BMJ Open

Meaning of Family Reported Outcome Measure (FROM-16) severity score bands: a cross-sectional online study in the UK

Rubina Shah ![ORCID],[1] Andrew Y Finlay ![ORCID],[1] Sam M Salek ![ORCID],[2,3] Stuart J Nixon,[4] Kennedy Otwombe ![ORCID],[5,6] Faraz M Ali ![ORCID],[1] John R Ingram ![ORCID] [1]

For numbered affiliations see end of article.

Correspondence to
Rubina Shah;
ShahR45@cardiff.ac.uk

ABSTRACT

Objective To assign clinical meanings to the Family Reported Outcome Measure (FROM-16) scores through the development of score bands using the anchor-based approach.

Design and setting A cross-sectional online study recruited participants through UK-based patient support groups, research support platforms (HealthWise Wales, Autism Research Centre-Cambridge University database, Join Dementia Research) and through social service departments in Wales.

Participants Family members/partners (aged ≥18 years) of patients with different health conditions.

Intervention Family members/partners of patients completed the FROM-16 questionnaire and a Global Question (GQ).

Main outcome measure Various FROM-16 band sets were devised as a result of mapping of mean, median and mode of the GQ scores to FROM-16 total score, and receiver operating characteristic-area under the curve cut-off values. The band set with the best agreement with GQ based on weighted kappa was selected.

Results A total of 4413 family members/partners (male=1533, 34.7%; female=2858, 64.8%; Prefer not to say=16, 0.4%; other=6, 0.14%) of people with a health condition (male=1994, 45.2%; female=2400, 54.4%; Prefer not to say=12, 0.3%; other=7, 0.16%) completed the online survey: mean FROM-16 score=15.02 (range 0–32, SD=8.08), mean GQ score=2.32 (range 0–4, SD=1.08). The proposed FROM-16 score bandings are: 0–1=no effect on the quality of life of family member; 2–8=small effect on family member; 9–16=moderate effect on family member; 17–25=very large effect on family member; 26–32=extremely large effect on family member (weighted kappa=0.60).

Conclusion The FROM-16 score descriptor bands provide new information to clinicians about interpreting scores and score changes, allowing better-informed treatment decisions for patients and their families. The score banding of FROM-16, along with a short administration time, demonstrates its potential to support holistic clinical practice.

INTRODUCTION

A person's health condition has a huge impact on their quality of life (QoL) but

---

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ Large sample size of >4400 participants.
⇒ Heterogeneous population involving family members of people with more than 200 health conditions across England, Northern Ireland, Scotland and Wales.
⇒ Online study with recruitment by self-selection has inherent selection bias.
⇒ Online nature of the study restricted recruitment to family member participants who could operate an electronic device.
⇒ As this study is restricted to the UK, the results may not be generalisable to other countries.

---

also on the QoL of their partner and other family members. Golics *et al*[1] demonstrated that family members of patients suffering from a wide range of different diseases are impacted in similar ways regarding psychological, emotional, personal and social well-being. This key finding from this work led to the development[2] of the Family Reported Outcome Measure (FROM-16), a generic family QoL (FQOL) questionnaire, which measures the impact of any disease, across all medical specialties, on the QoL of family members or partners of patients. The FROM-16 is a user-friendly and relatively simple questionnaire with a 2-minute completion time, making it a practical tool for use in a clinical setting. The FROM-16 has been translated and validated in various languages, indicating the high interest in its use across many parts of the world.[3–7]

Although a higher score indicates a greater impact on family members' QoL, for FROM-16 to be useful in a clinical or research setting, there is a need for scores to have practical meaning. The utility of QoL questionnaires can be maximised if a clinical meaning is assigned to the questionnaire scores.[8] This is important as in the absence of such

interpretation, scores are just arbitrary numbers, leaving clinicians to guess the magnitude of effect or importance of score change in response to treatment. The ability to interpret questionnaire scores is essential if the questionnaire is to be of value in shared decision-making or monitoring clinical change. For example, the score banding of the Dermatology Life Quality Index (DLQI) was clinically useful to dermatologists and has facilitated the integration of the DLQI into national guidelines in over 45 countries.[9] [10]

Descriptive score banding, therefore, gives vital meanings to absolute scores. The development of score bands for FROM-16 would create cut-off points, making it easier for clinicians to identify at-risk and high-risk family members and to direct them to the appropriate support services. Score band development would transform FROM-16 from being primarily a research tool to being of practical benefit to clinicians across all medical specialties. There is a need to conduct a comprehensive study to examine such benefit. Therefore, this study aims to develop score descriptor bands for the FROM-16, using the anchor-based approach.

METHODS
Settings and participants
This was an online cross-sectional study conducted between April and November 2021 involving family members/partners of patients with a wide range of health conditions. The participants were recruited through UK-based patient support groups, through research support platforms (HealthWise Wales (HWW),[11] Autism Research Centre-Cambridge University database (ARC), Join Dementia Research (JDR)) and through social service departments (SSDs) in Wales.

Convenience sampling was used for recruitment of the study participants. The study participants in this study were all family members/partners of patients. The study participants were all aged ≥18 years. Patients who were aged 18 years or above also contributed to this study by providing demographic information and by inviting a family member to take part in the study. Some of the study participants were family members/parents of patients under the age of 18 years. These patients were not involved at all in the study. The exclusion criteria included family members of deceased patients, family members/partners under 18 years of age, patients and family members not living in the UK and those not capable of operating an electronic device.

Patient and public involvement
Two patients and one family member were involved as integral study research partners, one of whom (SJN) is a coauthor. They were involved in reviewing the study protocol, drafting the survey, reviewing the manuscript and providing suggestions from the patient and family perspective. SJN contributed to weekly research team meetings.

Survey structure
The survey was carried out using the Jisc survey platform,[12] which is General Data Protection Regulation compliant. The survey was available in two formats:
- *Patient and family member survey*: this survey was directed to patients registered with various patient support groups who then provided consent to involve their family members in the study.
- *Family member-only survey:* this survey was directed to the family members of patients. The family member-only survey was also used for participant recruitment through HWW, JDR and SSDs.

The survey had two sections. Section one was completed by the patient. The patients' role was limited to completing some basic information about themselves (including gender, age, occupation, health condition and country of residence) and choosing and allowing their family member/partner to take part in the survey. The designated family member/partner had the choice to participate or not in the study. In the 'family member-only' survey, patient demographic information was completed by the family member.

Section two was completed by the family member/partner who provided some basic demographic information (age, gender, occupation and relationship to the patient) and answered FROM-16 and a Global Question (GQ). The participants were provided with information about the study in a Participant Information Sheet embedded in the survey, and electronic informed consent was sought from the participants.

Survey distribution
Before starting the study, a pilot was run with 15 patients registered with Acute Leukemia Advocates Network and their family members/partners to test whether there was any difficulty or ambiguity in the wording of the anchor-based question as well as the other survey questions. None of the participants reported any difficulties in understanding the anchor, although there were a few suggestions about the general format of the questions. The survey questions were revised based on collective feedback.

The survey was distributed by various local and national patient groups and associations. The UK-based patient support groups for a wide range of health conditions were approached. Of the 106 patient support groups invited to contribute, 58 (55%) participated in the survey. The survey was also distributed through the research support platforms of HWW,[11] ARC and JDR, and SSDs in Wales.

Measurement tools
Family Reported Outcome Measure
The FROM-16 is an FQoL questionnaire which measures the impact of any disease, across all medical specialties, on the QoL of adult family members or partners of patients of any age.[2] The FROM-16 comprises 16 items, each with three response options: 'not at all' (scoring 0), 'a little' (scoring 1) and 'a lot' (scoring 2). The 16 items

are divided into two categories (domains): emotional (comprising 6 items, maximum score of 12) and personal and social life (comprising 10 items, maximum score of 20). The lowest possible score of the FROM-16 is 0, and the highest is 32. The higher the total score, the greater the effect on the family member's QoL. All impacts are not necessarily negative to family members.[1] However, the purpose of the FROM-16 is to identify negative impacts which may require improvement by appropriate intervention.

Global Question

The GQ used was: 'How much is your life being affected by your family member's or partner's health condition at the moment?'

This measured the overall impact of a person's health condition on their family member/partner on a 5-point Likert scale: 0=no effect; 1=small effect; 2=moderate effect; 3=very large effect and 4=extremely large effect. The family member/partner was asked to tick one of the five options.

The single-item GQ was used as an anchor to assess the meaningfulness of total FROM-16 scores by being mapped against the multidimensional total FROM-16 scores.[2 13]

Both the FROM-16 and GQ questionnaires are validated tools that have been previously used in this setting.

Outcome

The primary outcome was to assign meanings to FROM-16 scores through the development of the score bands.

Exposure

Impact of patients' disease on their partners and family members.

Covariates

The covariates included family members' relationship to patients, country of residence, type of health condition, age and sex of family members and patient.

Missing data

There were no missing data.

Data processing and statistical analysis

Frequencies were determined for categorical variables. The FROM-16 scores were correlated with GQ scores, using Spearman's rank correlation coefficient (r_s). For each FROM-16 score from 0 to 32, the distribution and the mean, mode and median of the corresponding GQ score were calculated. Mean (rounded off to the nearest whole number), median and mode were used to group the FROM-16 scores into five discrete bands, with each FROM-16 band corresponding to a GQ band. Numerical cut-off points were considered based on FROM-16 scores that corresponded to a one-step increase in mean, median and/or mode on the anchor. In the case of an overlap between some of the possible discrete categories, where a number of FROM-16 scores could have fitted into one of two categories, different cut-offs of the FROM-16 score

were tested against the severity indicated on the anchor questions by using a weighted kappa (wκ), a coefficient of agreement to determine the level of agreement between FROM-16 and GQ bands. The cut-off value showing a greater agreement between FROM-16 and GQ bands, indicated by a higher value of wκ, was chosen as a final cut-off point.

For those family members/partners whose GQ score disagreed in a major way (by two or more bands) with that predicted from the devised FROM-16 banding score, subscore comparisons were made with those family members whose GQ scores agreed with the FROM-16 banding.

Receiver operating characteristic (ROC) curve analysis was used as another method to determine the optimum FROM-16 cut-off scores between successive GQ bands.[14] The optimum cut-off is the point on ROC curve where the sensitivity and specificity are maximised and the area under the curve (AUC) values were ≥0.7; the point on the curve with minimum distance from the left-upper corner of the unit square; and the point where the Youden's index is maximum.[15] There are different methods to identify optimal cut-off value for ROC curve, which include minimum p value approach, Youden index, minimum Euclidean distance, Concordance Probability Method and Index of Union.[16] The Youden Index, the most widely used method to detect cut-off points, and the 'minimum Euclidean distance measure' were used to determine the optimal balance between sensitivity (true positive rate) and specificity (true negative rate) in the estimation of the FROM-16 cut-off scores. Cut-off scores were rounded to zero decimal places. The level of agreement for each FROM-16 band with GQ was calculated using the wκ coefficient, which takes into account the magnitude of disagreement between different categories.[17]

A gender and age comparison was made using the Mann-Whitney U test, which compared the hypothesis of no difference in the mean of the ranks. Statistical analysis was done using SPSS V.27.

RESULTS
Sociodemographic characteristics of the study participants

A total of 4469 family members/partners of people with health conditions completed the FROM-16 questionnaire and GQ. Fifty-six responses were discarded as the respondents were not family members of people living with a health condition, hence not relevant. The final analysis included responses from 4413 adult family members (male=1533, 34.7%; female=2858, 64.8;%; Prefer not to say=16, 0.4%; other=6, 0.14%) of people with over 200 health conditions (male=1994, 45.2%; female=2400, 54.4%; Prefer not to say=12, 0.3%; other=7, 0.16%) representing 27 medical specialties across the UK (England=42.9%; Wales=51.8%; Scotland=4.2% and Northern Ireland=1.1%) (table 1).

The mean age of family members/partners was 57 years (range=18–95 years, SD=14.3), and that of patients was 61

Table 1 Demographic characteristics of the family members and patients

Variables	Categories	N (%) or N (SD)
People with health conditions (n=4413)		
Gender	Male	1994 (45.2)
	Female	2400 (54.4)
	Prefer not to say	12 (0.3)
	Other	7 (0.2)
Age (years)	Mean (SD)	61.47 (20.3)
	Median	66
	Range (IQR)	2–100 (26)
Occupation	In paid work	881 (20)
	Part-time job	165 (3.7)
	Unemployed	324 (7.3)
	In unpaid work	22 (0.5)
	Education/training	100 (2.3)
	Homemaker	151 (3.4)
	Retired	2557 (57.9)
	Rather not say	68 (1.5)
	Not applicable	145 (3.3)
Medical specialties	Audiology	19 (0.4)
	Cardiology	241 (5.5)
	Chronic pain	7 (0.2)
	Critical care	1 (0.02)
	Dermatology	138 (3.1)
	Endocrinology	271 (6.1)
	Gastroenterology	153 (3.5)
	Genetic/rare disease	44 (1)
	Gynaecology	38 (0.9)
	Haematology	183 (4.1)
	Hepatology	11 (0.2)
	Immunology	13 (0.3)
	Infectious diseases	10 (0.2)
	Movement disorder	10 (0.2)
	Nephrology	58 (1.3)
	Neurology	1620 (36.7)
	Oncology	251 (5.7)
	Ophthalmology	89 (2)
	Orthopaedics	24 (0.5)
	Otolaryngology	6 (0.1)
	Rehabilitation medicine	30 (0.7)
	Paediatrics	145 (3.3)
	Psychiatry	325 (7.4)
	Respiratory medicine	267 (6.1)
	Rheumatology	310 (7)
	Urology	21 (0.5)
	Wound healing	2 (0.05)
	Multiple health conditions	95 (2.2)
	Not stated	31 (0.7)
Place of residence in the UK	England	1895 (42.9)
	Northern Ireland	48 (1.1)
	Scotland	185 (4.2)
	Wales	2285 (51.8)

Continued

Table 1 Continued

Variables	Categories	N (%) or N (SD)
Family members (N=4413)		
Gender	Male	1533 (34.7)
	Female	2858 (64.8)
	Prefer not to say	16 (0.4)
	Other	6 (0.1)
Age (years)	Mean (SD)	57 (14.3)
	Median	60
	Range (IQR)	18–95 (20)
Occupation	In paid work	1728 (39.2)
	Part-time job	368 (8.3)
	Unemployed	118 (2.7)
	In unpaid work	52 (1.2)
	Education/training	74 (1.7)
	Homemaker	211 (4.8)
	Retired	1808 (41)
	Rather not say	54 (1.2)
Relationship to the person affected with health condition	Spouse/partner	2631 (59.6)
	Son/daughter	973 (22)
	Parent	523 (11.9)
	Other (brother/sister, father/mother-in-law, grandparent, uncle/aunt, grandson/granddaughter, brother/sister-in-law, nephew/niece, cousin, friend)	286 (6.5)
FROM-16 score	Mean (SD)	15.02 (8.08)
	Range	0–32
GQ score	Mean (SD)	2.32 (1.08)
	Range	0–4
Correlation: FROM-16 and GQ score	r_s (p value)	0.79 (0.001)

FROM-16, Family Reported Outcome Measure; GQ, Global Question.

years (range=2–100 years, SD=20.3). The family members were mostly spouse/partner (60%), followed by son/daughter (22%) and parent (12%). Forty-one per cent of the family members/partners were retired, 39% were in paid jobs and 8% were in part-time jobs. Of the people with health conditions, 58% were retired, 20% were in paid jobs and 4% were in part-time jobs (table 1).

The overall mean FROM-16 score was 15.02 (range 0–32; SD=8.08), and the mean GQ score was 2.32 (range 0–4, SD 1.08). The mean FROM-16 score for women (16.13, SD 8.04) was higher than that for men (12.92, SD 7.70, mean rank p=0.01, Mann-Whitney U test), as was the mean GQ score (female=2.46, SD=1.06; male=2.04, SD=1.03, mean rank p=0.01, Mann-Whitney U test) (online supplemental table 1). The mean FROM-16 score of those participants who identified their gender as 'other' was 18.7 (SD=9.9, n=6), and for those who stated 'prefer not to say', the mean score was 16.9 (SD=11.5, n=16). The mean GQ score for 'other' was 2.33 (SD=1.03) and for 'prefer not to say' was 2.25 (SD=1.48).

The mean FROM-16 score for age group 1 (under 60 years) was higher than that for age group 2 (over 60 years) (mean rank p=0.001, Mann-Whitney U test); however, there was no significant difference between the mean GQ score between the two age groups (mean rank p=0.391, Mann-Whitney U test) (online supplemental table 1).

FROM-16 score banding

There was a strong correlation between FROM-16 scores and GQ scores (r_s=0.79, p=0.001), a prerequisite for using the anchor-based method (table 1).

For each score of the FROM-16 from 0 to 32, the number of family members with that score and their corresponding GQ mode, mean and median score is shown in table 2 and figure 1, and these were used as the basis for grouping the FROM-16 scores together into a set of five discrete bands so that each band would correspond to a single GQ score.

There were a few FROM-16 scores that could possibly be included in either of the two adjacent bands, such as FROM-16 scores of 1, that could either be in the bands

Table 2 Number of family members with each FROM-16 score and details of the corresponding GQ score, with the mean, mode and median of the GQ scores

FROM-16 score	GQ score*					Mean	Median	Mode	Family member totals
	0	1	2	3	4				
0	34	18	2			0	0	0	54
1	27	41	1			1	1	1	69
2	28	65	7	2		1	1	1	102
3	12	90	13			1	1	1	115
4	15	84	32	2		1	1	1	133
5	10	101	37	3		1	1	1	151
6	3	99	65	5		1	1	1	172
7	5	73	60	5	1	1	1	1	144
8	1	79	65	17		2	2	1	162
9	5	60	98	17	1	2	2	2	181
10	2	49	93	23	1	2	2	2	168
11	3	49	83	25	5	2	2	2	165
12	1	38	98	37	5	2	2	2	179
13	1	24	93	44	7	2	2	2	169
14		19	109	49	9	2	2	2	186
15		16	103	61	7	2	2	2	187
16		13	88	82	14	2	2	2	197
17		5	70	71	20	3	3	3	166
18		4	64	79	14	3	3	3	161
19		3	55	85	17	3	3	3	160
20		4	46	107	26	3	3	3	183
21			35	83	44	3	3	3	162
22			34	63	30	3	3	3	127
23			17	83	56	3	3	3	156
24		1	16	73	38	3	3	3	128
25	1	6	49	44		3	3	3	100
26	1	9	33	61		3	4	4	104
27		1	40	61		4	4	4	102
28		2	27	60		4	4	4	89
29		2	21	54		4	4	4	77
30		1	18	41		4	4	4	60
31		1	11	45		4	4	4	57
32			5	42		4	4	4	47
Family member totals	149	935	1406	1220	703				4413

*GQ score: 0=no effect; 1=small effect; 2=moderate effect; 3=very large effect; 4=extremely large effect.
FROM-16, Family Reported Outcome Measure; GQ, Global Question.

corresponding to a GQ of 0 or 1. Similarly, FROM-16 scores of 7 and 8 could either be included in the GQ bands 1 or 2. Subsequently, separate sets of bands were therefore produced with different groupings of FROM-16 scores, and the wκ coefficient of agreement was calculated for each set of bands (table 3).

Based on the ROC-AUC analysis, FROM-16 cut-off scores between GQ bands 0–1, 1–2, 2–3 and 3–4 were ≥4 (sensitivity 86.7%, specificity 59.7%, AUC 79.4%), ≥8 (sensitivity 79.9%, specificity 69.5%, AUC 82.7%), ≥16 (sensitivity 76.2%, specificity 68.2%, AUC 79.2%) and ≥23 (sensitivity 71.4%, specificity 70.5%, AUC 77.6%), respectively (wκ=0.574) (figure 2).

The wκ coefficients of various banding sets of FROM-16 scores ranged from 0.574 to 0.596, implying a moderate strength of agreement between the banding sets and GQ
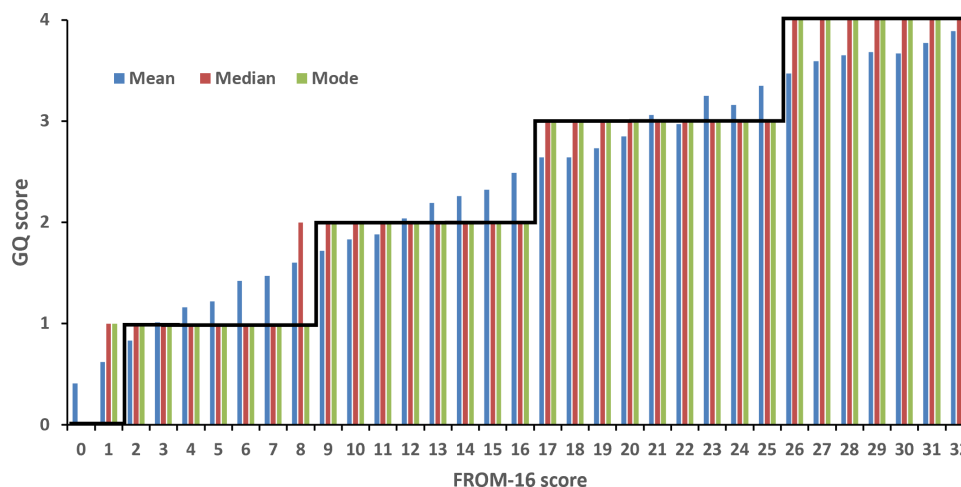
Figure 1 Relationship between the Family Reported Outcome Measure (FROM-16) and the mean, mode, and median of the Global Question (GQ) score. The proposed banding of FROM-16 scores (0–1, 2–8, 9–16, 17–25, 26–32) is also shown.

(table 3). The kappa coefficient is a measure of the level of agreement beyond that which could be expected by chance.

The maximum level of agreement is a kappa of 1.0, and values of 0.41–0.60 are considered a moderate strength of agreement.[18] The banding set that we propose for FROM-16 is 0–1, 2–8, 9–16, 17–25, 26–32 based on the highest value of the wκ (wκ=0.596) (figure 3).

Subscore analysis

There were a total of 87 family members (high—Extreme Value Group-'EVG') whose actual GQ score was 2 or more points higher than the FROM-16 band would have predicted from their FROM-16 score (table 4 and online supplemental table 2).

In contrast, there were only 47 family members (the 'low EVG') whose actual GQ score was 2 or more points lower than the FROM-16 band would have predicted. All the family members whose GQ scores were 2 or more points away from the banding allocation were compared with those family members whose GQ scores agreed with the FROM-16 banding (Normal Value Group ('NVG')) by carrying out a subscore analysis of the 16 individual questions on the FROM-16 (online supplemental table

2). The subscore comparison was carried out within each FROM-16 band. There was a higher proportion of 'maximum subscores' (subscores of 2; a lot) in the high EVG compared with the NVG (8.8% vs 2.7% for band 1; and 25.8% vs 15.8% for band 2) (online supplemental table 3). The opposite was true for EVG with low scores compared with the NVG (13% vs 15.8% for band 2; 36.5% vs 44.2% for band 3; and 73.5% vs 80.9% for band 4) indicating lower proportion of maximum subscores (a lot) in low EVG than NVG (online supplemental table 3).

DISCUSSION

There has been great interest in measuring FQoL in recent years. Several instruments have been developed to assess the impact of a person's health condition/disability on family members/partners to understand this secondary and often unrecognised burden of disease.[19] Although the reliability and validity of several FQoL instruments have been established, they are not yet accepted in clinical practice. The assessment of the meaning of scores of an instrument and the practicality of its use in a clinical setting are more important than just

Table 3 Weighed kappa coefficients of agreement for separate possible sets of bands of the FROM-16 scores

Banding set	Assignment of FROM-16 scores into bands					Weighted kappa coefficient of agreement
	Band 0	Band 1	Band 2	Band 3	Band 4	
Set A	0	1–7	8–16	17–25	26–32	0.588
Set B	0	1–8	9–16	17–25	26–32	0.594
Set C	0–1	2–7	8–16	17–25	26–32	0.590
Set D	0–1	2–8	9–16	17–25	26–32	0.596
Set E	0–2	3–7	8–16	17–25	26–32	0.588
Set F	0–2	3–8	9–16	17–25	26–32	0.595
ROC-AUC	0–4	5–9	10–16	17–23	24–32	0.574

AUC, area under the curve; FROM-16, Family Reported Outcome Measure; ROC, receiver operating characteristic.
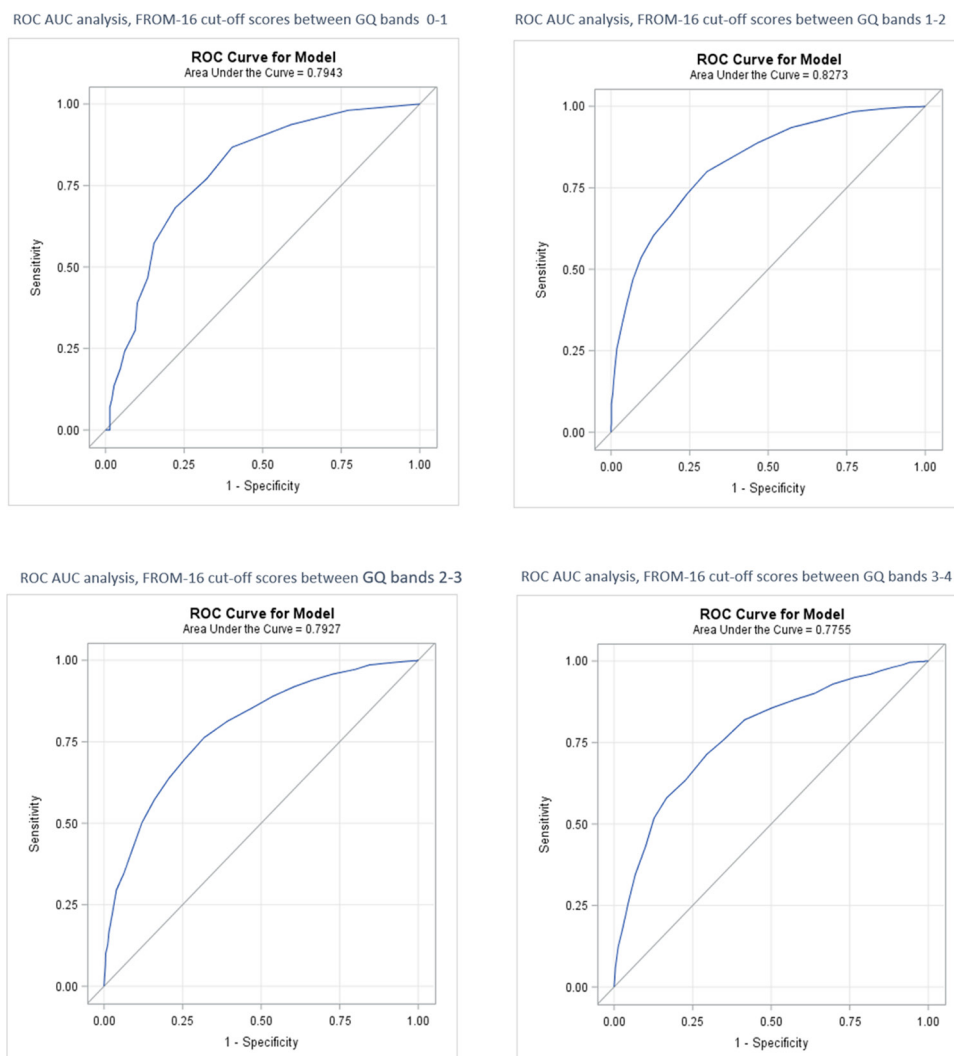
Figure 2 The receiver operating characteristic (ROC) curve and area under the curve (AUC) analysis. FROM-16 cut-off scores between GQ bands 0–1, 1–2, 2–3, 3–4. FROM-16, Family Reported Outcome Measure; GQ, Global Question.

meeting conventional psychometric requirements.[9] The FROM-16 is the only generic user-friendly FQoL instrument: it has a 2–3 min completion time, making it a practical tool for use in a clinical setting to measure the family impact of a person with a disease. However, to support the use of FROM-16 across all disciplines of medicine, there is a need for the scores to be meaningful and be easily interpreted in the context of individual patient to support holistic clinical decision-making. Such information will assist clinicians to identify at-risk and high-risk family members, and provide, or direct them to, appropriate support services. Therefore, the purpose of this study was to establish score bands for FROM-16, making overall scores more meaningful while complementing the

information that can be gained by examining the detailed subscores of FROM-16.

The score interpretation of QoL measures can be carried out either using the distribution-based method or anchor-based methods.[20 21] In the distribution-based technique, interpretations are based on the statistical distributions of scores in a given population, such as SD or SE of measurements.[22 23] On the other hand, in the anchor-based method, interpretations are made when the scores are compared or anchored with some theoretically related external measures.[24] A commonly used anchor is a global rating question that itself is easily interpreted. This study used an anchor-based approach since it is most appropriate for short, relatively simple questionnaires.[9]

FROM-16 Banding	0-1	2-8	9-16	17-25	26-32
	No effect	Small effect	Moderate effect	Very large effect	Extremely large effect

Figure 3 Proposed FROM-16 score banding. The effect mentioned is the level of adverse effect on the quality of life of a patient's partner or family member.

Table 4 Proposed banding of the FROM-16 with the distribution of GQ scores for the bands 0–1, 2–8, 9–16, 17–25, 26–32 (wκ coefficient of agreement=0.596)

Set D	FROM-16 score band	GQ score					Family member totals (%)
		0	1	2	3	4	
Band 0	0–1	61	59	3			123 (2.8)
Band 1	2–8	74	591	279	34	1	979 (22.2)
Band 2	9–16	12	268	765	338	49	1432 (32.5)
Band 3	17–25	1	17	343	693	289	1343 (30.4)
Band 4	26–32	1		16	155	364	536 (12.1)
Family member totals		149	935	1406	1220	703	4413

GQ score: 0=no effect; 1=small effect; 2=moderate effect; 3=very large effect; 4=extremely large effect.
FROM-16, Family Reported Outcome Measure; GQ, Global Question; wκ, weighted kappa.

The anchor-based method is based on patients' ratings (in this study, patient's family members' rating) and is therefore thought to provide the best estimate of an individual's perspective.[8] An anchor-based approach using the GQ has been used for score interpretation of various QoL measures in dermatology, including the DLQI,[9] Eczema Area and Severity Index,[25] Patient-Oriented Eczema Measure (POEM)[26] and Vitiligo Impact Scale-22 (VIS-22)[17] as well as in other areas such as nephrology (Renal Quality of Life profile (RQLP)).[27]

Anchor-based methods have two requirements.[23] The first is that the anchor must be interpretable, and this requirement was met as the GQ was clearly understood by family members of patients with leukaemia in the pilot study. Second, there must be a reasonable degree of association between the target instrument and the anchor, and this study demonstrated a very close correlation between the FROM-16 and the GQ. The correlation between anchor question GQ and FROM-16 (r=0.79) was comparable with other score banding studies (DLQI r=0.83; VIS-22 r=0.76; POEM r=0.77; RQLP r=0.73) indicating equally reliable score banding.[9 17 26 27] Although the strong correlation between the FROM-16 and the anchor makes the descriptive banding more accurate, it could be argued that the anchor question itself would be sufficient. However, unlike FROM-16, an anchor question is a single item measuring the overall perception of the impact of the patient's condition on the QoL of the family member and does not provide any information about what aspects of QoL are impacted. This information is vital in providing support to impacted family members.

In this study, the anchoring GQ was used as the basis for devising a 'banded scale' of the total FROM-16 scores. Multiple FROM-16 band sets were devised as a result of mapping of GQ summary scores to total FROM-16 score, and ROC-AUC analysis. The band set with the best agreement with GQ based on wκ was selected. The wκ coefficients of various banding sets of FROM-16 scores ranged from 0.574 to 0.596, implying a moderate strength of agreement between the banding sets and GQ. The wκ coefficient of agreement for the proposed banding was moderate (K=0.596) and this was similar with other

measures' development (POEM,[26] RQLP,[27] VIS-22[17] and DLQI).[9] Clinicians and researchers using FROM-16 can therefore have confidence that the bands accurately reflect the impact on family members.

The banding set proposed for FROM-16 (0–1, 2–8, 9–16, 17–25, 26–32) is robust, pragmatic as well as easy to remember, making it suitable for routine use in clinical settings. Furthermore, it is also easy to remember that once a FROM-16 score goes above the halfway point of 16, this suggests a person's health condition is having a 'very large effect' on their family member's QoL.

Although there was substantial agreement between the family members' global rating of overall impact and predicted FROM-16 banding, 3% of the family members' FROM-16 scores fell outside the proposed banding by 2 points (table 4). The anonymous nature of the study does not allow us a more detailed analysis of factors that contributed to these outliers. A few family members recorded a high GQ score, but a low FROM-16 score. One reason that could explain this anomaly could be that family members might have been overwhelmed by one or more FROM-16 items that impacted their life negatively, reflecting in high GQ scores. This was further evident from the subscore comparison of individual FROM-16 items of family members, which showed that family members who had higher GQ scores than their FROM-16 scores when responding were more likely to mark FROM-16 items with an extreme value (a lot) (online supplemental table 3).

Men and women may perceive QoL differently.[28 29] In our study, although there was a significant difference in total FROM-16 mean scores and GQ scores between the genders, and a significant difference in total FROM-16 mean scores between two age groups (18–59 and 60–95 years), the difference was not obvious in the banding (data not shown). Therefore, separate banding based on genders and age was not considered. The mean FROM-16 scores of those participants who identified their gender as 'other' and for those who stated 'prefer not to say' were higher than for male and for female participants: the reason for this is not clear but deserves further investigation.

People aged under 60 years had a higher mean FROM-16 score than older respondents but a similar GC score. This could be explained on the basis that 47.5% of family members caring for their relative were in paid employment and possibly overburdened by work, family duties and caring. As FROM-16 items allowed family members to express this impact, this may have contributed to mean FROM-16 scores being higher for people aged under 60 than for those over 60.

Strengths and limitations

This study has several strengths. One of the key strengths is the large sample size involving family members/partners of people with over 200 different health conditions. Another strength is the heterogeneous population involving family members of people with health conditions across England, Northern Ireland, Scotland and Wales. Both add to the generalisability of the findings; however, as the study was restricted to the UK, results may not be generalisable to other countries. The study may have potential selection bias. Only those family members who were registered with patient groups, and who were able to operate an electronic device to answer the survey, took part in the study. This may have resulted in recruiting more severely impacted or more motivated family members of people with health conditions. However, as we recruited the family members of patients across a wide range of health conditions and severities, we were able to recruit family members who experienced a wide range of impacts from 'no effect' to 'extremely large effect', which was critical for this type of study.

Conclusions

In conclusion, we present data to support validated clinical meaning of FROM-16 scores. The FROM-16 severity banding proposed in this study provides new information to clinicians and researchers to interpret scores and score changes allowing better shared decision-making for patients and families and to facilitate the wider assessment of burden of disease. The FROM-16 cut-off scores will help clinicians to identify and support severely impacted family members/partners of patients leading to more holistic clinical practice. This banding score meaning can now be applied in retrospect to previously published FROM-16 data and prospectively in future research studies.

Author affiliations
[1] Division of Infection and Immunity, Cardiff University, School of Medicine, Cardiff, UK
[2] School of Life & Medical Sciences, University of Hertfordshire, Hatfield, UK
[3] Institute of Medicines Development, Cardiff, UK
[4] Multiple Sclerosis Society, London, UK
[5] Statistics and Data Management Centre, Perinatal HIV Research Unit, Chris Hani Baragwanath Academic Hospital, University of the Witwatersrand, Johannesburg, South Africa
[6] 6School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

ORCID iDs
Rubina Shah http://orcid.org/0000-0001-8158-712X
Andrew Y Finlay http://orcid.org/00000003-2143-1646
Sam M Salek http://orcid.org/00000002-4612-5699
Kennedy Otwombe http://orcid.org/0000-0002-7433-4383
Faraz M Ali http://orcid.org/0000-0002-4184-2023
John R Ingram http://orcid.org/0000-0002-5257-1142

REFERENCES

1 Golics CJ, Basra MKA, Salek MS, *et al*. The impact of patients' chronic disease on family quality of life: an experience from 26 specialties. *Int J Gen Med* 2013;6:787–98.
2 Golics CJ, Basra MKA, Finlay AY, *et al*. The development and validation of the family reported outcome measure (FROM-16) © to assess the impact of disease on the partner or family member. *Qual Life Res* 2014;23:317–26.
3 FROM-16. Family reported outcome measure [Online]. 2021. Available: https://www.cardiff.ac.uk/medicine/resources/quality-of-life-questionnaires/family-reported-outcome-measure
4 Elsner SA, Salek SS, Finlay AY, *et al*. Validation of the German version of the family reported outcome measure (FROM-16) to assess the impact of disease on the partner or family member. *Health Qual Life Outcomes* 2021;19:106.
5 Oh HY, Shin YS. Psychometric properties of the Korean family reported outcome measure for family members of patients with acquired brain injury. *J Neurosci Nurs* 2021;53:256–61.
6 Wójcik E, Reszke R, Otlewska A. Family reported outcome measure-16 (FROM-16): creation, reliability and reproducibility of the polish language version. *Acta Derm Venereol* 2020;100.
7 Chantarasap P, Johns NP, Pairojkul S, *et al*. Validation of the Thai version of the family reported outcome measure (FROM-16) © to assess the impact of disease on the partner or family members of patients with cancer. *Health Qual Life Outcomes* 2019;17:32.
8 Rogers A, DeLong LK, Chen SC. Clinical meaning in skin-specific quality of life instruments: a comparison of the dermatology life quality index and skindex banding systems. *Dermatol Clin* 2012;30:333–42,
9 Hongbo Y, Thomas CL, Harrison MA, *et al*. Translating the science of quality of life into practice: what do dermatology life quality index scores mean? *J Invest Dermatol* 2005;125:659–64.
10 Singh RK, Finlay AY. Dermatology life quality index use in skin disease guidelines and registries worldwide. *J Eur Acad Dermatol Venereol* 2020;34:e822–4.
11 Hurt L, Ashfield-Watt P, Townson J, *et al*. Cohort profile: healthwise wales. A research register and population health data platform with linkage to national health service data sets in wales. *BMJ Open* 2019;9:e031705.
12 Jisc. Online surveys. Jisc B, UK. The online survey tool designed for academic research, education and public sector organisations. 2021. Available: https://www.onlinesurveys.ac.uk/
13 Hyland ME, Sodergren SC. Development of a new type of global quality of life scale, and comparison of performance and preference for 12 global scales. *Qual Life Res* 1996;5:469–80.
14 Prinsen CAC, Lindeboom R, Sprangers MAG, *et al*. Health-Related quality of life assessment in dermatology: interpretation of skindex-29 scores using patient-based anchors. *J Invest Dermatol* 2010;130:1318–22.
15 Habibzadeh F, Habibzadeh P, Yadollahie M. On determining the most appropriate test cut-off value: the case of tests with continuous results. *Biochem Med (Zagreb)* 2016;26:297–307.
16 Unal I. Defining an optimal cut-point value in ROC analysis: an alternative approach. *Comput Math Methods Med* 2017;2017:3762651.
17 Gupta V, Sreenivas V, Mehta M, *et al*. What do vitiligo impact scale-22 scores mean? studying the clinical interpretation of scores using an anchor-based approach. *Br J Dermatol* 2019;180:580–5.
18 Altman DG. *Practical statistics for medical research*. London: Chapman & Hall, 1991.
19 Shah R, Ali FM, Finlay AY, *et al*. Family reported outcomes, an unmet need in the management of a patient's disease: appraisal of the literature. *Health Qual Life Outcomes* 2021;19:194.
20 Marquis P, Chassany O, Abetz L. A comprehensive strategy for the interpretation of quality-of-life data based on existing methods. *Value Health* 2004;7:93–104.
21 Wyrwich KW, Norman GR. The challenges inherent with anchor-based approaches to the interpretation of important change in clinical outcome assessments. *Qual Life Res* 18, 2022.
22 Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 2003;56:395–407.
23 Guyatt GH, Osoba D, Wu AW, *et al*. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002;77:371–83.
24 Chren M-M. Interpretation of quality-of-life scores. *J Invest Dermatol* 2010;130:1207–9.
25 Leshem YA, Hajar T, Hanifin JM, *et al*. What the eczema area and severity index score tells us about the severity of atopic dermatitis: an interpretability study. *Br J Dermatol* 2015;172:1353–7.
26 Charman CR, Venn AJ, Ravenscroft JC, *et al*. Translating patient-oriented eczema measure (POEM) scores into clinical practice by suggesting severity strata derived using anchor-based methods. *Br J Dermatol* 2013;169:1326–32.
27 Aawar N, Moore R, Riley S, *et al*. Interpretation of renal quality of life profile scores in routine clinical practice: an aid to treatment decision-making. *Qual Life Res* 2016;25:1697–702.
28 Harlow D, Poyner T, Finlay AY, *et al*. Impaired quality of life of adults with skin disease in primary care. *Br J Dermatol* 2000;143:979–82.
29 Wijnhoven HAH, Kriegsman DMW, Snoek FJ, *et al*. Gender differences in health-related quality of life among asthma patients. *J Asthma* 2003;40:189–99.