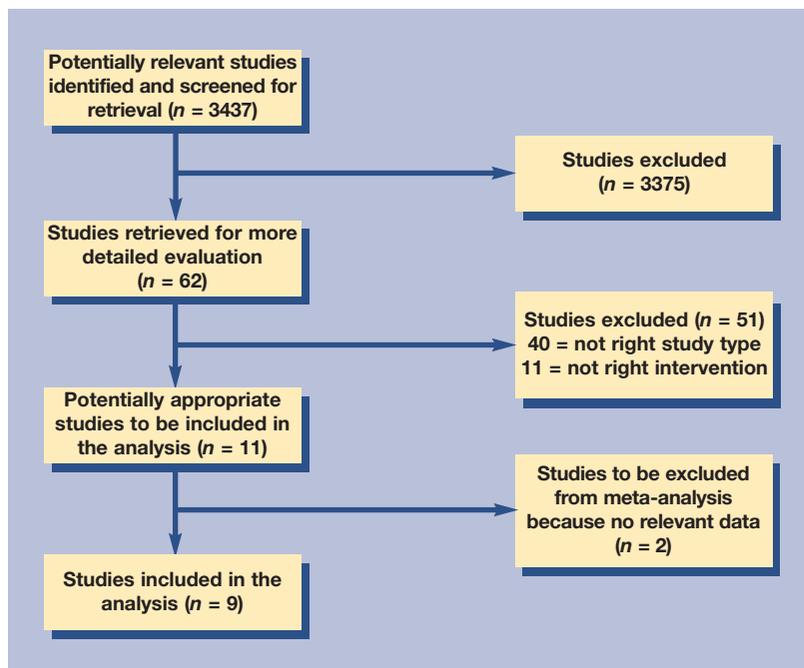


Figure 1. Flow chart showing how studies were identified.

confidence intervals (CIs) or *P*-values. For interrupted time series, where possible, we have calculated a change in the level of outcome at the first point after the introduction of the intervention, and estimated a change in the slopes of the regression line (calculated as post-intervention minus pre-intervention slope).

RESULTS

We identified 11 studies that met the inclusion criteria (Figure 1). However, two^{22,23} did not present relevant data and were excluded, leaving nine studies. Five were randomised controlled trials,^{19,24-27} one a controlled trial,²⁸ and three were interrupted time series,²⁹⁻³¹ one of which³¹ was a population-based study. Two of the randomised controlled trials^{19,27} were parallel trials using the same methodology; six were set in general practice,^{19,25,27,29-31} all except one of them³¹ in the UK. Four studies concerned out-of-hours care.^{19,24,27,31} In all studies, where a nurse delivered telephone consultation, algorithms or protocols were used. More information about individual studies can be found in Supplementary Table 1.

In the controlled studies, allocation concealment was adequate in three,^{19,25,27} inadequate in one,²⁸ and unclear in two.^{24,26} Four studies reported adequate follow up of patients^{19,25-27} and all five randomised controlled trials and the one controlled study had blinded assessment of the primary outcome. In all three interrupted time series the intervention was independent of other changes; they had blinded assessment of the primary

outcome and complete data sets. In two of the interrupted time series the data were analysed appropriately. However, in the third³¹ the researchers did not look for serial correlation and the analysis was redone using time series regression techniques. In one³¹ a change from manual to electronic recording after the start of the intervention may have led to detection bias.

When interpreting the data, it should be noted that for many of the outcomes equivalence was regarded as desirable. Researchers were normally concerned about determining whether telephone consultation or triage was as safe and effective as existing services. Numerical data from individual studies are presented in Supplementary Tables 2 and 3.

Telephone consultation versus normal care

Six studies compared telephone consultation with standard management that did not include telephone consultation (Supplementary Table 2). In four of these the telephone consultation was by a doctor,^{24,25,29,31} in one a nurse³⁰ and in one a clinic clerk.²⁶

Routine GP appointments. Of three routine GP appointments, one²⁹ found a significant reduction of 39% in GP visits ($P < 0.001$). However, one telephone consultation by a doctor²⁵ found that although same-day appointments had decreased, there was an increase of visits in the 2-week follow-up period (mean difference 0.2 [95% CI = 0.0 to 0.3]). The other telephone consultation (by a nurse)³⁰ also found that although there was a significant reduction in immediate visits (range difference [RD] -0.23 [95% CI = -0.26 to -0.20]) there was an increase in return consultations (RD = 0.32 [95% CI = 0.22 to 0.41]).

Calls handled by telephone advice alone. Calls handled by telephone advice alone ranged from 25.5% in the study of telephone consultation by nurses³⁰ to 52%²⁹ and 72%²⁵ by doctors.

Visits to A&E. With regard to visits to A&E, in the three studies of telephone consultation by a doctor, two^{24,25} found no significant difference between telephone consultation and face-to-face appointments (RD = -0.04 to 0). The other study³¹ found a significant increase in contacts with A&E but, given the constant rise in contact rates, the authors performed a regression model that showed the increase was not statistically significant. The study of nurse telephone consultation³⁰ found a significant rise in the number of visits to A&E (mean difference = 0.023 [95% CI = 0.015 to 0.032]).

Hospital admissions. When hospital admissions ($n = 2$) were compared, the trial of telephone consultation by doctors²⁴ found no significant difference between the intervention and control groups (adjusted risk difference at 2-year follow up = 0.03). However, the trial using clinic clerks to run a specialised telephone service²⁶ found a significant reduction in hospitalisations at 12 months (mean difference = 0.17; $P < 0.05$).

Home visits by GPs. One study reported the number of home visits by a GP³⁰ and found a non-significant reduction in the number of visits (RD = -0.02 [95% CI = -0.04 to 0.00]).

Out-of-hours contacts. One trial of telephone consultation by a doctor²⁵ found no difference in out-of-hours contacts between the two groups (mean difference = 0). However, the other, an interrupted time series of nurse telephone consultation,³⁰ found a significant increase in the number of out-of-hours contacts in the intervention group (mean difference = 0.04 [95% CI = 0.01 to 0.07]).

Patient satisfaction. Two randomised controlled trials compared satisfaction in intervention and control groups. One²⁵ found no significant difference in satisfaction between telephone and face-to-face consultations (difference = -8.4% [95% CI = -23.1 to 6.4%]) and the other²⁶ found that patients in the intervention group were more satisfied ($P < 0.05$). Satisfaction was high in the two other studies. In one of these²⁴ 78% of those interviewed were satisfied with length of time before the doctor responded, length of consultation, and care provided; in the other,²⁹ 98% were satisfied or very satisfied with the outcome of the telephone consultation and 84% happy to receive the service again in the future. However, the data regarding satisfaction needs to be interpreted cautiously. In one study²⁵ there was a response rate of less than 50%, and in two^{24,29} there was no comparison group — one because it was an interrupted time series²⁹ and the other because the researchers only collected data on a subset of intervention patients.

Cost. The study³⁰ that carried out an economic evaluation found little difference in cost between the intervention and control groups (mean difference = 1.48 [95% CI = -0.19 to 3.15]). In the other,²⁹ the researchers looked at cost of phone calls only, and found that telephone bills increased by 26%.

Telephone consultations compared by type of healthcare worker

Three studies compared telephone consultation by

one type of healthcare worker with another (Supplementary Table 3). Two randomised controlled trials compared nurse telephone consultation with telephone consultation by a doctor in an out-of-hours deputising service^{17,25} and one controlled trial compared telephone consultation by a health assistant with telephone advice from a doctor or a nurse.²⁶

Routine GP appointments. Two trials^{19,27} reported less GP appointments in surgery in the intervention group during the trial period. However, this was only significant in one¹⁹ (relative risk = 0.62 [95% CI = 0.58 to 0.66]).

Calls handled by telephone advice alone. In one study,¹⁹ both doctors and nurses handled 50% of calls by telephone advice alone. In the other,²⁷ 59% of calls in the nurse consultation group and 62% of calls in the GP group were managed by telephone advice alone.

Visits to A&E. All three studies found a slight increase in number of visits to A&E in the intervention group (range = 0.3–2% increase), but results were not significant.

Hospital admissions. Two studies^{19,27} found no significant difference between the intervention and control groups regarding the number of hospital admissions at 24 hours and 3 days after contact with out-of-hours services (RD at 3 days = -0.01 [95% CI = -0.02 to 0.00] and -0.02 [95% CI = -0.08 to 0.05]).

Out-of-hours contacts. Two studies²⁷ found a significant reduction in the number of home visits by the deputising service (RD = -0.06 [95% CI = -0.07 to -0.04] and -0.12 [95% CI = -0.24 to -0.11]).

Cost. In the trial with an economic evaluation,¹⁹ the cost of providing nurse telephone consultation was £81 237 a year. However, there was a reduction in overall costs of over £100 000.

Death. Neither randomised controlled trial^{19,27} found a significant difference in deaths between nurse telephone triage and triage by a doctor for patients who had been in contact with the out-of-hours service within the previous 7 days (RD = 0 [95% CI = 0.00 to 0.00] and RD = 0 [95% CI = -0.03 to 0.04]). However, one²⁷ was underpowered to detect mortality.

DISCUSSION

Summary of main findings

This systematic review found that telephone consultation and triage reduce immediate GP or

home visits and that, in general, at least 50% of calls can be handled by telephone advice alone. However, it is unclear if, in some instances, triage is just delaying visits as two studies^{25,30} showed an increase in return consultations. We found no evidence of an increase in adverse effects or use of other services and patients were satisfied. However, data on some important outcomes, in particular patient satisfaction, cost, and adverse events, were reported by few of the included studies. Initially, we felt there might be a distinction between telephone consultation and triage systems; in reality, we found that these terms were used interchangeably. The majority of studies in this review (five out of nine) were set in UK general practice.

One of the aims of this review was to compare telephone consultation by different groups of healthcare professionals. Only three of the included studies directly compared one group of healthcare worker with another.^{19,27,28} The two studies comparing nurse telephone consultation with a GP deputising service^{19,27} were good quality randomised controlled trials and found nurses could reduce GP workload without an increase in adverse events. Two of the older studies^{26,28} used unqualified staff to deliver telephone consultation and are, therefore, perhaps less relevant to present-day systems where the emphasis seems to be on consultation by qualified staff. In the other included studies the type of healthcare professional delivering the intervention did not appear to affect outcome, although one study — of nurse telephone consultation — found a small but significant increase in out-of-hours contacts and visits to A&E.³⁰ Although other uncontrolled studies have found high levels of satisfaction with nurse telephone consultation,^{11,12} we have no way of assessing this important outcome as none of the studies of nurse telephone consultation in this review reported it adequately.

Comparison with existing literature

This review supports previous estimates that at least 50% of calls can be handled by telephone advice alone.⁵⁻⁷ In addition, findings from an observational study of the impact of NHS Direct,⁸ showing that there was no decrease in the use of A&E departments but an impact on the use of GP cooperatives, are similar to the results of this review. Previous studies have highlighted the potential for errors or mismanagement with telephone consultation;¹⁵⁻¹⁶ however, few studies in this review reported adverse outcomes. The two that did^{19,27} found no increase in adverse events, although one²⁷ was underpowered to detect mortality.

Strengths and limitations of the study

We used systematic and rigorous methods to synthesise the current evidence on telephone consultation and highlight areas for further research.

However, there are a number of methodological issues that could have an important bearing on the validity of these results. Publication and other selection biases are a potential threat to validity in all systematic reviews, but this is a particular problem when searching for non-randomised studies. Non-randomised studies are more difficult to identify than those that are randomised because there is a variety of study designs, no standardised terminology, and they may not be keyworded according to study design.³² Despite our efforts to identify all eligible studies, published and unpublished, we cannot exclude the possibility that some studies were missed. In addition, no studies met all of the methodological criteria on the EPOC checklist (which may adversely affect the validity of the results) and the diversity of study types, interventions, and outcomes measured makes meaningful comparison between studies problematic.

Another limitation of this review concerns the identification of the most appropriate outcome measures. We chose service use as one of our major outcomes, as did the majority of studies included in the review. However, it could be argued that reducing service use should not be the aim of telephone consultation. Preventing patients from consulting GPs for minor illnesses is not necessarily desirable and may also discourage those with severe or treatable problems from attending — on the other hand, telephone consultation may have the potential to increase access for those who are unable or reluctant to present in person. In addition, although 50% of calls may be dealt with by telephone advice alone this does not necessarily equate to a 50% drop in workload. Indeed, a new service such as telephone consultation may attract patients who would previously have dealt with their problem without recourse to a healthcare professional. This may be a particular issue with a telephone advice and information service such as NHS Direct. Patient satisfaction and safety may, therefore, be the most important outcomes. However, there was a lack of data on both these outcomes and over half the studies in the review were randomised controlled trials, which are generally too small to detect rare adverse events.³³

Implications for clinical practice and future research

The increase in the use of telephone consultation is,

at least partially, a response to increased workloads for GPs and attempts to manage requests for same-day appointments. In addition the current government agenda is promoting the use of alternative technologies to improve access to health care. The largest telephone consultation service within the UK is now NHS Direct, which is presently staffed by qualified nurses. However, we found no controlled studies of this service that met our inclusion criteria. Therefore, although telephone consultation appears to have the potential to reduce GP workload, further rigorous evaluation is needed with emphasis on service use, safety, cost and patient satisfaction.

Supplementary information

Additional information is available online at <http://www.rcgp.org.uk/journal/supp/index.asp>

Competing interests

None

Acknowledgements

We would like to thank the Cochrane Effective Practice and Organisation of Care Group, in particular Phil Alderson and Jeremy Grimshaw for editorial support and Craig Ramsay for statistical advice.

REFERENCES

- Lattimer V, George S. Nurse telephone triage in out-of-hours primary care. *Primary Care Management* 1996; **6**(5): 3–6.
- Audit Commission. *Review of national findings: accident and emergency*. London: The Audit Commission, 2001.
- Christensen MB, Olesen F. Out of hours service in Denmark: evaluation five years after reform. *BMJ* 1998; **316**: 1502–1505.
- Dale J, Crouch R, Lloyd D. Primary care: nurse-led telephone triage and advice out of hours. *Nurs Stand* 1998; **12**(47): 41–45.
- Marsh GN, Horne RA, Channing DM. A study of telephone advice in managing out-of-hours calls. *J R Coll Gen Pract* 1987; **37**: 301–304.
- Department of Health. *The new NHS: modern, dependable*. London: The Stationery Office, 1997.
- Calman K. *Developing emergency services in the community. The final report*. London: NHS Executive, 1997.
- Munro J, Nicholl J, O’Cathain A, Knowles E. Impact of NHS direct on demand for immediate care: observational study. *BMJ* 2000; **321**: 150–153.
- Munro J, Nicholl J, O’Cathain A, et al. *Evaluation of NHS Direct first wave sites. Second interim report to the Department of Health*. Sheffield: Medical Care Research Unit, University of Sheffield, 2000.
- O’Cathain A, Munro JF, Nicholl JP, Knowles E. How helpful is NHS direct? Postal survey of callers. *BMJ* 2000; **320**: 1035.
- George S. NHS Direct audited. *BMJ* 2002; **324**: 558–559.
- Crouch R, Dale J. Telephone triage — how good are the decisions? (Part 2). *Nurs Stand* 1998; **12**(35): 33–39.
- Verdile VP, Paris PM, Stewart RD, Verdile LA. Emergency department telephone advice. *Ann Emerg Med* 1989; **18**(3): 278–282.
- Edwards B. Telephone triage: how experienced nurses reach decisions. *J Adv Nurs* 1994; **19**(4): 717–724.
- Salk ED, Schriger DL, Hubbell KA, Schwarz BL. Effect of visual clues, vital signs and protocols on triage: a prospective randomised crossover trial. *Ann Emerg Med* 1998; **32**(6): 655–664.
- Moriarty H, McLeod D, Dowell A. Mystery shopping in health service evaluation. *Br J Gen Pract* 2003; **53**: 942–946.
- Florin D, Rosen R. Evaluating NHS Direct. Early findings raise questions about expanding the service. *BMJ* 1999; **319**: 5–6.
- South Wiltshire Out of Hours Project (SWOOP) Group. Nurse telephone triage in out of hours primary care: a pilot study. *BMJ* 1997; **314**: 198–199.
- Lattimer V, George S, Thompson F, et al. Safety and effectiveness of nurse telephone consultation in out of hours primary care: randomised controlled trial. The South Wiltshire Out of Hours Project (SWOOP). *BMJ* 1998; **317**: 1054–1059.
- Poole SR, Schmitt BD, Carruth T, et al. After-hours telephone coverage: the application of an area-wide telephone triage and advice system for pediatric practices. *Pediatrics* 1993; **92**(5): 670–679.
- Cochrane Collaboration on Effective Professional Practice. The Cochrane Effective Practice and Organisation of Care Review Group. Data collection checklist. In: *The Cochrane Library*, Issue 2, Chichester, UK: John Wiley & Sons, Ltd, 2005.
- Dale J, Higgins J, Williams S, et al. Computer assisted assessment and advice for ‘non-serious’ 999 ambulance service callers: the potential impact on ambulance despatch. *Emerg Med J* 2003; **20**(2): 178–183.
- Lee TJ, Guzy J, Johnson D, et al. Caller satisfaction with after-hours telephone advice: nurse advice service versus on-call pediatricians. *Pediatrics* 2002; **110**(5): 865–872.
- Darnell JC, Hiner SL, Neill PJ, et al. After-hours telephone access to physicians with access to computerized medical records. Experience in an inner-city general medicine clinic. *Med Care* 1985; **23**(1): 20–26.
- McKinstry B, Walker J, Campbell C, et al. Telephone consultations to manage requests for same-day appointments: a randomised controlled trial in two practices. *Br J Gen Pract* 2002; **52**: 306–310.
- Stirewalt CF, Linn MW, Godoy G, et al. Effectiveness of an ambulatory care telephone service in reducing drop-in visits and improving satisfaction with care. *Med Care* 1982; **20**(7): 739–748.
- Thompson F, George S, Lattimer V, et al. Overnight calls in primary care: randomised controlled trial of management using nurse telephone consultation. *BMJ* 1999; **319**: 1408.
- Strasser PH, Levy JC, Lamb GA, Rosekrans J. Controlled clinical trial of pediatric telephone protocols. *Pediatrics* 1979; **64**(5): 553–557.
- Jiwa M, Mathers N, Campbell M. The effect of GP telephone triage on numbers seeking same-day appointments. *Br J Gen Pract* 2002; **52**: 390–391.
- Richards DA, Meakins J, Tawfik J, et al. Nurse telephone triage for same day appointments in general practice: multiple interrupted time series trial of effect on workload and costs. *BMJ* 2002; **325**: 1214–1217.
- Vedsted P, Christensen MB. The effect of an out-of-hours reform on attendance at casualty wards. *Scand J Prim Health Care* 2001; **19**(2): 95–98.
- Ernst E, Pittler MH. Assessment of therapeutic safety in systematic reviews: literature review. *BMJ* 2001; **323**: 546.
- Peersman G, Harden A, Oliver S, Oakley A. *Reviews of effectiveness in health promotion: A report for the Department of Health, England*. London: EPI-Centre, 1998.