ORIGINAL ARTICLE

Development of the osteoporosis assessment questionnaire—physical function (OPAQ-PF): an osteoporosis-targeted, patient-reported outcomes (PRO) measure of physical function

A. N. Naegeli · A. Nixon · R. Burge · D. T. Gold · S. Silverman

Abstract

Summary We have developed a short, patient-reported outcome questionnaire—the Osteoporosis Assessment Questionnaire—Physical Function (OPAQ-PF)—that assesses the impact of osteoporosis on physical function. OPAQ-PF contains 15 items in three domains (mobility, physical positions, and transfers) and has content validity in osteoporosis patients with and without a history of fracture.

Introduction This paper describes the development of the Osteoporosis Assessment Questionnaire—Physical Function (OPAQ-PF), a patient-reported outcome (PRO) questionnaire based on OPAQ v.2.0 (60 items, 14 domains) that assesses the impact of osteoporosis on physical function.

Methods OPAQ v.2.0 was administered to patients with osteoporosis. Item response theory methodology and clinical judgment were used to retain/eliminate items. The resulting instrument was modified during two sets of concept elicitation and cognitive debriefing interviews with osteoporosis patients.

Results Item response theory-based analysis of OPAQ v.2.0 ($n=1,478$) coupled with clinician input resulted in the generation of a 21-item, six-domain instrument with a frequency response format. Interview data from 32 participants were used to modify this version and led to generation of the final instrument, OPAQ-PF. This final version has a severity response format and contains 15 items in three domains (mobility, physical positions, and transfers) that group together to provide an overall assessment of physical function in patients with osteoporosis. Twenty-two of the 32 interview participants (69 %) had previously sustained a fracture. Symptoms occurred primarily in these patients.

Conclusions OPAQ-PF represents a brief, focused, PRO instrument that assesses physical function in patients with osteoporosis, specifically related to mobility, physical positions, and transfers. This questionnaire has content validity in osteoporosis patients who have, and have not, sustained a prior fracture.

Keywords Fracture · Low bone mass density · Osteoporosis · Patient-reported outcomes · Physical function · Questionnaire

A. N. Naegeli (✉) · R. Burge
Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, USA
e-mail: naegeli_april_n@lilly.com

A. Nixon
Oxford Outcomes, an ICON PLC company, Oxford, UK

D. T. Gold
Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA

D. T. Gold
Department of Sociology, Duke University Medical Center, Durham, NC, USA

D. T. Gold
Department of Psychology and Neuroscience, Duke University Medical Center, Durham, NC, USA

S. Silverman
Cedars-Sinai Medical Center, Los Angeles, CA, USA

Introduction

Osteoporosis is a systemic skeletal disease characterized by micro-architectural deterioration of bone with resultant low bone mass, bone fragility, and increased fracture risk [1]. Osteoporosis-related fractures, which most commonly occur at the hip, spine, and wrist, may be followed by full recovery or by chronic pain, disability, and death [1].

Osteoporosis is most prevalent in middle-aged and elderly adults, and currently affects approximately 10 million individuals in the USA [2]. It is estimated that up to 50 % of women and 25 % of men over the age of 50 years will experience an osteoporotic fracture in their remaining lifetime [2].

The effects of osteoporotic fracture on morbidity and mortality are significant. In a retrospective US Medicare claims

🖄 Springer

database analysis of over 97,000 patients with vertebral compression fractures, the hazard ratio for mortality vs. control patients was 1.83 (95 % confidence interval [CI], 1.80–1.86) [3]. Similarly, the prospective US Study of Osteoporotic Fractures found that, compared with women without vertebral fracture, women with ≥1 vertebral fracture had a 1.23-fold greater age-adjusted mortality rate (95 % CI, 1.10–1.37) [4]. Mortality increased with the number of vertebral fractures, rising from 19 per 1,000 woman-years in those without fractures to 44 per 1,000 woman-years in those with ≥5 fractures (p for trend <0.001).

Osteoporotic fracture-associated morbidity may impact on patients in several ways, including impaired physical functioning, disability, depression, social isolation, pain, loss of independence, and decreased quality of life [5–7]. Many such consequences can be measured using an appropriate specific patient-reported outcome (PRO) instrument.

The Osteoporosis Assessment Questionnaire (OPAQ) versions 1.0, 2.0, and short version are validated, reliable PRO measures used extensively in clinical trials to assess patient outcomes in individuals with osteoporosis [8–14]. The instruments were developed as disease-targeted questionnaires that would discriminate between postmenopausal women with and without osteoporotic fracture [11], and were also intended to be used as evaluative instruments in clinical trials [11].

The OPAQ v.1.0 contained 84 questions in 18 domains and four dimensions (physical function, emotional status, symptoms, and social interactions), plus 18 questions measuring satisfaction with each of the domains [11]. In 2000, Silverman modified the OPAQ and created v.2.0, a 14-domain, 60-item questionnaire that retained the same four dimensions as v.1.0 [11]. Our aim was to create a short, osteoporosis-targeted instrument, based on OPAQ v.2.0 and known as OPAQ—Physical Function (OPAQ-PF), which could be used in clinical trials to evaluate the impact of new osteoporosis treatments on patients' outcomes. Initially, we sought to develop a measure of the impact of osteoporosis on the dimensions of physical functioning, fear of falling, independence, and symptoms. However, this objective was re-evaluated and modified based on the interim results, and the instrument was refocused on physical function only. This paper describes the development of OPAQ-PF.

Methods

The study was conducted in two phases.

Phase 1: item elimination

Phase 1 consisted of a post hoc analysis of data generated when the 60-item OPAQ v.2.0 was administered, at the study baseline visit, to 1,478 patients enrolled in the Multiple

Outcomes of Raloxifene Evaluation (MORE) trial [15]. This phase 3, multicenter, double-blind, placebo-controlled, randomized clinical trial enrolled ambulatory, postmenopausal women aged ≤80 years with a diagnosis of osteoporosis (defined as the presence of vertebral fractures or a femoral neck or vertebral spine T-score of ≤−2.5) [15].

Each of the 60 items was analyzed using item response theory (IRT) methodology. First, exploratory factor analysis was used to confirm unidimensionality of each of the 14 domains independently. For each domain, a scree plot was used to determine whether only one construct was being measured in MORE clinical trial population.

Next, two sets of graphs (item characteristic curves [ICCs] and item information curves [IICs]) were generated to demonstrate how well items reflected the concept being measured, to provide graphical representations of the floor and ceiling effects of patient responses to each item, and to act as a focus for discussing the clinical relevance of the measured concepts (data not shown). The ICCs were used to assess each item's ability to discriminate across the continuum of the underlying construct experienced by patients. The extent to which each item was related to the underlying construct, and the range over which the item could distinguish responses, were determined using the IICs. Analyses were conducted using Mplus (Muthén and Muthén, Los Angeles, CA, USA) statistical software. More information on IRT methodology can be found in the article by Edelen and Reeve [16].

Items and responses were modified or subdivided, if necessary, and new items and responses could be added. Criteria for retaining items included: good IRT item performance (based on visual assessment of ICCs and IICs); good discrimination within a wide range of the construct; clinical relevance as assessed by two of the authors (SS, DTG); and construct relevance. The aim of this study phase was to identify items that performed well according to IRT criteria, and to use IRT results as a major factor in reducing the initial set of 60 items to one more closely focused on physical function (a dimension including mobility, physical positions, and transfers), fear of falling, independence, and symptoms, and amenable to qualitative investigation based on concept elicitation and cognitive debriefing interviews with osteoporosis patients.

Phase 2: qualitative research

Phase 2 comprised a cross-sectional, iterative, qualitative investigation to determine the impact of osteoporosis on patients' lives, to evaluate the suitability of the interim version of OPAQ generated in phase 1 as an endpoint in clinical trials, and to clarify the conceptual focus of the final instrument. This involved conducting concept elicitation and cognitive debriefing interviews on the interim version of the OPAQ. Interviews were conducted in 2010 and 2011 according to a semi-structured guide. This study phase was conducted in two

discrete stages ('first stage' and 'second stage'), each with a separate recruitment process. Substantial modifications were made to the instrument between these stages so that it focused solely on physical function in the second stage.

Interviews were audio-recorded and transcribed to facilitate data analysis. Methodology and analyses were in line with the US Food and Drug Administration (FDA)'s guidance on PRO instrument development and modification, and recent International Society for Pharmacoeconomics and Outcomes Research (ISPOR) recommendations [17–19]. Demographic and interview data are reported separately for the two stages of this phase.

Study population

Participants were recruited through three clinical sites in the USA. Full Institutional Review Board (IRB) approval was obtained from the Western IRB, Olympia, Washington, USA (first stage), and the Independent IRB, Plantation, Florida, USA (second stage), prior to patient recruitment.

Inclusion criteria were: female gender, ≥50 years of age, ambulatory (able to walk with or without assistance), and diagnosis of osteoporosis at least 1 year previously. The final study population included patients with osteoporosis of differing degrees of severity (who were therefore at different levels of fracture risk), and patients who had, and had not, experienced an osteoporosis-related fracture. Patients were allocated by a clinician to one of three diversity groups, according to the following inclusion criteria:

- Diversity group 1: (T-score between −1.0 and −2.5 at the femoral neck or spine and 10-year probability of hip fracture ≥3 %) or 10-year probability of major osteoporosis-related fracture ≥20 %.

- Fracture probabilities were based on the World Health Organization's 'FRAX®' algorithm, which estimates 10-year probabilities of hip fracture and major osteoporotic fracture (defined as clinical vertebral, hip, forearm, or proximal humerus fracture) based on the patient's femoral neck bone mineral density and clinical risk factors [1].

- Diversity group 2: T-score≤−2.5 and fragility fracture of the usual osteoporosis fracture sites (e.g., spine, wrist, hip, rib, or pelvis) in the past 12 months, as determined by the attending physician.
- Diversity group 3: T-score≤−2.5 and fragility fracture of the usual osteoporosis fracture sites (e.g., spine, wrist, hip, rib, or pelvis) >12 months previously.

The study protocol mandated the exclusion of patients with comorbid conditions that would affect their ability to differentiate any symptoms and impacts of osteoporosis from symptoms/impacts of other conditions. No patient participated in both stages.

Demographic and medical history data were provided on structured forms completed by the patient or clinical site staff and were summarized using descriptive statistics.

Interviews: concept elicitation

Interviews commenced after patients had provided written informed consent. Semi-structured, qualitative, one-on-one concept elicitation interviews involved the interviewer asking each participant questions about osteoporosis-related symptoms and impacts that were important to them. Patients were excluded from the analysis if they were unable to differentiate between osteoporosis and comorbid conditions as the cause of symptoms or impacts throughout the discussion. Patients who could discuss symptoms/impacts of osteoporosis specifically at some point in the interview were retained in the full analysis, but any symptoms/impacts that they had difficulty in attributing specifically to osteoporosis were excluded. Issues related to OPAQ dimensions/domains of interest gathered during these interviews provided evidence for content validity of the new instrument.

The resulting data were analyzed using a thematic analysis approach [20]. This involved reading and re-reading the data to identify themes and categories that centered on particular phrases, incidents, and types of behavior, in line with concepts and themes outlined in the interview guide. The codes used were captured in a codebook and in an evidence-based coding frame that were continuously updated as new categories and codes emerged. As each interview transcript was analyzed, the number of new codes generated by that transcript was recorded and used to determine saturation (the point at which no new categories, concepts, dimensions, or incidents emerged during the theory development process) [21]. Qualitative data analysis was assisted by using ATLAS.ti software version 5.7.1 (Cleverbridge, Chicago, IL, USA).

Interviews: cognitive debriefing

Following concept elicitation, participants were asked to complete the interim version of the OPAQ. The interviewer then asked participants for their thoughts and opinions on the general design of the instrument, item semantics, applicability and interpretation, response options, and recall period. Analysis of cognitive debriefing interviews was conducted on an overall questionnaire basis and on an item-by-item basis [22], with the goal of evaluating and improving the instrument's content validity. This included identifying items that presented cognitive challenges.

The questionnaire remained open to modification throughout the debriefing process. New items were added, redundant items removed, and minor changes made to the introduction, item wording, response options, and layout.

Results

Phase 1

Unidimensionality was confirmed for each domain of the OPAQ v.2.0. Information generated by the ICCs and IICs (available from the corresponding author) was used in conjunction with expert opinion (SS and DTG are both globally renowned key thought leaders on quality of life issues and measurement in osteoporosis) to make decisions regarding item deletion, retention, modification, or subdivision (e.g., "How often did you have trouble either walking one block or climbing one flight of stairs?" was divided into two questions: "How often did you have trouble walking one block?" and "How often did you have trouble climbing stairs or steps?"). Items were included in the interim version of OPAQ only if deemed relevant to the overall concepts of physical function, fear of falling, independence, and symptoms that were the original intended focus of the final questionnaire. The primary reason for item retention was good endorsement of the concept by IRT curves. However, some items that measured a clinically important aspect of the underlying construct were retained based on expert opinion, even if their ICCs and IICs did not show well-distributed responses. Slight modifications to the wording of items and responses were based solely on expert opinion.

The resulting interim version of OPAQ contained 21 items in six domains: walking and bending (six items); sitting and standing (three items); transfers (four items); back ache and pain (two items); fear of falling (three items); and independence (three items). Slight modifications to item wording and response option content (e.g., 'very often' changed to 'often', and 'almost never' changed to 'seldom') were necessary to focus concepts on domains of interest, to improve clinical relevance, and to describe concepts as depicted by patients per expert opinion. Resulting response formats were: 'all days', 'most days', 'some days', 'few days', 'no days' for 15 questions, and 'always', 'often', 'sometimes', 'seldom', 'never' for the remaining six questions.

Phase 2

This phase involved analysis of concept elicitation and cognitive debriefing data from 32 patients (first stage, 14 patients; second stage, 18 patients). All patients were receiving at least one prescription or non-prescription treatment for osteoporosis. Non-prescription treatments included calcium and vitamin D supplements.

First stage: patient demographics

Twenty-one patients (eight in diversity group 1, five in group 2, and eight in group 3) were recruited for the first stage of phase 2. However, data from seven of these participants were excluded from the analysis because of poor mastery of English ($n=1$) or because they were unable to distinguish the symptoms and impacts of osteoporosis from those of other comorbid conditions ($n=6$). These seven patients were white, with a mean (±standard deviation [SD]) age of 77.1±10.2 years and a mean disease duration of 9.7±7.9 years.

Demographic data for the 14 remaining patients (seven in diversity group 1, four in group 2, three in group 3) are shown in Table 1. This cohort was predominantly white (86 %) and had a mean (±SD) age of 68.0±11.3 years and a mean disease duration of 5.9±5.3 years. Seven patients were recruited at each of the two clinical sites. In total, 14 fractures had been sustained by ten of the 14 patients. Five of these fractures affected the spine. Remaining fractures were distributed among hip ($n=2$), wrist ($n=1$), shoulder ($n=1$), ribs ($n=2$), femur ($n=1$), and foot/toe ($n=2$). It proved impossible to recruit patients who were free of comorbid conditions that might be associated with fatigue, poor sleep, pain, or limited mobility, and comorbid conditions affecting these patients included Parkinson's disease, polymyalgia rheumatica, breast cancer, hyperlipidemia, osteoarthritis, rheumatoid arthritis, and diabetes.

First stage: concept elicitation

In this part of the interview, participants were asked about: (1) impacts osteoporosis had on their lives; (2) activities they were able/unable to do or avoided; and (3) any symptoms of which they were aware. The interviews therefore had a broader focus than the content of the instrument administered at that stage. We report here only the findings of relevance to the content of the final version of OPAQ-PF.

Relevant concept elicitation data from the first stage interviews are presented in conjunction with concept elicitation data from the second stage interviews in Table 2, and described in the section titled "Second stage: concept elicitation". In the first stage of phase 2, no new codes were added after the 12th concept elicitation interview, demonstrating that data saturation was achieved.

First stage: cognitive debriefing

Cognitive debriefing data showed that the interim version of OPAQ was well received but that a number of modifications were required. These included: (1) moving from a frequency response format to a severity response format; (2) making the introduction more informative and less likely to be overlooked; (3) adding a stem to the questionnaire to ensure participants responded specifically according to their osteoporosis and not another comorbid condition; (4) removing groups of items that did not yield information regarding the impact of osteoporosis on physical function; (5) improving item wording; (6) subdividing items that asked about more than one issue (e.g.,

Table 1 Participant characteristics, phase 2 (qualitative research)

Characteristic	First stage (n=14)	Second stage (n=18)
Age (years; mean±SD)	68.0±11.3	70.0±9.2
Ethnicity (n [%])		
White	12 (85.7)	15 (83.3)
Black/African American	1 (7.1)	0
Asian	1 (7.1)	0
Hispanic/Latino	0	1 (5.6)
Middle Eastern	0	1 (5.6)
Mixed	0	1 (5.6)
Main activity (n [%])		
Employed full time	2 (14.3)	4 (22.2)
Employed part time	0	2 (11.1)
Self-employed	1 (7.1)	0
Looking after home	4 (28.6)	2 (11.1)
Retired	5 (35.7)	8 (44.4)
Disabled	2 (14.3)	2 (11.1)
Disease duration (years; mean±SD)	5.9±5.3	6.0±4.1
Diversity group (n [%])		
Group 1	7 (50.0)	8 (44.4)
Group 2	4 (28.6)	5 (27.8)
Group 3	3 (21.4)	5 (27.8)
T-score		
Total hip (median [range])	−2.2 (−3.3 to −0.7)	−2.3 (−3.1 to −1.1)
Femoral neck (median [range])	−2.5 (−3.8 to −0.7)	−2.6 (−3.3 to −1.0)
Lumbar spine (median [range])	−2.2 (−3.7 to −0.4)	−2.1 (−3.9 to −0.6)
Fracture site (number of fractures)		
Hip	2	5
Spine	5	3
Wrist	1	1
Ankle	0	1
Distal forearm	0	1
Shoulder	1	0
Humerus	0	2
Ribs	2	1
Pelvis	0	1
Femur	1	0
Foot/toe	2	1

SD standard deviation

bending, lifting, and stooping); (7) adding new items identified as being of importance to osteoporosis patients; and (8) removing items considered irrelevant to osteoporosis patients. All modifications were tracked in an item-tracking matrix.

The change in response option format was introduced because some participants found it difficult to determine how best to respond when the recall period was limited to 7 days and the options were limited to the two sets of responses that were used in the interim version of OPAQ. As a result of this feedback, later interviews discussed the possibility of moving from a frequency response set to a severity-focused response set. Participants generally felt that a severity response format would be more appropriate.

Following completion of the first-stage cognitive debriefing interviews, the research team decided to focus the content of OPAQ-PF on physical function as a measure of the impact of osteoporosis, concentrating on the domains of mobility (walking, carrying, and climbing), physical positions (bending, reaching, picking up, standing, and sitting), and transfers (getting in and out of bed, chairs, and vehicles, and on and off the toilet). This led to the removal of items addressing fear of falling, independence, and symptoms. As a result, the instrument generated at the end of the first stage of phase 2 had 16 items in three domains (mobility, physical positions, and transfers) and included a five-point scale that was used throughout the questionnaire: 'no difficulty'; 'a little difficulty'; 'some difficulty'; 'a lot of difficulty'; and 'severe difficulty'. This instrument was used in the second stage of phase 2.

Second stage: patient demographics

Demographic data for the 18 participants (eight in diversity group 1, five in group 2, and five in group 3) recruited for this stage of the study are shown in Table 1. As in the first stage, this cohort was predominantly white (83 %), with a mean (±SD) age of 70.0±9.2 years and a mean disease duration of 6.0±4.1 years.

Twelve of the 18 patients had sustained a total of 16 fractures. The predominant fracture site in this cohort was the hip (n=5). The remaining fractures were distributed among spine (n=3), wrist (n=1), ankle (n=1), distal forearm (n=1), humerus (n=2), ribs (n=1), pelvis (n=1), and foot/toe (n=1). Comorbid conditions included osteoarthritis, inflammatory arthritis, rheumatoid arthritis, diabetes, hypercholesterolemia, asthma, chronic obstructive pulmonary disease, hypertension, and restless legs syndrome.

Second stage: concept elicitation

In the second stage of phase 2, saturation was achieved after the 13th concept elicitation interview. Concept elicitation data supporting the final version of OPAQ-PF are summarized in Table 2. First- and second-stage interview data are presented together. The results demonstrate widespread support for all items in the domains of mobility, physical positions, and transfers.

Second stage: cognitive debriefing

Cognitive debriefing results obtained in the first stage of phase 2 reflect participants' thoughts regarding the design of the

Table 2 Support for OPAQ-PF from concept elicitation data

Domain/item no.	Item wording	Support from concept elicitation data
Mobility	Relevant to all mobility domain items: 1. Walking to do your daily chores or errands (e.g., grocery shopping, taking out garbage, housework, going to post office, walking the dog)? 2. Walking unaided so you can do your day-to-day activities? 3. Carrying objects in order to perform your day-to-day activities (e.g., a bag of groceries, a bag of garbage)? 4. Walking one block? 5. Climbing one flight of stairs or steps?	- Household activities and walking identified as a cause of pain. - Pain reported as affecting usual activities inside and outside the home. - Fractures as a result of osteoporosis can affect the ability to walk unaided and to complete daily activities unaided. Participants reported being unable to complete/needing help completing basic activities and self-care activities, even after the fracture had healed. - Large number of mobility problems reported, including needing to walk with a cane, walking more slowly. Particularly relevant after a fracture. - 16 of the 32 analyzed participants reported problems walking. - Avoiding or limiting the time spent walking as a result of pain.
3.	Carrying objects in order to perform your day-to-day activities (e.g., a bag of groceries, a bag of garbage)?	- Reported losing balance when getting things out of a closet or carrying things. - A few participants reported being given a weight restriction by their doctors. - Avoiding or limiting the time spent on carrying objects as a result of pain.
5.	Climbing one flight of stairs or steps?	- Managing stairs a lot more difficult because of a combination of not being able to walk quickly, being off-balance, and/or feeling weak.
Physical positions	Relevant to all physical positions domain items: 6. Bending or stooping to do your daily chores or errands (e.g., grocery shopping, taking out garbage, housework, going to post office, walking the dog)? 7. Lifting objects in order to perform your day-to-day activities (e.g., a bag of groceries, a bag of garbage)? 8. Reaching overhead in order to perform your day-to-day activities? 9. Picking things up from the floor? 10. Standing as much as you needed to in order to perform your day-to-day activities? 11. Sitting as much as you needed to in order to perform your day-to-day activities?	- Extending/stretching/leaning forward identified as a cause of pain. - Pain reported as affecting usual activities inside and outside the home. Fractures as a result of osteoporosis can affect the ability to walk unaided and to complete daily activities unaided. Participants reported being unable to complete/needing help completingbasic activities and self-care activities, even after the fracture had healed.
6.	Bending or stooping to do your daily chores or errands (e.g., grocery shopping, taking out garbage, housework, going to post office, walking the dog)?	- 7 of the 32 analyzed participants reported problems bending down towards the floor.
7.	Lifting objects in order to perform your day-to-day activities (e.g., a bag of groceries, a bag of garbage)?	- 9 of the 32 analyzed participants reported problems lifting. - Ability to lift sometimes limited as a result of lack of strength or fear of injury.
8.	Reaching overhead in order to perform your day-to-day activities?	- 6 of the 32 analyzed participants reported problems reaching.
9.	Picking things up from the floor?	- 7 of the 32 analyzed participants reported problems bending down towards the floor.
10.	Standing as much as you needed to in order to perform your day-to-day activities?	- Stiffness occurring if the patient is in one position for too long. - Avoiding or limiting the time spent standing as a result of pain.
11.	Sitting as much as you needed to in order to perform your day-to-day activities?	- Sitting for too long identified as a cause of pain. - 8 of the 32 analyzed participants reported problems sitting. - Avoiding or limiting the time spent sitting as a result of pain. - Stiffness occurring if the patient is in one position for too long.
Transfers	Relevant to all transfers domain items: 12. Getting in or out of bed? 13. Getting in or out of a chair? 14. Getting on or off the toilet? 15. Getting in or out of cars on your own?	- Pain reported as affecting usual activities inside and outside the home. - Fractures as a result of osteoporosis can affect the ability to walk unaided and to complete daily activities unaided. Participants reported being unable to complete/needing help completing basic activities and self-care activities, even after the fracture had healed. - 11 of the 32 analyzed participants reported problems getting up.

questionnaire, the language used, its applicability, the ease with which the instructions could be interpreted, response options, and the recall period. The questionnaire underwent further iterative modifications during the second stage of phase 2 as a result of participants' feedback. These modifications included removing one item, re-wording of items, and the addition of examples for clarification. As in the first stage of phase 2, all modifications were tracked in an item-tracking matrix.

Participants in the cognitive debriefing interviews thought the length of the questionnaire was good. However, the overall response to the questionnaire was mixed, some participants rating it as very good and others expressing concerns. Examples of concerns included: questionnaire not sufficiently 'in-depth'; some aspects of questionnaire being ambiguous; some items appearing at first glance to ask the same thing. In general, however, the cognitive debriefing results showed that the modifications made to the interim version of OPAQ during the first stage of phase 2 represented an improvement. The change to a severity format was generally preferred and items in the 'mobility' and 'physical positions' domains performed well following modification during the course of the interviews. However, items in the 'transfers' domain attracted some criticism from patients, several participants expressing concerns about the relevance of some items for all osteoporosis patients (e.g., getting in and out of bed). One participant commented that there should be an 'unable to do' response option and several participants commented during the concept elicitation interviews that they avoided certain activities. As a result, the response option 'completely avoided doing this' was added to the instrument.

The final changes made to the OPAQ resulted in an instrument with 15 items in three domains (mobility, physical positions, and transfers), and a single six-point response scale for each item ('no difficulty'; 'a little difficulty'; 'some difficulty'; 'moderate difficulty'; 'severe difficulty'; and 'completely avoided doing this') (Table 3).

Discussion

This report summarizes the two-phase, iterative process by which OPAQ v.2.0 was modified into a short questionnaire (OPAQ-PF) focusing specifically on the impacts of osteoporosis on physical function. Although health-related quality of life (HRQoL) has traditionally been measured in interventional research for osteoporosis to evaluate health status and economic value, this new instrument represents a useful advance in osteoporosis research as clinicians and researchers now have a concise tool that focuses exclusively on mobility, physical conditions, and transfers in individuals with osteoporosis both before and after a fracture event. The measurement of HRQoL provides important information on the health

Table 3 Osteoporosis Assessment Questionnaire-Physical Function (OPAQ-PF)

Mobility
Walking to do daily chores or errands
Walking unaided so day-to-day activities can be carried out
Carrying objects in order to perform day-to-day activities
Walking one block
Climbing one flight of stairs or steps
Physical positions
Bending or stooping to do daily chores or errands
Lifting objects in order to perform day-to-day activities
Reaching overhead in order to perform day-to-day activities
Picking things up from the floor
Standing as much as needed in order to perform day-to-day activities
Sitting as much as needed in order to perform day-to-day activities
Transfers
Getting in or out of bed
Getting in or out of a chair
Getting on or off the toilet
Getting in or out of cars unaided

The questionnaire asked participants to evaluate the impact of osteoporosis on their ability to perform day-to-day activities during the previous 7 days using a 6-point severity response scale: 'no difficulty'; 'a little difficulty'; 'some difficulty'; 'moderate difficulty'; 'severe difficulty'; 'completely avoided doing this'. The 15 items were presented in three domains (mobility, physical positions, and transfers) as shown above

impact of osteoporosis including subcomponents of physical functioning; however, HRQoL and the dimension of physical functioning are by themselves complex, multi-domain concepts. Generic and/or disease-targeted instruments commonly used in osteoporosis research include the EuroQoL instrument (EQ-5D), SF-36® Health Survey, or Health Utilities Index (HUI), and Quality-of-Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO), OPAQ, or Osteoporosis Quality-of-Life Questionnaire (OQLQ), respectively [23]. These instruments would most likely not substantiate claims of treatment benefit for medical product labeling as they may be viewed by regulatory bodies as not adequate for assessing such a broad concept as HRQoL and physical functioning [17].

FDA guidance states that qualitative techniques may be used to develop and validate a modification of an existing PRO instrument if the modifications involve deletion of portions of the questionnaire or changes to the target patient population, patient instructions, order of items, item wording, response options, or recall period [17]. The methodology used in this study to confirm the adequacy of OPAQ-PF is consistent with these recommendations, and with those of the ISPOR task force papers [17–19], thereby providing more reliable evidence of the ability of the instrument to substantiate claims of treatment benefit for the specific concept of ability to perform daily activities of physical function.

A number of unforeseen problems arose during the first stage of phase 2; however, the iterative nature of the protocol allowed us to counter these problems during the second stage. One issue was the high prevalence of comorbidity, a common problem in this patient population because osteoporosis typically affects older adults [2]. Substantial comorbidity prevalence made it difficult for many patients to distinguish between osteoporosis and one of their comorbid conditions as the cause of symptom experiences. It also caused difficulties with data interpretation and necessitated discarding information regarding symptoms or impacts that the patient could not specifically relate to osteoporosis. This led to a more focused attempt to recruit patients without significant comorbidities in the second stage of phase 2. Papers that describe the creation of previous versions of OPAQ do not mention a confounding effect of comorbidity during instrument development [11]. However, Silverman does note that it is routine during analysis of OPAQ data to adjust for a number of factors, including concomitant medication use, this factor being used as a surrogate marker for comorbidity [11]. Likewise, data analyses for the OPAQ-PF may need to be adjusted for presence of musculoskeletal or other comorbidities (based on clinical examination or self-report).

Given the focus of previous versions of OPAQ on the ability to detect change in patient outcomes in association with fracture, it was expected that fracture and nonfracture patients would give different responses to the questionnaire. Therefore, we anticipate that the OPAQ-PF will be able to distinguish between these patient groups, and will be well placed to capture the decline of osteoporosis patients as they enter the phase of the disease in which they experience fractures, and related symptoms and impacts. It is also likely that OPAQ-PF will be able to document improvements in patient outcomes associated with fracture healing. This will be further explored through an ongoing psychometric validation study.

This study was subject to a number of limitations. First, content validity of the OPAQ-PF was established in a specific patient population that was exclusively female, predominantly white, and already receiving therapy for osteoporosis. Therefore, validity may not necessarily be assumed for all races/ethnicities, for men, or for untreated individuals. Second, because postmenopausal osteoporosis is largely asymptomatic [24], OPAQ-PF, in common with all other osteoporosis-specific PRO questionnaires, may provide more useful information when used in a population with a history of fracture than when used in a population without such history. Moreover, assessing women soon after a fracture event may be particularly informative. Recent data collected during the Fracture Reduction Evaluation of Denosumab in Osteoporosis Every 6 Months (FREEDOM) study show that, in women with incident clinical fractures, the largest deterioration in PROs is observed when patients are assessed <3 months post fracture [14]. This type of event-prompted assessment may allow researchers to document any differences in postfracture recovery between patients who are receiving therapy and those receiving placebo.

A third limitation of the study is the somewhat historical nature of the data used in the IRT analysis. The data in question were generated during the baseline visit of a 3-year clinical trial (MORE) conducted between 1994 and 1998 [15]. These data were therefore generated approximately 15 years before the current study was performed, when available therapeutic options were more limited than they are today. Responses to OPAQ provided by patients enrolled in MORE in the 1990s may differ from those of a more contemporary population receiving current treatments for osteoporosis. A further limitation regarding the IRT analysis relates to the criteria used to delete items. Rather than use goodness-of-fit analyses for this purpose (as is commonly done [16]), we chose to use ICCs and IICs to drive discussion and decisions pertaining to item elimination. However, subsequent use of interview data yielded feedback from patients and was useful in determining whether any relevant concepts were missing from the instrument.

One final point relates to the number of participants interviewed in phase 2 of this study. This number ($n=32$) might appear to be low, but this is typical of qualitative research. Moreover, the crucial criterion for achieving an acceptable value of 'n' in this type of research is the demonstration of data saturation having been achieved [21]. This was the case in both the first and second stages of phase 2.

The version of OPAQ that we have developed represents a useful advance for both researchers and patients. For researchers, a PRO instrument now exists that focuses solely on the mobility, physical position, and transfer aspects of physical function in patients with osteoporosis or low bone mass density. For patients, the small number of items reduces the burden associated with completion of the instrument compared with others that may be used. The use of IRT and qualitative concept elicitation, and cognitive debriefing interviews, combined with the clinical expertise of two of the authors (DTG, SS), resulted in the development of a concise PRO instrument that focuses on the impact of osteoporosis on physical function before and after a fracture event. Content validity of the OPAQ-PF has been established in fracture and nonfracture osteoporosis patients in the USA. We are currently conducting a psychometric validation study using OPAQ-PF to evaluate validity (including the ability of the OPAQ-PF to discriminate between those with fracture vs. those without), reliability, and sensitivity to change. Additionally, due to the comorbidities often seen in patients with osteoporosis that are associated with older age, it may be necessary to adjust for the presence of musculoskeletal or other related comorbidities when conducting analyses of OPAQ-PF data. Further research is needed to confirm the need for statistical adjustments.

Conclusions

The OPAQ-PF represents a new PRO tool that is uniquely tailored to the assessment of physical function in osteoporosis patients. The cohort used to develop the instrument included patients both with and without a history of fracture, and content validity was established in this patient group. This provides evidence that OPAQ-PF has relevance in a combined fracture/nonfracture population. Once psychometrically validated in a range of osteoporotic patient populations, OPAQ-PF will offer researchers a valid, reliable, and sensitive instrument that will be useful in clinical trials to evaluate pharmacological therapies that aim to reduce fracture risk and promote bone formation following fracture.

References

1. National Osteoporosis Foundation (2010) Clinician's Guide to Prevention and Treatment of Osteoporosis. National Osteoporosis Foundation, Washington, DC
2. National Osteoporosis Foundation (2012) Bone health basics: Get the facts. National Osteoporosis Foundation. http://www.nof.org/node/40. Accessed 6 December 2012
3. Lau E, Ong K, Kurtz S, Schmier J, Edidin A (2008) Mortality following the diagnosis of a vertebral compression fracture in the Medicare population. J Bone Joint Surg Am 90:1479–1486
4. Kado DM, Browner WS, Palermo L, Nevitt MC, Genant HK, Cummings SR (1999) Vertebral fractures and mortality in older women: a prospective study. Study of Osteoporotic Fractures Research Group. Arch Intern Med 159:1215–1220
5. Johnell O (1996) Advances in osteoporosis: better identification of risk factors can reduce morbidity and mortality. J Intern Med 239:299–304
6. Silverman SL (2005) Quality-of-life issues in osteoporosis. Curr Rheumatol Rep 7:39–45
7. Gold DT, Solimeo S (2006) Osteoporosis and depression: an historical perspective. Curr Osteoporos Rep 4:134–139
8. Lips P, van Schoor NM (2005) Quality of life in patients with osteoporosis. Osteoporos Int 16:447–455
9. Silverman SL, Piziak VK, Chen P, Misurski DA, Wagman RB (2005) Relationship of health related quality of life to prevalent and new or worsening back pain in postmenopausal women with osteoporosis. J Rheumatol 32:2405–2409
10. Oglesby AK, Minshall ME, Shen W, Xie S, Silverman SL (2003) The impact of incident vertebral and non-vertebral fragility fractures on health-related quality of life in established postmenopausal osteoporosis: results from the teriparatide randomized, placebo-controlled trial in postmenopausal women. J Rheumatol 30:1579–1583
11. Silverman SL (2000) The Osteoporosis Assessment Questionnaire (OPAQ): a reliable and valid disease-targeted measure of health-related quality of life (HRQOL) in osteoporosis. Qual Life Res 9:767–774
12. Silverman SL, Minshall ME, Shen W, Harper KD, Xie S, Health-Related Quality of Life Subgroup of the Multiple Outcomes of Raloxifene Evaluation Study (2001) The relationship of health-related quality of life to prevalent and incident vertebral fractures in postmenopausal women with osteoporosis: results from the Multiple Outcomes of Raloxifene Evaluation Study. Arthritis Rheum 44:2611–2619
13. Tosteson AN, Do TP, Wade SW, Anthony MS, Downs RW (2010) Persistence and switching patterns among women with varied osteoporosis medication histories: 12-month results from POSSIBLE US. Osteoporos Int 21:1769–1780
14. Silverman S, Viswanathan HN, Yang YC, Wang A, Boonen S, Ragi-Eis S, Fardellone P, Gilchrist N, Lips P, Nevitt M, Palacios Gil-Antuñano S, Pavelka K, Revicki D, Simon J, Macarios D, Siris ES (2012) Impact of clinical fractures on health-related quality of life is dependent on time of assessment since fracture: results from the FREEDOM trial. Osteoporos Int 23:1361–1369
15. Cummings SR, Eckert S, Krueger KA, Grady D, Powles TJ, Cauley JA, Norton L, Nickelsen T, Bjarnason NH, Morrow M, Lippman ME, Black D, Glusman JE, Costa A, Jordan VC (1999) The effect of raloxifene on risk of breast cancer in postmenopausal women: results from the MORE randomized trial. Multiple outcomes of raloxifene evaluation. JAMA 281:2189–2197
16. Edelen MO, Reeve BB (2007) Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. Qual Life Res 16(Suppl 1):5–18
17. Food and Drug Administration (2009) Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. U.S. Department of Health and Human Services; Food and Drug Administration; Center for Drug Evaluation and Research (CDER); Center for Biologics Evaluation and Research (CBER); Center for Devices and Radiological Health (CDRH). http://www.fda.gov/downloads/ Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ UCM193282.pdf. Accessed 6 November 2012
18. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, Ring L (2011) Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1—eliciting concepts for a new PRO instrument. Value Health 14:967–977
19. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, Ring L (2011) Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research

Practices Task Force report: part 2—assessing respondent understanding. Value Health 14:978–988

20. Joffe H, Yardley L (2004) Content and thematic analysis. In: Marks D, Yardley L (eds) Research methods for clinical and health psychology. Sage, London, pp 56–68

21. Kerr C, Nixon A, Wild D (2010) Assessing and demonstrating data saturation in qualitative inquiry supporting patient-reported outcomes research. Expert Rev Pharmacoecon Outcomes Res 10:269–281

22. Willis GB (2005) Cognitive interviewing: a tool for improving questionnaire design. Sage, Thousand Oaks

23. Tosteson AN, Hammond CS (2002) Quality-of-life assessment in osteoporosis: health-status and preference-based measures. Pharmacoeconomics 20:289–303

24. Lewiecki EM (2009) Current and emerging pharmacologic therapies for the management of postmenopausal osteoporosis. J Womens Health (Larchmt) 18:1615–1626