

Diagnostic Performance and Repeatability of a Novel Game-Based Visual Field Test for Children

Tariq Mehmood Aslam,^{1,2} Zaria Christine Ali,¹ Yanfang Wang,² Cecilia Fenerty,¹ Susmito Biswas,¹ Emmanouil Tsamis,² and David Barry Henson^{1,2}

¹Manchester Royal Eye Hospital, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, United Kingdom

²School of Pharmacy and Optometry, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom

Correspondence: Tariq Mehmood Aslam, Manchester Royal Eye Hospital, Oxford Road, Manchester M13 9WL, UK; tariq.aslam@mft.nhs.uk.

Submitted: December 7, 2017

Accepted: February 7, 2018

Citation: Aslam TM, Ali ZC, Wang Y, et al. Diagnostic performance and repeatability of a novel game-based visual field test for children. *Invest Ophthalmol Vis Sci.* 2018;59:1532-1537. <https://doi.org/10.1167/iops.17-23546>

PURPOSE. To demonstrate utility of a game-based test (“Caspar’s Castle”) for the detection of visual field defects in children.

METHODS. A validity and reliability study was carried out at Manchester Royal Eye Hospital Pediatric Ophthalmology Outpatients Department. We recruited 108 children with no eye pathology (aged 4–12 years) and examined a single eye with the Caspar’s Castle system using either normal thresholds or thresholds artificially adapted to recreate defects to assess diagnostic utility. Number of peripheral stimuli missed was used to determine sensitivity and specificity of artificial defect detection and to plot receiver-operator characteristic curves. A further 21 children (aged 4–16 years) with pathology were recruited and Caspar’s fields compared qualitatively with established field testing. A total of 106 of the Caspar’s Castle examinations were able to be performed twice and repeatability was determined through coefficient of repeatability and Bland–Altman chart.

RESULTS. In diagnostic testing using children with no eye pathology, 45 children completed a test using normal thresholds and 43 with tests using artificial defects. Area under receiver-operator characteristic curves for artificial defect detection was 0.895. Of the 21 children with pathology, seven had completed standard Humphreys field testing and Caspar’s Castle fields corresponded with each of these by expert opinion. Coefficient of repeatability for number of points missed across all cohorts of children (106 patients) was 6.9 (95% confidence interval: 6.16–8.07).

CONCLUSIONS. The Caspar’s Castle system of assessing visual fields using novel game-based strategies demonstrates encouraging levels of sensitivity, specificity, and reliability. It could help address current difficulties in perimetry in young children.

Keywords: perimetry, gamification, glaucoma, neurology

Although more commonly performed in an adult population, perimetry can be of vital importance for the early detection and monitoring of neurological disease and glaucoma in children. Despite challenges associated with inattention and poor compliance,¹ Tschopp et al.² concluded that reliable results can be obtained in children as young as 5 years old once a familiarization procedure has been conducted and if the duration of examination does not exceed the child’s capacity to remain task focused. Morales and Brown³ have demonstrated the feasibility of perimetry using an Octopus automated perimeter in children aged 6 to 12, and normal values for Octopus tendency-oriented perimetry in children 7 through 13 years old have been recorded.⁴ However, it was noted that when using the Octopus perimeter in younger children aged 6 to 7 years old, significant interindividual variability was present, with testing success dependent on the child’s maturity and ability to concentrate.³ The OPTIC group similarly concluded that good quality perimetry was feasible in children as young as 5 years, but required an experienced examiner with care to ensure balance between performing the test quickly and ensuring children were not overwhelmed.⁵ They found Gold-

mann perimetry to be the most effective strategy in children aged 5 to 8 years, but since this perimeter was no longer available, further research was recommended for a suitable alternative.⁵ More recent findings indicate a simple static perimetry approach potentially yields high-quality results in children younger than 10 years.⁶ In general, there has been an increasing interest in perimetry in children over the last few years with expert commentary highlighting the importance of continuing to attempt such measurement.⁷

Our work aims to contribute to this impetus for obtaining easier and higher quality fields in children. We previously presented a game-based visual field test, which demonstrated feasibility in testing children over 4 years of age with a high degree of acceptability.⁸ It was not dependent on expert, experienced administration and was designed to be as enjoyable as possible. In more recent studies, we developed the physical, psychophysical, and psychological aspects of the test further and obtained normative values for 138 children of different ages.⁹ This was a key step as although many visual fields are performed in children every year, most tests are



designed for adults and few instruments have a normative database for children.^{10,11}

The current version of our test incorporates our recent developments and normative values into a system of game-based perimetry designed specifically for children and known as Caspar's Castle.

METHODS

In this study, we assess the diagnostic performance and repeatability of the Caspar's Castle supra-threshold game-based field test. The study was approved by a local Ethics Committee and informed consent was obtained from all participants and parents. The research adhered to the tenets of the Declaration of Helsinki.

The game was programmed with Unity (version 4.3.4f1, Unity Technologies, San Francisco, CA, USA) and is described in detail in previous reports.⁹ In brief, the system consists of a host laptop (Thinkpad, Lenovo, China) and a calibrated OLED display monitor (Sony PVM2541A, Sony Corporation, Japan) encased within a model castle with a viewing window. A button control that is connected to the host laptop allows children to respond to central and peripheral targets of a game. The game involves a central character, Prince Caspar, tasked with a hoover and brush to sweep up central and peripheral "googlies" that have "invaded the castle." Further details including the development of the test hardware and software are given in previous publications.^{8,9,12}

We recorded details of consenting children and pertinent game data including visual field outcomes and reliability indices. One eye of each child was used for the study, determined by subject preference. Outcomes from children completing the game were recorded in spatial plots and quantified in terms of numbers of points missed. Reliability indices were based upon number of times the central "googlie" was missed (indicating poor fixation), the number of times peripheral points from previously seen locations were missed (false negatives), and the number of times the button was pressed without central or peripheral stimuli (false positives).

The main study cohort was used to quantitatively determine accuracy of Caspar's Castle game-based test to detect field defects according to guidelines set by the STARD committee.¹³ Children with reliably proven visual field defects are not common,¹⁴ and we addressed this methodological challenge by using simulated defects,¹⁵⁻¹⁸ with psychometric functions to predict how patients should respond to the altered stimulus presentations.^{19,20} In effect, we decreased the intensity of some of the age-adjusted peripheral stimuli such that in each age group normal children should be expected to miss the attenuated targets if they were performing the field test accurately and appropriately. The precise nature of these simulated defects were extracted from a previously generated pool of glaucomatous visual field defects consisting of early/moderate cases (from Brusini Glaucoma Staging System).²¹ The defect values were then transformed into location sensitive, age-matched defects for normal children for our Caspar's Castle test using values reported in a previous study.⁹ In this way, we were able to assess number of peripheral stimuli missed, anticipating that this should ideally match their (simulated) outcome field defect. This type of procedure has previously been successfully demonstrated in testing for driving visual fields for adults.²²

For this diagnostic accuracy cohort, we included children aged 4 to 12 years, who were physically able to play the game and had no clinical history of any relevance including symptoms or signs or history of eye disease or neurological disease. These children were recruited as siblings of children

who were attending the Ophthalmology Outpatients Department of Manchester Royal Eye Hospital and from a nursery school local to the hospital.

Sample size estimates for such a diagnostic accuracy study to detect simulated field defects with 90% sensitivity and specificity (within 10% of true value and 95% confidence) were 70 eyes (35 each with simulated abnormal and normal visual fields).²³ To achieve the minimum numbers for the power calculation, we planned to recruit 80 children with no eye pathology; 40 to be given a normal field test and 40 a simulated abnormal glaucomatous field test. Once the tests were completed, we determined the sensitivity and specificity of using numbers of points missed to detect the simulated field defects and constructed a receiver operating characteristic (ROC) curve.

In addition to the diagnostic accuracy study, we endeavored to explore utility of the Caspar's fields in children with real pathology. We included all children with conditions known to be associated with visual field defects who were able and willing to play the game and who were aged 4 to 16 years. These children were recruited from Manchester Royal Eye Hospital pediatric ophthalmology out-patients. Their clinical details were recorded, in particular noting results of any existing conventional field tests. Methodological concerns such as the differing outcome measures of conventional fields meant that this patient group was not merged with the existing diagnostic accuracy study. The main outcome measure for this aspect of the study was therefore a qualitative empirical comparison by clinical pediatric glaucoma experts, who assessed the derived Caspar's Castle visual fields to determine if they corresponded with any existing conventional field tests.

Finally, we determined the repeatability of the Caspar's Castle fields test. All Caspar's Castle fields from both cohorts were to be performed twice, approximately 30 minutes apart, and quantified to enable this aspect of the study. We compared the total number of points missed in the first and second visual field tests and used the methods of Bland-Altman to derive mean of difference between measures and limits of agreement. According to these authors, 100 patients would be an adequate sample size to determine acceptable 95% confidence intervals (CIs) for the limits of agreement to $\pm 0.34 s$, where s is the standard deviation of the differences between measurements.²⁴ For this study, we planned to recruit a total of 100 children, with either normal, simulated abnormal, or potentially pathological fields.

All statistics for diagnostic accuracy and repeatability were performed with MedCalc Statistical Software version 17.6 (MedCalc Software bvba, Ostend, Belgium; <http://www.medcalc.org>; 2017).

RESULTS

In total, 126 children were recruited to examination with the new Caspar's Castle game-based supra-threshold field test. Of these, 105 children with no eye pathology were recruited for the objective diagnostics study and 21 children with pathology for potential qualitative comparison of Caspar's Castle field test with conventional fields.

All children with pathology completed the Castle tests appropriately, but 17/105 children with no eye pathology were excluded from further analyses; three were unable to understand the game (ages 4, 4, and 5 years). Six had lack of concentration to prevent them from completing the game (ages 4, 5, 5, 6, 6, and 9 years). The 9-year-old had previously been diagnosed with learning difficulties. Four were excluded due to poor reliability indices (ages 4, 4, 5, and 6 years). Four

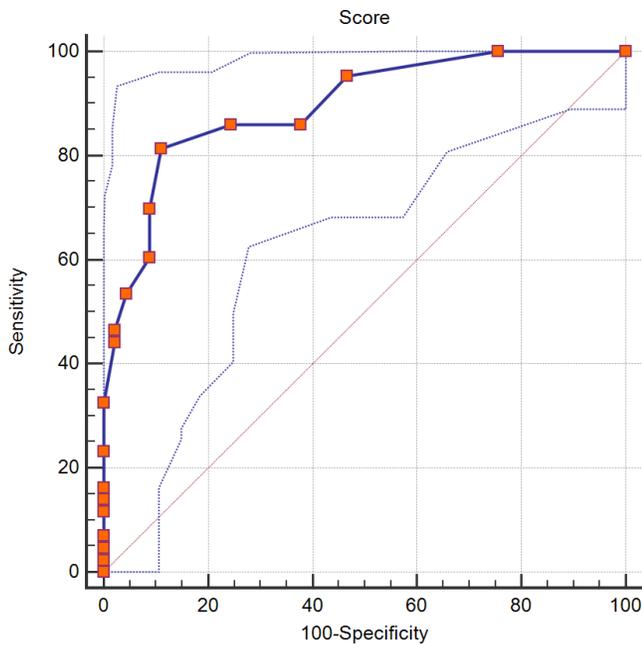


FIGURE 1. ROC curve for Caspar's Castle game-based field test.

children did not want to play as they were too tired or scared of the castle (ages 4, 4, 4, and 5 years).

Therefore, data from 88 normal children were used for the diagnostic analysis, involving 45 children completing visual fields with normal threshold levels (defect absent) and 43 children with thresholds that had been manipulated to reflect glaucomatous field defects (defect present). Median age was 6 years (range, 4–10). We determined the total number of points missed in each of these patients' first Caspar's Castle fields game and plotted a receiver operator characteristic curve (Fig. 1) using the grouping of defect present or absent as the classification variable. Area under the curve (AUC) was 0.895 (SE 0.0332, 95% CI: 0.812–0.950, $P < 0.0001$). Youden index J was 0.7028 (associated criterion >4 points missed), with sensitivity 81.40 and specificity of 88.89. A table of summary statistics is provided in the Table. In this table, we have included estimates of sensitivity for high levels of specificity. When used as a screening test, it is important to have high levels of specificity to ensure that very few normal subjects are classified as pathological.

We were able to recruit 21 children with established pathology where visual field changes might be expected and all children completed the test satisfactorily. Median age was 12 (range, 4–16). Thirteen children had congenital glaucoma, seven had secondary glaucoma, and one had neurological damage to a temporal lobe. Of all patients with pathology, seven had been able to complete conventional field testing, with a combination of fast and standard 24-2 Swedish Interactive Thresholding Algorithm (SITA) threshold strategies (Humphrey visual field [HVF], Carl Zeiss Meditec, Inc., Dublin, CA, USA). For this small sample, we compared the different outputs of the conventional versus game-based field tests qualitatively by ophthalmologists specialized in pediatric glaucoma. Of the seven patients who had achieved conventional SITA field plots, four were pathologically abnormal, two were normal, and one had functional ("clover-leaf" shaped fields). All the pathologically abnormal conventional tests had corresponding visual field loss seen on the Caspar's Castle game-based test. The two children with normal conventional fields had normal

TABLE. Summary of Values for Different Fixed Sensitivity and Specificity

Sensitivity	Specificity	95% CI	Criterion
Estimated specificity at fixed sensitivity			
80.00	89.16	60.30–97.97	>4.12
90.00	58.44	38.37–87.44	>1.58
95.00	53.67	33.69–78.87	>1.03
97.50	39.97	19.47–60.08	>0.54
99.00	30.66	14.46–47.66	>0.22
Estimated sensitivity at fixed specificity			
80.00	84.50	69.06–93.02	>3.33
90.00	75.58	43.75–90.90	>4.5
95.00	54.36	31.02–83.30	>6.88
97.50	47.38	0.00–0.00	>7.88
99.00	37.79	0.00–0.00	>9.55

CI, confidence interval.

Caspar's Castle fields. One child who had functional ("clover-leaf" shaped fields) had normal visual fields in the Caspar's Castle test. Examples of these plots are shown in Figures 2 through 5.

Participants from all cohorts were included in the repeatability analyses; 21 with pathology, 45 normal children with normal fields, and 40 normal children with simulated defects. (Of the 43 normal children who completed fields with simulated field defects, only 40 completed the game twice due to parents' desire to leave the clinic quickly.) Thus a total of 106 children were included in the repeatability study. Median age of subjects for the overall repeatability study was 7 years (range, 4–16) and overall average time to complete a game-based field test was 6.5 minutes.

Coefficient of repeatability for number of points missed was 6.9 (95% CI: 6.16–8.07), and a Bland-Altman chart is presented in Figure 6. The arithmetic mean of difference between first and second test was 0.61 (95% CI: 0.065–1.29). Lower limit of agreement was -6.29 (95% CI: -7.46 to -5.13) and upper limit of agreement 7.52 (95% CI: 6.36–8.69).

DISCUSSION

In this paper, we present evidence for the utility of a novel, game-based method of perimetry for children. It involves a specifically designed computerized system using a child-friendly castle structure, coherent storyline, animations, graphics, and music all written specifically to appeal to children.

We have been able to demonstrate acceptability and ability to complete visual fields tests in 109/126 children. On diagnostic evaluation with ROC curves, the AUC was found to be 0.895, which compares favorably with other novel visual field screening tools,^{25,26} especially when we consider the challenges we have set ourselves of detecting only mild to moderate defects in young children. The repeatability of the test appears satisfactory with variability of scatter that is consistent across the Bland-Altman chart, around a mean difference of around zero.

Although many challenges have been addressed in development of the current test, improvements to diagnostic performance and repeatability could still be made. Contrary to expectations, many of the instances where the game did not appear to perform well involved older children, aged 9 years or older, some of whom became able to play the central game so adeptly they were perhaps more tempted to scan the periphery as well as responding to central demands. The game was tailored toward younger children, and it may be that

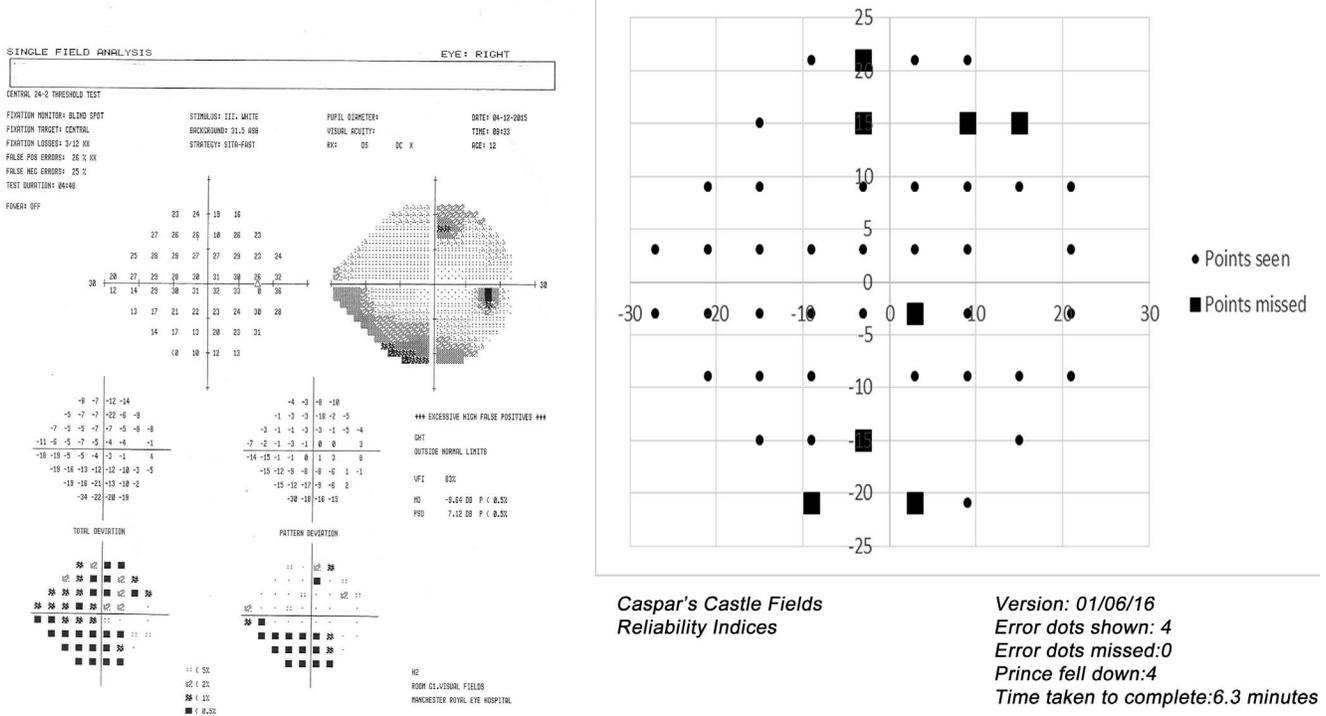


FIGURE 2. Congenital glaucoma in a 12-year-old child. Plot of Humphrey 24-2 (left) and Caspar's Castle fields (right).

more complex versions should be developed for older children. However, to achieve optimal clinical impact, test development should remain focused on younger children (e.g., 3–5 years old). Improvements using an eye tracker to monitor fixation should increase sensitivity and specificity, improving clinical utility for all ages. In future developments, we also plan to produce versions that should appeal to girls and boys

equally with the main character option to be played by a princess as well as a prince.

Assessment of the system's diagnostic accuracy was methodologically challenging. Children with known and clinically established defects were relatively rare, even in our large tertiary center. Furthermore, the variety of methods used to represent standard field test outcomes make statistical

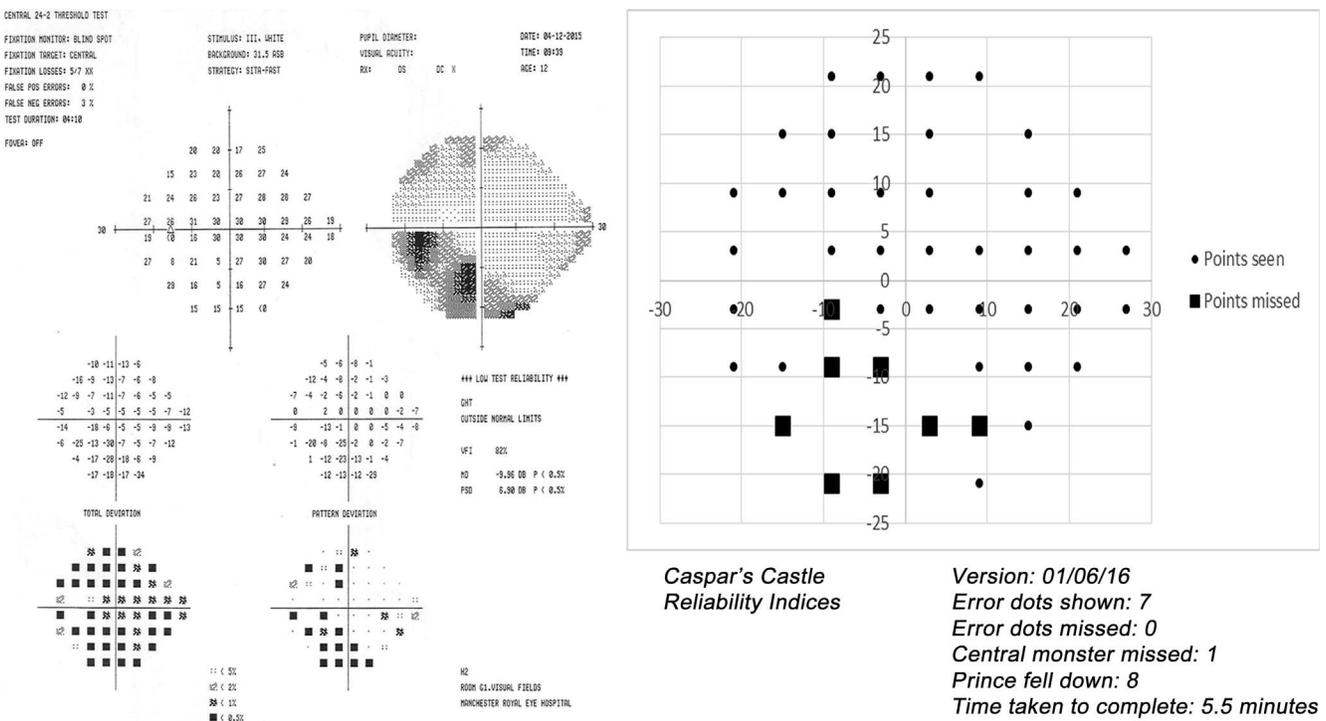


FIGURE 3. Congenital glaucoma in a 12-year-old child. Plot of Humphrey 24-2 (left) and Caspar's Castle fields (right).

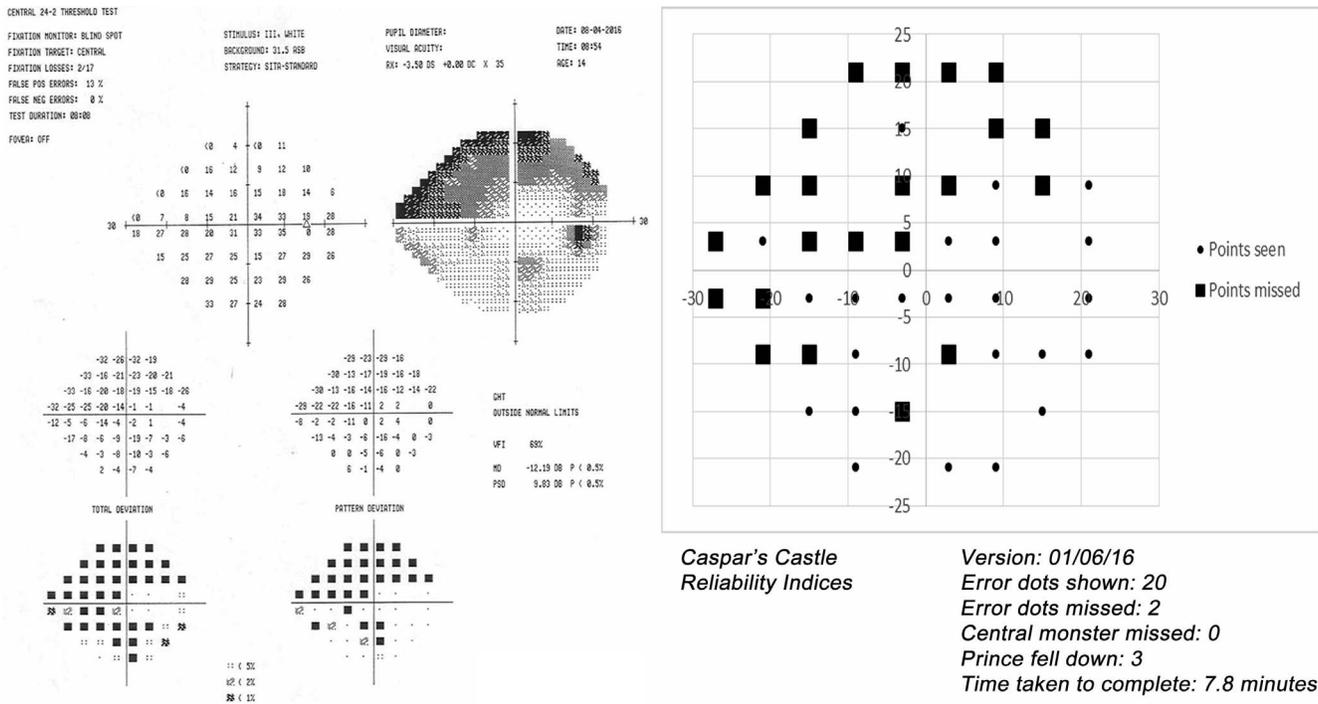


FIGURE 4. Congenital glaucoma in a 14-year-old child. Plot of Humphrey 24-2 (left) and Caspar's Castle fields (right).

comparisons difficult. We therefore chose to demonstrate utility using subjective assessments of children with pathology but also more robust comparisons of children using simulated field defects. However, this use of artificially created field defects might not transfer entirely to real defects in patients and future studies may attempt to compare more rigidly to existing perimetry such as HVFs.

A key feature of the presented Caspar's Castle system is that it is compatible with existing hardware that is already in widespread international use; a computer running Windows 7 service pack 1 or above, suitable screen and trigger device. The surrounding hardboard castle structure was simple to construct and the whole system could be easily available and is inexpensive. Further strengths include independence from

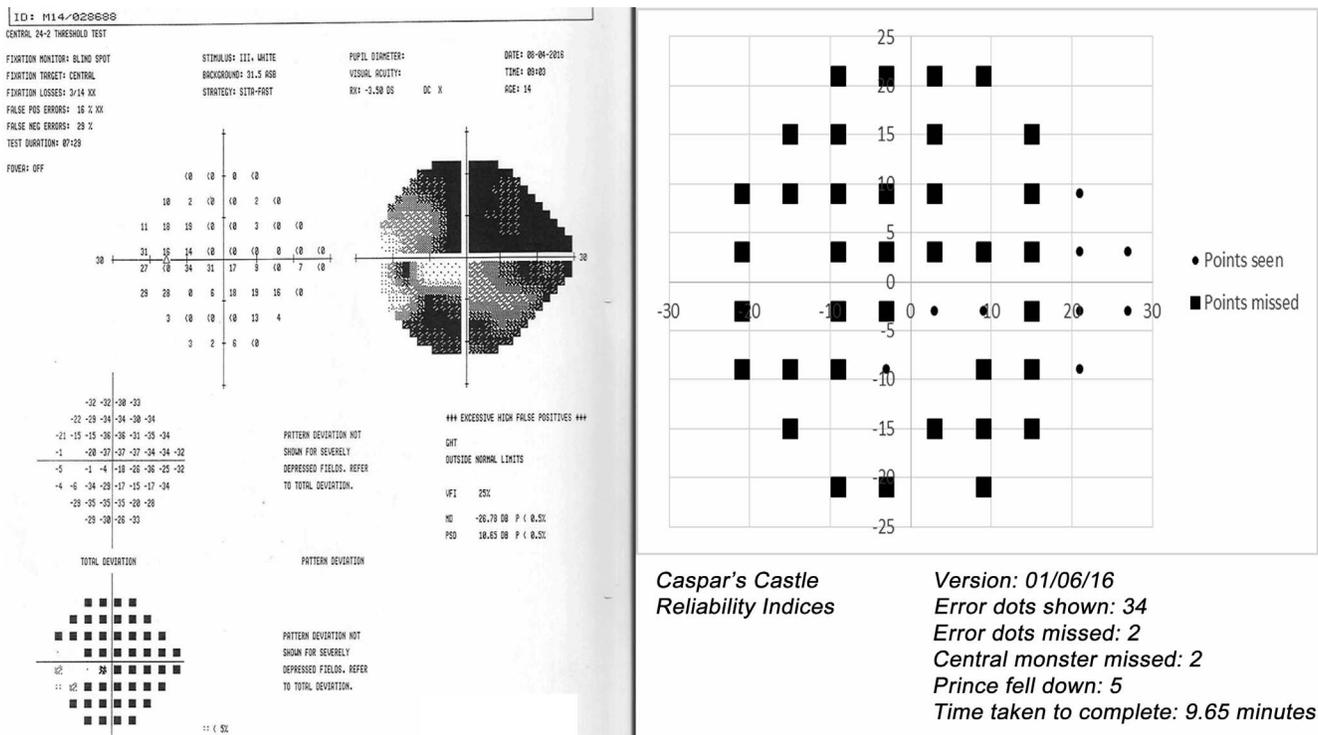


FIGURE 5. Severe congenital glaucoma in a 14-year-old child. Plot of Humphrey 24-2 (left) and Caspar's Castle fields (right).

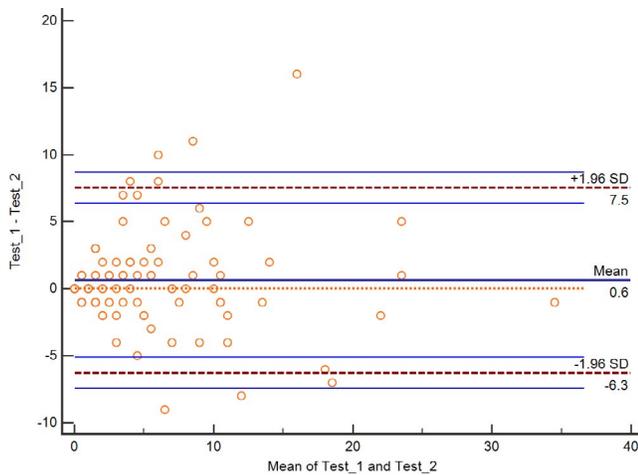


FIGURE 6. Bland-Altman chart for repeatability of number of points missed on Caspar's Castle game-based field test.

expert supervision requirement as well as ease of use and attractiveness to children. These features suggest potential utility particularly as a screening tool outside of specialized hospital settings. Examples of children aged 5 and 7 playing the test can be seen in Supplementary Videos S1 and S2, respectively.

In summary, Caspar's Castle represents a novel, affordable, noninvasive and entertaining means of obtaining visual field results from younger children with acceptable validity and reliability. It could be a useful tool in clinical practice to assist with the challenges of pediatric perimetry.

Acknowledgments

This paper presents independent research funded by the National Institute for Health Research (NIHR) under its Research for Patient Benefit (RfPB) Programme (Grant Reference Number PB-PG-0211-24064) and by an International Glaucoma Association Grant, IGA/UKEGS2015 (YW). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.

Disclosure: **T.M. Aslam**, Novartis (C, R), Bayer (C, R), Thea Pharmaceuticals (C, R) Bausch & Lomb (C, R), Oraya (C, R); **Z.C. Ali**, None; **Y. Wang**, None; **C. Fenerty**, None; **S. Biswas**, None; **E. Tsamis**, None; **D.B. Henson**, None

References

- Walters BC, Rahi JS, Cumberland PM. Perimetry in children: survey of current practices in the United Kingdom and Ireland. *Ophthalmic Epidemiol.* 2012;19:358-363.
- Tschopp C, Safran AB, Laffi JL, Mermoud C, Bullinger A, Viviani P. Automated static perimetry in the child: methodologic and practical problems [in German]. *Klinische Monatsblätter für Augenheilkunde.* 1995;206:416-419.
- Morales J, Brown SM. The feasibility of short automated static perimetry in children. *Ophthalmology.* 2001;108:157-162.
- Brown SM, Bradley JC, Monhart MJ, Baker DK. Normal values for Octopus tendency oriented perimetry in children 7 through 13 years old. *Graefes Arch Clin Exp Ophthalmol.* 2005;243:886-893.
- Patel DE, Cumberland PM, Walters BC, Russell-Eggitt I, Rahi JS; OPTIC Study Group. Study of Optimal Perimetric Testing in Children (OPTIC): feasibility, reliability and repeatability of perimetry in children. *PLoS One.* 2015;10:e0130895.

- Patel DE, Cumberland PM, Walters BC, et al. Comparison of quality and output of different optimal perimetric testing approaches in children with glaucoma. *JAMA Ophthalmol.* 2018;136:155-161.
- Caprioli J. Visual field assessment in children. *JAMA Ophthalmol.* 2018;136:162-163.
- Aslam TM, Rahman W, Henson D, Khaw PT. A novel paediatric game-based visual-fields assessor. *Br J Ophthalmol.* 2011;95:921-924.
- Wang Y, Ali Z, Subramani S, et al. Normal threshold size of stimuli in children using a game-based visual field test. *Ophthalmol Ther.* 2017;6:115-122.
- Molina-Martin A, Pinero DP, Perez-Cambrodi RJ. Normal values for microperimetry with the MAIA microperimeter: sensitivity and fixation analysis in healthy adults and children. *Eur J Ophthalmol.* 2017;27:607-613.
- Patel DE, Cumberland PM, Walters BC, et al. Study of Optimal Perimetric Testing in Children (OPTIC): normative visual field values in children. *Ophthalmology.* 2015;122:1711-1717.
- Miranda MA, Henson DB, Fenerty C, Biswas S, Aslam T. Development of a pediatric visual field test. *Trans Vis Sci Tech.* 2016;5(6):13.
- Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open.* 2016;6:e012799.
- Murray IC, Cameron LA, McTrusty AD, et al. Feasibility, accuracy, and repeatability of suprathreshold saccadic vector optokinetic perimetry. *Trans Vis Sci Tech.* 2016;5(4):15.
- Bebie H, Fankhauser F, Spahr J. Static perimetry: strategies. *Acta ophthalmologica.* 1976;54:325-338.
- Spenceley SE, Henson DB. Visual field test simulation and error in threshold estimation. *Br J Ophthalmol.* 1996;80:304-308.
- Spry PGD, Bates AB, Johnson CA, Chauhan BC. Simulation of longitudinal threshold visual field data. *Invest Ophthalmol Vis Sci.* 2000;41:2192-2200.
- Gardiner SK, Crabb DP. Examination of different pointwise linear regression methods for determining visual field progression. *Invest Ophthalmol Vis Sci.* 2002;43:1400-1407.
- Gardiner SK, Swanson WH, Goren D, Mansberger SL, Demirel S. Assessment of the reliability of standard automated perimetry in regions of glaucomatous damage. *Ophthalmology.* 2014;121:1359-1369.
- Henson DB, Chaudry S, Artes PH, Faragher EB, Ansons A. Response variability in the visual field: comparison of optic neuritis, glaucoma, ocular hypertension, and normal eyes. *Invest Ophthalmol Vis Sci.* 2000;41:417-421.
- Brusini P, Filacorda S. Enhanced Glaucoma Staging System (GSS 2) for classifying functional damage in glaucoma. *J Glaucoma.* 2006;15:40-46.
- Glen FC, Smith ND, Jones L, Crabb DP. "I didn't see that coming": simulated visual fields and driving hazard perception test performance. *Clin Exp Optom.* 2016;99:469-475.
- Schafer H. Efficient confidence bounds for ROC curves. *Stat Med.* 1994;13:1551-1561.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1:307-310.
- Olsen AS, Alberti M, Serup L, la Cour M, Damato B, Kolko M. Glaucoma detection with damato multifixation campimetry online. *Eye.* 2016;30:731-739.
- Yamada N, Chen PP, Mills RP, et al. Screening for glaucoma with frequency-doubling technology and Damato campimetry. *Arch Ophthalmol.* 1999;117:1479-1484.