

RESEARCH ARTICLE

Open Access

A short generic patient experience questionnaire: *howRwe* development and validation

Tim Benson^{1,2*} and Henry WW Potts²

Abstract

Background: Patient experience is a key quality outcome for modern health services, but most existing survey methods are long and setting-specific. We identified the need for a short generic questionnaire for tracking patient experience.

Methods: We describe the development and validation of the *howRwe* questionnaire. This has two items relating to clinical care (*treat you kindly; listen and explain*) and two items relating to the organisation of care (*see you promptly; well organised*) as perceived by patients. Each item has four responses (*excellent, good, fair and poor*). The questionnaire was trialled in 828 patients in an orthopaedic pre-operative assessment clinic (PAC).

Results: The *howRwe* questionnaire is shorter (29 words) and more readable (Flesch-Kincaid grade score 2.2) than other questionnaires with broadly similar objectives. Psychometric properties in this sample are good with Cronbach's $\alpha=0.82$. Following a change to the appointments system in the clinic, *howRwe* showed improvement in promptness and organisation, but not in kindness and communication, showing that it can distinguish between the clinical and organisational aspects of patient experience.

Conclusions: *howRwe* meets the criteria for a short generic patient experience questionnaire that is suitable for frequent use. In the validation study of PAC patients, it showed good psychometric properties and concurrent, construct and discriminant validity.

Keywords: Patient experience questionnaire, *howRwe*, Patient-reported outcome measures, Patient-reported experience measures

Background

Patient experience is a key quality outcome for health services and can be used to improve quality, governance, public accountability and patient choice [1]. Since the pioneering work during the 1980s [2,3], the use of patient experience surveys has grown enormously, but there is little evidence of their impact on quality improvement at the local level [4].

Large-scale national surveys address the needs of policymakers for accountability and transparency [5], but there is said to be a "chasm" between the views of senior managers and clinicians at the front line [6]. Traditional methods used have been criticised for survey length, infrequent sampling frequency, slow feedback and failure to use results to improve care [7].

In England the NHS undertakes national surveys of patient experience for inpatient, outpatient, accident and emergency, maternity, mental health and general practice care sectors. Each questionnaire is around 3000 words long and mailed to several hundred patients from each provider. Response rates vary from 34% (General Practice patient survey with one reminder [8]) to 49% for the Inpatient survey (two reminders) [9]. Response rates have declined over the years from 64% in 2001 [10].

At the opposite extreme from long surveys, the Friends and Family Test (FFT) is being introduced across all NHS services [11]. This has a single global question (*how likely are you to recommend this provider to friends of family if they needed similar care or treatment?*) with six possible responses (from *extremely likely* to *extremely unlikely* plus *don't know*) and a free text comment box. A review of the first year of operation showed that the FFT (and in particular the free text comments) could be useful for service improvement by promoting a culture of increased

* Correspondence: tim.benson@r-outcomes.com

¹R-Outcomes Ltd, 14 Pinewood Crescent, Hermitage, Berkshire RG18 9WL, UK

²CHIME, University College London, 222 Euston Road, London NW1 2DA, UK

responsiveness if patient feedback is provided in near real time [12]. However, statistical and methodological problems mean that the FFT should not be regarded as a survey instrument, is not suitable as a comparator across organisations [13], or as a basis for incentive payments. As a single global rating, the FFT cannot show differences between aspects of patient experience and summary scores may have advantages [14].

Some providers have invested in survey systems to support locally designed questionnaires, but these require substantial local effort in design, data collection, analysis and interpretation. Unless questions are standardised, the results are of limited value for comparison, benchmarking and tracking progress. The total costs are often underestimated or unknown.

There is little consensus about what patient experience is [15]. Patient experience, satisfaction, perception, engagement, participation, preferences and outcome are distinct concepts, but only experts appreciate the distinctions. The Beryl Institute's broad definition of patient experience as: "the sum of all interactions, shaped by an organization's culture, that influence patient perceptions, across the continuum of care" [16] does not help much when it comes to measure it.

Short survey instruments reduce the users' burden and various efforts have been made to create short form versions from longer survey instruments, although the difficulties are often underestimated [17].

Most patient experience questionnaires are specific to a particular setting, such as general practice, inpatient, outpatient, maternity, care home or domiciliary care, which limits their use in evaluation across different settings. On the other hand, generic instruments allow comparisons between settings along the patient pathway. However, the use of generic patient experience instruments remains rare.

We identified the need for a short generic patient experience measure to capture patients' perception of their experience with minimal effort and to provide rapid feedback to all stakeholders in a way that is comparable, scalable and economic.

This paper describes the development and testing of this new instrument, called *howRwe*.

Methods

Development of questionnaire

The development of *howRwe* began in 2009. The design criteria were similar to those required for patient-reported quality of life measures [18,19], and in particular that it should be clear, brief, generic, suitable for frequent use, support multi-modal data collection, be responsive and have good psychometric properties.

1. *Clear* – the wording should be simple and unambiguous, so that the instrument can be readily

understood by vulnerable people and translated accurately into other languages.

2. *Brief* – the instrument should be short and hence quick to use by patients or their proxies, if patients are too ill to complete it themselves.
3. *Generic* – the instrument should be generic, applicable without change across all patient categories and care settings, including primary, secondary, community, emergency, domiciliary and social care.
4. *Frequent use* – the instrument should be suitable for frequent and repeated use.
5. *Multimodal* – data collection modalities should include paper, touchscreen devices such as kiosks, smartphones and tablets, web browsers and telephones including automated interactive voice response (IVR) systems.
6. *Responsive* – the instrument should be sensitive to changes and only include items under the day-to-day control of local staff and management. It should exclude aspects, such as location, transport, car-parking, payments and other regulations that cannot easily be changed.
7. *Psychometrics* – the instrument should have good psychometric properties, including validity and reliability.

In addition to these criteria, we wanted the instrument to provide scores for each dimension and a summary score. Results should be easy to understand and interpret by all stakeholders. Feedback should be provided in near real time to enable immediate remedial action by clinicians and managers.

Finally, we wanted the instrument to have a broadly similar look and feel to that of our *howRu* patient-reported outcome measure (PROM) [20].

The core premise of *howRwe* is that all patients want high quality service from staff and from the organisation as a whole. Patient experience can be classified in terms of relationships with staff and system function [21]. The EUROPEP project used the terms *clinical behaviour* and *organisation of care* in evaluating general practice care [22]. Clinical behaviour covers interactions with staff such as kindness and communication; patients are good judges of these relationships. Organisation of care covers access, waiting times, reliability and efficiency; good staff may be let down by poor systems.

The methods used to develop and test the wording included extensive desk research and literature reviews, informal focus groups with patients and staff, and pilot studies over a five-year period across a range of health and social care settings including hospitals, GP surgeries, community services and care homes. Our approach was influenced by agile software development methods [23]. Prototypes were tested and numerous improvements

made. This process evolved through more than 50 distinct versions, with numerous minor changes, testing and refinements.

One of the challenges was to find short generic phrases applicable to all types of patient and setting. For example, we often refer to health and care professionals using setting-specific terms such as GP, physician, surgeon, dentist, nurse, paramedic and social worker. We avoided this issue by not referring explicitly to any staff.

Building on previous experience with *howRu*, we adopted the same pattern of using four short questions, each with four responses. We began by using the same responses as *howRu* (*none, a little, quite a lot and extreme*) focusing on patient concerns. However, feedback from patients and clinicians suggested that this seemed

negative, so we changed the focus to excellence, with choices *excellent, good, fair* and *poor*.

The *howRwe* questionnaire is shown in Figure 1.

The recall period applies to the present current service, because people's memory is less reliable over longer periods [24].

The core question – *How are we doing?* – gives rise to the name of the instrument.

The descriptive system has four items. The items are short and inclusive, rather than restrictive. The items are displayed as in Figure 1 without further guidance, but we give below the intended coverage of each:

1. *Treat you kindly* (kindness) covers how you are treated as a person including compassion, empathy,

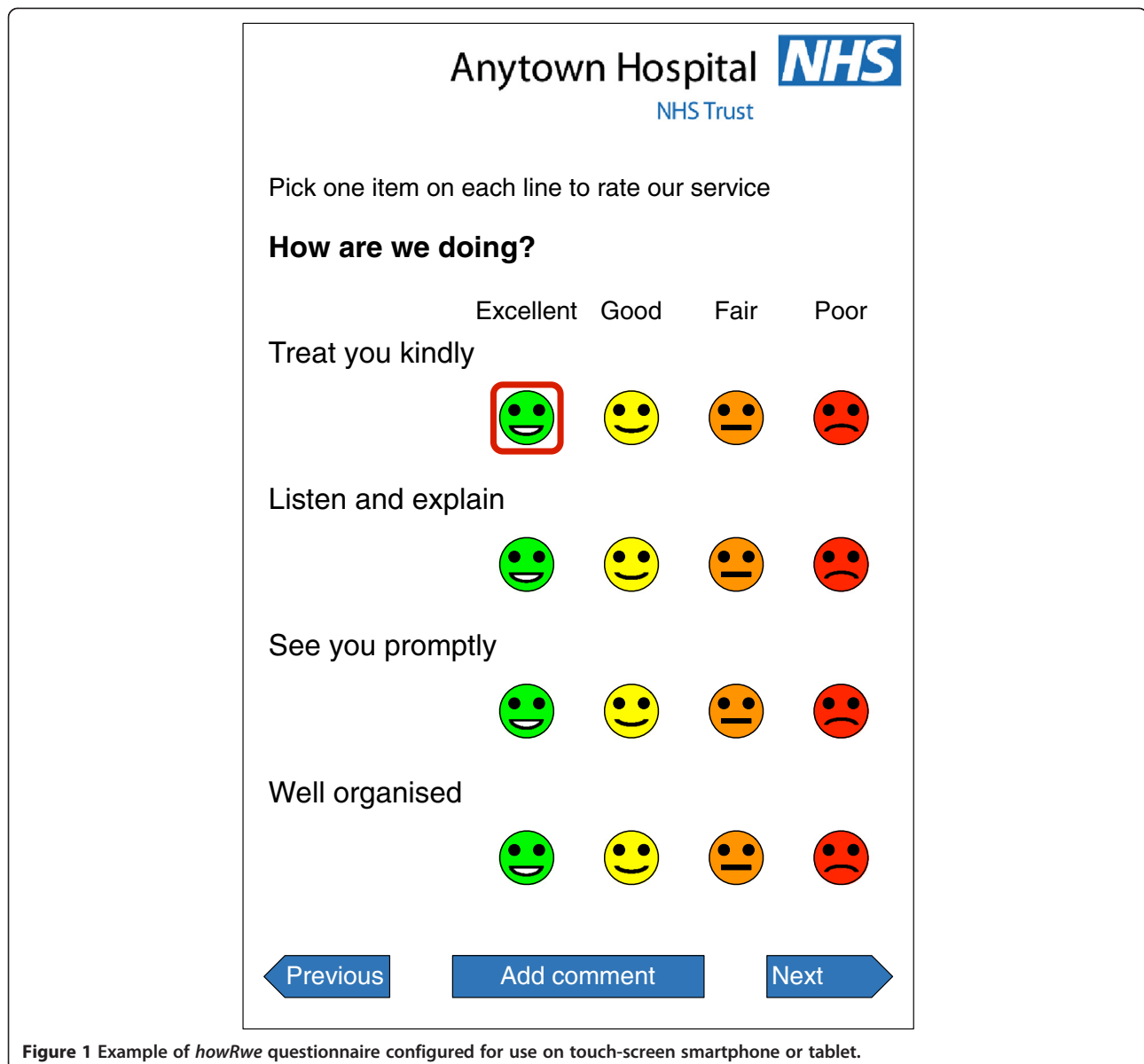


Figure 1 Example of *howRwe* questionnaire configured for use on touch-screen smartphone or tablet.

emotional support, politeness, dignity, respect and privacy.

2. *Listen and explain* (communication) covers all aspects of communication with health staff including patient engagement, information, education, choice, consent, shared decision-making and empowerment.
3. *See you promptly* (promptness) covers delays, waiting, access, cancellations and responsiveness, such as the delay from referral to being seen, waiting to see a clinician, or the time taken to answer a call bell.
4. *Well organised* (organisation) covers how well managed patients perceive the unit to be, including safety, reliability, efficiency, and whether information is available when and where needed and acted on appropriately.

The strength of each item is rated using four levels:

Excellent
Good
Fair
Poor

Each level may be indicated in four mutually supporting ways to minimise cognitive effort, for face validity and avoid the need for training:

Written labels: excellent, good, fair, poor.
Colour: green, yellow, orange and red.
Position: decreasing in excellence from left to right.
Pictographs: based on smiley faces.

Colour, position and pictographs are optional. For example, we propose that *howRwe* can be used in voice-based systems such as interactive voice response (IVR).

The combination of four items with four levels each creates a 4×4 matrix with 256 (4^4) combinations, although many of these are likely to be rare.

For analysis and reporting, each response level for each item is allocated a score on a 0–3 scale:

Excellent: 3
Good: 2
Fair: 1
Poor: 0

The summary *howRwe* score is calculated for individual respondents by adding the scores for each item, giving a scale with 13 possible values from the floor, 0 ($4 \times$ *poor*) to the ceiling, 12 ($4 \times$ *excellent*).

When reporting the results for a group comprising more than one respondent, mean scores are transformed arithmetically to a 0 to 100 scale, where 100 indicates that all respondents rated *excellent* and 0 that all rated

poor. This allows the mean item scores to be compared with the summary *howRwe* score on a common scale.

The *howRwe* questionnaire is generic (i.e. not condition- or domain-specific) and can be used by all types of patients and citizen. If the patient is not able to complete the form personally (e.g. through dementia), a proxy such as a relative may assist or complete it on their behalf, but this should be recorded (using categories such as: *unaided, with help from staff, with help from family, completed by staff as proxy, completed by family as proxy*).

The *howRwe* form usually includes a comment button or text box, providing a way for respondents to add free-text comments to expand on their answers.

In this paper, we set out to test:

- *Internal consistency:* assessed by whether correlations between the four *howRwe* items are moderate to strong, with the strongest correlation between the pairs of items on clinical behaviour and organisation of care (convergent validity), and Cronbach's α is between 0.7 and 0.9.
- *Concurrent validity:* assessed by correlation between the overall *howRwe* score and the NHS Friends and Family Test raw question.
- *Construct validity:* assessed by the measure being sensitive to system change, and system change impacting system function more than relational aspects.
- *Discriminant validity:* assessed by showing low correlation with *howRu*, a patient-reported outcome measure with a similar format.

The *howRwe* instrument is a measure of excellence. We should expect to see a ceiling effect, where service is excellent and the instrument is not able to detect further improvement. We should not expect to find a floor effect, where the service is poor and the measure cannot detect further deterioration.

Validation

For validation, we used data from a pathfinder study in a pre-operative assessment clinic (PAC) for patients scheduled for major orthopaedic operations. This location was chosen because there were known issues and a plan was to be implemented shortly to alleviate these, providing the opportunity for before and after comparison.

Before their operation patients attend the PAC and see up to six different members of the team in the course of a morning. The purpose is to help patients prepare for their operation, discharge home and recovery. Patients are assessed for medical fitness for anaesthetic and have blood, urine and ECG tests. The surgeon explains the proposed operation. Patients also see a physiotherapist

Figure 2 Image of the form used on iPad in validation study.

and occupational therapist and complete consent and other paperwork. All this can take several hours.

Data collection was conducted over seven months from June 2013 to January 2014 as part of routine use of the instrument. Changes to the appointment scheduling system and the physical layout of the clinic were made at the end of the third month (August).

A member of staff asked each patient to complete the questionnaire using an iPad towards the end of their visit and recorded which surgical team and subspecialty was seen. Patients completed the NHS Friends and Family Test, *howRwe*, *howRu* and optional free-text comments. The system recorded the date and time. Data collection used a dedicated iPad linked to Optimum Health Technology's Meridian server. An image of the form used is shown in Figure 2.

Data was exported for analysis using Excel and SPSS. We used Pearson's correlations in the analysis.

Table 1 Length and readability

Instrument	No of items	No of words	FKG readability grade	Reading age
<i>howRwe</i>	4	29	2.2	7.2
NHS Friends & Family Test	1	44	6.6	11.6
GS-PEQ*	10	150	8.8	13.8
EUROPEP 2006*	23	214	8.1	13.1
PPE-15	15	467	7.1	12.1
NHS adult inpatient Survey 2013	76	3,353	7.3	12.3
GP patient survey 2014	62	2,922	6.8	11.8

*Note the word count and readability for GS-PEQ and EUROPEP 2006 are based on the translations of questions and instructions as presented in the original papers. For other surveys, we used the text from actual questionnaires, including instructions.

Table 2 Responses by sub-specialty

Sub-specialty	n	%
Hip and knee replacement	354	43%
Foot and ankle	137	17%
Spinal	140	17%
Sarcoma	72	9%
Shoulders	45	5%
Other	80	10%
Total	828	100%

Research ethics approval was not required because the data was collected anonymously as part of routine use of the instrument for service monitoring. No identification, demographic or medical information was collected on individual patients. All participants freely consented to complete the task.

Readability

Questionnaire readability was measured using the Flesch-Kincaid Readability Grade (FKG) provided in Microsoft Word. The applicability of the FKG and other readability measures to questionnaire items has been questioned, but it remains a widely used tool [25]. We also offer word count as a more robust measure. As a general rule, patients should not be asked to complete questionnaires with a reading age of more than ten [26], which corresponds roughly to readability grade FKG=5.

We compared the length and readability of *howRwe* with the NHS FFT and five other patient experience questionnaires: GS-PEQ [27], EUROPEP 2006 [28], Picker PPE-15 [29], NHS Adult Inpatient Survey 2013 [30] and the GP Patient Survey 2014 [31]. For GS-PEQ and EUROPEP 2006, we used the English translations of questions and instructions as presented in original papers, which may misrepresent their performance in their original language. For the other questionnaires, we used the full text, including instructions, framing statements, questions and responses as used in surveys.

Results

Table 1 shows the number of items, the number of words, the FKG readability grade and approximate reading age for the questionnaires assessed. *howRwe* has FKG=2.2 (reading age 7.2). The other measures have

Table 3 Distribution of responses for each item (%)

Item	Excellent	Good	Fair	Poor
Treat me kindly	671 (81.0%)	148 (17.9%)	7 (0.8%)	2 (0.2%)
Listen and explain	609 (73.6%)	202 (24.4%)	16 (1.9%)	1 (0.1%)
See me promptly	447 (54.0%)	273 (33.0%)	90 (10.9%)	18 (2.2%)
Well organised	530 (64.0%)	233 (28.1%)	51 (6.2%)	14 (1.7%)

Table 4 Distribution of *howRwe* aggregate scores

<i>howRwe</i> score	n	%
12	404	48.8%
11	107	12.9%
10	77	9.3%
9	56	6.8%
8	123	14.9%
7	25	3.0%
6	21	2.5%
5	8	1.0%
4	5	0.6%
3	1	0.1%
2	0	0.0%
1	0	0.0%
0	1	0.1%
Total	828	100.0%

FKG in the range 6.6 to 8.8 (reading age 11.6 to 13.8), each of which is substantially greater than the criterion of FKG less than 5 (reading age 10).

In total, 828 respondents completed all parts of the *howRwe* questionnaire. Five respondents (0.6%) answered “Don’t know” to the FFT, so their FFT scores were coded as missing. We do not have a record of how many patients were asked but declined to participate. The numbers for each orthopaedic sub-specialty are shown in Table 2.

The distribution of responses for each item is shown in Table 3. 68% of all item responses were *excellent*, 26% *good*, 5% *fair* and 1% *poor*. As expected we found a ceiling effect with a large proportion of responses being *excellent*. The differences between items help identify aspects that need improvement.

The *howRwe* score is the aggregate of the four *howRwe* items. Table 4 shows the distribution of responses. 404 respondents (48.8%) rated the service as *excellent* in all respects (ceiling score), while only one (0.1%) rated it as *poor* in all respects (floor score).

Table 5 shows the mean patient score (raw data), the item score on a 0–100 scale, 95% confidence limits and standard deviation for each item and the aggregate

Table 5 Mean scores for each item and *howRwe* score

Item	Mean score (raw data)	Mean score (0–100 scale)	95% confidence limits	St dev
Treat me kindly	2.80	93.2	92.2 – 94.2	14.7
Listen and explain	2.72	90.5	89.3 – 91.6	16.7
See me promptly	2.39	79.6	77.9 – 81.3	25.5
Well organised	2.54	84.8	83.3 – 88.1	22.9
<i>howRwe</i> score	10.44	87.0	85.9 – 88.1	16.4

Table 6 Intra-item correlation matrix (95% confidence intervals)

	Listen and explain	See me promptly	Well organised
Treat me kindly	0.71 (0.67, 0.74)	0.39 (0.33, 0.44)	0.51 (0.46, 0.56)
Listen and explain		0.47 (0.42, 0.52)	0.56 (0.51, 0.60)
See me promptly			0.70 (0.66, 0.73)

howRwe score. *Treat me kindly* has the highest item score (93.2) and *See me promptly* the lowest (79.6).

The inter-item correlation matrix is shown in Table 6. The correlation between the clinical care items, *Treat me kindly* and *Listen and explain* is high ($r=0.71$), as is the correlation between the two organisation of care items, *See me promptly* and *Well organised* ($r=0.70$). The correlations between the other items are in the range $r=0.39$ to $r=0.56$.

A factor analysis of the four items found a single factor explaining 67% of the variance, demonstrating unidimensionality. The four eigenvalues were 2.67, 0.76, 0.30 and 0.28. The internal consistency reliability was satisfactory (Cronbach's $\alpha=0.82$; 95% CI: 0.79, 0.83).

The correlation between each item and the sum of the other three items is shown in Table 7. These all lie in the range $r=0.60$ to $r=0.74$. The correlation of each item to the FFT question responses is also shown. These correlations are negative due to the way that the FFT is marked (good is low). These all lie in the range $r=-0.36$ to $r=-0.48$. The correlation between the aggregate *howRwe* score and the FFT question is $r=-0.53$ (-0.58, -0.48).

The *howRwe* summary score and the *howRu* summary score (a measure of patient health status) show no significant correlation ($r=0.02$; 95% CI: -0.04, 0.09). Table 7 also shows the correlations between the individual *howRwe* items and the *howRu* summary score, which are also minimal.

Table 8 and Figure 3 show the *howRwe* item scores (on 0–100 scale) before and after changes to the appointments system. There is no significant change to the clinical care items (*Treat me kindly* and *Listen and explain*), but significant improvements in the organisation of care items (*See me promptly* and *Well organised*).

Table 7 Correlations between each *howRwe* items and the sum of the other three items, the Friends and Family Test (FFT) question and the *howRu* summary score

howRwe item	Sum of the other three <i>howRwe</i> items (<i>r</i>)	Item to FFT question (<i>r</i>)	Item to <i>howRu</i> summary score (<i>r</i>)
Treat me kindly	0.60 (0.56, 0.64)	-0.36 (-0.42, -0.30)	0.08 (0.01, 0.15)
Listen and explain	0.66 (0.62, 0.70)	-0.42 (-0.47, -0.36)	0.08 (0.01, 0.15)
See me promptly	0.64 (0.60, 0.68)	-0.44 (-0.49, -0.38)	-0.06 (-0.13, 0.01)
Well organised	0.74 (0.71, 0.77)	-0.48 (-0.53, -0.43)	0.02 (-0.05, 0.09)
<i>howRwe</i> summary score	-	-0.53 (-0.58, -0.48)	0.02 (-0.05, 0.09)

Discussion

howRwe is the first short generic patient experience measure we know of that has been designed for use across all health and social care sectors. Its practicality at scale and in social care was demonstrated in a survey in 360 care homes in the UK, Australia and New Zealand, completed by over 10,000 residents [32]. When we began this work, we were not aware of any other short generic patient experience questionnaire, but we have since become aware of the GS-PEQ (Generic Short Patient Experience Questionnaire), which has been developed in Norway as a generic short form of the Norwegian national sector-specific surveys [27].

This is the first published report of the development of *howRwe*. We have shown that *howRwe* is shorter than other measures, has good readability statistics and internal validity.

A Cronbach's α of 0.82 in this sample suggests that it is appropriate to use the overall *howRwe* score, as well as individual item scores [33]. Construct validity and sensitivity were shown by how scores responded to a change in the appointments system. The instrument shows minimal correlation with the *howRu* measure of quality of life, despite the shared layout.

Development

The usual method of developing new measures is to set out a development protocol for a funded piece of work using a methodology that allows people without specific domain knowledge to develop instruments within a specified time scale.

The development of short patient experience questionnaires often start with a long measure or long sets of statements, which are refined to create a short form with far fewer items. This is how GS-PEQ and PPE-15 were developed. A different short form approach, which does not result in a new questionnaire, is to derive a scale from secondary analysis of a long form data set. The Oxford Patient Involvement and Experience scale (OxPIE) was derived in this way from the NHS Inpatient Survey 2011 [34].

Our approach was different, but allowed a new type of short generic measure to evolve.

Table 8 *howRwe* item scores before and after changes to appointments system

<i>howRwe</i> Item	Before change (95% confidence limits)	After change (95% confidence limits)	Mann-Whitney test
Responses (n)	278	550	
Treat me kindly	93.9 (92.2 – 95.6)	92.9 (91.7 – 94.1)	$z=0.9, p=0.4$
Listen and explain	89.4 (87.5 – 91.4)	91.0 (89.6 – 92.4)	$z=1.0, p=0.3$
See me promptly	71.5 (68.5 – 74.5)	83.7 (81.6 – 85.8)	$z=6.1, p<0.0001$
Well organised	78.9 (76.2 – 81.6)	87.8 (85.9 – 89.7)	$z=4.7, p<0.0001$

Study limitations

Without a gold standard for patient experience instruments and with limited consensus about most important dimensions of patient experience, it is difficult to test content validity.

The data was collected as part of routine anonymous use, rather than as a special validation study, so we do not have additional demographic or clinical data that could be used for construct validation.

The patients in this sample were undergoing a distinct episode of care. In other settings, care is extended over time and multiple healthcare staff, which may make responses from the questions asked in *howRwe* more difficult to interpret. *howRwe* has also been used successfully with hospital inpatients, general practice and community service patients and care home residents.

Although *howRwe* was designed to be applicable and comparable across multiple care sectors, this study used a relatively homogeneous respondent sample of orthopaedic patients attending a pre-operative assessment clinic, with more than 40% due to have hip or knee replacements. We recognise the long-standing debates about the relative value of generic and specific measures (e.g. [35,36]) and it is important to test *howRwe* in further groups.

The data was collected using a questionnaire on an iPad, which included the FFT, *howRwe* and *howRu* questions on a single screen (Figure 2). The wording of each item is identical to that shown in Figure 1, but the order is different. The reason for changing the order is to put the items about clinical care and organisation of care together. We do not consider that this change in order impacts any of the conclusions drawn.

We were not able to measure response rates in the study. Individual patients were asked to complete the questionnaire by a member of staff, who tended not to do this when very busy. We do not know how many patients declined the offer to take part, or who only completed part of the survey, because data was only submitted to the database when the questionnaire was complete. A comparable paper-based study using *howRu* obtained high completion rates [37].

In the FFT, the mode of administration and patient demographics has an impact on both response rates and scores [12]. This study did not provide any way of testing for these effects in *howRwe*.

It would be valuable to investigate the instrument's test-retest reliability and to investigate further discriminant validity against, for example, measures of personality.

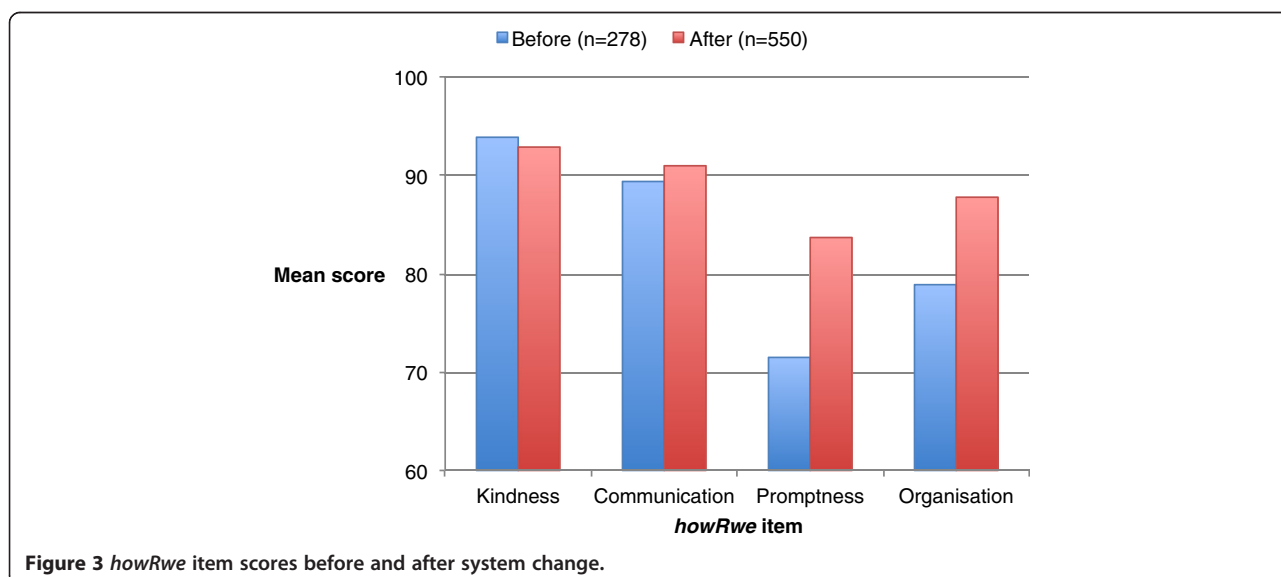


Figure 3 *howRwe* item scores before and after system change.

Implications for practice

The time and effort involved in monitoring patient experience is a source of concern. Response rates for long questionnaires are falling. Patients complain of survey fatigue and the surveys are expensive to administer. For example the Guidance Manual for the 2013 Inpatient Survey runs to 34,600 words [38]. Response rates for NHS national surveys are all below 50%, in spite of up to two reminders.

The local impact of large national surveys has been less than might have been hoped. Feedback needs to be quicker, ward-specific, include patient comments and offer staff an opportunity to discuss it [39]. Existing patient experience instruments are also specific to the mode of care, making it difficult to compare across modes of treatment.

The correlation between *howRwe* and the FFT was $r = -0.53$, indicating that they are not measuring quite the same things. The FFT is a global rating of recommendation, which is related to morale, optimism and loyalty, while *howRwe* measures patient perceptions of different aspects of service. Overall summary scores, such as the *howRwe* score, may perform better than global ratings (such as the FFT) as a way of summarising patients' experiences [14].

Conclusions

howRwe is a short generic patient experience measure. The questionnaire is shorter (29 words) and more readable (Flesch-Kincaid grade score 2.2) than other widely used instruments. It minimises respondent burden, allows rapid feedback and comparisons to be made between different care settings either within an organisation or across the patient pathway. Psychometric properties are good.

howRwe is being used in the UK in secondary, primary, community and social care providers. It adds value by distinguishing between the aspects of patient experience that relate to clinical care (kindness and communication) and organisation of care (promptness and organisation).

We hope that future work will investigate further psychometric properties of the instrument and test its role in promoting change and improvement of quality.

Abbreviations

FFT: Friends and Family Test; GP: General Practice/Practitioner; IVR: Interactive voice response; NICE: National Institute for Health and Care Excellence; NHS: National Health Service; PREM: Patient-reported experience measure; PROM: Patient-reported outcome measure; PAC: Pre-operative assessment clinic.

Competing interests

TB owns the intellectual property rights for *howRwe*. He is also a director of R-Outcomes Ltd, which provides quality improvement services using *howRwe*. The authors declare that they have no other competing interests.

Authors' contributions

TB designed the study, collected the data and wrote the first draft of the paper. TB developed *howRwe*. Both authors designed the analysis. HWWP

further analysed the data and both authors contributed to the final text. Both authors read and approved the final manuscript.

Authors' information

Those wishing to use the *howRwe* should contact R-Outcomes Ltd (tim.benson@r-outcomes.com).

Acknowledgement

We are grateful to the management, staff and patients at the Nuffield Orthopaedic Centre, Oxford University Hospitals, for helping with and participating in the data collection, and to Optimum Contact Ltd for supplying the equipment and software used to collect the data. We thank our reviewers for their useful comments and suggestions.

Received: 8 May 2014 Accepted: 6 October 2014

Published online: 22 October 2014

References

1. Ahmed F, Burt J, Rowland M: **Measuring patient experience: concepts and methods.** In *The Patient – Patient-Centered Outcomes Research*. doi:10.1007/s40271-014-0060-5. published online: 16 May 2014.
2. Gerteis M, Edgman-Levitan S, Daley J, Delbanco T (Eds): *Through the Patient's Eyes: Understanding and Promoting Patient-Centered Care*. San Francisco: Jossey-Bass; 1993.
3. Press I: **Concern for the patient's experience comes of age.** *Patient Exp J* 2014, **1**:4–6.
4. Haugum M, Danielsen K, Iversen HH, Bjertnaes O: **The use of data from national and other large-scale user experience surveys in local quality work: a systematic review.** *International J Qual Health Care* 2014:1–14. doi:10.1093/intqhc/mzu077.
5. Coulter A, Locock L, Ziebland S, Calabrese J: **Collecting data on patient experience is not enough: they must be used to improve care.** *BMJ* 2014, **348**:g2225. doi:10.1136/bmj.g2225.
6. Rozenblum R, Lisby M, Hockey P, Levtzion-Korach O, Salzberg C, Efrati N, Lipsitz S, Bates D: **The patient satisfaction chasm: the gap between hospital management and frontline clinicians.** *BMJ Qual Saf* 2013, **22**:242–250.
7. Robert G, Cornwell J: **Rethinking policy approaches to measuring and improving patient experience.** *J Health Serv Res Policy* 2013, **18**(2):67–69.
8. Ipsos MORI: *The GP Patient Survey: National Summary Report*. London: 2014. http://gp-survey-production.s3.amazonaws.com/archive/2014/July/1301375001_Y8W2%20National%20Summary%20Report_FINAL%20v1.pdf accessed 15-9-2014.
9. Care Quality Commission (CQC): *Technical Details – Patient Survey Information 2013 Inpatient Survey*. 2014. http://www.cqc.org.uk/sites/default/files/inpatient_survey_technical_document.pdf accessed 15-9-2014.
10. Department of Health: *Acute Inpatient Survey: National Overview 2001/02*. London: 2003. http://webarchive.nationalarchives.gov.uk/+www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsStatistics/DH_4066592 accessed 15-9-2014.
11. NHS England: *The Friends and Family Test*, Publications Gateway Ref No 01787. 2014. http://www.nhs.uk/NHSEngland/AboutNHSServices/Documents/FFTGuide_Final_1807_FINAL.pdf accessed 15-9-2014.
12. NHS England: *Review of the Friends and Family Test*. 2014. <http://www.england.nhs.uk/wp-content/uploads/2014/07/fft-rev1.pdf> accessed 15-9-2014.
13. Sizmur S, Graham C, Walsh J: **Influence of patients' age and sex and the mode of administration on results from the NHS Friends and Family Test of patient experience.** *J Health Serv Res Policy* 2014, doi:10.1177/1355819614536887. Published online before print.
14. Krol M, de Boer D, Rademakers J, Delnoij D: **Overall scores as an alternative to global ratings in patient experience surveys; a comparison of four methods.** *BMC Health Serv Res* 2013, **13**:479.
15. LaVela S, Gallan A: **Evaluation and measurement of patient experience.** *Patient Exp J* 2014, **1**:28–36.
16. Wolf J, Niederhauser V, Marshburn D, LaVela S: **Defining patient experience.** *Patient Exp J* 2014, **1**:7–19.
17. Smith G, McCarthy D, Anderson K: **On the sins of short-form development.** *Psychol Assess* 2000, **12**(1):102–111.
18. Fitzpatrick R, Fletcher A, Gore S, Jones D, Spiegelhalter D, Cox D: **Quality of life measures in health care. I: Applications and issues in assessment.** *BMJ* 1992, **305**(6861):1074–1077.

19. Eton D, Beebe T, Hgen P, Halyard M, Montori V, Naessens J, Sloan J, Thompson C, Wood D: **Harmonizing and consolidating the measurement of patient-reported information at health care institutions: a position statement of the Mayo Clinic.** *Patient Relat Outcome Meas* 2014, **5**:7–15.
20. Benson T, Whatling J, Arikan S, Sizmur S, McDonald D, Ingram D: **Evaluation of a new short generic measure of health status: howRu.** *Inform Prim Care* 2010, **18**:89–101.
21. Robert G, Cornwall J, Brearley S, Foot C, Goodrich J, Joule N, Levenson R, Maben J, Murrells T, Tsianakas V, Waite D: *What Matters to Patients; Developing the Evidence Base for Measuring and Improving the Patient Experience*, A report for the Department of Health and NHS Institute for Innovation and Improvement. Warwick: NHS Institute for Innovation & Improvement; 2011. http://www.institute.nhs.uk/images/Patient_Experience/Final%20Project%20Report%20pdf%20doc%20January%202012.pdf accessed 15-9-2014.
22. Wensing M, Mainz J, Grol R: **A standardised instrument for patient evaluations of general practice care in Europe.** *Eur J Gen Pract* 2000, **6**:82–87.
23. Beck K, Beedle M, van Bennekum A, Cockburn A, Cunningham W, Fowler M, Grenning J, Highsmith J, Hunt A, Jeffries R, Kern J, Marick B, Martin RC, Mellor S, Schwaber K, Sutherland J, Thomas D: *Manifesto for Agile Software Development*. Agile Alliance; 2001. <http://agilemanifesto.org/> accessed 13-9-2014.
24. Stull D, Kline Leidy N, Parasuraman B, Chassany O: **Optimal recall periods for patient-reported outcomes: challenges and potential solutions.** *Curr Med Res Opin* 2009, **25**:929–942.
25. Lenzner T: **Are readability formulas valid tools for assessing survey question difficulty?** *Sociol Methods Res* 2013, doi:10.1177/0049124113513436. Published online before print December 10, 2013.
26. Paz S, Jiu H, Fongwa M, Morales L, Hays R: **Readability estimates for commonly used health-related quality of life surveys.** *Qual Life Res* 2009, **18**:889–900.
27. Sjetne I, Bjertnaes O, Olsen R, Iversen H, Bukholm G: **The Generic Short Patient Experiences Questionnaire (GS-PEQ): identification of core items from a survey in Norway.** *BMC Health Serv Res* 2011, **11**:88.
28. Wensing M, Baker R, Vedsted P, Heje H, Klingenberg A, Broge B, Kersnik J, Seuntjens L, Kunzi B, Milano M, Mola E, Elwyn G, Grol R: *EUROPEP 2006 Revised Europep Instrument and User Manual*. TOPAS Europe Association; 2006. http://www.topaseurope.eu/files/Europep%202006rapport_0.pdf accessed 12-9-2014.
29. Jenkinson C, Coulter A, Bruster S: **The Picker Patient Experience Questionnaire: development and validation using data from in-patient surveys in five countries.** *International J Qual Health Care* 2002, **14**:353–358.
30. Picker Institute Europe: *NHS Adult Inpatient Survey 2013*. Oxford: 2013. http://www.nhssurveys.org/Filestore//Inpatient_2013/IP13_Core_Questionnaire_v1.pdf accessed 15-9-2014.
31. Ipsos MORI: *The GP Patient Survey*. London: 2014. <http://gp-survey-production.s3.amazonaws.com/archive/2014/July/qandletter/1301375001%20Y8W2%20Questionnaire%20FINAL.pdf> accessed 15-9-2014.
32. Benson T, Bowman C: **Health-related quality of life, outcome and patient experience in care homes: a study in three countries.** *Medicine* 2013, <http://www.medicine20congress.com/ocs/index.php/med/med2013/paper/view/1365> accessed 15-9-2014.
33. Streiner D, Norman G: *Health Measurement Scales: A Practical Guide to their Development and Use*. 4th edition. Oxford: Oxford University Press; 2008.
34. Hewitson P, Skew A, Graham C, Jenkinson C, Coulter A: **People with limiting long-term conditions report poorer experiences and more problems with hospital care.** *BMC Health Serv Res* 2014, **14**:33.
35. Singh S, Sodergren S, Hyland M, Williams J, Morgan M: **A comparison of three disease-specific and two generic health-status measures to evaluate the outcome of pulmonary rehabilitation in COPD.** *Respir Med* 2001, **95**(1):71–77.
36. Amadio P, Silverstein M, Ilstrup D, Schleck C, Jensen L: **Outcome assessment for carpal tunnel surgery: The relative responsiveness of generic, arthritis-specific, disease-specific, and physical examination measures.** *J Hand Surg* 1996, **21**(3):338–346.
37. Benson T, Potts HWW, Whatling JM, Patterson D: **Comparison of howRU and EQ-5D measures of health-related quality of life in an outpatient clinic.** *Inform Prim Care* 2013, **21**(1):12–17.
38. Picker Institute Europe: *Guidance Manual for the NHS Adult Inpatient Survey 2013*. Oxford: 2013. http://www.nhssurveys.org/Filestore//Inpatient_2013/IP13_Guidance_v12.pdf accessed 15-9-2014.
39. Reeves R, West E, Barron D: **Facilitated patient experience feedback can improve nursing care: a pilot study for a phase III cluster randomised controlled trial.** *BMC Health Serv Res* 2013, **13**:259.

doi:10.1186/s12913-014-0499-z

Cite this article as: Benson and Potts: A short generic patient experience questionnaire: howRwe development and validation. *BMC Health Services Research* 2014 **14**:499.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

