Title:

Are patient self-reported outcome measures (PROMs) sensitive enough to be used as endpoints in clinical trials? Evidence from the United Kingdom Glaucoma Treatment Study

Authors:

Lee Jones[1], MRes, David F. Garway-Heath[2], MD, FRCOphth, Augusto Azuara-Blanco[3], PhD, David P. Crabb[1], PhD, on behalf of the United Kingdom Glaucoma Treatment Study Investigators

Corresponding Author:

Professor David Crabb, Division of Optometry and Visual Science, School of Health Sciences, City, University of London, Northampton Square, London EC1V 0HB. Email: david.crabb.1@city.ac.uk. Phone: +44 207 040 0191

Affiliations:

1. Division of Optometry and Visual Science, School of Health Sciences, City, University of London, London, UK.
2. National Institute for Health Research Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK.
3. School of Medicine, Dentistry and Biomedical Sciences, Institute for Health Sciences, Centre for Public Health, Queen's University Belfast, Belfast, UK.

Meeting presentation:

The Association for Research and Vision in Ophthalmology, Seattle WA, 2016.

United Kingdom and Eire Glaucoma Society, Cheltenham UK, 2016.

Financial support:

DFG-H has received unrestricted research funding from Allergan, Pfizer, the Alcon Research Institute and Santen, equipment loans from Carl Zeiss Meditec, Heidelberg Engineering OptoVue and Topcon; has had advisory roles (compensated) with Aerie, Alcon, Alimera, Allergan, Bausch & Lomb, CenterVue, Quethera, Quark, Santhen and Santhera; and honoraria for lecturing from Allergan Pfizer and Santen. DPC has received unrestricted research funding from Allergan and Santen, and honoraria for lecturing from Allergan and Santen. AA-B declares no competing interests. LJ declares no competing interests.

The sponsor or funding organizations had no role in the design or conduct of this research.

Conflict of interest:

No conflicting relationship exists for any author.

Running head:

Patient reported outcome measures in glaucoma clinical trials

Keywords:
glaucoma; PROMs; visual fields; clinical trials

Abbreviations:

UKGTS = United Kingdom Glaucoma Treatment Study; **PROM** = Patient Reported Outcome Measure; **GPA** = Guided Progression Analysis; **CI** = Confidence interval; **EQ-5D** = European Quality of Life in 5 dimensions; **SF-36** = Short Form-36; **GQL-15** = Glaucoma Quality of Life-15; **GAL-9** = Glaucoma Activity Limitation-9; **VAS** = Visual Analogue Scale; **AFREV** = Assessment of Function Related to Vision; **IOP** = Intraocular Pressure; **MD** = Mean Deviation; **dB** = Decibels; **VA** = Visual Acuity.

Purpose: The UK Glaucoma Treatment Study (UKGTS) demonstrated the effectiveness of an intraocular pressure-lowering drug in patients with glaucoma using visual field progression as a primary outcome. We now test the hypothesis that responses on patient reported outcome measures (PROMs – secondary outcome measure) differ between patients receiving a topical prostaglandin analogue (Latanoprost) or placebo eye drops in UKGTS.

Design: Multi-centre, randomised, triple-masked, placebo-controlled trial.

Participants: Newly diagnosed glaucoma patients recruited into the UKGTS with baseline and exit PROM data (n= 182 and n=168 patients from the treatment and placebo group, respectively).

Methods: The UKGTS was a multi-centre, randomised, triple-masked, placebo-controlled trial, where patients with newly diagnosed open angle glaucoma were allocated to receive Latanoprost (treatment) or placebo (trial registration number: ISRCTN96423140); the observation period was 24-months. Patients completed general health PROMs (EQ-5D and SF-36) and PROMs specific to glaucoma (GQL-15 and GAL-9) at baseline and at exit from the trial. Percentage change between baseline and exit measurement on PROMs were calculated for each patient and compared between treatment arms. In addition, differences between stable patients (n=272) and those with glaucomatous progression (n=78), as determined by visual field change (primary outcome), were assessed.

Main Outcome Measure: PROMs on health-related and vision-related quality of life.

Results: Average percentage change on PROMs was similar for patients in both arms of the trial with no statistically significant differences between treatment and placebo

groups (EQ-5D, p = 0.98; EQ-5D VAS, p = 0.88; SF-36, p = 0.94, GQL-15, p = 0.66; GAL-9, p = 0.87). There were statistically significant differences between stable and progressing patients, as determined by visual fields, on glaucoma-specific PROMs (GQL-15, p = 0.02; GAL-9, p = 0.02) but not on general health PROMs (EQ-5D, p = 0.62; EQ-5D VAS, p = 0.23; SF-36, p = 0.65)

Conclusions: Average change in PROMs on health-related and vision-related quality of life was similar for the treatment and placebo group in the UKGTS. PROMs, specifically those used in the UKGTS, may not be sensitive enough to be used as a primary endpoint in clinical trials when participants have newly diagnosed early stage glaucoma.

Intraocular pressure (IOP) is currently the only modifiable risk factor for disease progression in glaucoma. All therapies approved for the treatment of glaucoma are licenced on their ability to reduce patients' IOP. Yet, the foremost outcome when treating glaucoma is to maintain what is most important to the patient, vision-related quality of life. [1] Randomised clinical trials have provided evidence for the visual field preserving benefit of reducing IOP. [2-12] Recently, the United Kingdom Glaucoma Treatment Study (UKGTS) evidenced the effectiveness of an IOP lowering treatment in patients with glaucoma using visual field deterioration determined by standard automated perimetry as the primary outcome measure over a two-year follow-up period. [12]

Typically, outcome measures in clinical trials are selected on their sensitivity to clinically meaningful changes in disease severity. However, diagnostic test measurements taken in the clinic do not directly capture the impact of glaucoma on the patient's life. [13] IOP is not a direct measure of glaucomatous optic neuropathy. Visual fields, however, indicate functional ability, and are therefore more closely associated with vision-related quality of life than IOP. Patient reported outcome measures (PROMs) are instruments derived from standardised, validated questionnaires that are used to measure perceived health status, functional status, or health-related quality of life. Asking a patient directly is an effective way to ascertain how someone feels about their condition and how it might be affecting their well-being. [14] PROMs can also be readily translated into measures of cost-effectiveness.

Use of PROMs in clinical research has increased in recent years, [15] and this is beginning to be mirrored in glaucoma research, [16] where a catalogue of vision-specific PROMs are now available. [17] PROMs are also becoming more frequently used in clinical trials, [18] including in ophthalmology trials, [19- 23]. Typically, PROMs are used to complement a more clinical primary outcome in trials. However, The United States Food

and Drug Administration endorses the use of PROMs as primary endpoints in glaucoma trials, [24] and this has been implemented in recent glaucoma trials. [25-27] An important attribute of a clinical trial outcome measure is to be sensitive enough to detect differences between a treatment and a control group. This is particularly true for glaucoma treatment trials because the disease process is slow and changes to vision can be challenging to measure. Moreover, disease progression in glaucoma is often unnoticeable to the patient in the early stages of disease. [28] A lack of sensitivity may necessitate prolonged trial duration which can add to the delay of drug development. For this reason, the sensitivity of PROMs when used as outcome measures in glaucoma trials should be scrutinised and this is the subject of our study. Specifically, we analyse PROM responses from patients in the UKGTS to test the hypothesis that these measures can determine differences between the groups randomised to treatment or placebo.

Methods

In this study, we analyse the responses on PROMs of patients enrolled into the UKGTS, a multi-centre, randomised, triple-masked, placebo-controlled trial assessing visual function preservation in newly diagnosed open-angle glaucoma patients (trial registration number: ISRCTN96423140). Patients recruited from ten eye clinics throughout the United Kingdom were randomly allocated to receive an IOP reducing prostaglandin analogue Latanoprost (0.005%) or placebo eye drops. The UKGTS, and the subsequent analysis of anonymised data in this study, adhered to the tenets of the Declaration of Helsinki and was approved by local institutional review boards (ethics approval reference: 09/H0721/56). Study participants provided written informed consent.

A total of 461 patients from 516 enrolled were analysed in the trial (Latanoprost N = 231, placebo N = 230). Patients in the UKGTS were scheduled to perform a series of 11 visual field examinations during a 2-year observation period. Visual field progression was used as the primary endpoint in the trial. Progression analysis was performed in the Humphrey Field Analyser Guided Progression Analysis (GPA) software; a sensitive technique that considers changes at individual points (test locations) in the visual field. Progression was defined as at least three visual field locations worse than baseline at the 5% levels in two consecutive reliable visual fields and at least three visual field locations worse than baseline at the 5% levels in the two subsequent consecutive reliable visual fields; the locations identified in the first and second pair were not required to be identical. Details of the trial design and the trial outcome are published elsewhere. [12; 29] In short, the risk of visual field progression was significantly lower in the treatment group than in the placebo group (adjusted hazard ratio 0.44 [95% confidence interval (CI) 0.28-0.69]).

PROMs were included as secondary outcome measures in UKGTS. PROMs were self-reported at patients' baseline and final visit and were administered by a trial researcher. In the event of a patient meeting the primary trial endpoint, PROMs were completed upon the patients' withdrawal from the trial. The PROMs used in UKGTS were as follows:

European Quality of Life in 5 dimensions (EQ-5D) is a classification of general health status. [30] EQ-5D assesses five attributes: mobility, self-care, usual activity, pain/discomfort, and anxiety/depression. We used the three-level measure meaning each dimension has three possible outcomes: no problems, some problems, and severe problems. Patients with no problems across all five attributes will produce a five-digit health status code of 11111. Patients with severe problems will score 33333. Five-digit codes were translated into a single health state score using an existing scoring system which is generated from a UK population sample. [30] Included in the EQ-5D is a visual analogue scale (**EQ-5D VAS**) where patients are asked to score their own health between 0 and 100 (where 0 and 100 are worst and the best imaginable health). EQ-5D is the most commonly used general health PROM and is recommended in The National Institute for Health and Care Excellence guidelines for health economic analysis in the United Kingdom. [31] Furthermore, following recommendations by the United States Public Health Service, [32] there now exists a large database of EQ-5D derived health statistics for the American population, too. [33]

Short Form-36 (SF-36) is another general health instrument featuring 36 items across eight domains relating to: physical functioning, role limitation due to physical problems, emotional problems, bodily pain, general health, social functioning, vitality, and mental health. [34] Responses are made on Likert-type scales and the 36 individual items can be translated to give a global score for general health (ranging 0-100) where

lower scores reflect poorer self-reported health. Following the International Quality of Life Assessment Project translation of SF-36 into several languages, [35] this PROM has become frequently used in cost-utility studies. [36]

Glaucoma Quality of Life (GQL-15) instrument has 15-items and is disease specific being designed to assess the impact of glaucoma on vision-related quality of life. [37] The GQL-15 was derived from an initial 62-item pilot questionnaire; the 15-items were included in the final instrument due to their strong relationship with visual field loss in glaucoma patients. [38] GQL-15 has four subscales: central and near vision, peripheral vision, mobility, and glare/dark adaptation. Scoring is based on five-point Likert-type scales where a response of 5 denotes severe difficulty and 1 indicates no difficulty. The measurement scale ranges from 15 to 75 where higher scores represent poorer vision-related quality of life. The instrument has been used in well-designed cross-sectional studies assessing the impact of glaucoma on patients' quality of life. [39, 40]

GQL-15 has previously been subjected to Rasch analysis to produce the 9-item **Glaucoma Activity Limitation (GAL-9)** PROM. [41] This instrument consists of a subset of nine items from the original GQL-15 and is considered to better reflect the effects of glaucoma on visual function. [41] GAL-9 has good external validity as scores from the instrument have been shown to correlate well with visual acuity and visual field scores. Furthermore, the GAL-9 is quicker to complete than the GQL-15 because it has fewer items. [41] In addition to our analysis of GQL-15 responses, we repeat the analysis on the items included in the GAL-9 for patients in the UKGTS.

For the data analysis, responses on the PROMs at baseline and exit were transposed into percentage scores. (The exit visit was at 24-months or, for progressing patients, at the visit when progression was confirmed). Differences between these scores

were used to detect the degree of change in each PROM between first and last trial visit. For example, no change is indicated by zero and scores greater than 0% indicate worsening on PROMs, i.e. patients report more problems on exit from the trial than at baseline; negative values indicate improvement from baseline. Two-sample independent t-tests were used to determine whether there was a statistically significant difference in change on PROMs between the two trial groups (treatment and placebo).

Additionally, we assessed whether statistically significant differences in PROM responses could be observed between patients who remained stable during the UKGTS and those who experienced the primary trial endpoint. We included this additional analysis as it was anticipated that the largest difference in score for health-related and vision-related quality of life would be observed between these two patient groups.

Results

Complete baseline and exit PROM data were available for n=182 (79%) and n=168 (73%) of patients with follow-up data in the treatment and placebo arm of the trial, respectively. Average change in scores was similar for both the treatment and placebo groups across all the PROMs (Table 1). There were no statistically significant differences between the trial groups on PROMs relating to general health. Furthermore, there remained no statistically significant differences between the two groups on the glaucoma-specific PROMs. In addition, the distribution in the baseline to exit scores were strikingly similar between the treatment and placebo groups (Figure 1).

PROM data were not available at the exit visit for a proportion of patients in the UKGTS. Further analysis of those with missing data indicates that these patients had a similar profile to those with complete data (Table 2). Specifically, as determined through two-sample t-tests, there were no statistically significant differences between these two groups on baseline better eye mean deviation (MD) ($p = 0.12$), worse eye MD ($p = 0.90$),

better eye visual acuity (p = 0.44), worse eye visual acuity (p = 0.56), and age (p = 0.27). As a group, patients without exit PROMs reported slightly worse average general and vision-related quality of life at baseline compared to those with exit PROMs. However, the magnitude of these differences was small; it might reflect some patients without exit PROMs being more likely to be people who were unwell at the start of the trial. For example, 32 patients had less than 21-months follow-up in the trial because of ill health and seven patients died during follow-up [12].

We assessed differences between stable patients (N=272) and patients with glaucomatous progression (N=78) as determined by the primary visual field outcome. Median (interquartile range) duration between baseline and progression confirmation visit was 465 (278, 553) days, in comparison to the 2-year (730 days) scheduled follow-up for patients remaining stable. No statistically significant differences were found between average responses from stable and progressed patients on PROMs relating to general health (EQ-5D, EQ-5D VAS and SF-36). Average differences between stable and progressed patients were statistically significant when assessing responses on glaucoma-specific PROMs (GQL-15 and GAL-9) (Table 3 and Figure 2). As a group, patients who had progressed on visual fields therefore reported a reduction in glaucoma-specific vision-related quality of life that was different to those who had remained stable on visual fields. Mean (95% CI) scores for the progression patients on the GAL-9 and GQL-15 was 6.5 (2.8–9.2) % and 3.9 (3.2–9.8) % respectively.

239

240 **Table 1.** Means (standard deviation) of percentage (%) change scores for the two trial
241 groups (treatment and placebo) on PROMs between baseline and trial exit in the UKGTS.
242 Mean (standard deviation) change in worse-eye mean deviation between baseline and
243 trial exit in the UKGTS. More negative MD indicates improved scores from baseline.

Table 1. Means (standard deviation) of percentage (%) change scores for the two trial groups (treatment and placebo) on PROMs between baseline and trial exit in the UKGTS. Mean [95% confidence interval] difference between the two samples. Mean (standard deviation) change in worse-eye mean deviation between baseline and trial exit in the UKGTS. More negative MD change indicates improved scores from baseline.

	Group			
PROM	*Treatment N = 182*	*Placebo N = 168*	Mean Difference [CI]	p-value
EQ-5D	1.7 (15.4)%	1.7 (10.6)%	0.0% [-2.8 to 2.8%]	0.98
EQ-5D VAS	2.1 (12.5)%	1.9 (12.0)%	0.2% [-2.8 to 2.4%]	0.88
SF-36	4.8 (19.8)%	5.0 (22.5)%	0.2% [-4.2 to 4.6%]	0.94
GQL-15	2.7 (7.7)%	3.2 (11.7)%	0.5% [-1.5 to 2.6%]	0.66
GAL-9	3.0 (8.5)%	3.2 (12.8)%	0.2% [-2.1 to 2.5%]	0.87
MD	-0.23 (1.9) dB	0.14 (2.0) dB		0.07

Change from baseline to exit is shown as a percentage (%). Percentages show the average amount of change on each PROM for treatment and placebo group. Positive percentages indicate worsening from baseline.
PROM = Patient reported outcome measure. CI = Confidence interval. EQ-5D = European quality of life in 5 dimensions. VAS = Visual analogue scale. SF-36 = Short from 36. GQL-15 = Glaucoma quality of life. GAL = Glaucoma activity limitation. MD = Mean deviation change in worse-eye. dB = Decibels.

244

245

246

247

248 **Figure 1.** Boxplots on the left show change in scores between baseline and exit PROMs
249 for patients in the placebo group (blue) and the treatment group (green) in the UKGTS.
250 Positive scores (higher than 0) indicate worsening from baseline. Boxplots on the right
251 show change in progressing/worse eye MD score between baseline and exit VFs for
252 placebo and treatment groups. (MD is a summary measure used to represent overall
253 reduction in visual field sensitivity relative to healthy aged-matched observers. Lower
254 MD values (more negative) are indicative of greater loss of vision). Boxplots give median,
255 interquartile range, 5^{th} and 95^{th} percentiles (whiskers). Due to large variability in
256 responses, 95^{th} percentile is capped at 40% change for SF-36 analysis (SF-36 placebo 95^{th}
257 percentile = 54.6%; SF-36 treatment 95^{th} percentile = 42.2%).
258
259
260
261
262
263
264
265
266
267
268

269 **Table 2.** Comparison of baseline characteristics between patients in the UKGTS with

270 PROM data (N=350) and those without PROM data at exit (N=166).

Table 2. Comparison of baseline characteristics between patients in the UKGTS with PROM data (N=350) and those without PROM data at exit (N=166).

	UKGTS patients with PROMs *N = 350*	UKGTS patients without PROMs *N = 166*		p-value
MD (dB)				
Better eye				
Mean	-0.5 (1.2)	-0.8 (1.8)		0.12
Median	-0.5 [-1.3, 0.4]	-0.6 [-1.4, 0.3]		
Worse eye				
Mean	-4.2 (3.3)	-4.3 (3.6)		0.90
Median	-3.3 [-5.6, -2.0]	-3.4 [-5.7, -1.7]		
Best-corrected VA				
Better eye				
Mean	1.0 (0.21)	1.0 (0.24)		0.44
Median	1.0 [1.0, 1.2]	1.0 [1.0, 1.2]		
Worse eye				
Mean	0.9 (0.24)	0.9 (0.25)		0.56
Median	1.0 [0.67, 1.0]	1.0 [0.67, 1.0]		
Age (years)				
Mean	65.8 (9.9)	67.4 (11.9)		0.27
Sex				
Male	188 (53.7%)	85 (51.2%)		
Female	162 (46.3%)	81 (48.8%)		
Baseline PROM			Mean difference [CI]	
Mean				
EQ-5D	5 (7.2) %	5 (6.5) %	0 [0 to 3%]	0.53
EQ-5D VAS	81 (15.1) %	75 (18.7) %	6 [2 to 13%]	0.03
SF-36	77 (17.2) %	70 (19.9) %	7 [3 to 14%]	0.002
GQL-15	7 (8.9) %	11 (12.7) %	4 [1 to 10%]	0.003
GAL-9	7 (9.9) %	11 (14.7) %	4 [1 to 10%]	0.01

Data are n (%) or mean (standard deviation) or median [interquartile range]. PROM = Patient reported outcome measure. MD = Mean deviation. dB = Decibels. VA = Visual acuity (decimal). CI = Confidence interval.

271

272 **Table 3.** Means (standard deviation) of percentage (%) change scores for stable and
273 progressed patients on PROMs between baseline and trial exit in the UKGTS. Mean
274 (standard deviation) change in worse-eye mean deviation between baseline and trial exit
275 in the UKGTS. More negative MD indicates improved scores from baseline.

Table 3. Means (standard deviation) of percentage (%) change scores for stable and progressed patients on PROMs between baseline and trial exit in the UKGTS. Mean [95% confidence interval] difference between the two samples. Mean (standard deviation) change in worse-eye mean deviation between baseline and trial exit in the UKGTS. More negative MD indicates improved scores from baseline.

	Outcome			
PROM	*Stable* *N = 272*	*Progressed* *N = 78*	Mean Difference [CI]	p-value
EQ-5D	1.5 (13.5)%	2.4 (12.5)%	0.9% [-2.5 to 4.3]	0.62
EQ-5D VAS	1.5 (11.8)%	3.6 (13.5)%	2.1% [-1.0 to 5.2]	0.23
SF-36	4.6 (20.3)%	6.0 (23.6)%	1.4% [-3.9 to 6.7]	0.65
GQL-15	2.1 (7.9)%	6.0 (14.3)%	3.9% [1.5 to 6.3]	0.02*
GAL-9	2.1 (9.1)%	6.5 (14.8)%	4.4% [1.7 to 7.1]	0.02*
MD	-0.22 (1.9) dB	0.55 (2.1) dB		0.003*

Change from baseline to exit is shown as a percentage (%). Percentages show the average amount of change on each PROM for stable and progressed trial outcomes. Positive percentages indicate worsening from baseline.
PROM = Patient reported outcome measure. CI = Confidence interval. EQ-5D = European quality of life in 5 dimensions. VAS = Visual analogue scale. SF-36 = Short from 36. GQL-15 = Glaucoma quality of life. GAL = Glaucoma activity limitation. MD = Mean deviation of worse-eye. dB = Decibels.

* = significant at 0.05 level

276

277 **Figure 2.** Boxplots on the left show change in scores between baseline and exit PROMs
278 for patients remaining stable (purple) and patients with visual field progression (red) in
279 the UKGTS. Positive scores (higher than 0) indicate worsening from baseline. Boxplots on
280 the right show change in progressing/worse eye MD score between baseline and exit VFs
281 for stable and progression groups. Boxplots give median, interquartile range, 5th and 95th

282 percentiles (whiskers). Due to large variability in responses, 95th percentile is capped at
283 40% change for SF-36 analysis (SF-36 stable 95th percentile = 42.4%; SF-36 progression
284 95th percentile = 53.8%).

285

Discussion

Results from this study show average changes in scores on general health-related PROMs (EQ-5D, EQ-5D VAS and SF-36) to be similar for patients receiving either Latanoprost or placebo eye drops in the UKGTS. Moreover, we did not find any evidence for differences between the two arms of the trial when analysing changes in PROMs specifically relating to vision and glaucoma (GQL-15 and GAL-9). Therefore, PROMs used in the UKGTS measured once at baseline and at 2-year follow-up (or final review, for those exiting early as a consequence of visual field progression) are not as sensitive as serial visual fields, taken over the same time course, in determining treatment differences in disease progression in a trial for glaucoma treatment.

There were other interesting findings from our study. Statistically significant differences were observed in average responses between stable and progressed patients on glaucoma-specific PROMs, but this was not the case for general health-related PROMs. This suggests general health-related PROMs are insensitive to treatment-induced changes in glaucoma progression, certainly in the population of patients represented in the UKGTS within the 24-month observation period. Another finding, not directly related to the aim of our study, concerns differences between GAL-9 and GQL-15. When comparing stable and progressing patients, GAL-9 yielded a marginally larger average effect (4.4%) when compared to the GQL-15 (3.9%). As such, we provide supporting evidence that the GAL-9 may be a satisfactory alternative to the GQL-15 when assessing glaucoma-specific vision-related quality of life. The GAL-9 has the added benefit of having fewer items and is therefore less burdensome for the patient to complete.

Our results have implications for trial design for glaucoma treatments. The UKGTS highlighted that a relatively short observation period could be implemented when

adopting a sensitive change-from-baseline event criterion to identify visual field progression. This was made possible by frequent visual field testing and sensitive statistical methods where measurements that were repeatedly worse than baseline were flagged. Our results suggest that PROMs may not be sensitive enough to be used as outcome measures in glaucoma treatment trials, especially over a relatively short follow-up. Yet, it is important to note in the UKGTS, patients only completed PROMs at baseline and exit visits. The difference in mean deviation (a global measure, in the same sense as a questionnaire score) of the visual fields taken at baseline and final review was also not sufficiently sensitive to identify differences between the treatment and placebo groups. Therefore, the explanation of the inability of the PROM scores to identify treatment differences is that either the PROM scores are insufficiently responsive to the small changes in disease observed over the short trial duration or that the scores are insufficiently precise, or both. Indeed, PROMs administered more frequently during the trial may have reduced the within person variability in responses and increase the likelihood of capturing significant changes. We are aware of at least two ongoing glaucoma trials that are doing this, albeit in different PROMS to the ones used in UKGTS. [26-27] Still, the relatively small effects and large variability in our PROM data indicate that even repeat measures may not provide adequate trial power. It is encouraging that our chosen primary end point for the UKGTS, namely visual field progression, was sensitive enough to detect changes that are likely imperceptible to most patients in the early stage of the disease. Longitudinal studies have revealed an association between visual field progression and changes in vision-related quality of life in glaucoma patients [42-45]. Yet, these studies have tended to use global or regional measures of visual field derived from binocular measures. We are unaware of any longitudinal studies reporting changes in quality of life measures that are associated with progression events detected at a visual

field test location level using GPA software. Ultimately, it makes sense that trial endpoints are aligned to relevant and meaningful outcomes for the patient, and we have highlighted that disease-specific instruments, like GAL-9 and GQL-15, can track visual field loss amongst glaucoma patients. Moreover, it remains important that all stakeholders are considered when deciding on outcome measures in clinical trials, and that includes the patients themselves. [46]

Other observations on our results are noteworthy. Average changes in PROMs, where they existed, were small and the variability in response between participants was large. For example, the average 6% decline on the GQL-15 in the N=78 patients who were progressing on visual fields is equivalent of a change from 'no difficulty' to 'a little bit of difficulty' on just four of the 15 items on the GQL-15. This small average change in vison-related quality of life suggests that patients experiencing the visual field endpoint do not perceive large changes in visual function, in this cohort with glaucoma mostly at its earliest stage. This is an interesting finding because it has been suggested that placebo-controlled clinical trials for glaucoma treatment can be harmful for those randomised to the placebo arm. [47] However, our findings certainly indicate that vision-related and health-related quality of life was similar between patients in the placebo group to those randomised to treatment over the course of the trial. In the case of the UKGTS, all patients were monitored closely over a short trial duration and the criterion for visual field deterioration was proven to be very sensitive. On average, patients progressing, based on visual fields, experience a small or unnoticeable reduction in vision-related quality of life. They certainly do not, on average, experience a change in general health as measured by the general-health PROMs considered in our study and this is particularly noteworthy. These findings support an argument for close monitoring being an alternative to medical treatment in the early stages of the disease, an observation made from the results of

previous clinical trials. [5, 8] As no statistically significant differences in PROM scores were observed between the treatment and placebo group in UKGTS, our findings might have implications for how health-related and vision-related quality of life are assessed in clinical trials. More objective or 'real-world' assessments of visual disability are emerging, and these have potential for use as trial outcomes that are meaningful to the patient. One such measure, the Assessment of Function Related to Vision (AFREV), requires users to perform visual tasks such as findings objects, using everyday technologies, and reading under various illuminations. [48] If used as an outcome measure, tools such as the AFREV may yield more discernible differences between treatment groups in glaucoma clinical trials, but this remains speculation until tested. An added advantage of such objective measures is that, unlike PROMs, they are less reliant on the functional literacy of the patient. Offering definitive guidance on the use of PROMs or visual fields, or a combination of the two, as outcome measures for glaucoma trials is beyond the remit of this study. These issues are complicated because, for example, PROMs are derived from the individual, who has two eyes, and the visual field outcome is derived from just one eye (the first showing progression), and in the UKGTS just 11% (n = 10) of progressing patients had visual field progression in both eyes. PROM performance in glaucoma is likely driven by the least affected eye but this is dependent on the stage of glaucoma [49, 50]; in the UKGTS, almost 50% of participants had glaucoma in only one eye. Furthermore, the visual field progression outcome occurred in one eye only in almost 90% of participants with identifiable progression (94 of 461 subjects) and in 73% of these, the progression was in the worse eye. Thus, the person-level PROM outcome would be expected to be less sensitive to glaucoma deterioration than eye-based measures of visual function. For example, standard automated perimetry will detect changes in sensitivity that may be unnoticed by the patient, whereas PROMs will likely be

more responsive to central visual field loss. This does not mean that PROMs do not have a role in treatment trials; they may have a more important role in identifying adverse (or even beneficial) effects of interventions on the person that they have in identifying disease modifying effects.

The study was not without limitations. In some cases, not all patients completed PROMs at baseline or exit from the trial and so no comparable data were available for analysis. Yet, patients with and without PROM data had similar demographic and visual function profiles. One key limitation comes from patients possibly being aware of the status of their glaucoma progression (stable or worsening) at the time of completing exit PROMs. This is certainly true for patients withdrawn early from the trial because visual field progression had occurred. If, for example, a patient was told they were exiting the trial because their clinically measured vision was getting worse, then that would likely influence self-report of quality of life. If this were the case, one might expect knowledge of glaucoma progression status to affect general health-related, as well as vision-related, quality of life, but there were no differences in the EQ-5D or SF36 between those who progressed and those who did not. As previously discussed, the design of the UKGTS meant that patients completed PROMs at only two time points. This is obviously different to the frequent collection of visual field data (primary outcome). Our results are also limited to apply to only a UK population of newly diagnosed patients, most of whom were at the earliest stage of the disease. We cannot say how PROMs may change over a period of 24-months in people with more advanced disease. Patient's vision-related quality of life may decrease more quickly when visual field loss is already quite advanced. [51]

In conclusion, patients randomised to treatment or placebo in the UKGTS returned similar responses to PROMs at baseline and final visits of the trial. It is accepted that no

single PROM covers all aspects of patients' vision-related quality of life, [52] and our findings at least emphasise the importance of appropriate PROM selection when designing and implementing clinical trials. Even if PROMs cannot capture the disease modification effect of an intervention, that certainly does not mean that they are not useful if they can capture other consequences of an intervention including, for example, side effects or inconvenience of treatment regimens. In the UKGTS differences in PROM responses only emerged when comparing stable and progressed patients on instruments that were specific to glaucoma. As such, we suggest PROMs alone, administered at the start and end of a 24-month trial assessing disease progression, may not be sensitive enough to be used as the primary endpoints in glaucoma clinical trials assessing disease progression.

421 References

422 1. European Glaucoma Society Terminology and Guidelines for Glaucoma, 4th Edition - Part
423 1 Supported by the EGS Foundation. *British Journal of Ophthalmology* 2017;101:1-72.
424 2. Holmin C, Thorburn W, Krakau CET. Treatment versus no treatment in chronic open angle
425 glaucoma. *Acta Ophthalmologica.* 1988;66:170-3.
426 3. AGIS Investigators. The Advanced Glaucoma Intervention Study (AGIS): 7. The
427 relationship between control of intraocular pressure and visual field
428 deterioration. *American Journal of Ophthalmology*. 2000;130:429-40.
429 4. Kass MA, Heuer DK, Higginbotham EJ, et al.. The Ocular Hypertension Treatment Study: a
430 randomized trial determines that topical ocular hypotensive medication delays or
431 prevents the onset of primary open-angle glaucoma. *Archives of Ophthalmology*.
432 2002;120:701-13.
433 5. Heijl A, Leske MC, Bengtsson B, et al. Reduction of intraocular pressure and glaucoma
434 progression: results from the Early Manifest Glaucoma Trial. *Archives of Ophthalmology.*
435 2002;120:1268-79.
436 6. Pajic B, Pajic-Eggspuehler B, Häfliger IO. Comparison of the effects of dorzolamide/timolol
437 and latanoprost/timolol fixed combinations upon intraocular pressure and progression
438 of visual field damage in primary open-angle glaucoma. *Current Medical Research and*
439 *Opinion*. 2010;26:2213-19.
440 7. Krupin T, Liebmann JM, Greenfield DS, et al. A randomized trial of brimonidine versus
441 timolol in preserving visual function: results from the Low-Pressure Glaucoma Treatment
442 Study. *American Journal of Ophthalmology*. 2011;151:671-81.
443 8. Group, C.N.T.G.S. Comparison of glaucomatous progression between untreated patients
444 with normal-tension glaucoma and patients with therapeutically reduced intraocular
445 pressures. *American Journal of Ophthalmology*. 1998;126:487-97.

9. Migdal C, Gregory W, Hitchings R. Long-term functional outcome after early surgery compared with laser and medicine in open-angle glaucoma. *Ophthalmology.* 1994;101:1651-57.
10. Jay JL, Murray SB. Early trabeculectomy versus conventional management in primary open angle glaucoma. *British Journal of Ophthalmology*. 1988;72:881-89.
11. Musch DC, Gillespie BW, Lichter PR, et al. Visual field progression in the Collaborative Initial Glaucoma Treatment Study: the impact of treatment and other baseline factors. *Ophthalmology*. 2009;116:200-7.
12. Garway-Heath DF, Crabb DP, Bunce C, et al. Latanoprost for open-angle glaucoma (UKGTS): a randomised, multicentre, placebo-controlled trial. *The Lancet*. 2015;*385*:1295-304.
13. Denniston AK, Kyte D, Calvert M, et al. An introduction to patient-reported outcome measures in ophthalmic research. *Eye.* 2014;*28*:637-45.
14. Deshpande PR, Rajan S, Sudeepthi BL, et al. Patient-reported outcomes: a new era in clinical research. *Perspectives in Clinical Research*. 2011;*2*:137.
15. Black N. Patient reported outcome measures could help transform healthcare. *BMJ*. 2013 346, p.f167.
16. Glen FC, Crabb DP, Garway-Heath DF. The direction of research into visual disability and quality of life in glaucoma. *BMC Ophthalmology*. 2011;11:19.
17. Hamzah JC, Burr JM, Ramsay CR, et al. Choosing appropriate patient-reported outcomes instrument for glaucoma research: a systematic review of vision instruments. *Quality of Life Research*. 2011;*20*:1141-58.
18. Vodicka E, Kim K, Devine EB, et al. Inclusion of patient-reported outcome measures in registered clinical trials: evidence from ClinicalTrials.gov (2007–2013). *Contemporary Clinical Trials*. 2015;43:1-9.

19. Chakravarthy U, Harding SP, Rogers CA, et al. Alternative treatments to inhibit VEGF in age-related choroidal neovascularisation: 2-year findings of the IVAN randomised controlled trial. *The Lancet*. 2013;*382*:1258-67.
20. Varma R, Bressler NM, Suñer I, et al. Improved vision-related function after ranibizumab for macular edema after retinal vein occlusion: results from the BRAVO and CRUISE trials. *Ophthalmology*. 2012;*119*:2108-18.
21. Mitchell P, Bandello F, Schmidt-Erfurth U, et al. The RESTORE study: ranibizumab monotherapy or combined with laser versus laser monotherapy for diabetic macular edema. *Ophthalmology*. 2011;*118*:615-25.
22. Sugar EA, Holbrook JT, Kempen JH, et al. Cost-effectiveness of fluocinolone acetonide implant versus systemic therapy for noninfectious intermediate, posterior, and panuveitis. *Ophthalmology*. 2014;*121:*1855-62.
23. Lois N, Burr J, Norrie J, et al. Internal limiting membrane peeling versus no peeling for idiopathic full-thickness macular hole: a pragmatic randomized controlled trial. *Investigative Ophthalmology & Visual Science*. 2011;52:1586-92.
24. Food and Drug Administration. Guidance for industry: Patient Reported Outcome measures: use in medical product development to support labeling claims. *Federal Register.* 2006;71:1–32.
25. Azuara-Blanco A, Burr J, Ramsay C, et al. Effectiveness of early lens extraction for the treatment of primary angle-closure glaucoma (EAGLE): a randomised controlled trial. *The Lancet*. 2016;*388*:1389-97.
26. King AJ, Fernie G, Azuara-Blanco A, et al Treatment of Advanced Glaucoma Study: a multicentre randomised controlled trial comparing primary medical treatment with primary trabeculectomy for people with newly diagnosed advanced glaucoma—study protocol. *British Journal of Ophthalmology.* 2018; 102: 922-8.

27. Vickerstaff V, Ambler G, Bunce C, et al. Statistical analysis plan for the Laser-1st versus Drops-1st for Glaucoma and Ocular Hypertension Trial (LiGHT): a multi-centre randomised controlled trial. *Trials.* 2015;16:517.
28. Crabb DP. A view on glaucoma—are we seeing it clearly? *Eye.* 2016;30:304-13.
29. Garway-Heath DF, Lascaratos G, Bunce C, et al. The United Kingdom Glaucoma Treatment Study: a multicenter, randomized, placebo-controlled clinical trial: design and methodology. *Ophthalmology*. 2013;120:68-76.
30. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16:199-208.
31. Devlin N, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. *Applied Health Economics and Health Policy*. 2017;15:127-37.
32. Siegel JE, Weinstein MC, Russell LB, et al. Recommendations for reporting cost-effectiveness analyses. *JAMA.* 1996;276:1339-41.
33. Sullivan PW, Lawrence WF, Ghushchyan V. A national catalog of preference-based scores for chronic conditions in the United States. *Medical Care*. 2005;43:736-49.
34. Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care*. 1992;30:473-83.
35. Aaronson NK, Acquadro C, Alonso J, et al. International quality of life assessment (IQOLA) project. *Quality of Life Research*. 1992;1:349-51.
36. Hall T, Krahn GL, Horner-Johnson W, et al. Examining functional content in widely used Health-Related Quality of Life scales. *Rehabilitation Psychology*. 2011;56:94.
37. Nelson P, Aspinall PA, O'Brien C . Patients' perception of visual impairment in glaucoma: a pilot study. *British Journal of Ophthalmology*. 1999;83:546–52.
38. Nelson P, Aspinall P, Papasouliotis O, et al. Quality of life in glaucoma and its relationship with visual function. *Journal of Glaucoma*. 2003;12:139–50.
39. van Gestel A, Webers CA, Beckers HJM, et al. The relationship between visual field loss in glaucoma and health-related quality-of-life. *Eye.* 2010;24:1759-69.

40. Goldberg I, Clement CI, Chiang TH, et al. Assessing quality of life in patients with glaucoma using the Glaucoma Quality of Life-15 (GQL-15) questionnaire. *Journal of Glaucoma.* 2009;18:6-12.
41. Khadka J, Pesudovs K, McAlinden C, et al. Reengineering the glaucoma quality of life-15 questionnaire with rasch analysis. *Investigative Ophthalmology & Visual Science.* 2011;52:6971-7.
42. Medeiros FA, Gracitelli CP, Boer ER, et al. Longitudinal changes in quality of life and rates of progressive visual field loss in glaucoma patients. *Ophthalmology*, 2015; 122: 293-301.
43. Abe RY, Diniz-Filho A, Costa VP, et al. The Impact of Location of Progressive Visual Field Loss on Longitudinal Changes in Quality of Life of Glaucoma Patients. *Ophthalmology*, 2016; 123: 552-7.
44. Peters D, Heijl A, Brenner L, et al. Visual impairment and vision-related quality of life in the Early Manifest Glaucoma Trial after 20 years of follow-up. *Acta Ophthalmologica*, 2015; 93: 745-52.
45. Diniz-Filho A, Abe RY, Cho HJ, et al. Fast visual field progression is associated with depressive symptoms in patients with glaucoma. *Ophthalmology*, 2016; 123: 754-9.
46. Dean S, Mathers JM, Calvert M, et al. 'The patient is speaking': discovering the patient voice in ophthalmology. *British Journal of Ophthalmology.* 2017;101:700-8.
47. Wegner A. Latanoprost for glaucoma: primum non nocere. *The Lancet.* 2015;386:651-52.
48. Altangerel U, Spaeth GL, Steinmann WC. Assessment of function related to vision (AFREV). *Ophthalmic Epidemiology*. 2006;13:67-80.
49. Skalicky SE, McAlinden C, Khatib T, et al. Activity limitation in glaucoma: objective assessment by the Cambridge Glaucoma Visual Function Test. *Investigative Ophthalmology & Visual Science*. 2016; 57:6158-66.
50. Arora KS, Boland MV, Friedman DS, et al. The relationship between better-eye and integrated visual field mean deviation and visual disability. *Ophthalmology.* 2013; 120: 2476-84.

550 51. Jones L, Bryan SR, Crabb DP. Gradually Then Suddenly? Decline in vision-related quality
551 of life as glaucoma worsens. *Journal of Ophthalmology*, vol. 2017, Article ID 1621640, 7
552 pages, 2017. doi:10.1155/2017/1621640.
553 52. Somner JE, Sii F, Bourne RR, et al. Moving from PROMs to POEMs for glaucoma care: A
554 qualitative scoping exercise. *Investigative Ophthalmology & Visual Science*. 2012;53:5940-
555 47.
556