
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/20245/

Link to published version: http://dx.doi.org/10.1136/bjophthalmol-2017-310440

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.

City Research Online: http://openaccess.city.ac.uk/ publications@city.ac.uk
Research article:
An example of monitoring measurements in a virtual eye clinic using “big data”

Authors:
Lee Jones¹, Susan R Bryan¹, Marco A Miranda²³, David P Crabb¹, Aachal Kotecha⁴

Affiliation:
1. Division of Optometry and Visual Science, School of Health Sciences, City, University of London.
2. NIHR Biomedical Research Centre (BRC) at Moorfields Eye Hospital and UCL Institute of Ophthalmology, London.

Corresponding Author:
Dr. Aachal Kotecha,
Glaucoma Service
Moorfields Eye Hospital NHS Foundation Trust
162 City Road, London, EC1V 2PD
Tel: +44 7608 4015
Email: Aachal.Kotecha@moorfields.nhs.uk

Word count: 2,493
Key words: virtual clinic, glaucoma, visual fields
Synopsis:

This study uses “big data” to assess outcomes of patients attending a virtual glaucoma monitoring service. The findings suggest this new model of service delivery is a viable means of monitoring low-risk glaucoma patients.
Abstract

**Aim:** To assess the equivalence of measurement outcomes between patients attending a standard glaucoma care service, where patients see an ophthalmologist in a face-to-face setting, and a glaucoma monitoring service (GMS).

**Methods:** The average mean deviation (MD) measurement on the visual field (VF) test for 250 patients attending a GMS were compared to a “big data” repository of patients attending a standard glaucoma care service (reference database). In addition, the speed of VF progression between GMS patients and reference database patients were compared. Reference database patients were used to create expected outcomes that GMS patients could be compared to. For GMS patients falling outside of the expected limits, further analysis was carried out on the clinical management decisions for these patients.

**Results:** The average MD of patients in the GMS ranged from +1.6 dB to -18.9 dB between two consecutive appointments at the clinic. In the first analysis, twelve (4.8%; 95% confidence interval [CI] 2.5 to 8.2%) GMS patients scored outside the 90% expected values based on the reference database. In the second analysis, 1.9% (95% CI 0.4 to 5.4 %) GMS patients had VF changes outside of expected 90% limits.

**Conclusions:** Using “big data” collected in the standard glaucoma care service, we found that patients attending a GMS have equivalent outcomes on the VF test. Our findings provide support for the implementation of virtual healthcare delivery in the hospital eye service (HES).
Introduction

The United Kingdom (UK) National Health Service (NHS) is facing unprecedented challenges. Although overall life expectancy is increasing, with it comes a greater prevalence of disease in the population, (1) and chronic disease management remains a significant burden on the NHS. (2) There is a need for the NHS to redesign its services to make a more efficient healthcare service provider.

A drive exists for the NHS to make more use of information technology (IT). (3) One such example is the development of virtual clinics, which remove the face-to-face doctor-patient consultation. Within the hospital eye service (HES), virtual clinics have not only been found to provide valuable additional out-patient capacity, but can also streamline referral rates, reduce costs, and improve the patients’ health care experience. (4-7) Improvements in disease detection by primary eye care service providers have meant that the HES has become one of the busiest health care providers in the UK. (8-11) As a result, the introduction of new methods to assist with the monitoring of patients with chronic ocular disease in the HES is a high priority.

Virtual clinics offer a viable means of monitoring glaucoma patients. (5, 12-14) To date, most studies have focused on the accuracy of disease staging, as well as patient satisfaction, cost reduction, and appointment durations. (4, 14-15) An important safety aspect of virtual clinics is whether disease progression can be identified and acted on effectively. By doing so, scrutiny can be placed on the extent to which virtual clinic patients differ from patients in consultant-led appointments when performing the same tests. This type of analysis can be conducted through an audit-style assessment using large scale data.

Following the development and expansion of the Internet, as well as the advent of new and innovative technologies, the use of large scale data, or “big data”, has increased dramatically in recent years. (16) Put simply, large databases from routine services can be used to compare individual or population results of patients attending a single hospital, practice, or clinic. This method has been used recently in the field of ophthalmology. (17-18)

A virtual clinic must be effective at identifying patients who have become unstable and are in need of closer observation. One method to assess this aspect of virtual clinics
is to use large scale data collected from consultant-led appointments as benchmarks for patient’s measurement results. This is the idea explored in this current work.

In this study, we examine the effectiveness of a virtual glaucoma monitoring service (GMS) at identifying unstable patients requiring closer observation. In addition, we assess whether “big data” analysis can be used to identify patients achieving visual field (VF) test scores outside of the expected range.
Materials and methods

Following authorisation from the Caldicott Guardian and Information Governance Lead at Moorfields Eye Hospital NHS Foundation Trust (MEH), anonymised VF results from the Humphrey Field Analyser (HFA; Carl-Zeiss Meditec, Dublin, CA, USA) of patients attending the glaucoma monitoring service (GMS) were analysed. The GMS, and criteria for patient inclusion into the service, is described in a previous publication from our group. (4) In brief, clinical examinations of ‘early’ and ‘moderate’ disease stage glaucoma patients are carried out by trained ophthalmic technicians and data are reviewed by two consultants and a senior glaucoma specialist optometrist on a different day for clinical management decisions. (4)

Data from VFs were extracted with the optical character recognition function using a purpose-written program authored in MatLab 2016b (Mathworks Inc.). Data were restricted to patient’s age, test date, test eye, test reliability, and mean deviation (MD). The latter is conventionally used in clinics; it is a summary measure of the overall reduction in VF sensitivity relative to a group of healthy age-matched observers, with more negative values indicating a worse VF. We used patients' worse eye (based on MD) at their first GMS visit as our study eye.

Inclusion criteria required patients to have at least two visits to the GMS no less than 4 months apart. The first 250 patients who entered the GMS since its start in 2014 who fulfilled this criterion were analysed. This sample represents approximately 15% of patients attending the GMS at the time of data collection.

Data Analysis

We used the difference in MD between GMS patients’ baseline and second appointment as a surrogate of VF stability. Large differences would suggest a change in the VF, or poor repeatability.

Limits were defined for change in MD from a database of 473,252 VF records (the reference database). These data are described elsewhere (19) and were pooled from 88,954 patients from four centres in the United Kingdom: Moorfields Eye Hospital NHS Foundation Trust, London; Cheltenham General Hospital Gloucestershire Eye Unit; Queen Alexandra Hospital, Portsmouth; and the Calderdale and Huddersfield NHS
Foundation Trust. Only patients tested using the 24-2 testing algorithm were included, resulting in a total of 83,794 patients.

Patients attending the GMS are all experienced in perimetry, but this may not necessarily be true for patients within the reference database. Thus, we excluded the first ever VF test for patients within the reference database to allow for perimetric 'learning'. Furthermore, eliminating the first visual field in a patient’s series would also exclude patients from the reference database who had a single, exploratory VF, thereby increasing the confidence that patients remaining in the reference database who had at least 2 subsequent VFs were being monitored for glaucoma. In this database, 41,048 patients (49%) were excluded based on this criterion. In addition, we restricted the age (minimum age of 20 years) for the reference database to ensure that these patients were age-related to patients seen in the GMS. Duration between appointments in the reference database was restricted to between 4 and 24 months to ensure similar time intervals for follow up between the two groups. After applying these criteria, 22,124 patients remained in the reference database. We then grouped the average MDs by VF defect severity using bins of 1dB width. We did this because VF measurement repeatability is strongly associated with VF severity. (20) The 5th, 25th, 75th and 95th percentile of the distribution of difference were then derived and plotted. Points were connected using a locally weighted smoothing operator (LOESS) to create a colour coded chart for the 50% and 90% limits of change for MD in the reference database (See figure 1). We would, for example, expect 10% of GMS patients to have repeat MD differences outside of the latter limits. By using this method, GMS patients who had VF results that were markedly different to those in the reference database (i.e. outside the 90% normal limit) could be identified.

For the second part of the analysis, we included a subset of the GMS patients who had attended three or more appointments (N=158). GMS patients with three VFs were compared to patients in the reference database with three VFs; this was repeated for patients with four and five VFs. Where two VFs were conducted within 4 months, the patients’ next measurement in their VF series within our inclusion criteria was used. Simple linear regression was used to calculate the rate of VF progression (MD dB loss per year). Regression lines for the reference database were plotted using a novel data visualisation tool, the Hedgehog Plot. This tool allows us to visualise the progression rates for all patients simultaneously. The reference database was used to determine the 90%
limits by computing the 5th and 95th percentiles of the estimated slopes. The regression lines for the GMS patients were then superimposed onto the Hedgehog plot and eyes which were found to be outside of the calculated limits were flagged.

The clinical management decisions for all 250 GMS patients at the time of their most recent field were also collected. All statistical analysis was done in R (www.R-project.org).

Results

Median (interquartile range; IQR) age of GMS patients at first visit was 65 (54, 72) years. Median (IQR) MD for GMS patients’ worse eye at baseline was -1.5 (-3.1, -0.3) dB. Median (IQR) number of months between the first and second GMS clinic appointments was 12 (10, 12) ranging from 4 months to 21 months. Average MD of the first and second appointments of the 250 GMS patients ranged between +1.6 dB and -18.9 dB (median -1.4 (-3.0, -0.4) dB). Median (IQR) age of patients from the reference database was 67 (57, 76) years.

Of the 250 GMS patients, 12 (4.8%; 95% confidence interval [CI] 2.5 to 8.2%) recorded values outside the 90% limits. This proportion was less than the expected value of 10% and was statistically significant (p = 0.003). (A post-hoc power calculation confirms our study to have had an adequate sample size. A total of 4.8% outside the 90% limit returns a power (beta) value of 0.86 when alpha is set at 0.05 and N = 250; Minitab 17 Statistical Software (2010); www.minitab.com).

Figure 2 shows the results of the GMS patients (points) compared to the reference database. We split the GMS patients into three equally sized groups. Plot A ranges from -18.9 dB to ≤ -2.4 dB (83 GMS patients), plot B ranges from > -2.4 dB to ≤ -0.8 dB (84 GMS patients), and plot C includes patients > -0.8 dB (83 GMS patients).

Figure 3 shows the rate of VF progression for patients in the reference database. Each line represents an eye, with the length of the line indicating the length of follow-up. The location of the line is aligned to the patient’s age (x-axis) and severity of initial loss.
(y-axis); steeply declining lines indicate rapidly progressing eyes. After applying our inclusion criteria, 18,414 reference database patients were included.

In Figure 4, GMS patients progressing (red lines) and improving (green lines) faster than the 90% limit in the reference database with the same number of VFs are highlighted. Three (1.9%; 95% CI 0.4 to 5.4%) patients are flagged as having VF changes outside of expected limits.

Table 1 shows the diagnoses of the 14 patients identified by both analyses as having VF MD changes outside the 90% limits of the reference database. Nine patients performed worse than patients in the reference database, although five of these had a positive MD at their baseline GMS visit. Of the remainder, one patient was judged to be progressing by the GMS reviewer, one patient had a retinal arterial occlusion unrelated to their glaucoma, one patient was exited from the GMS due to suspected unreliable VF performance, and one was deemed stable by the GMS reviewer and kept in the clinic.

Figure 1. Distribution of expected limits for size of difference in MD (dB) index between two appointments in the reference database (N = 22,124) based on the patient’s average
MD. Areas at the upper and lower most part of the plot (red) show results outside of 90% normal limit. The lower most part of the plot indicates worsening VF results.

**Figure 2.** Virtual GMS patient (points) data compared against the reference database. Our total sample (N=250) is divided into 3 groups based on average MD. Plot A ranges from -18.9 dB to ≤ -2.4 dB (83 virtual GMS patients), plot B ranges from >-2.4 dB to ≤-0.8 dB (84 virtual GMS patients), and plot C includes patients > -0.8 dB (83 virtual GMS patients).

**Figure 3.** Hedgehog plot showing the rate of VF progression in the reference database (N=18,414). Three patients have been highlighted. For each patient, a point represents a score on the VF test and patients’ age at time of test. A regression line is fitted for each.
patient using all of the points in their series. Steeply declining lines indicate faster VF progression. The blue shaded area denotes likely visual impairment. In this example, patient 1 has the most VF tests in their series and shows a faster rate of progression than patient 2 or 3.

**Figure 4.** Of the 250 virtual GMS patients, 158 (63.2%) had 3 or more VFs in their series. These patients are superimposed on the reference dataset Hedgehog Plot. Grey lines show rates of progression for patients in the reference database. Darker lines show virtual GMS patients. Red lines highlight the virtual GMS patients outside the 90% limits for progression. The green line highlights the virtual GMS patient outside the 90% limits for ‘improvement’. 
Table 1. Outcomes of GMS patients identified as falling outside of the 90% limits in the reference database.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Diagnosis</th>
<th>MD at GMS (dB)</th>
<th>Progressing (P) or 'improving' (I)</th>
<th>GMS outcome recorded in notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>OHT</td>
<td>+1.03</td>
<td>P</td>
<td>Remains in GMS; 12-month review</td>
</tr>
<tr>
<td>52</td>
<td>OHT</td>
<td>-22.59</td>
<td>I</td>
<td>Exited GMS in January 2015; noted to be poor VF performer; moved to consultant clinic for further follow up.</td>
</tr>
<tr>
<td>63</td>
<td>OHT</td>
<td>-2.60</td>
<td>I</td>
<td>Remains in GMS; 9-month review</td>
</tr>
<tr>
<td>65</td>
<td>OHT</td>
<td>-2.08</td>
<td>I</td>
<td>Remains in GMS; 12-month review</td>
</tr>
<tr>
<td>67</td>
<td>OHT</td>
<td>+1.45</td>
<td>P</td>
<td>Remains in GMS; 6-month review</td>
</tr>
<tr>
<td>71</td>
<td>OHT</td>
<td>+1.56</td>
<td>P</td>
<td>Remains in GMS; 12-month review</td>
</tr>
<tr>
<td>75</td>
<td>OHT</td>
<td>+0.03</td>
<td>P</td>
<td>Remains in GMS; 12-month review</td>
</tr>
<tr>
<td>77</td>
<td>OHT</td>
<td>-2.02</td>
<td>P</td>
<td>Remains in GMS; 18-month review</td>
</tr>
<tr>
<td>78</td>
<td>OHT</td>
<td>-3.24</td>
<td>I</td>
<td>Remains in GMS; 12-month review</td>
</tr>
<tr>
<td>42</td>
<td>Glaucoma suspect</td>
<td>+1.55</td>
<td>P</td>
<td>Remains in GMS; 12-month review</td>
</tr>
<tr>
<td>55</td>
<td>Glaucoma suspect</td>
<td>-4.17</td>
<td>P</td>
<td>Exited GMS April 2016; discharged from service- no evidence of glaucoma, poor VF performer.</td>
</tr>
<tr>
<td>77 *</td>
<td>Glaucoma suspect</td>
<td>-2.16</td>
<td>P</td>
<td>Retinal arterial occlusion; detected in March 2016 (i.e. pre-glaucoma service visit); glaucoma stable.</td>
</tr>
<tr>
<td>80</td>
<td>Glaucoma suspect</td>
<td>-13.78</td>
<td>I</td>
<td>Exited GMS June 2016; no evidence of glaucoma, poor VF performer; moved to consultant clinic to assess suitability for discharge from glaucoma service.</td>
</tr>
<tr>
<td>82</td>
<td>POAG</td>
<td>-7.97</td>
<td>P</td>
<td>Exited GMS April 2016 as evidence of progression. Review in consultant clinic.</td>
</tr>
</tbody>
</table>
**Key.** MD = mean deviation, GMS = glaucoma monitoring service, OHT = ocular hypertension, POAG = primary open angle glaucoma, VF = visual field.

* shown to be significantly progressing with both analyses 1 & 2
Discussion

Our study exploited a “big data” approach to investigate whether patients in a GMS score similarly on a measure of vision loss (i.e. visual field MD) to patients who attend consultant-led appointments (reference database). Our results show the difference in MD values between two hospital appointments for patients attending the GMS is similar to those in the reference database.

Using “big data” we created ranges of expected change in MD over a similar follow up interval, using patients with a similar profile to those attending the GMS. Our findings also showed that 12 patients in the GMS scored outside the expected range on the VF test. The proportion of GMS patients outside the expected range (4.8%) is smaller than the 10% (25 patients) we allowed for. Similarly, when compared to the reference database, there were fewer GMS patients with unusually fast progression. The results of our study indicate that the number of patients in a GMS performing better or worse than expected on the VF test is smaller than anticipated.

The results of our study are relevant to current clinical practice with regard to monitoring patients with glaucoma. Our findings suggest that patients attending a GMS are no ‘worse-off’ than those attending the standard-care appointments. Specifically, when using VF data from a large reference database as benchmarks for expected changes in MD score, GMS patients’ VF test results tended to be as expected. In the few cases where GMS patients’ MD scores were outside the expected results, further analysis on these patients was carried out. It was found that 5 patients scoring outside the expected results had been highlighted as ‘improving’. Of the 7 patients showing a worse performance compared to the reference database, 3 had been picked up by the GMS reviewer, with the remainder being deemed stable. It should be noted that those deemed clinically stable had a diagnosis of either ocular hypertension or suspected glaucoma with no significant VF defect.

A strength of our study is the number of patients we included in our analysis. Access to a wealth of VF entries in the reference database, even after applying sensible selection criteria, meant that we closely matched patients in the GMS to patients attending standard, consultant-led appointments. The total number of patients in the reference database for the first part of our analysis was 22,124 and 18,414 for the second
part. These large numbers allowed limits for variability to be stratified by disease severity (21-23).

It should be noted that the number of patients in the reference database used to create our expected limits is not equal across all average MD bins on the x-axis of our plots. For example, the number of reference database patients creating the -15 dB average MD limit was 257, whereas the number of patients creating the -5 dB average MD limit was 1,350. However, we wished to include as much data as possible and so, given that there were simply more patients with average MD of -5 dB than -15 dB, this disparity is to be expected. Additionally, the -15 dB limits were where the fewest reference database patients were included (N=257), but this number of patients remains substantial. A further point to consider is that in Figure 4 it appears older patients have less stable VFs in the GMS. This could be due to these patients presenting at a later stage in the disease, precipitating more VF variability, or they have had the disease for a longer period of time, or they are worse test takers.

A limitation to our study is the inclusion criteria we used to construct the reference data percentiles (Figure 1). We match GMS patients to reference database patients using baseline MD, age, and interval between clinic visits. We did not have access to reference database patients’ diagnoses. GMS patients are a highly selected sub-group of glaucoma patients attending the Moorfields Eye Hospital glaucoma outpatients service; some reference database patients would not be suitable for virtual monitoring. For example, glaucoma patients with a coexisting ocular comorbidity would not be suitable for GMS but may be present in the reference database and this represents a possible confounder. Furthermore, we anticipated that patients in the GMS would be experienced in performing the VF test. However, some GMS patients appeared to show improvement in their MD scores; these patients may be unreliable at performing the VF test or are continuing to have perimetric learning effects despite being experienced test-takers. We did not have data for variables such as intraocular pressure, or optic nerve assessment which may influence progression. This is a key limitation. Further analysis adjusting for these factors would be a valuable addition to the literature. Patients in the “big data” (reference) group are simply defined as having measurable glaucoma-like VF loss who are attending glaucoma clinics. Therefore, for example, we cannot rule out some patients having optic neuropathies that produce glaucoma-like VF deficits, but the number would
be insignificant given the sheer number of records in the reference database. Moreover, for example, patients with sudden onset retinal vein occlusions or unstable aggressive glaucoma may skew the expected parameter limits in the reference database. However, as the reference database is comprised of patients attending glaucoma clinics, the number of those with VF loss due to non-glaucomatous comorbidities is likely to be smaller than that reported in general population prevalence estimates. (24) But these examples do highlight some limitations of the "big data" approach. A final limitation surrounds the method used to assess change between VFs. Here we have used a VF index (MD) and alternative methods using all the points in the VF might offer more sensitivity to change. (25)

The average number of appointments for GMS patients in our study was three. Further analysis where patients attending a GMS are followed longitudinally may provide more information regarding the suitability of virtual monitoring. Given that GMS are a relatively new addition to the HES, this idea should be revisited in future research.

To our knowledge, this is the first study that utilised “big data” to evaluate outcomes of patients in a GMS. The utility of pooling large databases together to identify trends and also predict future risks to health is recognised elsewhere. (26) In the presented study, we have utilised “big data” to assess whether a new model of service delivery results in equivalent outcomes to that of the standard out-patient model, and for the metric we used (i.e. MD), we found that it did. The digital nature of VF test results lends itself to “big data” analysis. Still, the VF result is but one measure of glaucoma status. However, we feel that this study has shown the potential of using “big data” in the ophthalmology setting to confirm the equivalence of care between a new and standard model of service delivery.

**Contributors:** LJ: Data analysis, data interpretation, manuscript preparation. SRB: Data analysis, data interpretation, manuscript critique. MAM: data analysis, manuscript critique. DPC: Study design, data interpretation, manuscript critique. AK: Study design, data interpretation, manuscript critique.

**Conflict of interest:** The authors declare no conflict of interest.
References


