

## *Structure and content of risk assessment proformas in mental healthcare*

CHRIS J. HAWLEY, BRIAN LITTLECHILD,  
THANUSHA SIVAKUMARAN<sup>1</sup>, HELEN SENDER, TIM M. GALE, &  
KATE J. WILSON

### *Abstract*

**Background:** The NSF specifies that mental health service providers should have a locally agreed proforma for assessing risk. Risk assessment proformas (RAPs) currently in use vary considerably in both structure and content. This study describes some similarities and differences.

**Method:** We requested a copy of each provider trust's RAP. These were assessed across a range of structural characteristics, including layout, design, data coding, and were also studied for their content.

**Results:** Some consistency was seen in the themes addressed: suicide/self harm, risk to others (notably violence) and vulnerability were well represented. Several additional themes were notable by their absence. There was general conformity in the way in which data was coded, with most using tick box categories. However, there were striking differences in layout quality and usability and many common problems were identified.

**Conclusions:** RAPs as currently used in NHS mental-health practice vary in structure, content, length and quality. We question (i) the most common approaches in recording risk data in RAPs, and (ii) whether the focus on three domains of risk acts to neglect other, less dramatic, areas. Risk assessment procedures need to be critically evaluated and the introduction of standardized tools would be advantageous in this respect.

**Declaration of interest:** This work was supported by NHS R&D Support funding awarded to Hertfordshire Partnership NHS Trust.

**Keywords:** *Risk, assessment, mental health*

### *Introduction*

Risk assessment has become a standard process in many mental-health care provider systems worldwide. In England it has acquired near-statutory force, being a matter of emphasis in such governmental documents as the National Service Framework for Mental Health (NSF: DoH, 1999). The NSF regards risk assessment as a core skill for mental health professionals and requires that all staff should be re-trained periodically. The NSF also specifies that mental health services should have a locally agreed proforma for risk assessment. What "locally" means has not been made explicit but, in practice, it has been

interpreted as meaning “within each provider Trust”. This has led to the majority of Mental Health Trusts designing and producing their own risk assessment proformas (henceforth, RAPs). As far as we are aware there has been no national guidance about the structure and content of RAPs although some non-governmental organizations have developed assessment tools suggested for general adoption, e.g., CARDS (Watts et al., 2004) CORE (Leach et al., 2005), HCR-20 (Belfrage et al., 2000).

In a previous study (Gale et al., 2002) we examined the inter- and intra-professional reliability of a simple RAP when applied to the assessment of real service users with enduring mental health problems. Irrespective of the professional background of the rater, the level of pair-wise rater agreement was only moderate and, in some cases, rather poor. Moreover, in a later study (Gale et al., 2003), we found that mental health professionals had profound difficulties with some simple arithmetic and probabilistic concepts that underlie risk measurement. These studies, amongst others (e.g., Montandon & Harding, 1984; Harriss & Hawton, 2005) raise important questions about the reliability and validity of risk assessment in current mental health practice.

The stimulus to the current study was that we wished to further investigate reliability and validity of assessments using RAPs that have been designed following NSF recommendations. But our initial enquiries showed that RAPs currently in use by different providers vary considerably in both structure and content. Hence it would be difficult to conduct a meaningful study of reliability and validity because findings based upon one RAP may not generalize to others. The purpose of the current study was, therefore, to describe the structure and content of current RAPs and ascertain to what extent they vary across mental healthcare service providers.

Hitherto there has been just one English study examining this issue in detail (Higgins et al., 2005). Higgins et al. asked consultants working in general adult psychiatry in 66 randomly selected mental healthcare providers to provide copies of their trust’s RAP and to answer a series of questions about risk assessment documentation, training and guidance. Data were obtained from 45 providers and a content analysis was performed on the RAPs. The authors report evidence of “striking variation” (p. 133): RAPs varied considerably in their content and complexity and, in some cases, the rationale behind scoring or grading was unclear and may have been a barrier to further communication of risk information. Although this study focused specifically on the assessment of violence risk (for example, risk of self harm or suicide was not discussed), the general conclusion was that “Current risk assessment practice is highly variable, indicating a lack of consensus about suitable methods” (p. 131). In this study, we adopt a broadly similar methodology but examine a wider range of risk issues and provide a more detailed assessment of RAP structure and grading.

The objectives in comparing and contrasting RAPs are as follows: (i) to suggest whether it is feasible to consider reliability and validity studies in the future; (ii) to allow Trusts to compare their own methods against those of other trusts, and (iii) to suggest whether there might be a consensus format that could be adopted as a template or, at least, whether future work might proceed in such a direction.

## *Methods*

### Obtaining Risk Assessment Proformas (RAPs)

We telephoned the risk manager of all the NHS trusts in England that provide secondary care mental health services (n=483) to explain the nature of our study and to request a copy of their current RAP. It was explained that all submitted RAPs would be carefully anonymized by our research assistant on receipt, so the study team would not be able to attribute any RAP to a specific service provider. This conversation was followed up by sending a written study information sheet and SAE to each risk manager. The Central Office of Research Ethics Committees (COREC) confirmed that ethical approval was not required for the study.

We envisaged that each provider might have several RAPs for different purposes (e.g., basic vs. advanced, screening vs. comprehensive, or perhaps a RAP that was applicable only to a particular sub-group of service users). We therefore asked risk managers to send us the RAP that most closely conformed, in their opinion, to the following definition: “that RAP which is the first applied to a person aged between 18 and 65 years, who is receiving medical-psychiatric attention and the attention of at least one other mental health professional, for a general-adult psychiatric disorder in a non-hospitalized setting”. The intention here was to focus attention on the RAPs that are used most generally in mental healthcare. We did not ask to see trust policies or training documentation that may have been associated with the RAP. Neither did we ask for risk managers’ opinions about the quality or appropriateness of their Trust’s RAP, nor their personal experiences in using the RAPs. Our focus of attention was purely on the operating properties of the RAPs themselves: attempts to understand individual conceptualization of risk or policy position would have been an overwhelming undertaking beyond our resources.

All submitted RAPs were duly anonymized by removing any identifying details (e.g., trust names, addresses, phone numbers, logos, footers, etc.). This information was cut out manually and the edited forms were photocopied for circulation to the study team. For administrative purposes, our research assistant kept a master list of those providers who had returned RAPs but this list was not made available to the rest of the study team.

### *Evaluation of RAPs*

Data collection procedures were defined clearly in advance and were not amended during RAP analysis. These included, as much as possible given that some data extraction was qualitative in nature, definitions of the properties against which the RAPs were evaluated. These properties (P1 – P12) are outlined fully below. These were chosen to reflect a range of physical, psychometric and qualitative properties of the RAPs. The physical properties were objective measures that described the length and structure of the RAP. The psychometric properties were the means by which information was encoded in the RAPs and the qualitative properties reflected the content of the RAPs. We tried to use objective, verifiable, measures wherever possible. These properties have some overlap with those used by Higgins et al. (2005) but permit a more detailed analysis.

P1: The number of A4 sides (a part filled side counted as one whole side).

P2: The total number of items. Any prompt requiring a response was counted as an item.

This included any tick box response (e.g., Is the patient currently experiencing hallucinations Y/N?) or a free-text field (e.g., Describe the patient’s current mental state).

P3: The distribution of response types, classified as follows:

- (a) Forced-choice dichotomy (e.g., yes/no, present/absent).
- (b) Forced-choice dichotomy with censored option (e.g., yes/no/not known, present/absent/insufficient info).
- (c) Ordinal responses (e.g., high, medium, low; or never, sometimes, frequently).
- (d) Interval responses (i.e., frequency counts, e.g., number of previous overdoses).
- (e) Free-text responses (e.g., Describe the patient's current mental state).

Risk assessment proformas in mental health 439

P4: The number of pages carrying an adequate identifier. Consistent with usual NHS practice, we regarded an adequate identifier as being the patient name plus one other identifying characteristic (e.g., DOB or hospital number).

P5: The number of significant typographical errors. These are errors that could potentially obscure, or render ambiguous, the meaning of a statement within a RAP. An example might be something like "risk of sever self-neglect". Minor typos and grammatical errors that would not lead to potential misinterpretation were not counted. An example would be something like "more than one incidents".

P6: Those items that captured historical features were classed as period-invariant or period dependent.

For example, the prompt "is there a history of homelessness?" would be classed as period-invariant because the fact that a service user has ever been homeless can never change. A period-dependent version might be "has the person been homeless within the past 5 years?" (i.e., the answer is dependent on the passage of time and can be updated).

P7: The visual appearance of the RAP was rated according to the following instruction: "Consider the visual appearance and appeal of the RAP. Do not consider the content of the text. Make a rating from 1 – 5 using the scale below. The rating should take into account variables such as legibility, consistency of layout, suitability of fonts and font size, sufficient space for responses, and unnecessary duplication".

- (1) Very poor layout: unattractive, hard to follow, hard to read, appears confusing; the quality of the layout would likely impede completion.
- (2) Poor layout: has significant deficiencies, may require extra effort on the user's part, but the layout is not a major impediment to use.
- (3) Satisfactory layout: the RAP may not be attractive, appealing or elegant but is free of obvious deficiencies.
- (4) Good layout: The layout is clear and precise, visually appealing, any flaws will be trivial in nature.
- (5) Very good layout: The RAP is not only free from any flaws, but looks highly professional and appears actively attractive to the user.

P8: The number of instructional statements were counted in each RAP. These were defined as statements that directed the completer to undertake a clinical action (e.g., "If high risk, then proceed to Mental Health Act assessment"). We did not count directives about the process of care (e.g., "If high risk, file form B to the CPA office"). Nor did we count items in risk management plans since these sections were a consequence of risk assessment, rather than part of the assessment process.

P9: We identified those examples where the RAP instructions invited completion of more detailed documentation depending upon information recorded in the initial RAP. This helped us to identify whether the RAP was considered to be a sufficient process of evaluation

on its own or, rather, a screening document that would necessitate further paperwork in cases where sufficient concern was raised.

P10: We assessed whether completers were asked to express an opinion, either in words or numbers, about the likelihood of a future state of affairs, for example, the chance of suicide, re-hospitalization, or violence. In each RAP we counted prompts that called upon the user to express such a predictive view. Our coding here was liberal: even a prompt as vaguely worded as “is there a high risk of self-neglect?”, was counted as a predictive statement. We noted whether such predictions required a lexical response (e.g., low risk, medium risk, high risk) or a numerical response (e.g., 10%, 20%, etc.) and also whether the prediction was time-limited (e.g., “over the next month”) or not.

P11: A Content Analysis of RAP items and instructions was performed. We did not know in advance what the common themes might be so the grouping of content by theme was necessarily a post-hoc qualitative process.

P12: Lastly, two comments about the RAP, one positive and one negative, were noted. Each of these was less than 25 words and the objective here was to record additional information that was worthy of remark.

Each RAP was evaluated by a pair of raters and their consensus view was recorded. The pairs were drawn from the following authors, their professional backgrounds being, psychiatry (CJH), experimental psychology (TMG), and health and social care (BL and HS).

We acknowledge the potential subjectivity associated with some of the above rating criteria. We minimized this by using rater pairs rather than individuals and by making group ratings (i.e., all 4 raters) on a subset of RAPs at beginning and middle of the study, to ensure that individuals were approaching the evaluation in a similar way. Nonetheless, we acknowledge the potential for subjective bias in this kind of study.

## *Results*

### *RAPs returned*

We contacted 83 provider organizations and 53 (64%) of these returned RAPs to us after one or more reminders. This rate of return is broadly similar to that found by Higgins et al. (2005). We did not explore reasons for non-returns.

### *RAP size and number of items (P1, P2)*

All returned RAPs varied in length between 1 and 6 sides (mean 2.2). The number of items ranged between 5 and 148 (mean 48+36.5, median 44). Table I displays the distribution of number of items.

### *Item structure (P3)*

Most of the RAPs in our study used forced-choice dichotomy responses (e.g., Yes/No, or tick-boxes). Summing across all RAPs there were some 2632 items of which 2217 (84.2%) used this format. Of these, 1508 (68%) had no censored option (i.e., responses took the form Yes/No rather than Yes/No/Don't know). Across all RAPs surveyed, 197 (7.5%) items were free-text responses, 216 (8.2%) items used a form of ordinal coding and 2 (50.1%) items used interval coding (i.e., frequency counts or measurements).

Table I. Distribution of “number of items” across 53 RAPs.

No. Items No. RAPS

1–30 19

31 – 60 17

61 – 90 8

91 – 120 7

121 – 150 2

Risk assessment proformas in mental health 441

In general, the RAPs with the greatest number of items were those relying on forced choice type questions while the shorter proformas with fewer items tended to use a greater number of free-text responses (i.e., they relied more on written accounts of salient detail).

Adequate identifier (P4)

Twenty-six (49%) of returned RAPs included one or more pages that did not have fields for adequate identifying information. Of these, most proformas carried adequate fields on page 1 but not on subsequent pages. However, in 5 (9.4%) cases, there were no fields for identifying information whatsoever.

Typographical errors (P5)

Ten RAPs (19%) contained at least one significant typographical error. Examples were typically of the form “has this person never harmed themselves”, when the context obviously would imply “ever”. We are unable to report these errors in detail since it would compromise anonymity.

Period invariance and dependence (P6)

Questions were generally of a period-invariant type (i.e., they did not specify any time period during which an action or event may have occurred). Across all RAPs, 68% questions were of this nature. In fact, very few of the period-dependent questions (i.e., the remaining 32%) went so far as to specify a precise time period. Nonetheless, if the question took a form such as, “does the patient currently have a problem with medication compliance”, we counted this as a period dependent question because it clearly relates to the here and now, rather than a previous period in the patient’s life.

Look and feel (P7)

Table II presents our subjective visual rating scores ascribed to the 53 RAPs. Six RAPs (11%) attracted the worst rating while none attracted the highest score (though it must be borne in mind that the criteria for a score of “high” were very exacting).

Problems in the layout and presentation of RAPs were common. Typical examples included (i) free-text boxes that were too small to accommodate a response, (ii) long columns of tick-boxes that were not aligned to their corresponding prompt statements, (iii) inconsistency in the order of presentation of items (e.g., working across the page for some, and down the page for other items), and (iv) use of multiple or inconsistent typefaces.

Table II. Distribution of ratings for visual “appeal” of RAPs.

Rating Category N %

1 Very poor 6 11

2 Poor 16 30

3 Satisfactory 21 40

4 Good 10 19

## 5 Very Good 0 0

### Directing further action (P8)

We envisaged that RAPs might direct the completer towards clinical actions following completion (for example, if the service user was found to be psychotic with a history of violence then admission to hospital might be advisable). However, this did not turn out to be the case. Although 18 (34%) of RAPs included one or more statements about the process of care, there were no clear instances where a specific clinical action was directed. A typical example was similar to: “‘if judged medium or high suicide risk, discuss with consultant’” (i.e., when directions were stated they concerned the processes surrounding care, not the care itself).

### Making predictions (P9)

To a large extent, RAP items recorded features that might inform judgement about risk as opposed to recording what the judgement about risk was. Fifty-seven per cent of RAPs assessed in this study did not include any items where the completer was required to comment on a future state of affairs. A hypothetical example of such an item might be, “‘based upon this assessment what is the chance of a suicidal act – high, medium or low?’”. Moreover, of those RAPs that did invite statements about future events, only 2 (4%) specified a period in question (e.g., “‘what is the risk of suicide in the next three months?’”). Thus the concept of risk as the likelihood of an outcome in a specified period of time does not appear to be a central consideration to the majority of RAPs considered here.

### More advanced RAPs (P10)

Twenty-two RAPs (42%) specifically mentioned that, if warranted by the assessment so far, the user should complete more advanced paperwork. Thus, for a number of Trusts at least, a multiple-stage risk assessment process is in evidence. One might predict that those RAPs which mentioned a more advanced assessment would be those with fewer items. In fact this was not the case. The RAPs that mentioned more advanced proformas had a mean of 53 items compared with a mean of 42 items for those RAPs that did not.

### Content analysis (P11)

Content analysis of the RAPs revealed a number of frequently occurring themes. Questions related to suicide appeared in 90% of all RAPs and these were often reflected in substantial check lists of demographic features that are known to be epidemiologically-determined risk factors in suicide (e.g., gender, age, occupational or marital status (Qin et al., 2003; Cheng et al., 2000)).

Risk to others was also strongly represented. Seventy-five per cent of RAPs included items pertaining to this theme, most usually with specific mention of the potential for violence to others (56%) and, more specifically, harm to children (62%). Other, strongly represented domains were vulnerability to, or exploitation by, others (50%); self harm, as distinct from suicide (81%); and substance misuse (69%).

### Qualitative comments (P12)

These comments were, of course, no more than the opinions of the researchers and so we did not attempt to group these into themes, nor analyse them in any way. However, to

provide a feel for the views raised in response to the RAPs we present a selection in Table III.

### Discussion

In this study we have assessed Risk Assessment Proformas (RAPs) from 53 mental health provider trusts in the UK. Our aim was to consider the similarities and differences that exist between different proformas currently in use.

One of the most immediately visible differences was in the length of the RAPs: the number of pages varied between 1 and 6 while the number of rateable items varied between 5 and 148. However, the number of items and page length did not co-vary. The reason for this is that where “tick box” responses are used, a great number of items can be accommodated on a single page (many items/few pages). Free-text responses, however, tend to require a lot more space (few items/many pages). It would appear then that different providers have varying expectations about how much “data” should be recorded in the process of risk assessment. Clearly there are considerations to be made about time-burden and cognitive load as the number of data items increases. While we are not necessarily advocating that “shorter is better” we do suggest that the time taken to complete the RAP will be a significant issue for most mental health professionals and may affect the quality of information recorded. It is plausible to suggest that there would be a trade-off between ease of completion and information quality; this would certainly be a fruitful issue for future research. Indeed, an extensive psychology literature convincingly demonstrates that human information-processing and decision-making is profoundly influenced by time-pressure and cognitive load.

Another area in which there was high variability between RAPs was in the quality of their layout and structure. While our comments about visual appeal may be subjective, we found many recurring problems that were, arguably, of a more intrinsic nature. For example, providing check boxes that are not aligned properly to their corresponding prompt statements, or are too small to mark clearly, is a fundamental breach of good questionnaire design. Similarly, providing free-text boxes that are too small to contain more than a single short sentence does nothing to improve the quality of information provided. Carelessness in design was also highlighted by a significant number of RAPs with typographic errors, or pages without an adequate patient identifier. Moreover some of the typefaces used were too small to read clearly, a problem that may be exacerbated further if the RAPs were photocopied, while others used typefaces inconsistently. There were, of course, some good examples that were very clearly laid out with good spacing and alignment and with plenty of space to provide detailed responses if required. Our promise to maintain full anonymity and confidentiality prevents us from reproducing any examples of good or weak examples here but we would stress that poor layout quality is not a trivial issue: completion of risk assessment documentation assumes high importance in mental healthcare and if the assessor’s attention is divided between, on the one hand, providing useful and detailed information and, on the other, navigating a badly laid-out proforma, the quality of the assessment can only suffer.

[Table III. Examples of “good” and “bad” points raised in relation to RAPs.

*Good Points Bad Points*

- Risk indicators clear and verifiable. Same questions asked repeatedly in different sections.
- Useful “opt out” option if it is self-evident there are no risks to consider.
- Some nonsensical questions, e.g., “is there a past history of significant relationships?”
- Useful attempts to operationalize what is meant by the questions.
- No instructions on how to complete the form or what to do with it.
- Comprehensive: gives a full profile of the patient.
- Little more than a listing of historical features from which some opinion is then just expected to emerge.
- Header statement acknowledges limitations of the risk assessment process.
- RAP only to be filled if there are “identified risks” – but if already identified why go through the process again?
- Plenty of free text space for the completer to capture and explain their opinions.
- Assessment and management plan mixed up. Incoherent and impossible to respond to.]

Across the other properties we investigated, there were actually many consistencies between RAPs. Looking at the way in which information was coded at an item level, we found a general reliance on forced-choice dichotomy responses (e.g., yes/no tickboxes) with only a minority of these including an additional option to indicate lack of certainty (e.g., yes/no/don’t know). This type of information coding raises some important questions about the very nature of the risk assessment process. For example, consider the issue of whether a hypothetical patient has violent thoughts: if we code this predictor simply in terms of “present” or “absent”, we fail to take into account the severity, frequency or duration of these thoughts, each of which may carry huge significance in predicting future behaviour. On the one hand, tick boxes are a neat and simple way of summarizing a risk profile but, on the other, they may force the completer to neglect important detail. In some cases, we found that tick boxes were augmented with free-text boxes, but this was relatively uncommon. In short, reducing qualitative information about a patient into a series of nominal categories may provide paper-trail evidence that a checking process has been carried out but one might question whether the quality of information recorded will actually be sufficient to inform future decision-making. A further problem here is that ticking a series of boxes to summarize a patient’s risk profile may lead to a sense of security that the risk assessment has been done and completed when, in all likelihood, completion of the RAP should be the first stage of a more comprehensive assessment.

A broadly similar consideration applies to the issue of time-period invariance. In general we found that questions about a patient’s history were not constrained within specific timeframes:

questions such as “within the last 10 years has this patient been homeless?” were very unusual indeed. Timing is an important issue because it arguably affects the weighting ascribed to a purported risk factor. For example, a behavioural feature that occurred 30 years ago in a 50-year-old patient, but which has never happened since, should carry less

significance than the same feature exhibited by a 20-year-old within the last month, perhaps for the third time. Combined with dichotomous coding, ignoring the time-history of salient features can only lead to a degradation of important information. The “reductive and stifling” effect of recording clinical information in this way has been commented on by Godin (2004).

None of the RAPs carried instructions about clinical actions that might be informed by the result of the assessment. The nearest examples we found were somewhat restrained statements such as “if high risk, consider discussing with supervisor”. Nonetheless, almost half the RAPs we saw gave instructions to complete a more advanced risk assessment document. One could therefore argue that the most common response to identification of risks through one paper process is to proceed to a yet more elaborate paper process. Of course one might argue that RAPs are not intended to direct clinical action. For example, the Clinical Assessment of Risk Decision Support (CARDS) approach of Watts et al. (2004) is designed to support clinical decisions rather than to Risk assessment proformas in mental health 445 make them. Whether this is the intention behind RAPs more broadly is unclear since the majority did not include information describing the scope and purpose of the assessment.

One striking similarity between RAPs in this study was the infrequency with which they called upon the completing clinician to make any predictive statements about future events or states of affairs. The majority of RAPs required nothing of the completer in this respect and appeared to be a means of simply accumulating descriptive data. Even of those RAPs that did require some kind of risk estimation from the user, very few seemed to encourage anything more than vague statements. This does raise the question of what we actually mean by “risk assessment”: for example are we using a RAP just to describe a person’s characteristics? Or is risk assessment about formulating, and then recording, opinion about the likelihood of future events or outcomes for an individual? If the intention is the latter then it appears to be poorly served by the majority of RAPs we have seen. If the intention is, rather, that RAPs should be used as supportive tools in the process of decision making about a given patient, then there is arguably a need to be specific about the information that is required and what it might be used for.

Turning to the thematic consistencies between RAPs, we found that item content could be broadly grouped into: (i) risk to self, most notably in relation to suicide, (ii) risk to others, usually in terms of violence, and (iii) vulnerability, usually meaning exploitation by others. There are many reasons why there may be such strong convergence on these themes: for example, they tend to be those aspects of risk that are frequently discussed in academic publications and DoH reports; similarly they are often the most newsworthy aspects of mental healthcare and, perhaps, are thereby over-inflated in the media. Moreover, we found many examples of RAPs that had been either partially or wholly copied from other Trusts and this would also tend to increase their overall thematic similarity. Thus it would appear that, by design or evolution, risk assessment tends to be viewed as encompassing these three concepts, a point also noted by Ryan (1998). Just as notable, however, are those concepts that are seldom mentioned. It is conceivable that clinicians deliberate about other “risks”

(e.g., socio-economic risks, pharmaco-therapeutic risks, risks associated with driving, etc.) outside the confines of a RAP framework. However, it is equally possible that such factors are completely neglected because they are not mentioned. In essence, it is quite possible that RAPs do not simply reflect prevailing beliefs about what matters in risk assessment; they may actually shape those beliefs as well. That the demand-characteristics of a task determine the performance of the task is well established in the psychology literature (e.g., Orne & Whitehouse, 2000).

One of the main stimuli to conducting this study was our interest in empirically testing the reliability and validity of RAPs as they currently stand. We suggest that this is not possible: there is too little commonality between the RAPs we examined and it is clearly implausible to suggest that each Trust's RAP should be separately subjected to rigorous research evaluation. We concur with Watts et al. (2004) that the evidence base to support the clinical effectiveness of risk assessment in general psychiatric populations is slender and that substantial development work and evaluation is required.

One possible recommendation from this work is that it would be more appropriate for the NHS to adopt a more unified, evidence-based, approach to RAP development. Standardized instruments, perhaps introduced provisionally, would enable field-testing across a number of treatment settings and providers. At very least, a standardized tool or template would relieve the dozens of providers from having to design and re-design RAPs for local use, a task that, as we have found, some are better at than others. But the optimal design of such an instrument would be a matter of considerable conjecture. An alternative is, of course, that, until we can demonstrate evidence that completion of RAPs leads to better clinical care, there should be no RAPs at all.

### *Conclusion*

We have found that RAPs as currently used in NHS mental-health practice vary greatly in structure, content, length and quality. We also have reservations about the most common approaches taken to recording data in RAPs and raise the question of whether the narrow focus on three domains of risk acts to the neglect of other, less dramatic, areas. Processes used for risk assessment need to be critically evaluated and the introduction of standardized tools would be advantageous in this respect.

### *Limitations*

The limitations of this study lie fundamentally with the process of RAP evaluation. While we tried to use criteria that were transparent, well-described and replicable, some of the judgments made were necessarily subjective in nature. Secondly, we were not able to include RAPs from all Mental health provider Trusts in this study since many did not respond to our request. Our findings are broadly consistent with those of Higgins et al. who had a similar response rate in their study but it remains possible that our sample is not fully representative of all NHS Trusts.

### *Acknowledgements*

We are grateful to two anonymous reviewers for their helpful comments on an earlier version of this manuscript.

## References

- Belfrage, H., Fransson, G., & Strand, S. (2000). Prediction of violence using the HCR-20: a prospective study in two maximum-security correctional institutions. *Journal of Forensic Psychiatry, 11*, 167 – 175.
- Cheng, T. A., Chen, H. H., Chen, C., & Jenkins, R. (2000). Psychosocial and psychiatric risk factors for suicide: case-control psychological autopsy study. *British Journal of Psychiatry, 177*, 360 – 365.
- DoH. (1999). National Service Framework for Mental Health: Modern standards and service models. London: DoH White paper.
- Gale, T. M., Hawley, C. J., & Sivakumaran, T. (2003). Do mental health professionals really understand probability? Implications for risk assessment and evidence-based practice. *Journal of Mental Health, 12*, 417 – 430.
- Gale, T. M., Woodward, A., Hawley, C. J., Hayes, J., Sivakumaran, T., & Hansen, G. (2002). Risk assessment for people with mental health problems: A pilot study of reliability in working practice. *International Journal of Psychiatry in Clinical Practice, 6*, 73 – 81.
- Godin, P. (2004). 'You don't tick boxes on a form': A study of how community mental health nurses assess and manage risk. *Health, Risk & Society, 6*, 347 – 360.
- Harriss, L., & Hawton, K. (2005). Suicidal intent in deliberate self-harm and the risk of suicide: The predictive power of the Suicide Intent Scale. *Journal of Affective Disorders, 86*, 225 – 233.
- Higgins, N., Watts, D., Bindman, J., Slade, M., & Thornicroft, G. (2005). Assessing violence risk in general adult psychiatry. *Psychiatric Bulletin, 29*, 131 – 133.
- Leach, C., Lucock, M., Narkham, M., Noble, R., Clarke, L., & Iveson, S. (2005). Assessing risk and emotional disturbance using CORE-OM and HoNOS outcome measures at the interface between primary and secondary care. *Psychiatric Bulletin, 29*, 419 – 422.
- Risk assessment proformas in mental health 447
- Montandon, C., & Harding, T. (1984). The reliability of dangerousness assessments: A decision making exercise. *British Journal of Psychiatry, 144*, 149 – 155.
- Orne, M. T., & Whitehouse, W. G. (2000). Demand characteristics. In A. E. Kazdin (Ed.), *Encyclopedia of Psychology*. New York: Oxford University Press.
- Qin, P., Agerbo, E., & Mortensen, P. B. (2003). Suicide risk in relation to socioeconomic, demographic, psychiatric, and familial factors: A National Register-Based Study of all suicides in Denmark, 1981 – 1997. *American Journal of Psychiatry, 160*, 765 – 772.
- Ryan, T. (1998). Perceived risks associated with mental illness: Homicide and suicide. *Social Science and Medicine, 46*, 287 – 297.
- Watts, D., Bindman, J., Slade, M., Holloway, F., Rosen, A., & Thornicroft, G. (2004). Clinical assessment of risk decision support (CARDS): The development and evaluation of a feasible violence risk assessment for routine psychiatric practice. *Journal of Mental Health, 13*, 569 – 581.