BMJ Open

Preference-based disease-specific health-related quality of life instrument for glaucoma: a mixed methods study protocol

Sergei Muratov,[1] Dominik W Podbielski,[2] Susan M Jack,[3] Iqbal Ike K Ahmed,[2,4] Levine A H Mitchell,[1] Monika Baltaziak,[2] Feng Xie[5]

CrossMark

For numbered affiliations see end of article.

Correspondence to
Dr Feng Xie;
fengxie@mcmaster.ca

ABSTRACT

Introduction: A primary objective of healthcare services is to improve patients' health and health-related quality of life (HRQoL). Glaucoma, which affects a substantial proportion of the world population, has a significant detrimental impact on HRQoL. Although there are a number of glaucoma-specific questionnaires to measure HRQoL, none is preference-based which prevent them from being used in health economic evaluation. The proposed study is aimed to develop a preference-based instrument that is capable of capturing important effects specific to glaucoma and treatments on HRQoL and is scored based on the patients' preferences.

Methods: A sequential, exploratory mixed methods design will be used to guide the development and evaluation of the HRQoL instrument. The study consists of several stages to be implemented sequentially: item identification, item selection, validation and valuation. The instrument items will be identified and selected through a literature review and the conduct of a qualitative study. Validation will be conducted to establish psychometric properties of the instrument followed by a valuation exercise to derive utility scores for the health states described.

Ethics and dissemination: This study has been approved by the Trillium Health Partners Research Ethics Board (ID number 753). All personal information will be de-identified with the identification code kept in a secured location including the rest of the study data. Only qualified and study-related personnel will be allowed to access the data. The results of the study will be distributed widely through peer-reviewed journals, conferences and internal meetings.

Strengths and limitations of this study

- Use of a mixed methods approach facilitates development of a health state descriptive system for the instrument.
- Reliance on patient inputs to identify and select items is likely to increase overall validity of the instrument.
- Glaucoma-specific measure of health-related quality of life and preference-based scoring enable this instrument to be used in clinical and economic evaluations.
- Incorporating patients' inputs supplemented by the evidence from literature into the instrument development is to some extent subjective.
- A topic of an ongoing debate, the use of patients as the source of preferences for health state valuation is subject to limitations that have been discussed in the literature.

INTRODUCTION

In a patient-oriented healthcare system, a primary objective of health services is to improve patients' health and health-related quality of life (HRQoL) while maintaining efficiency of the system. Glaucoma is an ocular condition that is characterised by increased intraocular pressure, retinal ganglion cell death and often blindness. It is the second leading cause of irreversible blindness in the world with 60 million people worldwide estimated to be suffering from the condition.[1] Topical drugs form the basis of current glaucoma management. Trabeculectomy, and other filtering glaucoma procedures, are used when drugs fail.[2] Intraocular devices that drain aqueous humour and microinvasive glaucoma surgery (MIGS) are being increasingly used in glaucoma patient management.[3 4] Although glaucoma or its treatment have no sizable impact on mortality, it does affect HRQoL of patients. Therefore, HRQoL has become an important outcome that can assist with glaucoma management when measured properly.

Health utility is a common method of measuring patients' health status and HRQoL using a single index score anchored at 0 for being dead and 1 for full health.[5]

BMJ

Health utility can be also used to measure population health. For instance, Statistics Canada uses utility values collected by the generic Health Utility Index as a measure of overall functional health.[6] Another common application of health utilities is the calculation of quality adjusted life years (QALYs) (ie, quantity of life weighted by quality of life as measured using health utility) for health economic evaluations (EE). When an array of treatment options is available, public and hospital administrators are faced with challenges in selecting treatments in order to use public funds efficiently. EEs synthesise evidence on costs and QALYs and are widely used to inform resource allocation decisions.[5] However, the contribution of EE to decision-making appears to be limited in the field of glaucoma due to the lack of a utility-based instrument that is validated, sensitive and easy enough to use in longitudinal studies.

Health utility in patients with glaucoma can be measured using direct and indirect approaches.[7–9] Both approaches however have drawbacks that render their use less suitable for glaucoma clinical practice or research. For instance, the direct use of established preference elicitation techniques such as the time trade off or the standard gamble is time consuming and cognitively challenging. One indirect option is to use predeveloped preference-based instruments (eg, EuroQol-5 dimensions). These existing instruments were developed for generic use so are limited in detecting important health changes in patients with glaucoma.[7 10] Existing vision or glaucoma specific instruments (eg, National Eye Institute Visual Function Questionnaire or Glaucoma Quality of Life-15) consist of multiple domains and multiple items per domain that are relevant to the disease (eg, vision) and produce a score by summing up the responses to the items.[11–13] Owing to their non-preference scoring method the summary scores are not on the health utility scale and thus cannot be used in economic evaluation. Developing a mapping algorithm to convert existing non-preference-based glaucoma-specific HRQoL scores to health utility through statistical models is an option.[14] Mapping results, however, seem inconclusive with studies often reporting poor model performances and limited validity.[10 15] Attempts to employ mapping with routine measures of vision have not been satisfactory either.[10 16] Moreover, existing glaucoma instruments have been reported to use attributes that do not cover the broader HRQoL aspects such as the impact of treatments.[17] They tend to focus on disease burden only (ie, visual acuity, loss of peripheral vision, etc). However, for many patients with glaucoma, use and adherence to eye drops and local ocular issues may have non-negligible impact on HRQoL.

There were previous attempts to develop preference-based HRQoL instruments to measure health utilities in ophthalmology. For example, the Vision and Quality of Life Index (VisQOL) was developed for patients with vision impairment.[18] However, only about 18% of the focus group participants that contributed to the VisQOL's item bank had glaucoma. Also, VisQoL cannot be used as a standalone instrument for generating utility values.[19] To the best of our knowledge, the only glaucoma-specific preference-based instrument (GPI) was developed by Burr et al.[20] The GPI has six dimensions with four response levels per dimension. Discrete choice experiment was then used for health state valuation. This work however has several limitations. First, psychometric properties of the GPI were not assessed neither prior to (preferably) conducting the valuation exercise nor afterwards reliability, responsiveness and validity of the instrument remain unknown. Second, the anchoring of the value set was performed on the 0 (worst state)—1 (best state) scale instead of the 0 (dead)—1 (perfect health). As a result of the latter, the measure in this form cannot be used to calculate QALYs for economic evaluations.[20 21] Our intention therefore is to address the current need for a preference-based glaucoma measure building on the body of experience and knowledge gained so far.

Aim: To develop a glaucoma-specific preference-based HRQoL instrument that can capture important impact of glaucoma disease and treatments on HRQoL and be used in economic evaluation of glaucoma-related interventions.

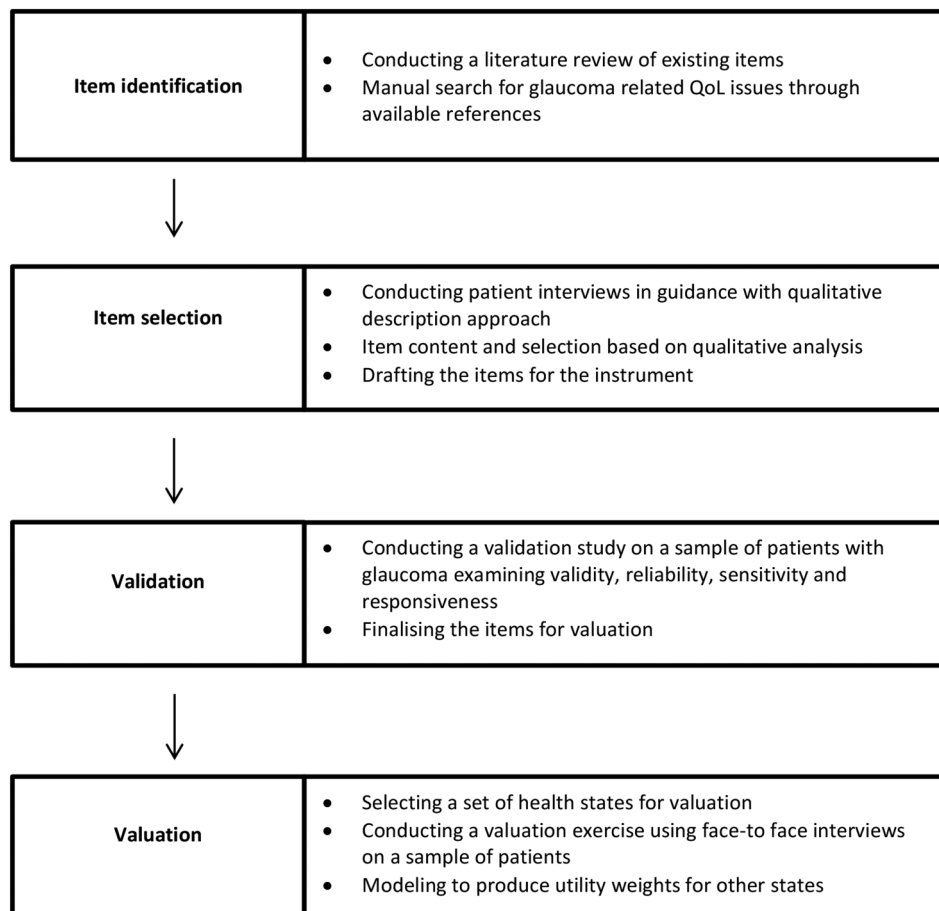*Objective(s)*: The specific objectives of the study are to:
1. Develop the descriptive system for this instrument using a mixed methods approach
2. Establish psychometric properties of this instrument
3. Develop preference-based scoring algorithm for this instrument.

METHODS AND ANALYSIS

A mixed methods approach combines the best features of qualitative and quantitative research in a rigorous manner. We will use a common mixed methods design, an exploratory sequential study that builds every next stage on the results of the preceding stages.[22] The design is well suited for developing a HRQL instrument as it allows for the identification of key domains/items using qualitative techniques and then testing them using quantitative techniques (psychometric and econometric).[22] The development of this instrument therefore consists of several distinct stages to be implemented sequentially: item identification, item selection, validation and valuation. Figure 1 presents a brief descriptive summary of the stages.

When developing a HRQoL instrument, it is important to reflect the full range of important patient experience and views to ensure that it is a valid, reliable, sensitive and responsive measure.[23] There are two main approaches. One is a top-down approach whereby the content is obtained from the literature, including existing instruments and surveys. The other is a bottom-up method that relies on patient inputs and uses qualitative techniques to generate items. Although more time consuming, the latter approach is likely to increase content

Figure 1 Summary of glaucoma HRQoL tool development by stages. HRQoL, health-related quality of life.

Item identification	• Conducting a literature review of existing items • Manual search for glaucoma related QoL issues through available references

↓

Item selection	• Conducting patient interviews in guidance with qualitative description approach • Item content and selection based on qualitative analysis • Drafting the items for the instrument

↓

Validation	• Conducting a validation study on a sample of patients with glaucoma examining validity, reliability, sensitivity and responsiveness • Finalising the items for valuation

↓

Valuation	• Selecting a set of health states for valuation • Conducting a valuation exercise using face-to face interviews on a sample of patients • Modeling to produce utility weights for other states

validity and improve responsiveness to change.[24] Another benefit of the bottom-up methodology is that it is well suited for developing an instrument that is amenable to valuation.[24] In order to ensure that we take into consideration both patient experience and existing evidence, we will use a combination of the two.

Stage 1: Item identification
A review of published studies will be undertaken using a comprehensive search strategy. The main objective of this stage is to identify items that have been in use to measure or describe various facets of HRQoL in patients with glaucoma and to inform the next qualitative stage. The following databases will be searched: MEDLINE In Process (1950 to present), EMBASE (1980 to present) and MEDLINE PubMed (up to present). Search filters will include the language (English only) and age (above 18). The key search terms will include glaucoma, quality of life, patient satisfaction, questionnaire, instrument, glaucoma drug effects and eye surgery. The title and abstract of all the retrieved articles will be screened for relevance and then full text review for eligibility. Publications will be included if they describe the general impact on HRQoL of the disease and treatments as perceived by patients. The identified items will be used to develop a patient interview guide.

Stage 2: Item selection
Qualitative description, a qualitative research approach, will inform all sampling, data collection and analytic decisions at this stage of the project. Use of this naturalistic mode of inquiry allows us to understand study participant's experiences and views related to HRQoL and glaucoma helps determine items of most importance and relevance to them and document information in the language used by the participants–which will be used in creating a sensitive and responsive tool that will resonate with this population.[25]

The study will take place on the premises of the Prism Eye Institute, Ontario, Canada. A purposeful sample of patients with glaucoma will be recruited at one of the institute's clinics. The following inclusion criteria will be applied to determine eligibility of participating patients:
1. Confirmed diagnosis of glaucoma in one or both eyes (according to the Canadian Ophthalmological Society's Glaucoma Guideline)[26]
2. No evidence of other eye disease significantly affecting vision
3. Age 18 years and above,
4. No cognitive impairment,
5. Able to speak and understand English
6. Able to consent

Study eligibility is determined during ophthalmology clinic visits. All visiting patients are provided a one-page

information sheet summarising the study. Interested individuals who meet study eligibility criteria are contacted by an interviewer to confirm participation and book an interview. Written consent is obtained prior to beginning of the interview.

Purposeful sampling refers to identifying individuals who have experienced the phenomenon under study and who can provide a rich description of their experiences.[27] Further, we will employ maximum variation (MV) as a strategy of purposeful sampling. MV aims to identify central themes that are common among the patients by sampling participants with diverse characteristics (eg, duration of glaucoma, its severity, history of glaucoma surgery and the number of times eye drops are used per day).[27] Such a small sample of great variation captures core shared items while also allowing for documentation of unique features of each case. The central themes that are common across all the patients but different in terms of intensity are likely to increase the validity and sensitivity of the instrument once converted into the items. A MV sampling matrix will be created to ensure that recruitment covers the desired variation: each person in the sample should be as different as possible from the others based on the aforementioned characteristics.

An estimate of the number of individuals required to capture a comprehensive set of health experiences in this patient population is determined. The sample sizes from published qualitative studies ranged from 6 to 50 participants with 15–20 participants on average.[28] For a descriptive qualitative study of this nature, employing maximum variation sampling, we estimate recruiting at least 30 participants.

Data collection

Each participant will undergo a face-to-face, semistructured interview, conducted by a researcher with experience in qualitative interviewing. Each interview will last up to 60 min and will be conducted in a private room located within the eye clinic.

Using the key items identified in stage 1 as probes, the interview will explore (1) the participants' experiences of living with glaucoma; (2) how living with glaucoma impacts their HRQoL; and (3) their experiences of undergoing glaucoma treatments. Responsive questioning will be employed throughout the interviews. Also, since the qualitative analysis of the emerging data will begin immediately after the first interviews and be simultaneous, new information that is discovered in the process can be added to the guide for further exploration in subsequent interviews.

At the end of each interview, the participant is asked to complete a short demographic questionnaire. In addition, the ophthalmologist fills out a form containing basic clinical information for each participant such as disease severity, vision parameters (eg, vision acuity, field measurements, recent intraocular pressure) and details of current and previous treatments.

Data analysis

The interviews with participants will be audio recorded and transcribed verbatim. Identifying information will be removed from interview transcripts and then uploaded into computer-assisted qualitative data analysis software to aid in data storage, management and coding.

Analysis of the transcribed interviews will be guided through the use of the Framework Method (FM). FM is a highly systematic method of qualitative content analysis gaining more popularity with developers of preference-based HRQoL instruments as the goal of the analysis is known at the beginning (ie, item generation and selection) and no underlying theories need to be induced in contrast to some other methods of qualitative analysis.[24 29 30]

Step 1: Familiarisation and coding

The researchers will familiarise themselves with the interview transcripts. Through this process the researchers will identify a number of common and recurring themes among all the items discussed during the interviews. These will include the themes introduced by the guide and the new issues raised by the participants. Each recurring theme will receive a code and a brief description.

Step 2: Development and application of a conceptual framework

The codes from the first few transcripts will be organised into themes and form the initial conceptual framework. Researchers will apply the framework to the subsequent transcripts as interviews continue and refine the coding by incorporating new items and grouping or regrouping the codes. The new emerging items will also be incorporated into the guide to be used as a probe in the following interviews. The refine-apply cycle will be repeated until no new information is identified. The final conceptual framework will be reapplied to each transcript.

Step 3: Charting a matrix

At this step the coded data will be summarised in a framework matrix: one row for each study participant and one column per code. The abstracted data of participants' expressions will be used to fill out the corresponding matrix cells. The matrix will help finalise the themes and retain the original wording for the items that describe each theme.

Step 4: Draft items for the instrument and pilot testing

In general, the number of items used in a preference-based instrument is usually limited (eg, ≤9) in order to ensure the feasibility of developing a preference-based scoring algorithm.[21 24] Candidate items that are deemed most pertinent for the purposes of the measurement and their corresponding response levels will be selected based on joint consideration of qualitative feedback from patients and clinical experts of the team. These items will form the items of the instrument.

To ensure that the items are understandable to patients, we will conduct a pilot test among a small group (10–15 patients) of not previously involved patients asking them to complete the instrument. We will evaluate the draft instrument on the following parameters: the patients' overall impression, the clarity of the instructions and of the items/response choices, the readability of the format, level of difficulty and whether any assistance is required.[31–33] Building on the use of a mixed methods approach, the instrument will be validated and valued following the pilot test.

Stage 3: Validation

The instrument's psychometric properties will be established through a prospective validation study. The results of the validation study such as validity, reliability and sensitivity will lead to finalisation of the instrument's health state descriptive system.

Design and sample

A longitudinal study will be conducted using the same inclusion criteria as for the qualitative stage 2. Participants will be recruited through stratified convenience sampling. Stratification will be performed by treatment type: only topical drops, filtering procedures (trabeculectomies, tubes, etc) and MIGS. Approximately 20% of patients from each treatment arm will represent patients who have not received prior glaucoma therapy ('virgin eye'): responses from these patients will be used to test responsiveness. Recruitment will also ensure that the sample is balanced by three disease severity levels: mild, moderate and advanced.

The instrument will be administered to the participants at baseline and at two follow-up points: at 2 weeks and at 3 months following treatment event/initiation in order to test key psychometric properties (figure 2). The National Eye Institute 25-Item Visual Function Questionnaire (NEI-VFQ-25) will be used throughout the study as the 'gold' standard.[11] The NEI-VFQ-25 is an abbreviated (25 item) version of the NEI-VFQ-51 vision specific tool to measure HRQoL in patients with chronic eye conditions. It covers 12 domains and is currently considered the standard for non-preference-based HRQoL tools in ophthalmology.[17]

For the required sample size, we will follow a general assumption: the number of participants in a validation study needs to exceed the number of items in the instrument by a factor of at least 5 for each treatment subgroup bringing the total to ~150 participants.[34 35]

Item and psychometric analysis
Missing responses

Item analysis will be carried out to assess completion rates and missing values (ie, the number of participants who completed the entire instrument, or missed one or more items). Further analysis of missing values will take account of the importance of missing data (important data, eg, no level is selected or less important, eg, a gap in demographic data).

Score distribution

'Ceiling' and 'floor' effects indicate that responses are rather skewed towards the extremes of the scale. A higher proportion of responses accumulating at the favourable end of the scale will indicate the 'ceiling' effect, while the opposite will point toward the 'floor' effects. This reduces the ability of the instrument to determine any changes among those who score on the extremes and if present requires further investigation. The percentages of the sample will be calculated: 10% or greater of responses scoring at either end of the scale will be suggestive of the ceiling or floor effects.[36]

Construct validity

Construct validity refers to the degree to which an instrument measures what it is designed to measure and is an important characteristic.[35] Construct validity will be measured through assessing convergent and divergent validity. Convergent validity shows correlation among attributes that should be correlated in theory whereas divergent (or discriminant) validity demonstrates the lack of correlation among attributes that in theory
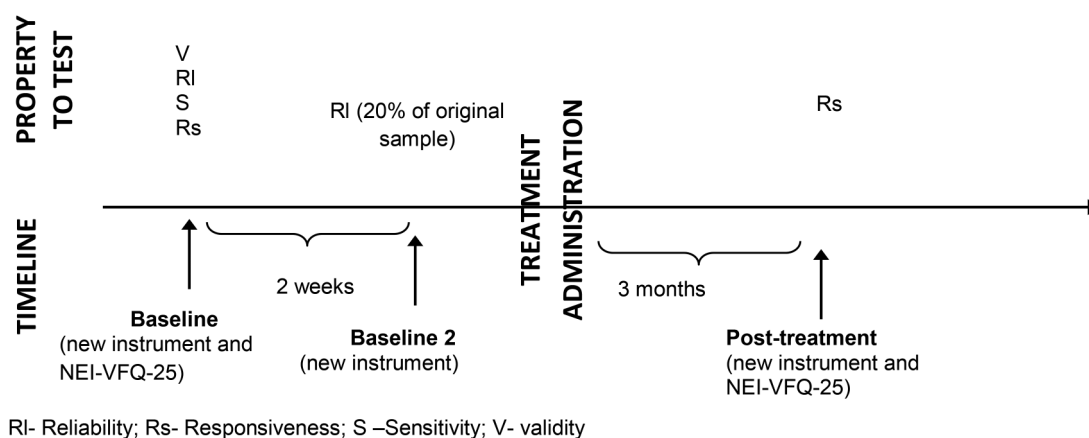

Figure 2 Schematic plan of validation study. NEI-VFQ, The National Eye Institute 25-Item Visual Function Questionnaire.

should not be associated. Both types of construct validity will be measured using the NEI-VFQ-25 as a reference instrument.

We will use Pearson correlation coefficients to evaluate construct validity. Moderate to strong correlations ($\geq$0.5) between similar attributes will support convergent validity, while weak correlations ($\leq$0.3) between dissimilar attributes will support the discriminant validity.[35 36]

Reliability

Test–retest reliability refers to the repeatability of a measurement administered on two occasions during which there is no significant change in the participant's status. We will readminister the instrument 2 weeks after the baseline to ensure that the participants do not recall their previous responses while remaining in a stable condition. For the second assessment, the instrument will be mailed to a random sample of 20% of the original group who are expected to have a stable condition.[35] Participants' condition will be considered stable if, within the 2-week period, they do not consult an ophthalmologist or another physician due to progression of any disease, there is no change in treatment and the participants experience no major traumatic event in their life.[37] Reliability will be measured by determining intraclass coefficient correlation (ICC) that calculates the proportion of the between-subject variance to the total variance. An ICC of 0.7 and greater will indicate good reliability.[36 37]

Sensitivity

When used in clinical trials, it is important for the instrument to be able to detect differences among patients with various disease severity levels. Sensitivity will be evaluated by cross-sectional comparison among participants with various severities using between group analysis of variance F statistics. Relative efficiency (RE) will be calculated based on a ratio of the F-statistic of one group to the F-statistic of a reference group. A larger RE value indicates higher sensitivity.[35]

Responsiveness

Responsiveness is the ability to detect within-patient change, often as a result of a treatment.[33 35] This will be assessed by applying the instrument to participants on two occasions: at baseline and after receiving treatments. Readministration of the instrument will occur at 3 months following treatment to ensure that potential treatment effects have set in. The results of NEI-VFQ-25 will be used as a reference of any changes in HRQoL. Standardised response mean (SRM), the mean change in score from baseline assessment to the 3-month reassessment divided by the SD of the change in scores, will be calculated. An SRM of 0.20 will be considered as small, one of 0.50 as moderate and one of 0.80 or greater as large. A moderate to large SRM indicates that the instrument is responsive.[31 35]

Stage 4: Valuation

Once the instrument is validated, the next step would be to derive utility values for the health states described by the instrument. The design of the valuation study will be determined once the descriptive system of the instrument is developed.

We expect to have 5–9 items, each with 3–5 levels of responses. Thus, the descriptive system will likely define over thousands of health states. It would not be feasible to value all health states as seen in other preference-based instruments' valuation studies.[13] A common practical solution is to choose a subset of health states for valuation.[38 39] We will choose from two main approaches to the health state selection: factorial design and orthogonal design.[40–42] The factorial design ensures that multiple combinations of response levels across the dimensions are included. Health states are classified into severity groups (eg, mild, moderate and severe) and then a subset is selected that includes the worst possible health state, and a number of health states from each of these severity groups.[39 42] The second approach assumes independence of the dimensions and constructs an orthogonal array from which to choose the minimum sample of health states. The resulting valuation set of health states will cover an adequate mix of mild, moderate and severe states. The number of states will be balanced with the number of observations per state. The common tendency is to trade off in favour of the number of states.[41] For example, an average of 15 observations per state was used to estimate the Short Form six dimension in a sample of 249 states.[40] An allocation procedure will be determined based on the number of health states selected and the targeted sample size. The sampling method will aim to achieve a representative sample of the glaucoma patient population and reflect geographical and socioeconomic characteristics of this population.

Time-trade off (TTO) has been used widely to elicit health utilities.[5 43] We will use face-to-face interviews to conduct valuation exercises using TTO to obtain utility values for the selected health states. Statistical modelling will then use the observed utilities of a subset of the states to predict utility values for all health states described by the instrument.

Modelling of health state valuations is almost always exploratory.[40] Regression techniques will be used to fit a number of alternative additive models onto the individual-level data. Consistent with extensive previous experience, the TTO values will serve as the dependent variable, whereas the independent variables will be derived from the instrument's descriptive system which we assume is ordinal in nature.[44–46] The independent variables will represent the items and response levels of the instrument's descriptive system.

The following criteria will be applied to select the preferred model: face validity and goodness-of-fit measures. Face validity implies logical consistency across generated utility values: a clinically 'better' state should have a

higher utility index.[46] For goodness-of-fit, mean absolute error (MAE) and mean-squared error (MSE) will be used in a leave-a-state-out cross validation approach.[46] Smaller MAE/MSEs are indicative of the best model fit. The algorithm from the preferred model will be used to obtain all utility values for the instrument.

Author affiliations
[1]Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada
[2]Department of Ophthalmology and Vision Sciences, University of Toronto, Toronto, Ontario, Canada
[3]School of Nursing, McMaster University, Hamilton, Ontario, Canada
[4]Trillium Health Partners, Mississauga, Ontario, Canada
[5]Department of Clinical Epidemiology and Biostatistics, McMaster University, Research Institute of St Joseph's Hamilton, and Program for Health Economics and Outcome Measures, Hamilton, Ontario, Canada

REFERENCES

1. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol* 2006;90:262–7.
2. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. *JAMA* 2014;311:1901–11.
3. Buys Y. Glaucoma in Canada: Challenges and Changing Paradigms. *Glaucoma Today* 2013. http://glaucomatoday.com/pdfs/gt0113_F_Buys.pdf (accessed 6 Jan 2016).
4. Saheb H, Ahmed II. Micro-invasive glaucoma surgery: current perspectives and future directions. *Curr Opin Ophthalmol* 2012;23:96–104.
5. Drummond M, Sculpher M, Torrance G, *et al. Methods for the economic evaluation of healthcare programmes*. 3rd edn. New York, USA: Oxford University Press, 2005.
6. Statistics Canada, Health Status. http://www.statcan.gc.ca/pub/82-221-x/2013001/def/def1-eng.htm (accessed 18 Mar 2016).
7. Bozzani FM, Alavi Y, Jofre-Bonet M, *et al.* A comparison of the sensitivity of EQ-5D, SF-6D and TTO utility values to changes in vision and perceived visual function in patients with primary open-angle glaucoma. *BMC Ophthalmol* 2012;12:43.
8. Gupta V, Srinivasan G, Mei SS, *et al.* Utility values among glaucoma patients: an impact on the quality of life. *Br J Ophthalmol* 2005;89:1241–4.
9. Paletta Guedes RA, Paletta Guedes VM, Freitas SM, *et al.* Utility values for glaucoma in Brazil and their correlation with visual function. *Clin Ophthalmol* 2014;8:529–35.
10. Browne C, Brazier J, Carlton J, *et al.* Estimating quality-adjusted life years from patient-reported visual functioning. *Eye (Lond)* 2012;26:1295–301.
11. Mangione CM, Lee PP, Gutierrez PR, *et al.* Development of the 25-item National Eye Institute Visual Function Questionnaire. *Arch Ophthalmol* 2001;119:1050–8.
12. Nelson P, Aspinall P, Papasouliotis O, *et al.* Quality of life in glaucoma and its relationship with visual function. *J Glaucoma* 2003;12:139–50.
13. Brazier J, Deverill M, Green C. A review of the use of health status measures in economic evaluation. *J Health Serv Res Policy* 1999;4:174–84.
14. Dakin H. Review of studies mapping from quality of life or clinical measures to EQ-5D: an online database. *Health Qual Life Outcomes* 2013;11:151.
15. Kay S, Ferreira A. Mapping the 25-item National Eye Institute Visual Functioning Questionnaire (NEI VFQ-25) to EQ-5D utility scores. *Ophthalmic Epidemiol* 2014;21:66–78.
16. Alavi Y, Jofre-Bonet M, Bunce C, *et al.* Developing an algorithm to convert routine measures of vision into utility values for glaucoma. *Ophthalmic Epidemiol* 2011;18:233–43.
17. Severn P, Fraser S, Finch T, *et al.* Which quality of life score is best for glaucoma patients and why? *BMC Ophthalmol* 2008;8:2.
18. Misajon R, Hawthorne G, Richardson J, *et al.* Vision and quality of life: the development of a utility measure. *Invest Ophthalmol Vis Sci* 2005;46:4007–15.
19. Richardson J, Iezzi A, Peacock S, *et al.* Utility weights for the vision-related Assessment of Quality of Life (AQoL)-7D instrument. *Ophthalmic Epidemiol* 2012;19:172–82.
20. Burr JM, Kilonzo M, Vale L, *et al.* Developing a preference-based Glaucoma Utility Index using a discrete choice experiment. *Optom Vis Sci* 2007;84:797–808.
21. Brazier JE, Rowen D, Mavranezouli I, *et al.* Developing and testing methods for deriving preference-based measures of health from condition-specific measures (and other patient-based measures of outcome). *Health Technol Assess* 2012;16:1–114.
22. Creswell J, Klassen A, Clark V. *Best practices for mixed methods research in the health sciences*. National Institute of Health, Office of Behavioral and Social Science Research, 2011.
23. Patrick DL, Burke LB, Gwaltney CJ, *et al.* 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* 2011;14:967–77.
24. Stevens K, Palfreyman S. The use of qualitative methods in developing the descriptive systems of preference-based measures of health-related quality of life for use in economic evaluation. *Value Health* 2012;15:991–8.
25. Neergaard MA, Olesen F, Andersen RS, *et al.* Qualitative description—the poor cousin of health research? *BMC Med Res Methodol* 2009;9:52.
26. Committee COSGCPGE. Canadian Ophthalmological Society evidence-based clinical practice guidelines for the management of glaucoma in the adult eye. *Can J Ophthalmol* 2009;44:S7–54.
27. Patton M. *Purposeful sampling. Qualitative evaluation and research methods*. Beverly Hills, CA: Sage, 1990:169–86.
28. Mason M. *Sample size and saturation in PhD studies using qualitative*. Vol 11. Forum: Qualitative Social Research, 2010. http://www.qualitative-research.net/index.php/fqs/article/view/1428 (accessed 6 Jan 2016).
29. Gale NK, Heath G, Cameron E, *et al.* Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol* 2013;13:117.
30. Ritchie J, Spencer L, O'Connor W. *Carrying out qualitative analysis. Qualitative research practice: a guide for social science students and researchers*. London, UK: Sage, 2003:219–62.
31. Spilker B. (Ed.) Quality of Life and Pharmacoeconomics in Clinical Trials. Lippincott – Raven Publishers, Philadelphia, 1996.
32. Sakane Y, Yamaguchi M, Yokoi N, *et al.* Development and validation of the Dry Eye-Related Quality-of-Life Score questionnaire. *JAMA Ophthalmol* 2013;131:1331–8.
33. Torrance GW, Keresteci MA, Casey RW, *et al.* Development and initial validation of a new preference-based disease-specific

health-related quality of life instrument for erectile function. *Qual Life Res* 2004;13:349–59.

34. Alrubaiy L, Hutchings HA, Williams JG. Protocol for a prospective multicentre cohort study to develop and validate two new outcome measures for patients with inflammatory bowel disease. *BMJ Open* 2013;3:pii: e003192.

35. Fayers P, Machin D. *Quality of life. Assessment, analysis and interpretation*. Chichester, UK: John Wiley & Sons, , 2016.

36. Walters S. *Quality of life outcomes in clinical trials and healthcare evaluation: a practical guide to analysis and interpretation*. Chichester, UK: John Wiley & Sons, , 2009.

37. Bechetoille A, Arnould B, Bron A, *et al*. Measurement of health-related quality of life with glaucoma: validation of the Glau-QoL 36-item questionnaire. *Acta Ophthalmol* 2008;86:71–80.

38. Brooks R, Rabin R, de Charro F. *The measurement and valuation of health status using EQ-5D: a European perspective. Evidence from the EuroQol BIOMED Research Programme*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2003.

39. Viney R, Norman R, King MT, *et al*. Time trade-off derived EQ-5D weights for Australia. *Value Health* 2011;14:928–36.

40. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;21:271–92.

41. Chuang LH, Kind P. The effect of health state selection on the valuation of EQ-5D. *Med Decis Making* 2011;31:186–94.

42. Dolan P, Gudex C, Kind P, *et al*. The time trade-off method: results from a general population study. *Health Econ* 1996;5: 141–54.

43. Dolan P, Gudex C, Kind P, *et al*. Valuing health states: a comparison of methods. *J Health Econ* 1996;15:209–31.

44. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;35:1095–108.

45. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care* 2005;43:203–20.

46. Xie F, Pullenayegum E, Gaebel K, *et al*. A time trade-off-derived value set of the EQ-5D-5L for Canada. *Med Care* 2016;54:98–105.