

The Bristol Shared Care Glaucoma Study: reliability of community optometric and hospital eye service test measures

P G D Spry, I C Spencer, J M Sparrow, T J Peters, S T Brookes, S Gray, I Baker,
J E Furber, D L Easty

Abstract

Background/aims—Primary open angle glaucoma patients and glaucoma suspects make up a considerable proportion of outpatient ophthalmological attendances and require lifelong review. Community optometrists can be suitably trained for assessment of glaucoma. This randomised controlled trial aims to assess the ability of community optometrists in the monitoring of this group of patients.

Methods—Measures of cup to disc ratio, visual field score, and intraocular pressure were taken by community optometrists, the hospital eye service and a research clinic reference "gold" standard in 405 stable glaucoma patients and ocular hypertensives. Agreement between and within the three centres was assessed using mean differences and intraclass correlation coefficients. Tolerance limits for a change in status at the level of individual pairs of measurements were also calculated.

Results—Compared with a research clinic reference standard, measurements made by community optometrists and those made in the routine hospital eye service were similar. Mean measurement differences and variability were similar across all three groups compared for each of the test variables (IOP, cup to disc ratio, and visual field). Overall, the visual field was found to be the most reliable measurement and the cup to disc ratio the least.

Conclusions—Trained community optometrists are able to make reliable measurements of the factors important in the assessment of glaucoma patients and glaucoma suspects. This clinical ability should allow those optometrists with appropriate training to play a role in the monitoring of suitable patients.

(Br J Ophthalmol 1999;83:707-712)

Primary open angle glaucoma (POAG) is an age related optic neuropathy of complex multifactorial aetiology. It is defined as a slowly progressive atrophy of the optic nerve, characterised by loss of peripheral visual function and an excavated appearance of the optic disc.¹ The presence of IOP above the statistically defined limits of normality without either other sign is referred to as ocular hypertension (OHT) and represents a powerful risk factor for POAG.²

POAG prevalence increases with age,³ from around 0.9% in the fifth decade to nearly 5% over 75 years of age.⁴ The estimated prevalence of OHT varies between 3.6% and 7.6% in the over 50 year age group.³⁻⁵ Estimates of the incidence of POAG among OHT populations vary, but is thought to be 1% to 2% per year.⁶⁻⁷

Once diagnosed, the chronic nature of POAG necessitates lifelong observation. Patients with OHT require regular follow up, the frequency of which depends upon a variety of factors, including magnitude of intraocular pressure elevation and other coexisting risk factors.

Unpublished data collected at the Bristol Eye Hospital indicate that about 23% of total outpatient attendances are for glaucoma follow up. A survey of consultant ophthalmologists in the south west of England⁸ found that almost two thirds of respondents estimated that glaucoma patients made up 10%–25% of their outpatient time, with a quarter of respondents estimating between 25% and 50%. The demography of an aging population and improved optometric case finding will increase the number of POAG and OHT patients requiring monitoring.

Community optometrists have the background knowledge, skills, and instrumentation necessary to carry out measurements applicable to glaucoma and with suitable training could have a role in undertaking glaucoma assessments currently performed by hospital based staff. The Bristol Shared Care Glaucoma Study was designed as a randomised controlled trial to investigate a model of "shared care" whereby community optometrists would monitor selected POAG and OHT patients.⁹ In the trial, care by community optometrists was compared with routine HES follow up. Such a shared care system has the potential to relieve overloaded outpatient departments and so make hospital resources more available for other aspects of ophthalmological care.

The aims of the study were:

(1) Assessment of the reliability of measurements made by community optometrists and the hospital eye service in comparison with standardised measures made by a research clinic gold standard assessment.

(2) Examination of patient satisfaction with the two methods of care (community based optometric follow up versus routine HES care).¹⁰

Department of
Ophthalmology,
University of Bristol,
Bristol

P G D Spry
I C Spencer
J M Sparrow
D L Easty

Discoveries in Sight,
Devers Eye Institute,
Portland, Oregon,
USA

P G D Spry

Department of Social
Medicine, University
of Bristol, Bristol

T J Peters
S T Brookes
S Gray

Avon Health Authority,
Bristol

I Baker

Bristol Eye Hospital,
Bristol
J E Furber

Correspondence to:
Dr P G D Spry, Discoveries
in Sight, Devers Eye
Institute, Legacy Clinical
Research and Technology
Center, 1225 NE 2nd
Avenue, Portland, Oregon,
97208, USA.

Accepted for publication
18 December 1998

Table 1 Recruitment criteria

Inclusion criteria	Exclusion criteria
Glaucoma suspects	Unstable glaucoma
Stable*:	Other glaucomas:
Primary open angle glaucoma	Normal tension glaucoma
Pigment dispersion glaucoma	Secondary glaucoma
Pseudoexfoliative glaucoma	Narrow angle glaucoma
Informed consent	Other coexisting ocular pathology
Ability to cooperate with examination	Extensive visual field loss†
Snellen visual acuity of 6/18 or better in both eyes	Best corrected visual acuity in either eye worse than 6/18
Aged 50 years and over	Age less than 50 years

*Stable defined as no observable change in the individual's visual field within the last year of follow up.

†Extensive visual field loss defined as loss greater than 66/132 points missed on the threshold related suprathreshold examination with the Henson CFA3000.

(3) Examination of the costs of providing glaucoma follow up by community optometrists and by traditional hospital based services, including costs borne by patients.

This paper will utilise the initial cross sectional multiple observer study design to examine measurement reliability, in terms of interobserver and intraobserver agreement and therefore addresses aim (1). The implications of the observed reliability in the context both of shared care models and general ophthalmological clinical environments will be considered. Previous publications from the Bristol Shared Care Glaucoma Study have reported on measurement validity and patient satisfaction¹⁰ and cost analysis.^{11 12}

Methods

The Bristol Glaucoma Shared Care Study model has been described in detail elsewhere.⁹ Briefly, patients were selected from those attending dedicated glaucoma clinics at the Bristol Eye Hospital using the inclusion and exclusion criteria shown in Table 1. Potential participants were invited to attend a research clinic where suitability was assessed and informed consent obtained from eligible individuals wishing to participate in the study. Participants were randomly allocated to receive either structured community optometric follow up (CO) or routine hospital eye service review (HES). For those allocated to CO, one of 12 specially trained participating optometrists in community practice was chosen for each patient according to convenience of location. As part of this process, observations made on all participants selected at the research clinic were termed research clinic reference standards (RCRS). Within 2 months of this visit similar assessments were also made in the HES and in one of the CO practices so that all

selected patients had measurements by all three centres on entry. The test measurements of interest are detailed in Table 2.

ANALYSIS

(1) Intercentre agreement (between RCRS, CO, and HES) was determined for the following measurement factors:

- Vertical cup to disc ratio (CDR); measured with the patient in attendance to the nearest 0.05.
- Visual field score (VF); points missed out of 132 (Henson suprathreshold score).
- IOP; mm Hg.

(2) Intracentre agreement (within each test centre) was assessed in a variety of ways:

- VF; between repeated measures taken on different days in each of the research clinic (RCRS) and the community optometric practices (CO)
- IOP; between triple RCRS measures (R1, R2, and R3) performed unmasked and in rapid succession at the same visit.

The study design did not provide for direct intraobserver comparisons for CDR measurements. Measures from stereoscopic disc photographs were therefore included to aid assessment of grading performance for this feature. Photographs were independently graded by two observers. Intracentre agreement for photographic grading was assessed by comparison of the grades made by the two observers. The validity of photographic grades was assessed by comparison against the direct observations made in the research clinic (RCRS; patient in attendance). The following comparisons are thus presented:

- CDR; intracentre agreement of photographic grades between two independent observers
- CDR; validity of photographic grades made by observer 1 in comparison with RCRS clinical observations also made by observer 1.

Table 2 Test measurements assessed

Research clinic reference standard	Hospital glaucoma clinic	Community based optometrists
Visual field analysis with Henson CFA3000 132 point threshold related suprathreshold examination	Visual field analysis with a Henson CFS2000/CFA3000	Visual field analysis with Henson CFA3000 132 point threshold related suprathreshold examination
Repeat visual field examination		Repeat visual field examination (50% of patients)
Triple IOP measurement by Goldmann applanation tonometry	Single IOP measurement by Goldmann applanation tonometry	Single IOP measurement by Goldmann applanation tonometry
Vertical cup to disc ratio assessment by binocular indirect ophthalmoscopy (dilated pupil)	Vertical cup to disc ratio assessment using direct ophthalmoscopy or binocular indirect ophthalmoscopy	Vertical cup to disc ratio assessment by binocular indirect ophthalmoscopy (dilated pupil)
Stereophotographic analysis of vertical cup to disc ratio by observer 1		
Stereophotographic analysis of vertical cup to disc ratio by observer 2		

- CDR; validity of photographic grades made by observer 2 in comparison with RCRS clinical observations made by observer 1.

It should be noted that the visual field quantification employed provided a measure of defect size but did not attempt to provide spatial information about the defect.

STATISTICAL METHODS

Reliability/agreement levels were assessed by two main approaches. In the first, attention was focused on score differences at the level of individual pairs of measures.¹³ In this approach, paired observations were examined for evidence of systematic bias between sets of observations (means and 95% confidence intervals of paired differences) and for the extent of disagreement within individual pairs. For the latter, from the standard deviation (SD) of the score differences, the “tolerance limits” were obtained as: plus or minus ($1.96 \times \text{SD}$). This statistic provides a quantity by which individual paired observations must differ (for example, change in score from baseline to follow up) to have evidence of a true change in the status of the factor under study, assuming no systematic difference between the two groups of observations used to obtain the paired differences. Where score differences on pairs of observations are less than the “tolerance limit for detecting change”, then this suggests that the observed score difference may have occurred as a result of measurement error and should therefore not be relied upon as evidence of a true change in status. Such values are therefore of use in subsequent review appointments. To aid further the interpretation of this factor and to facilitate comparisons across different variables, the extent of the variation associated with paired differences (the tolerance limit for detecting change) was compared with overall variation. These were calculated as the tolerance limit as a percentage of the range of the relevant original measurement.

The second approach involved calculation of the intraclass correlation coefficient (ICC). This coefficient is equivalent to a quadratic weighted kappa statistic, which is a chance corrected measure of agreement, weighting degrees of discrepancies according to the square of the difference between the (paired) measurements.¹⁴ There are no universally applicable standard values for the ICC which represent adequate reliability, but to aid presentation the following convention is followed here: ICC <0.20 “slight agreement”; 0.21–0.40 “fair agreement”; 0.41–0.60 “moderate agreement”; 0.61–0.80 “substantial agreement”; and above 0.80 “almost perfect agreement”.

The ICC is preferable to the usual (Pearson) correlation coefficient since the latter strictly speaking measures association rather than agreement. Unlike the Pearson correlation, the ICC only indicates perfect agreement if the two assessments are numerically equal—that is, if a plot of the two measurements has zero intercept and a slope of unity. However, the (crude) ICC is affected specifically,

reduced by any systematic differences between the observations within the pairs. In other words, even perfect “agreement” in the context of such systematic differences will result in an ICC less than 1. To investigate whether this was the case, the measurements were effectively recalibrated by subtracting the mean difference from the higher of the two. The ICC was then recalculated, yielding a measure corrected for systematic bias. In this way, the adjusted ICC represents the reliability/agreement correcting for both chance agreement and systematic bias, and the impact on the ICC of the adjustment reflects the magnitude of this bias. The same influences are therefore represented in these statistics as in the summary measures for the paired differences in the first approach.

Results

A total of 2780 glaucoma patients’ notes were reviewed of which 674 (23%) met the inclusion criteria and appeared suitable for recruitment. Of this number, 405 (60%) were willing to participate.

INTERCENTRE MEASUREMENT RELIABILITY

Intercentre comparisons of the various test measurements are shown in Table 3. The between centres comparison demonstrated that the visual field was the most reliable of the three measures of interest, with moderate to substantial agreement and minimal mean differences being observed. Similarly, the tolerance limits expressed as a percentage of the range was lowest for visual field, illustrating that the spread for visual field differences was least in comparison with the spread of the original measures. IOP and CDR exhibited fair to moderate agreement, with mean differences being small in comparison with standard deviations of the differences. Typical mean differences were around or under $0.5 \times \text{SD}$ of the differences (Table 3). The level of agreement between CO measures and the RCRS was at least equal to that between HES based measures and the RCRS. In general, measurements made by the CO were closer to those of the RCRS than the HES. Systematic bias appears to have affected measurement of both CDR and IOP. This is illustrated by the 95% CI for the mean and the improvement in the ICC following adjustments for systematic bias. Visual field measurement did not appear to exhibit any systematic bias as illustrated by the adjusted and unadjusted ICCs being the same.

It is of interest to note that the use of the mean of two repeated visual field measures enhanced the level of agreement between centres from substantial to almost perfect and considerably narrowed tolerance limits for change. It should be noted that the sample size for this comparison is reduced because only the patients randomised to community follow up had test retest visual field data available for community optometrists.

Table 3 Inter (between) centre agreement analysis (direct clinical measurements with patients in attendance)

Groups compared	No	Mean difference	95% CI for mean difference	SD	95% TL	TL as % of range	Crude ICC	Adjusted ICC
<i>C/D ratio:</i>								
RCRS v HES R	360	0.10	0.08, 0.12	0.17	0.33	39	0.43	0.52
RCRS v HES L	358	0.05	0.03, 0.07	0.16	0.32	33	0.46	0.48
RCRS v CO R	360	0.05	0.03, 0.07	0.17	0.33	39	0.42	0.44
RCRS v CO L	358	0.10	0.08, 0.12	0.19	0.36	44	0.36	0.46
HES v CO R	360	-0.05	-0.03, -0.07	0.18	0.35	39	0.47	0.50
HES v CO L	358	0.05	0.03, 0.07	0.18	0.35	44	0.50	0.54
<i>IOP:</i>								
RCRS2 v HES R	388	2.4	1.95, 2.85	4.50	8.8	26	0.33	0.44
RCRS2 v HES L	388	2.8	2.36, 3.24	4.46	8.7	26	0.34	0.48
RCRS2 v CO R	388	2.0	1.57, 2.43	4.36	8.6	22	0.45	0.52
RCRS2 v CO L	388	2.2	1.75, 2.65	4.48	8.8	21	0.42	0.50
HES v CO R	388	0.4	-0.05, 0.85	4.50	8.8	23	0.45	0.45
HES v CO L	388	0.6	0.13, 1.07	4.69	9.2	27	0.40	0.40
<i>VF:</i>								
RCRS1 v CO1 R	366	-0.4	-1.53, 0.73	11.03	22	29	0.64	0.64
RCRS1 v CO1 L	366	0.1	-0.93, 1.13	10.12	20	21	0.73	0.73
HES v CO R	287	1.1	-0.38, 2.58	12.80	25	33	0.54	0.55
HES v CO L	287	0.7	-0.80, 2.20	12.95	25	27	0.61	0.61
RCRSM v HES R	308	0.6	-0.59, 1.69	10.66	21	28	0.64	0.64
RCRSM v HES L	305	0.3	-1.03, 1.63	11.88	23	25	0.61	0.61
RCRSM v COM R	170	0.8	-0.32, 1.92	7.44	15	21	0.80	0.80
RCRSM v COM L	170	0.2	-0.87, 1.27	7.33	14	17	0.84	0.84

RCRS = research clinic reference standard observation; RCRS1 = research clinic reference standard first observation; RCRS2 = research clinic reference standard second observation; RCRSM = mean research clinic reference standard observation; HES = hospital eye service assessment; CO = community optometric assessment; CO1 = community optometric first assessment (where repeated assessments made); COM = mean community optometric assessment; R = right eye; L = left eye; SD = standard deviation of paired differences; CI = confidence interval (for the mean); TL = tolerance limit for change; ICC = intraclass correlation coefficient.

INTRACENTRE MEASUREMENT RELIABILITY

Intracentre comparisons of the various measurements are shown in Table 4. Reliability within a single centre was better than that between centres for both IOP and VF measurements, as indicated by lower mean measurement differences and standard deviations and reduced tolerance limits for detection of change (Table 4). Agreement levels were graded as almost perfect for unmasked rapid sequence measurement of IOP and substantial/almost perfect for VF. Mean differences were

again small in comparison with SDs. Adjusted and unadjusted ICCs were the same, indicating no systematic bias.

INTRACENTRE VALIDITY

CDR assessment comparison between the photographic and clinical methods demonstrated greater variability (Table 5). Considerable systematic bias existed between photographic assessments and those made clinically with the patient in attendance (illustrated by 95% CIs for the means and the fact that

Table 4 Intra (within) centre agreement analysis (direct clinical measurements with patients in attendance)

Groups compared	No	Mean difference	95% CI for mean difference	SD	95% TL	TL as % of range	Crude ICC	Adjusted ICC
<i>IOP*:</i>								
RCRS1 v RCRS2 R	388	0.1	0.01, 0.19	0.88	1.7	5	0.98	0.98
RCRS1 v RCRS2 L	388	0.1	0.00, 0.20	0.98	1.9	6	0.98	0.98
RCRS1 v RCRS3 R	388	0.4	0.31, 0.49	0.91	1.8	5	0.97	0.98
RCRS1 v RCRS3 L	388	0.3	0.20, 0.40	1.04	2.0	6	0.97	0.97
RCRS2 v RCRS3 R	388	0.3	0.22, 0.38	0.83	1.6	4	0.98	0.98
RCRS2 v RCRS3 L	388	0.1	0.01, 0.19	0.87	1.7	5	0.98	0.98
<i>VF:</i>								
RCRS1 v RCRS2 R	400	0.1	-0.65, 0.85	7.14	14	17	0.83	0.83
RCRS1 v RCRS2 L	400	0.3	-0.58, 1.18	8.58	17	27	0.78	0.78
CO1 v CO2 R	166	0.2	-1.03, 1.43	8.25	16	22	0.80	0.80
CO1 v CO2 L	166	0.8	-0.13, 1.73	6.15	12	13	0.89	0.89

*Unmasked measurements performed in rapid sequence; RCRS1 = research clinic reference standard first observation; RCRS2 = research clinic reference standard second observation; RCRS3 = research clinic reference standard third observation; CO1 = community optometric first assessment; CO2 = community optometric second assessment; R = right eye; L = left eye; SD = standard deviation of paired differences; CI = confidence interval (for the mean); TL = tolerance limit for change; ICC = intraclass correlation coefficient.

Table 5 Intra (within) centre validity for C/D ratio (direct clinical and photographic measurements)

Groups compared	No	Mean difference	95% CI for mean difference	SD	95% TL	TL as % of range	Crude ICC	Adjusted ICC
<i>C/D ratio:</i>								
RCRS v Obs 1 R	381	-0.15	-0.14, -0.16	0.14	0.27	33	-0.04	0.39
RCRS v Obs 1 L	384	-0.15	-0.14, -0.16	0.15	0.28	40	0.10	0.44
RCRS v Obs 2 R	381	-0.15	-0.13, -0.17	0.15	0.30	40	0.12	0.38
RCRS v Obs 2 L	384	-0.10	-0.08, -0.12	0.16	0.31	32	0.21	0.42
Obs 1 v Obs 2 R	381	0.05	0.04, 0.06	0.12	0.23	38	0.34	0.38
Obs 1 v Obs 2 L	384	0.05	0.03, 0.07	0.15	0.30	32	0.24	0.25

RCRS = research clinic reference standard observation; Obs 1 = stereo photograph assessment by first observer; Obs 2 = Stereo photograph assessment by second observer; R = right eye; L = left eye; SD = standard deviation of paired differences; CI = confidence interval (for the mean); TL = tolerance limit for change; ICC = intraclass correlation coefficient.

adjusted ICCs were higher than unadjusted). When independently grading stereo retinal photographs, the two observers did not systematically disagree. However, the level of agreement was only fair to moderate, indicating the difficulty of grading CDR from retinal photographs.

Discussion

This paper provides data which inform the debate on the appropriateness of involving optometrists in shared care for patients with glaucoma. The focus here is on aspects of the reliability of glaucoma related measurements. In the model under study, measurement reliability appears acceptable, which opens the way to examination of other aspects of shared care, such as convenience and acceptability to the patient plus cost considerations. Additional advantages of such a shared care model may include reduction of HES outpatient load allowing medically qualified staff to focus on activities which they alone can perform; and improvements in quality of primary referrals for suspected glaucoma by participating optometrists.

Comparison of measures performed by community optometrists and routine HES clinics in relation to a research clinic reference or "gold" standard show that levels of agreement achieved between the three centres are similar for each factor. This suggests that community optometrists are able to provide clinical measurement information of a quality equivalent to that of current HES outpatient services.

Of the three factors studied, VF measurement has been observed as the most reliable, with agreement levels ranging from moderate to almost perfect. IOP agreement levels ranged from fair to moderate, although it is of note that the tolerance limits for detecting change for IOP were high when these measurements were performed on different days by different individuals. These wide limits (around 9 mm Hg, Table 3) reflect diurnal and day to day variation in addition to measurement error. Repeated measures made (unmasked) on the same occasion in the research clinic were much less variable (tolerance limit around 2 mm Hg, Table 4).

Analysis of CDR measures found low levels of agreement between centres, agreement ranging from slight to moderate. This factor was also shown to be of least value in differentiating true variable change from measurement "noise", illustrated by tolerance limits in the region of 0.35, representing 40% of the entire range of measures. This easily exceeded that of both IOP and visual field, their tolerance limits comprising approximately 25% of their measurement ranges. This lack of scale sensitivity for CDR to detect change means that in order to be 95% certain that a true change in status has occurred, an individual pair of measurements must differ (increase or decrease) by 0.35. This result may be surprising to many clinicians, who generally (and probably falsely) believe that they are able to judge optic disc cupping more accurately than this. This formal quantification of the interobserver measure-

ment noise contradicts the accepted traditional criterion of a CDR change of ≥ 0.20 representing a "clinically significant" change in glaucomatous cupping, since this level of change falls well within the range of measurement error. This finding will be particularly relevant to the "multiobserver" environment typified by NHS outpatient departments.

Previous studies have suggested that stereophotographs are more valuable than clinical examination in evaluation of possible glaucomatous progression.^{15,16} The present study has found only a "slight" agreement for intraobserver CDR measurement when using stereophotographs. Agreement improved following "calibration" for a systematic bias between clinical and photographic CDR assessment in which observations from stereophotographs assessed CDRs up to 0.15 larger on average compared with a clinical measure made through dilated pupils using a stereoscopic viewing system. This result conflicts with previous work,¹⁶ where the opposite bias was observed and reinforces the notion that the grading of stereo disc photographs is a difficult and subjective task.

Available comparisons for CDR suggest that great caution is required when considering mixed photographic and clinical measurements. In this study clinical measurements of CDR were generally more reliable than photographically based or mixed comparisons.

The removal of stable POAG patients from routine hospital outpatient clinics into non-ophthalmologist shared care environments can only be considered a viable option if the standard of shared care received is at least as good as that of the current care system. A prerequisite for such a standard is accuracy of clinical measurement. This accuracy may be perceived as "acceptable agreement" between variables measured by participating observers. Although local requirements will determine the mode of implementation of shared care schemes, our measurement findings should be equally applicable to shared care models based around hospital or community optometrists although model safety should be considered. The use of optometrists in a community setting is one example of the sharing of care between traditional primary and secondary forms of delivery. Strict return referral criteria should be implemented for return ophthalmological review at the hospital. These criteria should identify likely disease progression but should also provide for intermittent HES review.

When there is uncertainty about progression of the disease, repeated measures¹⁷⁻¹⁹ or more frequent optometric review may be necessary. In this context our findings strongly support the usefulness of repeated visual field measures. Using the mean of two measurements (Table 3), the between centre reliability and tolerance limits for change are enhanced to a level comparable with that found within a single "gold standard" setting. It should be noted that the analysis uses a range of visual field defect sizes. If defects are graded by size, variability within certain subgroups may be further

reduced,^{20 21} although larger visual field defects can be expected to exhibit more variability.

In conclusion, the results indicate that when using specified measurement techniques, optometrists trained in assessment of glaucoma related measures perform as reliably as traditional methods of HES glaucoma review when monitoring patients with ocular hypertension or stable primary open angle glaucoma. Of the three relevant measures, visual field assessment was the most reliable factor.

We thank all members of the Bristol Shared Care Glaucoma Study Group, Mrs Nancy Thomas, consultant ophthalmologists, and staff of Bristol Eye Hospital. Grateful thanks are extended to participating optometrists for their invaluable contributions.

The Bristol Shared Care Glaucoma Study was supported by the South & West Regional Research and Development Directorate, Medical Research Council, International Glaucoma Association and Avon Health Authority.

- 1 Quigley HA. Number of people with glaucoma worldwide. *Br J Ophthalmol* 1996;**80**:389–93.
- 2 Sommer A, Tielsch JM, Katz J, et al. Relationship between intraocular pressure and primary open angle glaucoma among white and black Americans. *Arch Ophthalmol* 1991;**109**:1090–5.
- 3 Coffey M, Reidy A, Wormald R, et al. Prevalence of glaucoma in the west of Ireland. *Br J Ophthalmol* 1993;**77**:17–21.
- 4 Klein BE, Klein R, Sponsel WE, et al. Prevalence of glaucoma. The Beaver Dam Eye Study. *Ophthalmology* 1992;**99**:1499–504.
- 5 Leibowitz HM, Krueger DE, Maunder LR. The Framingham eye study monograph. *Surv Ophthalmol* 1980;**24**(suppl):335–610.
- 6 Kitazawa Y, Horie T, Aoki S, et al. Untreated ocular hypertension. *Arch Ophthalmol* 1977;**95**:1180–4.
- 7 David R, Livingston DG, Luntz MH. Ocular hypertension—a long term follow-up of treated and untreated patients. *Br J Ophthalmol* 1977;**61**:668–74.
- 8 Spencer IC. *How do we manage chronic open angle glaucoma? A validation study in the south-west region of Britain*. MSc thesis. Bristol: University of Bristol, 1992.
- 9 Spencer IC, Spry PGD, Gray SF, et al. The Bristol Shared Care Glaucoma Study: study design. *Ophthalmic Physiol Optics* 1995;**15**:391–4.
- 10 Gray SF, Spencer IC, Spry PGD, et al. The Bristol Shared Care Glaucoma Study—validity of measurements and patients satisfaction. *J Public Health Med* 1997;**19**:431–6.
- 11 Spencer IC, Coast J, Spry PGD, et al. The cost of monitoring glaucoma patients by community optometrists. *Ophthalmic Physiol Opt* 1995;**15**:383–6.
- 12 Coast J, Spencer IC, Smith L, et al. Comparing the costs of monitoring for glaucoma patients: ophthalmologists in the hospital verses optometrists in the community. *J Health Serv Res Policy* 1997;**2**:19–25.
- 13 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**28**:307–10.
- 14 Streiner DL, Norman GR. *Health measurement scales*. Oxford: Oxford Medical Publications, 1995:104–26 and 163–80.
- 15 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74.
- 16 Tielsch JM, Katz JM, Quigley HA, et al. Intraobserver and interobserver agreement of optic disc characteristics. *Ophthalmology* 1988;**95**:350–6.
- 17 Henson DB, Darling MN. Detecting progressive visual field loss. *Ophthalmic Physiol Opt* 1995;**15**:387–90.
- 18 Henson DB, Spry PGD, Spencer IC, et al. Variability in glaucomatous visual fields: implications for shared care schemes. *Ophthalmic Physiol Opt* 1998;**18**:120–5.
- 19 Shultzer M. The Normal-tension Glaucoma Study Group. Errors in the diagnosis of visual field progression in normal-tension glaucoma. *Ophthalmology* 1994;**101**:1589–95.
- 20 Spry PGD, Henson DB, Sparrow JM. Detection of progressive visual field loss by suprathreshold perimetry. *Invest Ophthalmol Vis Sci* 1997;**38**:S2656.
- 21 Bailey IL, Bullimore MA, Raasch TW, et al. Clinical grading and the effects of scaling. *Invest Ophthalmol Vis Sci* 1991;**32**:422–32.