Auditing service delivery in glaucoma clinics using visual field records: a feasibility study

Stephen R Kelly, Susan R Bryan, John M Sparrow, David P Crabb

ABSTRACT

Objective This study aimed to demonstrate that large-scale visual field (VF) data can be extracted from electronic medical records (EMRs) and to assess the feasibility of calculating metrics from these data that could be used to audit aspects of service delivery of glaucoma care.

Method and analysis Humphrey visual field analyser (HFA) data were extracted from Medisoft EMRs from five regionally different clinics in England in November 2015, resulting in 602 439 records from 73 994 people. Target patients were defined as people in glaucoma clinics with measurable and sustained VF loss in at least one eye (HFA mean deviation (MD) outside normal limits ≥2 VFs). Metrics for VF reliability, stage of VF loss at presentation, speed of MD loss, predicted loss of sight years (bilateral VF impairment) and frequency of VFs were calculated.

Results One-third of people (34.8%) in the EMRs had measurable and repeatable VF loss and were subject to analyses (n=25 760 patients). Median (IQR) age and presenting MD in these patients were 71 (61, 78) years and −6 (−10, −4) dB, respectively. In 19 264 patients with >4 years follow-up, median (IQR) MD loss was −0.2 (−0.8, 0.3) dB/year and median (IQR) intervals between VF examinations was 11 (8, 16) months. Metrics predicting loss of sight years and reliability of examinations varied between centres (p<0.001).

Conclusion This study illustrates the feasibility of assessing aspects of health service delivery in glaucoma clinics through analysis of VF databases. Proposed metrics could be useful for blindness prevention from glaucoma in secondary care centres.

INTRODUCTION

Electronic medical records (EMRs) have potential to form a repository of data on patient encounters that can be directly used in research and clinical audit. The latter can be used as a first step towards improving health service delivery and improving patient care. Indeed, it is only possible to gauge improvements in a process after it has been measured in the first place.

Glaucoma clinics in Hospital Eye Services (HES) in England handle more than 1 million patient visits per year. Most of these visits are for monitoring people with established glaucoma. Once diagnosed, all patients with glaucoma require lifelong monitoring so that any worsening of disease can be detected and treatment intensified accordingly. Patient management focuses on monitoring the visual field (VF), assessment of the optic nerve and measurement of the intraocular pressure (IOP). The latter is critical because it is the only modifiable risk factor for the condition worsening (progression). Research evidence from clinics in England suggest measures of IOP are the main determinant of how often a patient is monitored over a period of time. In contrast to IOP, a measurement of the VF is the closest surrogate to what matters to the patient in the glaucoma clinic, that is, preservation of their vision. VF monitoring is therefore recognised as being critically important for the clinical management of the patient but it is also perceived to be difficult to do unless it is well implemented. The computerised technology used to measure VFs (standard
automated perimetry (SAP)) has been in clinics for 20–30 years and has remained largely unchanged. Data from SAP are stored electronically, often in an EMR. These records, which are likely to be historically rich, should be amenable to easy electronic auditing. Therefore, at different clinical centres it might be possible, for example, to audit measures of disease severity (VF loss) of patients at diagnosis. Moreover, it might be possible to audit speed at which patients in different clinics might be losing vision and, for example, whether frequency of VF monitoring is consistent across clinics. It is these ideas that we explore in this report.

The National Ophthalmology Database (NOD) was established under the auspices of the Royal College of Ophthalmologists (RCOphth) in 2010. The NOD aimed to collate pseudonymised data collected as a by-product of routine clinical care using EMR systems for the purposes of national audit, research and establishing meaningful measures for revalidation of ophthalmologists. In 2014, the Healthcare Quality Improvement Partnership (HQIP) commissioned NOD to do a feasibility study to investigate the use VF data to audit activity in glaucoma clinics. The results of the study are reported here with the specific aim of examining the viability of extracting meaningful metrics of health service delivery that might in the future allow comparison between glaucoma clinics.

METHODS

VF data were extracted from the Medisoft EMR system (Medisoft, Leeds, UK) from five regionally different National Health Service (NHS) Hospital Trust glaucoma clinics in England. The extraction was done in November 2015 and data transferred to the RCOphth NOD. All patient data were anonymised and subsequently transferred to a single secure database held at City university. For the purpose of this report, the five centres are anonymised. No other clinical data were used in this study apart from patient’s age, gender and the dates of the VF examinations. Subsequent analyses of the data were apart from patient’s age, gender and the dates of the mised. No other clinical data were used in this study. For the purpose of this report, the five centres are anony-

For the purpose of this report, our study population was defined as people in glaucoma clinics with measurable and sustained VF loss in at least one eye. This definition aims to exclude people suspected of having possible glaucoma (glaucoma suspects) and people with normal VFs and raised eye pressure (ocular hypertension (OHT)). Therefore, patients were only included if they had a VF with an HFA mean deviation (MD) flagged as outside the 95% normative limits in the HFA VF analysis software in at least one eye. (MD is a standard measure of the overall severity of VF loss, relative to healthy peers, with more negative values indicating greater VF loss.) Moreover, this proxy criterion for measurable VF loss had to be satisfied for both of the first two VFs recorded in the clinic; this was done in order to improve the precision of the estimate of an individual likely to have real VF loss at their presentation to secondary care. The number of patients satisfying these criteria expressed as a percentage of the total number of people with a VF record was calculated. This can be thought of as a simple count of people in clinics with actual VF loss at presentation to secondary care (diagnosis) as opposed to, for example, being a glaucoma suspect, a false-positive referral or having OHT.

### Metrics for assessing service delivery

Six different metrics were calculated to characterise and estimate aspects of patient monitoring and outcomes in the clinics.

**Age at presentation** was estimated by the age of the patient (years) at the time of their first VF record. **Reliability of VFs** was estimated by using the HFA false-positive (FP) measure. It is accepted that HFA FP is a useful measure of a reliable examination. The HFA flags VFs as unreliable if there are more than 15% FP errors. Percentage of all VFs considered as unreliable due to FP errors was therefore determined for each centre.

### Inclusion and exclusion criteria

SAP in these clinics, and most others in England, is routinely performed on a Humphrey visual field analyser (HFA, Carl Zeiss Meditec, Dublin, California, USA). Only VFs recorded on the HFA using a Goldmann size III stimulus with a 24-2 test pattern acquired with the Swedish Interactive Testing Algorithm (SITA Standard or SITA Fast) were included, reducing the aggregated database to 576 615 VFs from 71 361 people. SITA fast is commonly used in clinics in England. (Although SITA Standard is a more precise testing algorithm than SITA Fast at lower VF sensitivities, it is unlikely to make a sizeable difference to improving the time to detect VF progression.) In our study population, 83% of the recorded VFs were SITA Fast and the rest were SITA Standard.

For the purpose of this report, the total number of visual fields per centre. Each centre is simply labelled 1–5 and represented by a specific colour. This colour coding is used throughout this report (every centre had data recorded between April 2000 and March 2015 apart from centre 5 where data were first recorded in May 2000).

<table>
<thead>
<tr>
<th>Centre</th>
<th>People, n</th>
<th>VF records, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centre 1</td>
<td>3423</td>
<td>16162</td>
</tr>
<tr>
<td>Centre 2</td>
<td>8459</td>
<td>65355</td>
</tr>
<tr>
<td>Centre 3</td>
<td>27 921</td>
<td>285 552</td>
</tr>
<tr>
<td>Centre 4</td>
<td>18 636</td>
<td>113 847</td>
</tr>
<tr>
<td>Centre 5</td>
<td>15 555</td>
<td>121 523</td>
</tr>
</tbody>
</table>

VF, visual field.
Stage of VF loss at presentation was estimated by MD in the worse eye (the one with the more negative MD) at the second VF examination. The second VF was used to ameliorate the bias of the perimetry learning effect.\textsuperscript{14,15} We chose the worse eye as a surrogate of the most ‘detectable’ level of VF loss at the stage of case finding in primary care. Patients with MD worse than $-12$ dB in this eye were defined as having advanced VF loss. Patients with MD better than $-6$ dB in this most affected eye were defined as having early VF loss, with all other patients classified as having moderate VF loss. These VF criteria have been used in health economic investigations of service delivery of glaucoma.\textsuperscript{16,17} The proportion of patients within each of these three categories (early, moderate, advanced) can be summarised in a traffic-light waffle plot (green, yellow, red) for each centre.

Next, a subset of the study patients with sufficient series of VF examinations were defined as those with at least five VFs recorded over a period of 4 or more years. This subset of data were used for three more metrics summarising patient follow-up activity at each clinic.

**Speed (rate) of VF loss** in clinics was determined by using simple linear regression of MD against time of follow-up (dB/year) and was only calculated in patients with series of data. The first VF examination in each series was removed to account for perimetric learning effects.\textsuperscript{14,15} The metric considers rate of VF loss in both eyes and the patient’s residual life expectancy based on age and sex as reported in UK Office of National Statistics.\textsuperscript{20} Residual life expectancy takes into account that a person aged, for example, 80 years is more likely to live to age 81 years than someone aged 70 years and is a useful measure of relative life expectancy. For patients with two eligible eyes with series of VF data, we determined whether LSY would be predicted to be longer than 3 years. We then calculated the percentage of patients in each centre with this attribute.

**Frequency of VF examination** in clinics was simply estimated as the average interval (months) between recorded VF examinations during the follow-up period in those patients with series of VF data.

Summary measures and distributions of these metrics were evaluated and compared for the five glaucoma clinics. We used medians and IQRs along with conservative non-parametric tests to make simple, illustrative comparisons between clinics. All analyses were done using R (R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, URL: http://www.R-project.org, 2008).\textsuperscript{21}

**RESULTS**

Application of the inclusion and exclusion criteria resulted in 22379 VFs from 25 760 patients. Therefore 65.2\% (n=48234) of people were excluded because they only had one VF examination or normal VFs in both eyes in their first or second VF examination. These people were excluded from further analysis. Series of VFs (more than 4 years of follow-up) were available for 19264 patients. Summary measures of the metrics for assessing glaucoma service delivery for the five centres (and the aggregate data) are given in table 2.

There was a statistically significant difference in the median age at presentation between the centres (p<0.001; Kruskal-Wallis test). Centre 1 had slightly older patients with little difference between the other centres. There was also a statistically significant difference between centres for percent of unreliable VFs (p<0.001; $\chi^2$ test). Centre 1 (despite having older test takers) returned the lowest proportion of unreliable VFs. Differences between centres were generally small. Yet, for example, the rate of unreliable VFs recorded at centre 3 was 1.7 times higher than that of centre 1 (95\% CI for the relative risk of 1.4 to 1.9). It is noteworthy that around 1 in 20 VF examinations were unreliable according to this proxy measure across all centres.

**Stage of VF loss at presentation** is also summarised as a series waffle (traffic light) plots for each centre in figure 1. Centre 1 had notably more patients presenting with advanced VF loss when compared with the other centres. (Note this value summarises percent with advanced loss for those who present with some sort of VF loss. The value would be lower if the denominator was all people in the clinics.)\textsuperscript{22}

**Frequency of VF examination** at the five clinics is summarised as distributions in figure 2. The distributions were generally similar with the exception of centre 1 (and to a lesser extent centre 2); these distributions exhibited much heavier tails which means there were more eyes experiencing faster speed of loss VF loss at these centres compared with the others.

**Risk of VF loss blindness** as estimated by percentage of patients predicted to have LSY >3 years was slightly greater in centre 2 (12.9\%) compared with the other centres and this was statistically significant (p<0.001; Kruskal-Wallis test).

**Frequency of VF examination** at the five clinics is summarised by distributions of the time interval in months betweenVF tests in figure 3. Centre 4 had a noteworthy longer median time interval between VF tests (15.4 months) compared with the other centres and this was statistically significant (p<0.001; Kruskal-Wallis test).

One way of visualising speed of VF loss in clinics is to use a Hedgehog plot.\textsuperscript{18} One of these is shown for centre 1 in figure 4.

**DISCUSSION**

This work confirms that routinely collected VF data from glaucoma clinics can be downloaded, aggregated and...
anonymised from an EMR and be made available for analysis. This has previously been done for research purposes yet, here we illustrate how it could be done to potentially monitor and compare health service delivery at different centres—that optimising use of VF resources could improve clinical management of patients and save money at the same time. For this exercise, assessments were done on anonymised centres with a key point of this work showing that this approach is feasible for future implementation. We did not assess the logistics of the VF data extraction, which in this instance was carried out as part of the NOD work commissioned by HQIP. We have shown that assessment of VF records in glaucoma clinics could provide a first step towards quality improvement of services; this is a novel idea. For example, we have demonstrated how VF metrics of late presentation of glaucoma, or speed of loss of VF in people in glaucoma clinics during follow-up could be easily summarised for a clinical centre. We also have shown how it might be possible to compare the reliability and use of VF measurement between centres—for instance, it is feasible to identify centres that are doing more or fewer VF examinations compared with others. The latter is important because it has been shown previously, via health economic modelling of retrospective data, that optimising use of VF resources could improve clinical management of patients and save money at the same time. Since VF data can be held in EMRs this makes them amenable to automated and live analysis. Our feasibility study indicates that this approach could be used to monitor what happens to people in different glaucoma clinics in real time. Assessment of quality improvement of glaucoma services with an implementation of this idea could be the subject of future work.

The results from this report illustrate the feasibility of calculating metrics for assessing service delivery between centres using VF records alone. Nevertheless, some discussion about the differences in these metrics between the five anonymous centres is worthwhile. For example, centre 1 had more patients presenting with advanced VF loss and had more patients losing VF at a faster speed than other centres. These variables are highly associated in people with glaucoma and are, in turn, positively associated with older age. It is noteworthy that patients in centre 1 were generally older than those in the other centres and this might, at least in part, explain these observations. Still, our chosen metric for risk of VF loss blindness as estimated by percentage of patients predicted to have LSY >3 years was 12.5% in centre 1. This observation underlines the importance of preventing late detection of VF loss for prevention of avoidable blindness.

The main metrics highlighted in the Methods section are given in bold font. Median (and IQR) values or percentages are reported as summary measures. Patients with VF loss, n

### Table 2

This table shows the summary measures of a number of metrics across the five centres (and a total column). The six main metrics highlighted in the Methods section are given in bold font. Median (and IQR) values or percentages are reported as summary measures.

<table>
<thead>
<tr>
<th>Centre 1</th>
<th>Centre 2</th>
<th>Centre 3</th>
<th>Centre 4</th>
<th>Centre 5</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with VF loss, n</td>
<td>1373</td>
<td>2404</td>
<td>11589</td>
<td>5713</td>
<td>4681</td>
</tr>
<tr>
<td>VFs, n</td>
<td>7813</td>
<td>19654</td>
<td>121939</td>
<td>36011</td>
<td>35737</td>
</tr>
<tr>
<td>Median (IQR) MD at presentation (dB)</td>
<td>76 (67, 80)</td>
<td>71 (61, 78)</td>
<td>71 (61, 78)</td>
<td>71 (62, 79)</td>
<td>69 (58, 77)</td>
</tr>
<tr>
<td>% of unreliable VFs</td>
<td>3.0</td>
<td>4.4</td>
<td>5.1</td>
<td>4.3</td>
<td>4.1</td>
</tr>
<tr>
<td>% patients with advanced VF loss at presentation</td>
<td>29.6</td>
<td>25.0</td>
<td>24.2</td>
<td>23.0</td>
<td>22.8</td>
</tr>
<tr>
<td>Patients* with VF series &gt;4 years, n</td>
<td>843</td>
<td>1786</td>
<td>9208</td>
<td>3917</td>
<td>3480</td>
</tr>
<tr>
<td>Median (IQR) MD loss per year (dB/year)*</td>
<td>-0.37 (-1.2, 0.32)</td>
<td>-0.26 (-0.90, 0.20)</td>
<td>-0.23 (-0.83, 0.21)</td>
<td>-0.19 (-0.81, 0.27)</td>
<td>-0.10 (-0.71, 0.37)</td>
</tr>
<tr>
<td>% Patients LSY &gt;3 years</td>
<td>12.5</td>
<td>12.9</td>
<td>11.8</td>
<td>10.1</td>
<td>10.0</td>
</tr>
<tr>
<td>Median (IQR) interval between VFs (months)</td>
<td>12.6 (8.6, 18.1)</td>
<td>10.3 (7.6, 13.6)</td>
<td>9.5 (7.2, 13.0)</td>
<td>15.4 (11.5, 21.1)</td>
<td>12.3 (9.6, 16.5)</td>
</tr>
</tbody>
</table>

**Note:**

- LSY, Loss of Sight Years; MD, mean deviation; VF, visual field.

Figure 1  A traffic-light (waffle) plot showing the classification of visual field (VF) loss of newly presenting patients in each of the five centres. Each square represents 1% of the patients being classified as early (green), moderate (yellow) or advanced (red) VF loss on presentation.

Figure 2  A plot showing the distribution of the speed of mean deviation (MD) loss (dB/year) in each of the five centres. Each line represents one of the centres. Centre 1 (red) has a heavier tail compared with the other centres, indicating a higher proportion of patients with higher speeds of loss. Centre 1 also has a lower kurtosis (tail-to-peak ratio) and more negative skewness than the others. The coloured symbols on the x-axis indicate the median value for each centre (see table 2).

Figure 3  A plot showing the distributions of the time interval between patient visits (in months) in each of the five centres. Each line represents one of the centres. Centre 4 is more positively skewed, indicating a higher proportion of patients with longer intervals between visits. The coloured symbols indicate the median time interval between visits for each centre (see table 2).

Figure 4  A hedgehog plot showing the mean deviation loss over time of each eye in centre 1. Eyes highlighted in red indicate a speed of loss worse than −1.5 dB/year. The region marked in blue is a threshold for a severely impaired visual field. More detail on these plots can be found in Bryan et al.18

in these data sets had single ‘one-off’ VFs or had normal VFs at presentation to the clinics. This figure illustrates the huge volume of likely false-positive glaucoma referrals, glaucoma suspects and ocular hypertensive that glaucoma clinics deal with on a daily basis. Moreover, we found that around 1 in 20 VFs are recorded with reliability indices outside normal limits. In addition, the median interval between VF tests in people being followed over time was 11.2 months and as high as 15.4 months in one centre. This supports previous findings that annual VF testing is the norm for most patients in glaucoma clinics in England.6 Heath economic modelling has highlighted the benefits of stratifying patients to more or less VF monitoring based on age and stage of disease at diagnosis; a prospective study is needed to prove these findings.3

Using EMRs for research or audit is not a new idea, having been implemented in many fields of medicine to study diseases such as diabetes, heart failure, cancer and asthma.26 In eye clinics EMRs have been used for audit of cataract surgical outcomes but importantly also have
potential for making healthcare delivery more efficient by facilitating more streamlined clinical workflow, better patient management and improved data tracking. However, ideas and implementation are different entities and meaningful use of EMRs in ophthalmology is still a work in progress. At present, certainly in clinics in England, there are challenges about how clinical data are recorded, archived and stored. Moreover, issues such as non-collated data sets, duplicate IDs and differing databases add to the challenge of moving towards comprehensive use of EMRs.

Our approach to the concept of assessing service delivery in glaucoma clinics has several strengths. In theory, VF records should be easily stored on EMRs, making them amenable to easy extraction, analysis and audit. Whether this happens in practice is dependent on the motivation of implementing electronic archiving of VF examinations and use of EMRs in HES. Another rational for using VF metrics is their relevance to measuring vision. IOP is the only modifiable risk factor for glaucoma progression and is crucial to patient management but VF metrics will best estimate the status of people’s vision loss in the clinic. Our approach does not consider everyone in glaucoma clinics but centres on those with a proxy measure of sustained VF loss. These patients are at higher risk of further significant vision loss in their lifetime compared, for example, with people with OHT.

There are limitations to the ideas that we report in this study. The main problem is reliance on the VF data alone. A more complete assessment of what is happening in a glaucoma clinic would be achieved by considering exact diagnosis, treatment regimens, IOP, optic nerve head characteristics, individual patient history or other risk factors. For example, much can be learnt about what is happening in glaucoma clinics by reviewing repositories of data on prescribed medications. However, EMRs need an established and standardised minimum data set for glaucoma care and this is the subject of future work. Our idea will also be limited to how well VF records can be archived at a centre. The five sites chosen for this study were all EMR enabled and known to run large glaucoma services with aggregated electronic VF databases. Alternative options could include separate data extractions from individual machines with subsequent aggregation into a single database, but this would be time consuming and carry significant cost, in particular if these services were delivered in different settings such as outreach clinics.

There are also limitations to some of the metrics we have proposed for assessing service delivery. First, while MD is a useful summary measure of how much sensitivity loss there is in a VF and particularly convenient to monitor changes over time, it is not a perfect measure for glaucomatous VF loss. MD can be affected by non-glaucomatous changes such as a general reduction inVF sensitivity caused by, for example, cataract. Second, as noted previously, we only used the FP reliability index as a measure of patient test taking performance. Of course, this measure, or any other similar measure, would not capture patients failing to complete an examination or those excluded because of a previous failure to reliably conduct a VF examination. Third, our measure of “risk of blindness” (LSY) makes a number of assumptions around residual life expectancy and progression of VF loss being constantly linear. Finally, when comparing metrics it will also be important to consider some centres simply differ in terms of population factors (eg, racial and socioeconomic profile) and audits using the methods proposed in our report would have to take this into account.

In conclusion, this study illustrates the feasibility of assessing some aspects of quality of care in glaucoma clinics through analysis of VF databases from EMR enabled centres. This approach, which is outcome focused, is a potentially useful method for assessing blindness prevention from glaucoma in secondary care centres. VF testing technology is standardised in the UK NHS, and although in many centres the electronic VF tests will be distributed across several VF testing machines, it is feasible to aggregate these fields into a central database located in each centre for central analysis. Ideally, such a central field database would reside within a specialty specific EMR implementation serving both clinical and quality assurance needs. Secondary benefits from such an approach would include the ability to more easily detect patients whose VF loss is progressing rapidly in order to intensify their treatment as well as detection of those patients whose VFs are stable who may require less intensive monitoring once VF stability has been documented. By shifting focus towards those in most need, health services resources can be more effectively used. In the current NHS digital environment, a variety of challenges would need to be overcome in order to extend this audit approach into a national audit of vision preservation in people with glaucoma.

Acknowledgements The authors would like to thank Beth Barnes (head of professional support), Paul Donachie (medical statistician) and the rest of the audit team at the national ophthalmology database https://www.nodaudit.org.uk/about/team.

Contributors SRK: design, analysis and interpretation of data, drafting the article, revising the article. SRB: design, analysis and interpretation of data, drafting the article. JS: conception, design, revising the article. DPC: conception, design, interpretation of data, drafting the article, revising the article.

Funding Data were electronically extracted from contributing hospitals by the Royal College of Ophthalmologists audit team as part of the Healthcare Quality Improvement Partnership commissioned National Ophthalmology Database Audit (which forms part of the National Clinical Audit and Patient Outcomes Programme). SRK received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Grant Agreement No. 675033. The listed funding organisations had no role in the design or conduct of this research.

Competing interests DPC reports speaker fees from Allergan, Bayer, Santen; unrestricted funding from Allergan, Roche, Santen; consultancy with Centervue—all outside the remit of the submitted work.

Patient consent for publication Not required.

Ethics approval All patient data were anonymised and transferred to a single secure database. No other clinical data were made available apart from each patient’s age. Subsequent analyses of the data were approved by a research ethics committee of City, University of London and this study adhered to the tenets of the Declaration of Helsinki.
Provenance and peer review  Not commissioned; externally peer reviewed.

Open access  This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) licence, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES


