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

Tariq Mehmood Aslam, Zaria Christine Ali, Yanfang Wang, Cecilia Fenerty, Susmito Biswas, Emmanouil Tsamis, and David Barry Henson

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

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
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 “gooogie”
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.6 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 Depart-
ment 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 ±0.34 s, where s is the
standard deviation of the differences between measure-
ments.24 For this study, we planned to recruit a total of 100
children, with either normal, simulated abnormal, or poten-
tially 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.med-
calc.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
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, which compares favorably with other novel visual field screening tools.

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 neurologic 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 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.

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 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

### Table: Summary of Values for Different Fixed Sensitivity and Specificity

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Specificity</th>
<th>95% CI</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated specificity at fixed sensitivity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80.00</td>
<td>89.16</td>
<td>60.30–97.97</td>
<td>&gt;4.12</td>
</tr>
<tr>
<td>90.00</td>
<td>84.12</td>
<td>38.37–87.44</td>
<td>&gt;4.57</td>
</tr>
<tr>
<td>95.00</td>
<td>43.59</td>
<td>33.69–78.87</td>
<td>&gt;1.05</td>
</tr>
<tr>
<td>97.50</td>
<td>39.97</td>
<td>19.47–60.08</td>
<td>&gt;0.54</td>
</tr>
<tr>
<td>99.00</td>
<td>30.66</td>
<td>14.46–47.66</td>
<td>&gt;0.22</td>
</tr>
<tr>
<td>Estimated sensitivity at fixed specificity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80.00</td>
<td>43.87</td>
<td>69.06–93.02</td>
<td>&gt;3.35</td>
</tr>
<tr>
<td>90.00</td>
<td>43.75</td>
<td>43.75–90.90</td>
<td>&gt;4.57</td>
</tr>
<tr>
<td>95.00</td>
<td>43.10</td>
<td>31.02–83.30</td>
<td>&gt;6.88</td>
</tr>
<tr>
<td>97.50</td>
<td>47.38</td>
<td>0.00–0.00</td>
<td>&gt;7.88</td>
</tr>
<tr>
<td>99.00</td>
<td>37.79</td>
<td>0.00–0.00</td>
<td>&gt;9.55</td>
</tr>
</tbody>
</table>

CI, confidence interval.
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...
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 4.** Congenital glaucoma in a 14-year-old child. Plot of Humphrey 24-2 (left) and Caspar’s Castle fields (right).

**Figure 5.** Severe congenital glaucoma in a 14-year-old child. Plot of Humphrey 24-2 (left) and Caspar’s Castle fields (right).
Diagnostic Performance of a Game-Based Field Test

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/UEKGS2015 (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); Thea Pharmaceuticals (C, R) Bausch & Lomb (C, R); Novartis (C, R), Bausch & Lomb (C, R); None; Z.C. Ali, None; Y. Wang, None; C. Fenerty, None; S. Biswas, None; E. Tsamis, None; D.B. Henson, None

References


