BACKGROUND: Cataract waiting lists are growing globally. Pragmatic, cost-effective methods are required to prioritise the most urgent cases. Here we investigate the feasibility of using a third-party pen-and-paper contrast sensitivity, CS, test (SpotChecks<sup>TM</sup>), delivered by mail, and performed by patients at home unsupervised, to flag eyes requiring surgery.

METHODS: Pen-and-paper CS tests were mailed to 233 people waiting for a cataract assessment, along with a prepaid return envelope (cross-sectional study). Response rates were tabulated (stratified by age, sex and socioeconomic status), and test scores analysed to see how well the home tests predicted which eyes were listed subsequently for surgery. A subset of patients (N = 39) also underwent in-person follow-up testing, to confirm the accuracy of the home data.

RESULTS: Forty-six percent of patients responded (216 eyes). No gross differences were observed between respondents and non-respondents, either in terms of age, sex, socioeconomic status, or geographic location (all P > 0.05). The home-test CS scores predicted which eyes were subsequently listed for surgery, with an AUROC (±CI<sub>95%</sub>) of 0.69 (0.61–0.76). Predictive performance was further-improved when machine learning was used to combine CS scores with letter acuity, extracted from patients’ medical records (AUROC (±CI<sub>95%</sub>) = 0.77 (0.70–0.83)). Among 39 patients who underwent follow-up testing, home CS scores were correlated with various measures made in clinic: biometry signal-to-noise (P = 0.032), LogMAR acuity, Pelli-Robson CS and SpotChecks CS (all P < 0.001).

CONCLUSIONS: Mailing patients pen-and-paper CS tests may be a feasible, ‘low-tech’ way of prioritising patients on cataract waiting lists.

Eye; https://doi.org/10.1038/s41433-024-03081-6

INTRODUCTION

Patients and physicians agree that cataracts should ideally be treated within 3 months of diagnosis and that waiting times longer than 6 months are excessive [1–4]. Waits longer than 6 months are also associated with reduced quality of life, and increased risk of depressive symptoms, falls, and other life-changing accidents [5–10]: often exacerbating the burden on healthcare services long-term [11].

Historically, many health services have struggled to meet these targets. In the last decade, for example, patients typically waited: 1.5 months (United States Medicare [12]) 1–6 months (mainland Europe [13]), 3 months (Scotland [14]), or 8 months (Australia [15]) for surgery—often with an additional 3–12 months [14, 16] wait for the initial pre-surgical assessment following referral.

And even the best-performing services will face unprecedented strain as societies age, with the demand for cataract surgery forecast to increase by 50% over the next 20 years [17]. Even with more efficient practices (e.g. via same-day assessments and surgery [18], simultaneous bilateral extractions [19–21], dedicated operating rooms [22], out-of-office hours slots for cataract surgery [23], or by foregoing surgery altogether in low-impact cases [24]), cataract waiting lists are only likely to grow globally.

In light of this, there have been calls to revisit the longstanding question of how best to manage cataract waiting lists [25, 26]. Clearly, more urgent cases should be prioritised for treatment [1]. And while precisely how urgency should be calculated is a complex and contentious topic [27], one key determinant must be the severity of vision loss that the patient is currently experiencing. The question then becomes how to quantify patients’ current level of vision loss, in a way that is scalable and cost-effective—and does not further burden already overstretched health services?

‘Telemedicine’ may provide an answer: enabling patients to assess their own vision at home. The logistical hurdles are considerable, however. Providing digital testing equipment to millions of patients would be prohibitively expensive, and asking patients themselves (two thirds of whom are over 60 years old [28] and many with limited vision) to access and learn to use custom software using their own devices is unlikely to prove feasible, and risks ‘Digital Exclusion’ for a subset of the most vulnerable individuals [29].

In the present study we, therefore, took a novel, ‘low tech’ approach: examining whether it is feasible to simply post out a pen-and-paper assay of contrast sensitivity [30] (CS) to patients currently waiting for a cataract assessment. CS was preferred over visual acuity since CS is thought to be a more sensitive marker of
Methods. A Contents of the test pack posted to each patient (see body text for details). B Close-up of the SpotChecks test, including the circular targets that the user must circle or tick. There were six variants in total (A–F; only 4 of which are shown here). Patients were posted a random pair, with no duplication within patient. Note that the image has been enhanced for visibility—the spots in the bottom rows are not normally visible even to a normally sighted observer.
RESULTS

Response rates
Of the 233 patients posted a test pack, 108 (46.4%) responded. All 108 respondents successfully enclosed two completed SpotChecks tests (one per eye; \(N = 216\) eyes total). As shown in Supplemental Fig. S1, there were no obvious demographic differences between respondents and non-respondents, either in terms of age, sex, socioeconomic status, or geographic distance (all \(P > 0.1\); see Table 1 for statistics). One patient who didn’t respond was subsequently reported as deceased, but no systematic inquiries were made into the circumstances of the other 124 non-respondents.

Predicting future listing for cataract surgery
Figure 2 shows how well the home CS scores predicted which eyes were subsequently listed for cataract surgery following the patient’s next consultation (NB: the consultant was blind to the results of the SpotChecks home-testing data when determining which eyes to list). Note that for these analyses only 95 of the 108 respondents were included (i.e. \(N = 190\) eyes), as 13 patients had yet to attend an ophthalmic consultation at the time of writing. Two eyes were not listed for surgery, but were nevertheless scored as listed, as surgery was recommended but subsequently postponed (once by patient request; once due to more urgent medical complications).

Eyes listed for surgery tended (Fig. 2A), on average, to score significantly more poorly on SpotChecks \((t_{188} = 4.79; P < 0.001)\), indicating that the result of the home test was associated with the need for cataract surgery (see Fig. 2C for raw scores).

More directly, Fig. 2B (black line) shows how well SpotChecks predicted which individual eyes were subsequently listed for surgery. The resultant classifier had a Sensitivity = 73% at a Specificity = 54%, with an overall Area Under the ROC (±CI95%) of 0.69 (0.61–0.76). We also considered whether this score could be further improved by combining the SpotChecks data with other information easily obtainable from a patient’s medical record. For example, the dashed blue line in Fig. 2B illustrates the results of a machine learning classifier (Support Vector Machine) that combined SpotChecks scores with Snellen acuity. The model was trained and evaluated using a leave-one-out technique and was found to improve performance by approximately 10%. Thus, the resultant classifier had a Sensitivity = 79% at a Specificity = 61%, with an overall Area Under the ROC (±CI95%) of 0.77 (0.70–0.83). Adding other additional factors to the model (e.g. age, sex, biometry scores, a history of ocular disease [yes/no]) did not appear to further improve the classifier, though we did not explore this question exhaustively given the limited size of the data set (see Supplemental Fig. S3 for additional analyses).

Agreement with other biomarkers
As shown in Fig. 3, SpotChecks performed at home was weakly correlated with biometry signal-to-noise ratio \([r_{171} = 0.16, \ P = 0.032; \ \text{Fig. 3A}]\), and negatively correlated with letter acuity \([r_{188} = -0.49, \ P = 0.001; \ \text{Fig. 3B}]\). A subset of patients \((N = 39)\) also underwent a more detailed follow-up assessment. These data confirmed that SpotChecks performed at home were positively correlated with SpotChecks performed in clinic under supervision \([r_{3} = 0.71, \ P < 0.001; \ \text{Fig. 3C}]\), and were also correlated with the

Table 1. Associated statistics for Supplemental Fig. S2.

<table>
<thead>
<tr>
<th>Characteristics of patients who did/did not respond</th>
<th>Mean (SD) or %</th>
<th>Group comparison stats</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-respondents</td>
<td>respondents</td>
</tr>
<tr>
<td>Age (years)</td>
<td>74.8 (12.6)</td>
<td>75.1 (9.3)</td>
</tr>
<tr>
<td></td>
<td>(t_{221} = -0.22, P = 0.823)</td>
<td></td>
</tr>
<tr>
<td>Distance from hospital (km)</td>
<td>12.4 (8.3)</td>
<td>11.6 (7.7)</td>
</tr>
<tr>
<td></td>
<td>(t_{221} = 0.79, P = 0.428)</td>
<td></td>
</tr>
<tr>
<td>Index of multiple deprivation (rank)</td>
<td>25211 (2097)</td>
<td>25451 (1906)</td>
</tr>
<tr>
<td></td>
<td>(t_{229} = -0.91, P = 0.364)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>53.6% female</td>
<td>54.6% female</td>
</tr>
<tr>
<td></td>
<td>(X^2_{(1, \text{df}=233)} = 0.03, P = 0.875)</td>
<td></td>
</tr>
</tbody>
</table>

No statistically significant differences were apparent between mean values from the two groups, either in terms of age, location, IMD or sex.

Fig. 2 Ability of SpotChecks, performed at home, to predict which eyes were subsequently listed for surgery. A Mean [±95% confidence intervals] SpotChecks scores for eyes that were/were not subsequently listed for surgery, plus associated t-test values (see also Fig S2 for further analysis). B Receiver Operating Characteristics (ROCs) showing the ability of SpotChecks to predict which eyes were/were not subsequently listed for surgery. Shaded regions indicate the 95% confidence intervals. The black line indicates when SpotChecks data alone were used. The blue dashed line indicates when a Support Vector Machine was used to combine SpotChecks data and Snellen acuity scores (model trained and evaluated using leave-one-out analysis, using the following Matlab functions: fitcsvm.m, fitSVMPosterior.m, kfoldPredict.m, perfcurve.m). Numerical values show the Area Under the ROC [AUROC], plus 95% confidence intervals. C Histograms showing the distributions of raw SpotChecks scores.
results of the Pelli Robson letter chart: the clinical reference standard for CS ($r_{24} = 0.69$, $P < 0.001$; Fig. 3D).

Patient feedback
Thirty-six of these patients were asked if they would have preferred the home test to be delivered digitally. 72% (26 of 36) said they preferred the pen-and-paper testing approach, while 14% would have preferred a digital test and 14% expressed no preference.

DISCUSSION
This pilot study demonstrates the feasibility of using a low-cost pen-and-paper CS test to prioritise those individuals most in need of cataract surgery.

Response rate
The response rate was 46%. Prima facie, this figure may appear low. However, given how the tests were administered (with no patient selection, pre-warning, incentive, follow-up, or support), we actually consider it remarkably high. Prior research suggests that the rate of return could be further increased through relatively inexpensive measures such as automated reminders or financial incentives (e.g. lotteries). It should also be noted that a high return rate is not necessarily a prerequisite for this approach to be viable. Thus, given its low per-patient cost, and given that non-respondents are not necessarily disadvantaged (e.g. rather than being put to the back of the queue, non-respondents could be assigned randomly generated scores, leaving them no worse off than in the present prioritisation ‘lottery’), it may be that the collection of additional data for patient prioritisation may be justified, even if the response-rate were low.

No gross differences were observed between respondents and non-respondents, either in terms of age, sex, socioeconomic status, or geographic location. However, this should be taken in the context of the relatively homogenous sample. Further research is required to identify whether specific demographics might be particularly well/ill-served by pen-and-paper home testing.

Cost
The cost of pen-and-paper home testing was around £3 ($4 USD) per patient (incl. postage fees and test materials), not accounting for the staff time taken to prepare the outbound packs and score the returned tests. To minimise these staffing costs, we also developed a means of automatically scoring and transmitting test results using a smartphone camera – potentially obviating the need for patients to even post backtests. We intend to publish technical details of this software at a future date (manuscript in preparation).

Other variables to consider when determining how to prioritise patients
Visual function is just one of the factors a clinician must consider when deciding how to prioritise patients. For example, when computing a prioritisation ‘score’ it may also be prudent to factor-in patient self-reports (e.g. general health, life expectancy, and the patient’s circumstances, including possible threats to independent living or employment. Exactly how to weigh these factors is outside the scope of the present work, and they are moral and political judgments as much as they are scientific questions.

Furthermore, even if just considering visual function, it is highly likely that other measures—in addition to the simple CS summary measure considered in the present study—would allow cataract severity to be more fully characterised. For example, in addition to
CS, low contrast acuity, [31] disability glare, [41] visual search performance, [42] and stereopsis [43] have all been shown to be associated with degraded quality of life due to cataracts (indeed, often much more so than conventional measures of visual acuity [41, 43]). It is therefore extremely likely that by also collecting such measures (and/or structural information from photographs [44]), would allow more accurate decisions regarding patient prioritisation to be made. Whether the benefits would justify the additional costs is unknown at present, however.

Study limitations and future work

The sample of the present study was relatively small \((N = 233;\) versus 400,000 surgeries performed annually in England [45]). And while patients were randomly selected, the sample population was not widely representative—all being residents of a disproportionately affluent/Caucasian suburb of Greater London. That said, there is no specific reason to think that the results of the present study would not generalise to a larger and more diverse sample, particularly given that the test itself poses relatively few linguistic or cognitive demands (e.g. 4-year-old children have been shown capable of performing the SpotChecks test competently [35]).

Limitations of the test

CS measurements can be affected by ambient lighting conditions [46]—and unlike with digital tests, there was no way of automatically recording what the illumination levels were, or of warning patients if their testing environment is inappropriate. [47] (In the present study patients were simply asked to perform the test in a ‘well lit room’.) As the test was performed unsupervised, there was also no way of ensuring that patients patched the fellow eye correctly (or at all), or even of ensuring that the correct person took the test. And since the patient was not ‘forced’ to mark every response box on the page (see Fig. 1B), individuals may have chosen to stop responding altogether when they could no longer clearly discern the target: confounding confidence with visual ability [48]. All of these factors may have affected the accuracy and reliability of the SpotChecks data to some degree. However, the fact that the home tests showed good agreement with those performed subsequently in clinic (supervised) is encouraging: Suggesting that either patients can be relied upon to perform the tests at home appropriately, and/or that variations in home testing environments do not deleteriously affect the quality of the data substantively.

A potential limitation of the concept (of using home measures of visual function to prioritise patients) is that ‘if such a system were to be implemented, there would be tremendous incentive for patients to artificially suppress their own visual function scores [49]. We did not see any evidence of such malingering in the present study, but if home testing were integrated into routine practice then careful efforts may indeed be required to detect and militate against anomalous results.

On the benefits of pen and paper testing

We believe that pen-and-paper testing was particularly well-suited to the present use-case, since testing was required to be one off, self-administered, and performed at scale and since even a moderate level of test accuracy was expected to be sufficient. However, this should not be taken to imply that pen-and-paper tests should always be preferred. For example, digital apps may be better suited in situations where more detailed assessments of vision need to performed, where performance needs to be tracked over time (e.g. for disease monitoring [50, 51]), or where it is important to refer a patient for further testing (e.g. for mass screening [47]). Nonetheless, the present data suggest that when it comes to prioritising cataract waiting list, a simple pen-and-paper test appeared to have many attractive qualities: being easily scalable, low maintenance, acceptable to patients (though see [52]), and avoiding issues of Digital Exclusion [29].

CONCLUSIONS

This study examined the feasibility of using a pen-and-paper CS test, administered by mail, and performed unsupervised at home, to help prioritise patients waiting for a pre-surgical cataract assessment. The data showed that around half (46%) of patients responded. And in those that did respond, the results of the home test were correlated with related measures made subsequently in clinic (biometry, acuity, CS). The home data were also reasonably predictive of which eyes were subsequently listed for surgery, particularly when combined with visual acuity scores extracted from patients’ medical records. Taken together, these results indicate that a low-tech, low-cost pen-and-paper test might feasibly be used to help inform the prioritisation of patients on cataract waiting lists, and complements a wider trend, both in ophthalmology and beyond, towards using ‘asynchronous testing’ to augment more conventional methods of patient assessment [53].

SUMMARY

What was known before

- It is known that cataract waiting lists are long and growing and that tools are needed to intelligently prioritise patients.

What this study adds

- This study demonstrates that mailing patients pen-and-paper vision tests may be a feasible, low tech way of prioritising patients on cataract waiting lists.
- This showcases a new, pragmatic means of managing cataract services.

DATA AVAILABILITY

Anonymised data will be made available online. This includes all of the data reported in the present manuscript, with the exception of patient identifying information (date of birth, home address, medical history). This study demonstrates that cataract patients are willing and able to perform pen-and-paper vision testing at home, and the data provided can be used to identify eyes in need of surgery (thereby suggesting a pragmatic means of managing overstretched eye care services).

REFERENCES

63. Fun WH, Tan EH, Sararakis S, Md Sharif S, Ab Rahim I, et al. Implications of dual versions of the manuscript.

ACKNOWLEDGEMENTS

Authors MR and PR were supported by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 955590. The listed funding organisation(s) had no role in the design or conduct of this research.

AUTHOR CONTRIBUTIONS

Study conception and design: EB, DL, DPC, PRJ; data collection: EB, PFR, MR; analysis and interpretation of results: EB, DL, DPC, PRJ; draft manuscript preparation: EB, PRJ. All authors reviewed the results and approved the final version of the manuscript.

COMPETING INTERESTS

E Bianchi, None; M Rathore, None; P Reddingsius, None; D Lindfelt, C. Allergan/Abbvie, Alcon, Sight Sciences, Ellios, Vision Engineering, Glakous, Santen, Medicon, EndoOptiks, Spectrum, Thea DP Crabb, C: Abbvie/Allergan; Apellis; Janssen; F: Abb vie/Allergan; Apellis; Santens; R: Abbvie/Allergan; Santen; Thea; Glakous; PR Jones, None; The authors declare no competing interests.

E. Bianchi et al.
ETHICS APPROVAL
This study was approved by the NHS Health Research Authority (IRAS ID: #300328).

ADDITIONAL INFORMATION
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41433-024-03081-6.

Correspondence and requests for materials should be addressed to Pete R. Jones.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.