



# City Research Online

## City, University of London Institutional Repository

---

**Citation:** Jones, L., Garway-Heath, D. F., Azuara-Blanco, A., Crabb, D. P. ORCID: 0000-0001-8754-3902, Bunce, C., Lascaratos, G., Amalfitano, F., Anand, N., Bourne, R., Broadway, D. C., Cunliffe, I. A., Diamond, J. P., Fraser, S. G., Ho, T. A., Martin, K. R., McNaught, A., Negi, A., Patel, K., Russell, R. A., Shah, A., Spry, P. G., Suzuki, K., White, E.T., Wormald, R., Xing, W. and Zeyen, T. G. (2018). Are Patient Self-Reported Outcome Measures Sensitive Enough to Be Used as End Points in Clinical Trials?: Evidence from the United Kingdom Glaucoma Treatment Study. *Ophthalmology*, doi: 10.1016/j.ophtha.2018.09.034

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

---

**Permanent repository link:** <http://openaccess.city.ac.uk/id/eprint/21780/>

**Link to published version:** <http://dx.doi.org/10.1016/j.ophtha.2018.09.034>

**Copyright and reuse:** City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

---

1 **Title:**

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

4 **Authors:**

5 Lee Jones<sup>1</sup>, MRes, David F. Garway-Heath<sup>2</sup>, MD, FRCOphth, Augusto Azuara-Blanco<sup>3</sup>, PhD, David P.  
6 Crabb<sup>1</sup>, PhD, on behalf of the United Kingdom Glaucoma Treatment Study Investigators

7  
8 **Corresponding Author:**

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

12  
13 **Affiliations:**

14 1. Division of Optometry and Visual Science, School of Health Sciences, City, University of London,  
15 London, UK.

16 2. National Institute for Health Research Biomedical Research Centre at Moorfields Eye Hospital  
17 NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK.

18 3. School of Medicine, Dentistry and Biomedical Sciences, Institute for Health Sciences, Centre for  
19 Public Health, Queen's University Belfast, Belfast, UK.

20 **Meeting presentation:**

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

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

23 **Financial support:**

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

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

32 **Conflict of interest:**

33 No conflicting relationship exists for any author.

34 **Running head:**

35 Patient reported outcome measures in glaucoma clinical trials

36 **Keywords:**

37 glaucoma; PROMs; visual fields; clinical trials

38 **Abbreviations:**

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

45

46

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

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

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

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

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

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

Patient reported outcome measures in glaucoma clinical trials

70 groups (EQ-5D,  $p = 0.98$ ; EQ-5D VAS,  $p = 0.88$ ; SF-36,  $p = 0.94$ , GQL-15,  $p = 0.66$ ; GAL-9,  $p$   
71  $= 0.87$ ). There were statistically significant differences between stable and progressing  
72 patients, as determined by visual fields, on glaucoma-specific PROMs (GQL-15,  $p = 0.02$ ;  
73 GAL-9,  $p = 0.02$ ) but not on general health PROMs (EQ-5D,  $p = 0.62$ ; EQ-5D VAS,  $p = 0.23$ ;  
74 SF-36,  $p = 0.65$ )

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

79

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

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

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

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

## 117 **Methods**

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

128 A total of 461 patients from 516 enrolled were analysed in the trial (Latanoprost  
129 N = 231, placebo N = 230). Patients in the UKGTS were scheduled to perform a series of  
130 11 visual field examinations during a 2-year observation period. Visual field progression  
131 was used as the primary endpoint in the trial. Progression analysis was performed in the  
132 Humphrey Field Analyser Guided Progression Analysis (GPA) software; a sensitive  
133 technique that considers changes at individual points (test locations) in the visual field.  
134 Progression was defined as at least three visual field locations worse than baseline at the  
135 5% levels in two consecutive reliable visual fields and at least three visual field locations  
136 worse than baseline at the 5% levels in the two subsequent consecutive reliable visual  
137 fields; the locations identified in the first and second pair were not required to be  
138 identical. Details of the trial design and the trial outcome are published elsewhere. (12; 29)  
139 In short, the risk of visual field progression was significantly lower in the treatment group  
140 than in the placebo group (adjusted hazard ratio 0.44 [95% confidence interval (CI) 0.28-  
141 0.69]).



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

147 **European Quality of Life in 5 dimensions (EQ-5D)** is a classification of general  
148 health status. <sup>(30)</sup> EQ-5D assesses five attributes: mobility, self-care, usual activity,  
149 pain/discomfort, and anxiety/depression. We used the three-level measure meaning  
150 each dimension has three possible outcomes: no problems, some problems, and severe  
151 problems. Patients with no problems across all five attributes will produce a five-digit  
152 health status code of 11111. Patients with severe problems will score 33333. Five-digit  
153 codes were translated into a single health state score using an existing scoring system  
154 which is generated from a UK population sample. <sup>(30)</sup> Included in the EQ-5D is a visual  
155 analogue scale (**EQ-5D VAS**) where patients are asked to score their own health between  
156 0 and 100 (where 0 and 100 are worst and the best imaginable health). EQ-5D is the most  
157 commonly used general health PROM and is recommended in The National Institute for  
158 Health and Care Excellence guidelines for health economic analysis in the United  
159 Kingdom. <sup>(31)</sup> Furthermore, following recommendations by the United States Public  
160 Health Service, <sup>(32)</sup> there now exists a large database of EQ-5D derived health statistics  
161 for the American population, too. <sup>(33)</sup>

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

167 lower scores reflect poorer self-reported health. Following the International Quality of  
168 Life Assessment Project translation of SF-36 into several languages, <sup>(35)</sup> this PROM has  
169 become frequently used in cost-utility studies. <sup>(36)</sup>

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

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

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

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

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

## 202 **Results**

203 Complete baseline and exit PROM data were available for n=182 (79%) and n=168  
204 (73%) of patients with follow-up data in the treatment and placebo arm of the trial,  
205 respectively. Average change in scores was similar for both the treatment and placebo  
206 groups across all the PROMs (Table 1). There were no statistically significant differences  
207 between the trial groups on PROMs relating to general health. Furthermore, there  
208 remained no statistically significant differences between the two groups on the  
209 glaucoma-specific PROMs. In addition, the distribution in the baseline to exit scores were  
210 strikingly similar between the treatment and placebo groups (Figure 1).

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

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

223 We assessed differences between stable patients ( $N=272$ ) and patients with  
224 glaucomatous progression ( $N=78$ ) as determined by the primary visual field outcome.  
225 Median (interquartile range) duration between baseline and progression confirmation  
226 visit was 465 (278, 553) days, in comparison to the 2-year (730 days) scheduled follow-  
227 up for patients remaining stable. No statistically significant differences were found  
228 between average responses from stable and progressed patients on PROMs relating to  
229 general health (EQ-5D, EQ-5D VAS and SF-36). Average differences between stable and  
230 progressed patients were statistically significant when assessing responses on glaucoma-  
231 specific PROMs (GQL-15 and GAL-9) (Table 3 and Figure 2). As a group, patients who had  
232 progressed on visual fields therefore reported a reduction in glaucoma-specific vision-  
233 related quality of life that was different to those who had remained stable on visual fields.  
234 Mean (95% CI) scores for the progression patients on the GAL-9 and GQL-15 was 6.5 (2.8–  
235 9.2) % and 3.9 (3.2–9.8) % respectively.

236

237

238

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.

PROM	Group		Mean Difference [CI]	p-value
	Treatment N = 182	Placebo N = 168		
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 form 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<sup>th</sup> and 95<sup>th</sup> percentiles (whiskers). Due to large variability in  
256 responses, 95<sup>th</sup> percentile is capped at 40% change for SF-36 analysis (SF-36 placebo 95<sup>th</sup>  
257 percentile = 54.6%; SF-36 treatment 95<sup>th</sup> 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).

<b>Table 2.</b> 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 <i>N = 350</i>	UKGTS patients without PROMs <i>N = 166</i>		p-value
<b>MD (dB)</b>				
<b>Better eye</b>				
Mean	-0.5 (1.2)	-0.8 (1.8)		0.12
Median	-0.5 [-1.3, 0.4]	-0.6 [-1.4, 0.3]		
<b>Worse eye</b>				
Mean	-4.2 (3.3)	-4.3 (3.6)		0.90
Median	-3.3 [-5.6, -2.0]	-3.4 [-5.7, -1.7]		
<b>Best-corrected VA</b>				
<b>Better eye</b>				
Mean	1.0 (0.21)	1.0 (0.24)		0.44
Median	1.0 [1.0, 1.2]	1.0 [1.0, 1.2]		
<b>Worse eye</b>				
Mean	0.9 (0.24)	0.9 (0.25)		0.56
Median	1.0 [0.67, 1.0]	1.0 [0.67, 1.0]		
<b>Age (years)</b>				
Mean	65.8 (9.9)	67.4 (11.9)		0.27
<b>Sex</b>				
Male	188 (53.7%)	85 (51.2%)		
Female	162 (46.3%)	81 (48.8%)		
<b>Baseline PROM</b>				
			Mean difference [CI]	
<b>Mean</b>				
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.

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.

PROM	Outcome		Mean Difference [CI]	p-value
	Stable N = 272	Progressed N = 78		
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 form 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, 5<sup>th</sup> and 95<sup>th</sup>



Patient reported outcome measures in glaucoma clinical trials

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

285

## 286 **Discussion**

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

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

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

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

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

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

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

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

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

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

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

420

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.



- 446 9. Migdal C, Gregory W, Hitchings R. Long-term functional outcome after early surgery  
447 compared with laser and medicine in open-angle glaucoma. *Ophthalmology*.  
448 1994;101:1651-57.
- 449 10. Jay JL, Murray SB. Early trabeculectomy versus conventional management in primary  
450 open angle glaucoma. *British Journal of Ophthalmology*. 1988;72:881-89.
- 451 11. Musch DC, Gillespie BW, Lichter PR, et al. Visual field progression in the Collaborative  
452 Initial Glaucoma Treatment Study: the impact of treatment and other baseline factors.  
453 *Ophthalmology*. 2009;116:200-7.
- 454 12. Garway-Heath DF, Crabb DP, Bunce C, et al. Latanoprost for open-angle glaucoma  
455 (UKGTS): a randomised, multicentre, placebo-controlled trial. *The*  
456 *Lancet*. 2015;385:1295-304.
- 457 13. Denniston AK, Kyte D, Calvert M, et al. An introduction to patient-reported outcome  
458 measures in ophthalmic research. *Eye*. 2014;28:637-45.
- 459 14. Deshpande PR, Rajan S, Sudeepthi BL, et al. Patient-reported outcomes: a new era in  
460 clinical research. *Perspectives in Clinical Research*. 2011;2:137.
- 461 15. Black N. Patient reported outcome measures could help transform healthcare. *BMJ*.  
462 2013 346, p.f167.
- 463 16. Glen FC, Crabb DP, Garway-Heath DF. The direction of research into visual disability and  
464 quality of life in glaucoma. *BMC Ophthalmology*. 2011;11:19.
- 465 17. Hamzah JC, Burr JM, Ramsay CR, et al. Choosing appropriate patient-reported outcomes  
466 instrument for glaucoma research: a systematic review of vision instruments. *Quality of*  
467 *Life Research*. 2011;20:1141-58.
- 468 18. Vodicka E, Kim K, Devine EB, et al. Inclusion of patient-reported outcome measures in  
469 registered clinical trials: evidence from ClinicalTrials.gov (2007–2013). *Contemporary*  
470 *Clinical Trials*. 2015;43:1-9.

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

- 496 27. Vickerstaff V, Ambler G, Bunce C, et al. Statistical analysis plan for the Laser-1st versus  
497 Drops-1st for Glaucoma and Ocular Hypertension Trial (LiGHT): a multi-centre  
498 randomised controlled trial. *Trials*. 2015;16:517.
- 499 28. Crabb DP. A view on glaucoma—are we seeing it clearly? *Eye*. 2016;30:304-13.
- 500 29. Garway-Heath DF, Lascaratos G, Bunce C, et al. The United Kingdom Glaucoma Treatment  
501 Study: a multicenter, randomized, placebo-controlled clinical trial: design and  
502 methodology. *Ophthalmology*. 2013;120:68-76.
- 503 30. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy*.  
504 1990;16:199-208.
- 505 31. Devlin N, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. *Applied*  
506 *Health Economics and Health Policy*. 2017;15:127-37.
- 507 32. Siegel JE, Weinstein MC, Russell LB, et al. Recommendations for reporting cost-  
508 effectiveness analyses. *JAMA*. 1996;276:1339-41.
- 509 33. Sullivan PW, Lawrence WF, Ghushchyan V. A national catalog of preference-based scores  
510 for chronic conditions in the United States. *Medical Care*. 2005;43:736-49.
- 511 34. Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I.  
512 Conceptual framework and item selection. *Medical Care*. 1992;30:473-83.
- 513 35. Aaronson NK, Acquadro C, Alonso J, et al. International quality of life assessment (IQOLA)  
514 project. *Quality of Life Research*. 1992;1:349-51.
- 515 36. Hall T, Krahn GL, Horner-Johnson W, et al. Examining functional content in widely used  
516 Health-Related Quality of Life scales. *Rehabilitation Psychology*. 2011;56:94.
- 517 37. Nelson P, Aspinall PA, O'Brien C. Patients' perception of visual impairment in glaucoma:  
518 a pilot study. *British Journal of Ophthalmology*. 1999;83:546-52.
- 519 38. Nelson P, Aspinall P, Papasouliotis O, et al. Quality of life in glaucoma and its relationship  
520 with visual function. *Journal of Glaucoma*. 2003;12:139-50.
- 521 39. van Gestel A, Webers CA, Beckers HJM, et al. The relationship between visual field loss in  
522 glaucoma and health-related quality-of-life. *Eye*. 2010;24:1759-69.

- 523 40. Goldberg I, Clement CI, Chiang TH, et al. Assessing quality of life in patients with glaucoma  
524 using the Glaucoma Quality of Life-15 (GQL-15) questionnaire. *Journal of Glaucoma*.  
525 2009;18:6-12.
- 526 41. Khadka J, Pesudovs K, McAlinden C, et al. Reengineering the glaucoma quality of life-15  
527 questionnaire with rasch analysis. *Investigative Ophthalmology & Visual Science*.  
528 2011;52:6971-7.
- 529 42. Medeiros FA, Gracitelli CP, Boer ER, et al. Longitudinal changes in quality of life and rates  
530 of progressive visual field loss in glaucoma patients. *Ophthalmology*, 2015; 122: 293-301.
- 531 43. Abe RY, Diniz-Filho A, Costa VP, et al. The Impact of Location of Progressive Visual Field  
532 Loss on Longitudinal Changes in Quality of Life of Glaucoma Patients. *Ophthalmology*,  
533 2016; 123: 552-7.
- 534 44. Peters D, Heijl A, Brenner L, et al. Visual impairment and vision-related quality of life in  
535 the Early Manifest Glaucoma Trial after 20 years of follow-up. *Acta Ophthalmologica*,  
536 2015; 93: 745-52.
- 537 45. Diniz-Filho A, Abe RY, Cho HJ, et al. Fast visual field progression is associated with  
538 depressive symptoms in patients with glaucoma. *Ophthalmology*, 2016; 123: 754-9.
- 539 46. Dean S, Mathers JM, Calvert M, et al. 'The patient is speaking': discovering the patient voice  
540 in ophthalmology. *British Journal of Ophthalmology*. 2017;101:700-8.
- 541 47. Wegner A. Latanoprost for glaucoma: primum non nocere. *The Lancet*. 2015;386:651-52.
- 542 48. Altangerel U, Spaeth GL, Steinmann WC. Assessment of function related to vision (AFREV).  
543 *Ophthalmic Epidemiology*. 2006;13:67-80.
- 544 49. Skalicky SE, McAlinden C, Khatib T, et al. Activity limitation in glaucoma: objective  
545 assessment by the Cambridge Glaucoma Visual Function Test. *Investigative*  
546 *Ophthalmology & Visual Science*. 2016; 57:6158-66.
- 547 50. Arora KS, Boland MV, Friedman DS, et al. The relationship between better-eye and  
548 integrated visual field mean deviation and visual disability. *Ophthalmology*. 2013; 120:  
549 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