Identifying content for the glaucoma-specific item bank to measure quality of life parameters

Jyoti Khadka PhD,[1] Colm McAlinden PhD,[1] Jamie E Craig PhD FRANZCO,[2] Eva K Fenwick MA,[3] Ecosse L Lamoureux PhD,[3,4] Konrad Pesudovs PhD.[1]

[1] NH&MRC Centre for Clinical Eye Research, Discipline of Optometry and Vision Science, Level 2, Sturt West Wing, Flinders University, Sturt Road, Bedford Park, Adelaide, South Australia, 5042, Australia.
[2] Department of Ophthalmology, Flinders Medical Centre, Flinders University of South Australia, Bedford Park, South Australia, 5042, Australia.
[3] Centre for Eye Research Australia, Department of Ophthalmology, University of Melbourne, Victoria, 8002, Australia.
[4] Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.

Corresponding author: Konrad Pesudovs, NH&MRC Centre for Clinical Eye Research, Discipline of Optometry and Vision Science, Level 2, Sturt West Wing, Flinders University, Sturt Road, Bedford Park, Adelaide, South Australia, 5042, Australia. Tel +618 8 7221 8401, Fax +618 7221 8399, email: Konrad.Pesudovs@flinders.edu.au.

Conflict of Interest: none declared

Funding source: NHMRC 1031838

Address for reprints: NH&MRC Centre for Clinical Eye Research, Discipline of Optometry and Vision Science, Flinders Medical Centre, Bedford Park, South Australia, 5042, Australia.

Word count
Abstract: 250
Text: 3998

Abstract

Purpose: Patient Reported Outcomes (PROs) have become essential clinical trial end points. However, a comprehensive, multi-dimensional, patient-relevant and precise glaucoma-specific PRO instrument is not available. Therefore, the purpose of this study was to identify content for a new, glaucoma-specific, quality of life (QOL) item bank.

Methods: Content identification was undertaken in five phases: 1. Identification of extant items in glaucoma-specific instruments and the qualitative literature, 2. Focus groups and interviews with glaucoma patients, 3. Item classification and selection, 4. Expert review and revision of items, and 5. Cognitive interviews with patients.

Results: A total of 737 unique items (extant items from PRO instruments, 247; qualitative articles, 14 items; focus groups and semi-structured interviews, 476 items) were identified. These items were classified into 10 QOL domains. Four criteria (item redundancy, item inconsistent with domain definition, item content too narrow to have wider applicability, and item clarity) were used to remove and refine the items. After the cognitive interviews, the final minimally representative item-set had a total of 342 unique items belonging to 10 domains: activity limitation (88), mobility (20), visual symptoms (19), ocular surface symptoms (22), general symptoms (15), convenience (39), health concerns (45), emotional well-being (49), social (23), and economic issues (22).

Conclusion: The systematic content identification process identified ten QOL domains which were important to patients with glaucoma. The majority of items were identified from the patient-specific focus groups and semi-structured interviews suggesting that the existing PRO instruments do not adequately address QOL issues relevant to individuals with glaucoma.

Introduction

Glaucoma is a chronic progressive disease in which optic nerve damage occurs leading to loss of peripheral vision; this damage is irreversible and may lead to blindness. The strongest risk factor for glaucoma is raised intraocular pressure (IOP), and the main management strategies are early detection and long-term IOP reduction with the goal of slowing the progression of the disease.[1] Glaucoma affects the lives of patients in a multitude of ways. Activity limitation tends to affect individuals towards the later stages of the disease. However, the disease may affect other areas of a patient's life from the time of diagnosis. Examples include the anxiety of having glaucoma, the inconvenience of instillation of eye drops, frequent changes in treatment regimen, treatment side effects, ocular surgery and its potential complications, regular hospital or clinic visits, and difficult testing such as visual field testing.[2] As a result, glaucoma can have a significant impact on issues such as, emotional well-being, social relationships, and financial constraints.[3-6] Therefore, traditional clinical measures of glaucoma (intraocular pressure, optic nerve head evaluation and visual field examination) fail to elicit the actual impact of glaucoma on a patient's quality of life (QOL).[7]

Over the years, several glaucoma-specific patient-reported outcome (PRO) instruments have been developed.[3-5, 8-23] While these instruments measure various QOL domains such as activity limitation, symptoms, convenience of eye-drop use, treatment of side effects and satisfaction; each instrument typically measures only one domain.[7, 24] For instance, the Glaucoma Quality of Life (GQL-15) measures only activity limitation,[3, 25] the Treatment Satisfaction Survey of Intra-ocular Pressure (TSS-IOP) was developed to assess treatment satisfaction and side effects of glaucoma medication,[13] and the Comparison of Ophthalmic Medication for Tolerability (COMTOL) measures tolerability and compliance of glaucoma eyedrops.[10]

Three instruments [Glaucoma Quality of Life-36 (Glau-QOL-36); Symptoms Impact Glaucoma Score (SIG); and a questionnaire by Uenishi et al. 2003[12]] have incorporated multiple QOL domains, but each instrument has only a small number of items in each domain.[4, 12, 14] This is problematic because, as has been frequently demonstrated with Rasch analysis, too few items result in inadequate measurement precision.[26-28] Rasch analysis is a psychometric assessment tool, which provides a more comprehensive insight into the psychometric properties of a PRO instrument.[29] Rasch analysis is increasingly being accepted as the standard method in the development and re-engineering of PRO instruments in ophthalmology.[29-34] Only

three glaucoma-specific PROs [Glaucoma Symptom Scale (GSS), Glaucoma Quality of Life (GQL), and Glaucoma Symptom Index (GSI)] have been evaluated with Rasch analysis. The GSS was found to have unsatisfactory measurement precision.[35] The GQL was shown to be a sound measure of activity limitation albeit with modification.[36] Similarly, the GSI demonstrated adequate psychometric properties, but is confined to measuring activity limitation and symptoms.[20] Therefore, a comprehensive, precise and valid glaucoma-specific PRO instrument which can be used as a clinical end point in glaucoma outcome studies remains elusive.[24]

A superior option to traditional paper-pencil based PRO instruments is to develop a glaucoma-specific item bank. An item bank is a large collection of calibrated items attempting to measure an underlying latent trait.[37] A computer adaptive testing (CAT) system can be used to implement the items from an item bank.[38, 39] The CAT system tailors calibrated items from the item bank to individual participants to produce highly accurate and precise measurement of QOL.[38, 40] By implementing items targeted to person ability and thus ential to revolutionize PRO measurement in glaucoma.

Our overarching aim is to develop a comprehensive and multi-domain Rasch scaled PRO in the form of a glaucoma-specific item bank. In this paper, we describe the content identification phase of this novel glaucoma-specific item bank development process.

Materials and Methods

A systematic step-wise item identification and review process was used to develop the content of the glaucoma-specific item bank which included: identification of extant items in currently available glaucoma-specific instruments, a qualitative literature review, patient focus groups and semi-structured interviews, item classification and selection, expert review and revision of items, and cognitive interviews with patients.

Identification of extant items and literature review

Rather than developing new items, an initial item pool was created from extant items obtained from available glaucoma-specific PRO instruments. An extensive literature review in Entrez Pubmed was carried out in May 2011. The key search terms used were ‘glaucoma’ and (‘questionnaire’ or ‘instrument’ or ‘patient reported outcome’ or ‘PRO’ or ‘quality of life’ or ‘symptom’ or ‘mobility’ or ‘satisfaction’ or ‘qualitative’). The

search was not limited to any time frame or type of PRO instrument. The search identified 1981 papers. Reference lists of the papers identified (including review papers) were also scanned manually for additional relevant papers. All identified papers were reviewed against the following selection criteria: PRO instruments specifically developed for patients with glaucoma, content available in English and developed using valid content development methods such as structured/semi-structured interviews and/or literature reviews. Nineteen glaucoma-specific PRO instruments fulfilled the selection criteria (Table 1). We also included six peer-reviewed articles that reported qualitative research exploring the impact of glaucoma on QOL and systematically reviewed them to identify potential concepts that were not covered by extant items.

Focus groups and semi-structured interviews

Focus groups and semi-structured interviews were used as a supplementary source of item identification to our literature review. More importantly, focus groups and semi-structured interviews were used to identify important areas of measurement of glaucoma-specific QOL that were not adequately covered by the extant items and the qualitative literature review.

Focus group and semi-structured interviews participants

Participants were specifically selected (purposive sampling) in order to recruit patients with different stages of the disease, ethnicity, age, and gender. Patients included were aged 18 years or above, those who had glaucoma as a primary diagnosis for at least six months, without other significant eye diseases, able to communicate in English, and without any cognitive impairment that could hamper their participation both in the group discussion or semi-structured telephone interviews.

Ten focus groups and 19 semi-structured telephone interviews were conducted with glaucoma patients. Four focus groups and one telephone interview were conducted at Flinders Medical Centre in South Australia, four focus groups and 18 interviews were conducted at the Royal Victorian Eye and Ear Hospital in Melbourne, Australia and two focus groups were conducted at centres for glaucoma located in Aberdeen and Leeds in the United Kingdom. In total, 72 patients with glaucoma took part in 10 focus groups and 19 semi-structured individual interviews. Among these, detail demographic data of 58 (80.6%) Australian participants were collected. However, demographic data of 14 (19.4%) UK (Leeds and Aberdeen) participants were less detailed.

Participants were divided into different groups on the basis of glaucoma severity using the Advanced Intervention Study (AGIS) visual field classification system.[41] The grouping was done in order to conduct focus groups with the participants having similar levels of the disease severity and also to compare and contrast focus groups/interviews data between and within the participants during analysis. According to the AGIS classification, participants' visual fields are sub-divided into 21 stages of glaucoma severity (0 = none, 1-5 = mild, 6-11 = moderate, 12-17 = severe and 18-20 = end stage).[41] These scores were obtained for the better eye and the worse eye from the participants' total deviation plot of Humphrey Visual Field Analyser. As the level of glaucoma severity can be different between two eyes, the participants were sub-divided into four groups on the basis of the AGIS classification in the better and worse eyes. These groups were, Group 1: ≤ Mild in better, ≤ moderate in worse; Group 2= ≤ Mild in better, ≤ moderate; Group 3= Moderate in better, ≥ Moderate in worse; Group 4 = ≥ Severe (both). Patients on a range of treatments were also included, namely glaucoma topical (drops) therapy, those who had undergone laser, and those who had incisional surgery for glaucoma (Table 2).

All the focus groups and interviews were conducted using a topic guide which was prepared based on the literature review and clinical experience. The focus groups and interviews were carried out in a semi-structured format with minimum involvement by the moderator. Informed consent was obtained from all participants. Ethical approval was obtained from the Flinders Clinical Research Ethics Committee, Royal Victorian Eye and Ear Hospital and the respective collaborative institutions, and the study adhered to the tenants of the Declaration of Helsinki.

All interviews were audio recorded and later transcribed verbatim. Thematic data analysis was carried out based on the constant comparative method with the assistance of qualitative data analysis software NVivo version 9 to sort and code the data.[42, 43] Specifically, we looked for keywords, phrases, and quotes regarding activity limitations, symptoms (visual, ocular surface and general), treatment effects, social, emotional, work and economic impact of glaucoma expressed by the participants. Two criteria were used to include any topic as a theme in the final analysis.[44] The first criterion was that at least two participants had to make substantive comment on the topic in a single group. The second selection criterion was that the topic was discussed by at least one participant in more than one group. The themes were reviewed by the

authors and rephrased to develop new items. These items were added to the initial item pool which consisted of the extant items from the literature review.

Identification of domains, item stem and response options

On the basis of the comprehensive literature review, input from patient focus groups, semi-structured interviews and consensus between the researchers, ten domains important to patients with glaucoma were identified. The domains were activity limitation, mobility, visual symptoms, ocular surface symptoms, general symptoms, convenience, health concerns, emotional well-being, social well-being and economic issues. Item stem and response options for each of these ten domains were also formulated on the basis of a comprehensive literature review and consensus between the researchers (JK, KP, EL and EF). The item stem is simply the question that states the particular item, for example *"How often do you experience…"* Response options are a set of categories defined by descriptive terms provided to the participants to endorse the item,[45] such as, *"Never, Occasionally, Quite often, Very often."* The literature review was carried out to explore the evidence on optimally functioning item stem and response option across all the health specific PRO instruments which were developed or validated using Rasch analysis. Rasch analysis provides an in-depth assessment of the appropriateness of response options and a framework to improve their functioning.[46]

Item classification and selection

Upon domain identification, items in the initial item pool were selected and classified under the ten domains on the basis of their colloquial meaning that represented the domain most accurately. A method of "binning and winnowing" was used to classify and refine items within each of the ten domains.

Binning

Binning is a systematic method of grouping (binning) items according to their meaning under a specific domain.[47] For example, "reading small print" became a bin within the activity limitation domain. The goal for this process was to identify items that capture the meaning of the domain and to eliminate obviously redundant items. For example, "reading small writing" was similar to "reading small print" and was therefore removed on the grounds of redundancy.

Winnowing

Winnowing is a systematic method of reducing the large item pool to a representative set of items according to a set of inclusion and exclusion criteria.[47] The following criteria were used for removing items from the initial item pool:

(1) item inconsistent with the domain definition

(2) item similar in meaning with other item(s)

(3) item content too narrow to have wider applicability

(4) item confusing or unclear.[47]

Expert review and revision of items

After binning and winnowing, authors JK, KP, EF and EL systematically refined the item pool again. These authors are either internationally recognised experts in PRO development and validation or have experience in PRO development. The authors reassessed the clarity and appropriateness of all remaining items based on the four criteria outlined above. Items that were ambiguous or confusing were rephrased and any discrepancies that occurred between the authors were resolved by discussion.

Cognitive interviews with patients

Cognitive interviews were conducted to assess the clarity of instructions, items and response options prior to pilot testing of the instrument.[48] These interviews also allowed us to assess whether the pilot instrument had adequate content coverage and the items were relevant to patients. The 341 items were administered to each participant and any problematic items and particular comments by the participants were noted. Feedback from the cognitive interviews was incorporated and necessary changes were made.

Results

Identification of extant items from published PRO instruments

Nineteen PRO instruments which were specifically developed for glaucoma patients were identified from the literature review (Table 1). A total of 530 items were extracted from the 19 instruments, out of which only 247 (46.6%) were semantically unique.

Identification of items from a qualitative literature review

A further literature review of peer-reviewed publications that reported qualitative research with glaucoma patients was also carried out to extract content which was not included in the extant PROs. Six publications were identified from which fourteen unique items were drawn.[3, 5, 6, 13, 49, 50]

Focus groups and semi-structured interviews

The demographic characteristics of the focus group and interview participants are presented in table 2. A total of 614 items were extracted from the focus groups and semi-structured interviews. Of these, 138 had already been identified in the extant PROs and qualitative literature review and therefore, 476 (77.5%) were unique items. Table 3 provides examples of items which were drawn from the focus groups and semi-structured interviews.

Item classification and selection

A total of 737 items were identified from extant instruments, the literature review, focus groups and semi-structured interviews. An initial item pool comprising 10 domains was organised using the binning method. Activity limitation (168) and health concerns (149) were the largest domains, while general symptoms (19), visual symptoms (31) and mobility (32) were the smallest domains. Data from the focus groups and semi-structured interviews significantly increased the number of items in all of the domains (Figure 1).

After four sessions of binning and winnowing, 388 items were removed, leaving 341 minimally representative, unique items distributed among the 10 domains (Table 4). Table 5 provides examples of items that were removed at the binning and winnowing stage. The domain “health concerns” had the highest number of items removed (70% from the initial pool removed) followed by activity limitation (48% removed).

Domain question format and response option

Table 5 shows the item stems and response options formulated for the ten domains. The generation of the item stem and response options were guided by empirical evidence on commonly occurring domains (activity limitation, mobility, symptoms, and emotional well-being). The review suggested that symptoms can be rated in three scales (frequency, severity and bothersome).[31] Therefore, three different question formats

and response options were used (table 6). There was no or little evidence available for the three novel domains (convenience, health concerns and economic), so the item stem and response options were generated by the researchers. The literature review also suggested that four to five response options are optimal and thus our item pools utilized four or five response options across the ten domains (table 6).[30, 51] A non-applicable option was also used when necessary. A preceding statement "*Because of your glaucoma and its treatment…"* was formulated for the overall item bank. The term 'glaucoma' was not defined further so as to appear inclusive for all people potentially completing the questions. By treatment we include people using medication or undergone laser or surgery for glaucoma.

Expert review and revision of the items

After binning and winnowing, the remaining 341 items were thoroughly reviewed by the authors for wording, fit to question format, and meaning. See table, supplement digital content file 1, which shows some examples of items that were rephrased in order to fit the domain question format.

Cognitive interviews with patients and the pilot instrument

After 10 cognitive interviews, the wordings of 4 items were modified and 1 item was added (Table, supplement digital content file 2). The participants did not have any issue with the clarity of instruction, item stem and response categories. Therefore, at the conclusion of the cognitive interviews the pilot instrument had 342 items. Table, supplement digital content file 3, provides some examples of items in the pilot instrument which were derived from the existing PRO instruments and patients' input (focus groups or semi-structured interviews).

Discussion

While many glaucoma-specific PRO instruments exist, there is not a single PRO that comprehensively measures QOL. Such an instrument is needed as an end-point for clinical research in order to provide a more comprehensive assessment of the disease or treatment impact from the perspective of the patient. Creating such an instrument is our overall aim and the starting point is to identify a comprehensive set of items within all QOL domains important to people with glaucoma. Hence, we adopted a systematic and multi-phased content development method to identify items.

A review of existing glaucoma-specific PRO instruments and associated literature can serve as a rich source of item content.[20, 52] Therefore, we created an initial item pool from existing PROs and qualitative literature review. We surveyed PRO instruments that have been specifically developed and validated for patients with glaucoma and peer-reviewed qualitative publications that explored issues important to people with glaucoma. By restricting our review to glaucoma specific material, the possibility of content contamination by issues that are not relevant to patients with glaucoma might have been avoided.

The initial item pool was supplemented by material derived from patients (focus groups and semi-structured interviews). The value of patient input in identifying content has been highlighted for several other vision-specific instruments.[53-56] This approach has allowed us to produce items that are important and relevant to people with glaucoma. In this study, the topic guide was designed specifically to identify new content (items) which was not covered by existing PRO instruments and the qualitative literature review thus ensuring that the item pools would cover the full range of glaucoma-specific QOL issues. It is also important to have a sufficient number of items and content coverage in a PRO or a particular domain to enable adequate measurement precision.[29, 57] For example, the Glaucoma Symptoms Scale (GSS) has been found to possess inadequate measurement precision due to insufficient items and content coverage.[35] Therefore, the interviews with patients were also targeted to explore, in more detail, issues in domains (i.e. health concerns, social and economic) which were under-represented in the initial item pool (Figure 1). Patients' input also served to test content saturation of the ten domains; that is to determine when all possible items of importance to patients were identified so the item identification process could stop.

Following content identification, the process of binning and winnowing provided a means to systematically classify items into different domains and the item pool to be refined. This method has been successfully used in developing and refining items during the development of the Patient-Reported Outcomes Measurement Information System item bank.[47, 58] Here, a total of 530 items were identified from the extant PROs. However, the binning process revealed that many of these items were similar and therefore the total number of unique items was less than 50%. Furthermore, when the items from the qualitative literature review, focus groups and semi-structured interviews were added to the initial item pool, the total number was 737. This number was reduced to 341 minimally representative unique items after four sessions of

binning and winnowing (Table 3). One new item was added after the cognitive interviews. Therefore, the final pilot instrument had 342 items.

Interestingly, the final item pool bears the majority of the items under “activity limitation”, “symptoms (visual, ocular surface and general)” and “emotional well-being”. Several other studies have also highlighted that restriction of daily living activities, treatment side-effects and emotional well-being are important issues to patients with glaucoma.[4-6, 17, 59] These domains are also those with the most item coverage in existing glaucoma PRO instruments. The remaining 5 domains had little input from existing instruments, in particular, the health concerns, social and economic domains. This is an important finding as it demonstrates that existing instruments fail to capture of the range of QOL domains important to people with glaucoma. Therefore, the item bank will provide new insight into the impact of glaucoma on QOL which has not been measured previously. This will be a valuable addition to clinical research which will lead to a better understand the efficacy of new and existing treatment options from the perspective of the patient. This will help clinicians to understand which treatment options best suit an individual patient’s needs.

In the final process of refining the items, we modified many extant items from their original format (wording). We have taken this approach for two reasons: firstly, to make the items more uniform within each domain to simplify interpretation by the patients and secondly, to keep item wording as close as possible to the original utterances of the focus groups and semi-structured interview participants. This process should enhance participants’ understanding of the item meaning and simplify implementation of the item bank. One potential limitation of our approach is that the focus group participants were not probability sampled. The problem with probability sampling is that larger numbers of participants are required in order to cover the breadth of the condition as many “average” patients are recruited who tend to offer similar input. Instead, participants were purposively sampled to cover the breadth of the disease severity with a minimum number of participants. Given that focus groups and interviews were conducted in four centres in two countries, it is reasonable to expect these groups represent a wider glaucoma population.

Noisy clinical trial endpoints measured from poor PRO instruments are a threat to research validity.[58] It has become apparent that more reliable PRO instruments are necessary to improve efficiency, precision and accuracy of clinical trials outcomes. We expect the new glaucoma-specific QOL item bank to be a precise

and efficient instrument which will outperform the existing glaucoma-specific PRO instruments. The developmental stage of the project reported in this paper has utilised comprehensive item identification and item refinement methods. The next step will be to collect data on the items in the item bank and to use Rasch analysis to refine and calibrate items under each of the ten domains.[29] The item bank will undergo rigorous validation and reliability testing before it is made available for use in clinical trials. Our group is currently developing item banks for all major ophthalmic diseases and conditions (the Eye-tem Bank).[37] This glaucoma-specific item bank will serve as the glaucoma module within the Eye-tem Bank project.

Acknowledgement/ Disclosure

The authors would like to thank Ms Catherine Hurley, Mr Thomas A. Wright and Ms Lynda Caudle for their contribution in conducting focus groups at Flinders Medical Centre, South Australia and acknowledge Ms Jennifer M. Burr for providing focus groups reports which were conducted in the United Kingdom. We would also like to thank Ms Sutha Sanmugasundram and Ms Theona Nicolaou for their contribution in conducting focus groups and semi-structured telephone interviews at Royal Victorian Eye and Ear Hospital, Melbourne. We would also like to recognize the contribution of focus group participants who generously shared their time and experience.

References

1. H. A. Quigley. Glaucoma. *Lancet* 2011;377:1367-1377.
2. G. C. M. Rossi, G. M. Pasinetti and P. E. Briola. Effect of Glaucoma Medications on Quality of Life Examined by Generic and Vision Specific Instruments *J Clinic Experiment Ophthalmol* 2010;1:106.
3. P. Nelson, P. Aspinall and C. O'Brien. Patients' perception of visual impairment in glaucoma: a pilot study. *Br J Ophthalmol* 1999;83:546-552.
4. N. K. Janz, P. A. Wren, P. R. Lichter, et al. Quality of life in newly diagnosed glaucoma patients : The Collaborative Initial Glaucoma Treatment Study. *Ophthalmology* 2001;108:887-897.
5. T. Odberg, J. E. Jakobsen, S. J. Hultgren, et al. The impact of glaucoma on the quality of life of patients in Norway. I. Results from a self-administered questionnaire. *Acta Ophthalmol Scand* 2001;79:116-120.
6. P. Ramulu. Glaucoma and disability: which tasks are affected, and at what stage of disease? *Curr Opin Ophthalmol* 2009;20:92-98
7. G. Spaeth, J. Walt and J. Keener. Evaluation of Quality of Life for Patients With Glaucoma. *Am J Ophthalmol* 2006;141:3-14.
8. R. P. Mills and S. M. Drance. Esterman disability rating in severe glaucoma. *Ophthalmology* 1986;93:371-378.
9. J. E. Ross, A. J. Bron and D. D. Clarke. Contrast sensitivity and visual disability in chronic simple glaucoma. *Br J Ophthalmol* 1984;68:821-827.
10. B. L. Barber, E. R. Strahlman, R. Laibovitz, et al. Validation of a questionnaire for comparing the tolerability of ophthalmic medications. *Ophthalmology* 1997;104:334-342.
11. S. Dunker, A. Schmucker and H. Maier. Tolerability, quality of life, and persistency of use in patients with glaucoma who are switched to the fixed combination of latanoprost and timolol. *Adv Ther* 2007;24:376-386.
12. Y. Uenishi, H. Tsumura, T. Miki, et al. Quality of life of elderly Japanese patients with glaucoma. *Int J Nurs Pract* 2003;9:18-25.
13. M. J. Atkinson, W. C. Stewart, J. M. Fain, et al. A new measure of patient satisfaction with ocular hypotensive medications: the Treatment Satisfaction Survey for Intraocular Pressure (TSS-IOP). *Health Qual Life Outcomes* 2003;1:67.
14. A. Béchetoille, B. Arnould, A. Bron, et al. Measurement of health-related quality of life with glaucoma: validation of the Glau-QoL© 36-item questionnaire. *Acta Ophthalmologica* 2008;86:71-80.
15. J. P. Nordmann, P. Denis, M. Vigneux, et al. Development of the conceptual framework for the Eye-Drop Satisfaction Questionnaire (EDSQ) in glaucoma using a qualitative study. *BMC Health Serv Res* 2007;7:124.
16. B. L. Lee, P. Gutierrez, M. Gordon, et al. The Glaucoma Symptom Scale. A brief index of glaucoma-specific symptoms. *Arch Ophthalmol* 1998;116:861-866.
17. T. Odberg, J. E. Jakobsen, S. J. Hultgren, et al. The impact of glaucoma on the quality of life of patients in Norway. II. Patient response correlated to objective data. *Acta Ophthalmol Scand* 2001;79:121-124.

18. M. A. Ruiz, A. Pardo, J. M. Martinez de la Casa, et al. Development of a specific questionnaire measuring patient satisfaction with glaucoma treatment: Glausat. *Ophthalmic Epidemiol* 2010;17:131-143.

19. A. C. Viswanathan, A. I. McNaught, D. Poinoosawmy, et al. Severity and stability of glaucoma: patient perception compared with objective measurement. *Arch Ophthalmol* 1999;117:450-454.

20. J. G. Walt, R. Rendas-Baum, M. Kosinski, et al. Psychometric Evaluation of the Glaucoma Symptom Identifier. *J Glaucoma* 2011;20:148-159.

21. G. F. Schwartz, K. S. Plake and M. A. Mychaskiw. An assessment of readiness for behaviour change in patients prescribed ocular hypotensive therapy. *Eye (Lond)* 2009;23:1668-1674.

22. J. M. Burr, M. Kilonzo, L. Vale, et al. Developing a preference-based Glaucoma Utility Index using a discrete choice experiment. *Optom Vis Sci* 2007;84:797-808.

23. B. Sleath, S. J. Blalock, A. Robin, et al. Development of an instrument to measure glaucoma medication self-efficacy and outcome expectations. *Eye (Lond)* 2010;24:624-631.

24. P. Severn, S. Fraser, T. Finch, et al. Which quality of life score is best for glaucoma patients and why? *BMC Ophthalmol* 2008;8:2.

25. P. Nelson, P. Aspinall, O. Papasouliotis, et al. Quality of life in glaucoma and its relationship with visual function. *J Glaucoma* 2003;12:139-150.

26. V. K. Gothwal, T. A. Wright, D. B. Elliott, et al. The refractive status and vision profile: Rasch analysis of subscale validity. *J Refract Surg* 2010;26:912-915.

27. V. K. Gothwal, T. A. Wright, E. L. Lamoureux, et al. Activities of Daily Vision Scale: what do the subscales measure? *Invest Ophthalmol Vis Sci* 2010;51:694-700.

28. K. Pesudovs, V. K. Gothwal, T. Wright, et al. Remediating serious flaws in the National Eye Institute Visual Function Questionnaire. *J Cataract Refract Surg* 2010;36:718-732.

29. K. Pesudovs, J. M. Burr, C. Harley, et al. The development, assessment, and selection of questionnaires. *Optom Vis Sci* 2007;84:663-674.

30. J. Khadka, B. Ryan, T. H. Margrain, et al. Development of the 25-item Cardiff Visual Ability Questionnaire for Children (CVAQC). *Br J Ophthalmol* 2010;94:730-735.

31. C. McAlinden, K. Pesudovs and J. E. Moore. The development of an instrument to measure quality of vision: the Quality of Vision (QoV) questionnaire. *Invest Ophthalmol Vis Sci* 2010;51:5537-5545.

32. C. McAlinden, E. Skiadaresi, K. Pesudovs, et al. Quality of vision after myopic and hyperopic laser-assisted subepithelial keratectomy. *J Cataract Refract Surg* 2011;37:1097-1100.

33. M. Lundstrom and K. Pesudovs. Questionnaires for measuring cataract surgery outcomes. *J Cataract Refract Surg* 2011;37:945-959.

34. V. K. Gothwal, T. A. Wright, E. L. Lamoureux, et al. Rasch analysis of the quality of life and vision function questionnaire. *Optom Vis Sci* 2009;86:E836-844.

35. E. L. Lamoureux, J. G. Ferraro, J. F. Pallant, et al. Are standard instruments valid for the assessment of quality of life and symptoms in glaucoma? *Optom Vis Sci* 2007;84:789-796.

36. J. Khadka, K. Pesudovs, C. McAlinden, et al. Reengineering the glaucoma quality of life-15 questionnaire with rasch analysis. *Invest Ophthalmol Vis Sci* 2011;52:6971-6977.

37. K. Pesudovs. Item banking: a generational change in patient-reported outcome measurement. *Optom Vis Sci* 2010;87:285-293.

38. J. B. Bjorner, C. H. Chang, D. Thissen, et al. Developing tailored instruments: item banking and computerized adaptive assessment. *Qual Life Res* 2007;16 Suppl 1:95-108.

39. D. Thissen, B. B. Reeve, J. B. Bjorner, et al. Methodological issues for building item banks and computerized adaptive scales. *Qual Life Res* 2007;16 Suppl 1:109-119.

40. H. Wainer. *Computerized adaptive testing. A primer*. 2nd ed. Mahway, NJ: Lawerance Erlbaum; 2000.

41. D. E. Gaasterland, F. Ederer, K. Sullivan, et al. Advanced Glaucoma Intervention Study. 2. Visual field test scoring and reliability. *Ophthalmology* 1994;101:1445-1455.

42. QSR International. NVIVO 9. Victoria: QSR International Pty Ltd; 2009.

43. B. M. Edhlund. *NVivo essentials : the ultimate help when you work with qualitative analysis*. Stallarholmen, Sweden: Form & Kunskap; 2007.

44. T. R. Walsh, D. E. Irwin, A. Meier, et al. The use of focus groups in the development of the PROMIS pediatrics item bank. *Qual Life Res* 2008;17:725-735.

45. G. T. Bond and C. M. Fox. *Applying the Rasch Model: Fundamental Measurement in the Human Science*. 2nd ed. Mahwah, New Jersey: Lawrence Erlbaum Associates; 2007.

46. J. M. Linacre. Optimizing rating scale category effectiveness. *J Appl Meas* 2002;3:85-106.

47. D. A. DeWalt, N. Rothrock, S. Yount, et al. Evaluation of item candidates: the PROMIS qualitative item review. *Med Care* 2007;45:S12-21.

48. J. B. Jobe and D. J. Mingay. Cognitive research improves questionnaires. *Am J Public Health* 1989;79:1053-1055.

49. J. Lacey, H. Cate and D. C. Broadway. Barriers to adherence with glaucoma medications: a qualitative research study. *Eye (Lond)* 2009;23:924-932.

50. S. A. Taylor, S. M. Galbraith and R. P. Mills. Causes of non-compliance with drug regimens in glaucoma patients: a qualitative study. *J Ocul Pharmacol Ther* 2002;18:401-409.

51. V. K. Gothwal, T. A. Wright, E. L. Lamoureux, et al. Multiplicative rating scales do not enable measurement of vision-related quality of life. *Clin Exp Optom* 2011;94:52-62.

52. D. Amtmann, K. F. Cook, M. P. Jensen, et al. Development of a PROMIS item bank to measure pain interference. *Pain* 2010;150:173-182.

53. K. S. Coyne, M. K. Margolis, T. Kennedy-Martin, et al. The impact of diabetic retinopathy: perspectives from patient focus groups. *Fam Pract* 2004;21:447-453.

54. R. Misajon, G. Hawthorne, J. Richardson, et al. Vision and quality of life: the development of a utility measure. *Invest Ophthalmol Vis Sci* 2005;46:4007-4015.

55. L. E. Caudle, K. A. Williams and K. Pesudovs. The Eye Sensation Scale: an ophthalmic pain severity measure. *Optom Vis Sci* 2007;84:752-762.

56. K. Pesudovs, E. Garamendi and D. B. Elliott. The Contact Lens Impact on Quality of Life (CLIQ) Questionnaire: development and validation. *Invest Ophthalmol Vis Sci* 2006;47:2789-2796.

57. E. L. Lamoureux, K. Pesudovs, J. F. Pallant, et al. An evaluation of the 10-item vision core measure 1 (VCM1) scale (the Core Module of the Vision-Related Quality of Life scale) using Rasch analysis. *Ophthalmic Epidemiol* 2008;15:224-233.

58. J. F. Fries, B. Bruce, J. Bjorner, et al. More relevant, precise, and efficient items for assessment of physical function and disability: moving beyond the classic instruments. *Ann Rheum Dis* 2006;65:iii16-iii21.
59. M. R. Wilson, A. L. Coleman, F. Yu, et al. Depression in patients with glaucoma as measured by self-report surveys. *Ophthalmology* 2002;109:1018-1022.

Figure legend:

Figure 1: Bar diagram showing the number of unique items identified from the literature review and qualitative sessions (focus groups and semi-structured interviews). (AL = Activity limitation, MB = Mobility, VS = Visual symptoms, OS = Ocular surface symptoms, GS = General symptoms, CV = Convenience, HC = Health concerns, EM = Emotional, SC = Social, and EC = Economic).

Table 1: Glaucoma-specific PRO instruments used to extract items for the initial item pool

	Instrument	No. of items	Basis of original content development
1	Ross et al (1984)[9]	16	Not reported
2	Mills et al (1986)[8]	15	Literature review
3	COMTOL - Barber et al (1997)[10]	50	Lay focus groups
4	GSS - Lee et al (1998)[16]	10	Literature review/other PROs
5	Viswanathan (1999)[19]	10	Literature review
6	SIG - Janz et al (2001)[4]	43	Literature review
7	Odberg et al (2001)[5]	33	Not reported
8	Uenishi et al (2003)[12]	31	Literature review/other PROs
9	GQL-15 - Nelson et al (2003)[25]	15	Literature review/other PROs
10	TSS-IOP - Atkinson et al (2003)[13]	15	Literature review/other PROs/lay focus groups
11	EDSQ - Nordman et al (2007)[15]	46	Expert interviews
12	Dunker et al (2007)[11]	29	Not reported
13	Glau-QOL - Bechetoille et al (2008)[14]	36	Literature review/ other PROs/ lay focus groups
14	Glausat - Ruiz et al (2010)[18]	22	Expert/lay focus groups
15	GSI - Walt et al (2011)[20]	32	Other instruments/ patients and clinicians input
16	Adherence Questionnaire-Schwartz et al (2009)[21]	56	Literature/expert input and opinion
17	Self-efficacy - Sleath et al (2010)[23]	35	Other PROs/peer-reviewed literature/ expert and patients input
18	Outcome Expectation Scale-Sleath et al (2010)[23]	4	Other PROs/peer-reviewed literature/expert and patients input
19	Burr et al (2007)[22]	32	Focus groups/other PROs/expert opinion

PRO: Patient reported outcome
COMTOL: Comparison of Ophthalmic Medication for Tolerability
GSS: Glaucoma Symptom Identifier
GQL-15: Glaucoma Quality of Life Questionnarie-15
VFDQ: Visual Field Disability Questionnaire
TSS-IOP: Treatment Satisfaction Survey of Intraocular Pressure
SIG: Symptoms Impact Glaucoma Score
EDSQ: Eye-drop Satisfaction Questionnaire
Glau-QOL: Glaucoma Quality of Life
Gluasat: Glaucoma Satisfaction Questionnaire
GSI : Glaucoma Symptom Index

Table 2: Demographic characteristics of the participants

Variable	
Age	
Mean± standard deviation (years)	63±13
Range (years)	32-88
Gender; n (%)	
Female	38 (52.8%)
Glaucoma type; n (%)	72 (100%)
POAG	51 (70.8%)
Secondary glaucoma	9 (12.5%)
ACG	7 (9.7%)
PXF	4 (5.6%)
NTG	1 (1.4%)
Glaucoma severity (AGIS classification; n (%)	58 (100%)
≤ Mild in better eye, ≤ moderate in worse eye	17 (29.3%)
≤ Mild in better eye, ≥ Severe in worse eye	27 (46.6%)
Moderate in better eye, ≥ Moderate in worse eye	7 (12.1%)
≥ Severe (both)	6 (10.3%)
Visual acuity (Better eye); LogMAR (Snellen equivalent)	
Median	0.1 (6/7.5)
Range	-0.2 to 1.30 (6/4 to 6/120)
Visual (worse eye; LogMAR (Snellen equivalent)	
Median	0.2 (6/9.5)
Range	-0.2 to PL (6/4 to PL)
Current management; n (%)	58 (100%)
Topical medications only	19 (32.8%)
Surgery with no topical medications	26 (44.8%)
Surgery with topical medications	7 (12.1%)
None	5 (8.6%)
Types of surgery among the surgical group (%)	
Selective laser trabeculoplasty	13%
Argon laser trabeculoplasty	5.1%
Trabeculectomy	38.5%
Trabeculectomy with mitomycin-C or 5-fluorouracil	23%
Molteno tube	5.1%
Number of topical medications among the medication user group (%)	
One	50%
Two	46.2%
Three	3.8%
Types of topical medications (%)	
Prostaglandin agonist	45%
Beta-blocker	27.5%
Carbonic anhydrase inhibitor	7.5%
Carbonic anhydrase inhibitor plus Beta-blocker	7.5%
Prostaglandin agonist plus Beta-blocker	5.0%
Prostaglandin agonist	2.5%
Muscarinic receptor agonist	2.5%
Selective alpha-2 adrenergic agonist	2.5%

* Percentages for some variables may not equal 100% due to missing data
POAG= Primary open angle glaucoma; ACG =Acute Angle closure glaucoma; PXF = Pseudo-exfoliation glaucoma; NTG = Normal tension glaucoma; AGIS = Advanced Glaucoma Intervention Study; Log MAR = Log of Minimum Angle of Resolution; PL = Light Perception

Table 3: Example of items extracted from the focus groups and semi-structured interviews

Participant's statement	Number of times the issue was discussed	Item stem
"No, I can't, I can't, I can see what they [price labels] are, but I can't read them you know"	3	…reading price labels in shops
"Ah, I found the last couple of weeks when I've been watching the television, in spite of the fact that the next program is something I want to watch, after some time my eyes feel tired and they feel sore. And I have had enough and I turn the box off – because I can't cope"	4	…watching television
"..sunlight during the day there's sun shining brightly you can't see and then you go into, down the side of the building, there is shadow, then you notice a difference then"	7	…seeing in bright sunlight And …difficulty seeing when going from bright sunlight to dark conditions
"Yes, tiredness… I forget which drops I got, but the first drops I got originally, one of the side effects was tiredness"	3	…tiredness
"Because you can not squeeze the droppers and get about 8 drops out and it flushes over your face – things like that – and they can erode your ability to treat yourself. And I have spoken to people who don't apply their own drops, they get the spouse…"	4	…administering eyedrops
"I used to work; I lost my job through my eye sight because I used to drive…"	2	…losing your job

Table 4: Number of items in the item pool before and after binning and winnowing

Domain	AL	MB	VS	OS	GS	CV	HC	EM	SC	EC	**Total no. of Items**
Initial Item pool	168	32	31	36	19	90	149	142	35	35	737
Final item pool	88	20	18	22	15	39	45	49	23	22	341

AL: Activity Limitation; MB: Mobility; VS: Visual Symptoms, OS: Ocular Surface Symptoms; GY: General Symptoms; CV: Convenience; HC: Health Concerns; EM: Emotional; SC: Social; EC: Economic

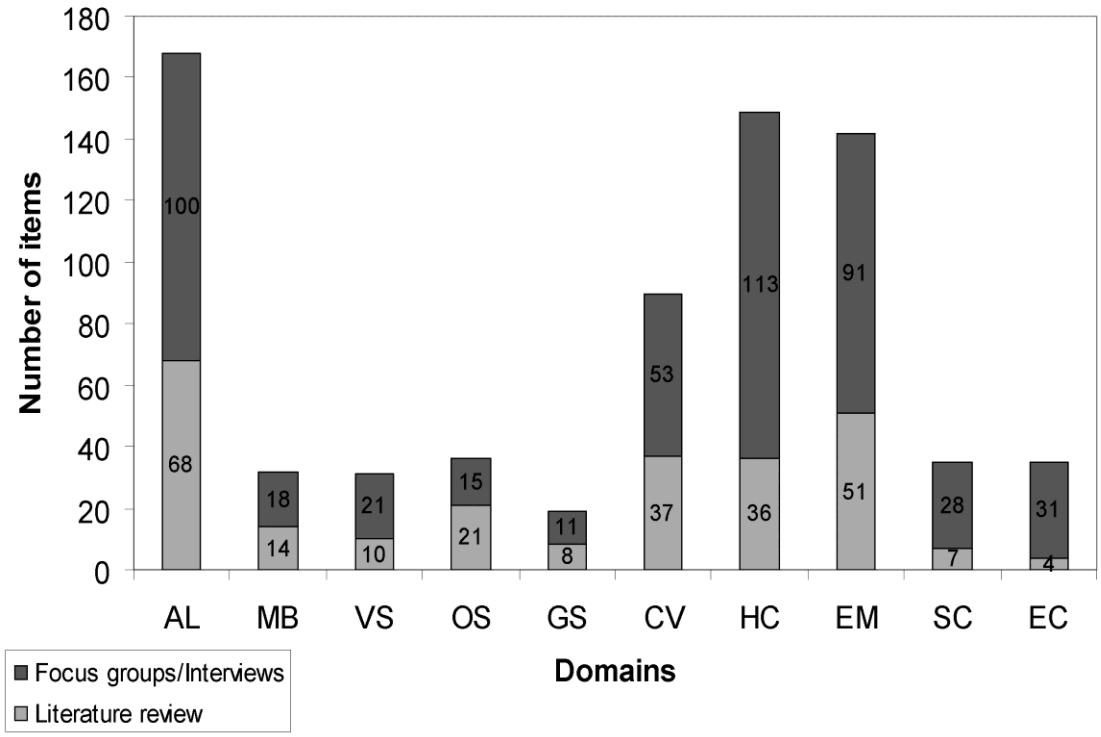
Table 5: Example of items deleted at the binning and winnowing stage

Item stem	Domain	Reason for removal
"driving in glare"	Activity limitation	Redundant item (captured by other items; *"driving towards oncoming headlights"* and *"driving towards the sun"*)
"seeing objects to the side of you"	Mobility	Redundant item (captured by other item; *"noticing things to the left or the right of you while you are walking"*)
"discomfort in your eye after an eye operation due to stiches"	Symptoms	Narrow item (too specific and refers to discomfort due to stiches) and *"discomfort in your eyes"* already included
"wish something could be done to cure glaucoma"	Emotional	Not really a measurable item, but the underlying issue is denial; *"have trouble accepting that your eye problems are permanent"* already included
"felt weaker, less strong than I used to be"	Emotional	Confusing item (double barrelled; item can refer to both physical or emotional weakness) and *"feel emotionally drained"* already included
"not being able to do things for as long as you want to, used to, e.g. reading"	Health concern	Inconsistent with domain definition (item does not refer to health related concerns, really an inconvenience)

Table 6: Domain question format and response options

Domain	Item stem ***Because of your glaucoma and its treatment…***	Response options
Activity limitation and Mobility	How much difficulty do you have…..?	*None to Unable to do because of my vision*
Symptoms	How often do you experience….?	*Never to Very often*
	How severe is the…?	*Not at all to Severe*
	How much of a problem is/are the…?	*None to A lot*
Convenience	How much trouble is…?	*None to Extreme*
Health concerns	How concerned are you about…?	*Not at all to Extreme*
Emotional	How often do you….?	*None of the time to All of the time*
Social	How much of a problem do you have with…?	*None to Unable to do because of my vision*
Economic	How concerned are you about…?	*Not at all to Extreme*

Figure 1:


180
160
140
120
100
80
60
40
20
0
Number of items
100
68
18
14
21
10
15
21
11
8
53
37
113
36
91
51
28
7
31
4
AL
MB
VS
OS
GS
CV
HC
EM
SC
EC
Domains
Focus groups/Interviews
Literature review