select statistical services
Oxygen House
Grenadier Road
Exeter Business Park
Exeter, Devon, EX1 3LH
tel: 01392 440426
email: info@select-statistics.co.uk
web: www.select-statistics.co.uk

Statistical Support for QuickScreen Dyslexia Test

Initial Analysis

Author: Sarah Marley
Reviewed by: Lynsey McColl
Revision Date: 10th November 2016
Prepared for: Pico Educational Systems Ltd
Reference Number: PICO001

Introduction

Select were pleased to be asked to help with the statistical analysis of Pico Educational Systems Ltd's QuickScreen dyslexia test, on behalf of Dr Dee Walker.

QuickScreen is an adult computerised screening test, developed with the aim of providing a reasonably in-depth assessment of dyslexia. The test delivers an indication of possible dyslexia without the need for users to undergo a costly professional assessment by an educational or occupational psychologist.

An essential step in the evaluation process of any diagnostic/screening test is to assess its accuracy via diagnostic accuracy measures. We agreed to produce these measures for QuickScreen, based on observational data compiled by Pico Educational Systems Ltd. These data were collected from participants completing the online assessment via three sources: a link offered on the British Dyslexia Association (BDA) website, personally sent links to individual email addresses, and some small university trials.

Data

The QuickScreen dyslexia test results were provided in comma separated value (csv) format in a number of separate files. These csv files all had a consistent layout and were combined prior to analysis to create a single dataset.

The data received included results for participants where an independent assessment of their dyslexia diagnosis was not available. Dr Dee Walker provided Word documents listing the participants with an available, independent diagnosis (positive or negative). Only these participants were included in the analysis (i.e., all those participants who had not been independently assessed were dropped). Test results were available for 245 participants with an independent dyslexia diagnosis; 193 (78.8%) had a positive diagnosis and 52 (21.2%) a negative diagnosis. The QuickScreen test reports the possibility of dyslexia assessment in terms of one of five possible indications: None, Borderline, Mild, Moderate, or Strong. Of the 245 participants included in the analysis, 40 (16.3%) received an indication of None; 71 (29.0%) an indication of Borderline; 65 (26.5%) Mild; 62 (25.3%) Moderate; and 7 (2.9%) Strong (as shown in the cross-tabulation in Table 1).

	None	Borderline	Mild	Moderate	Strong	Total
Negative	29	23	0	0	0	52 (21.2%)
Positive	11	48	65	62	7	193 (78.8%)
Total	40 (16.3%)	71 (29.0%)	65 (26.5%)	62 (25.3%)	7 (2.9%)	245 (100%)

Table 1: Cross-tabulation of the dyslexia diagnosis (Negative/Positive) versus the QuickScreen test result (None/Borderline/Mild/Moderate/Strong) for the full 245 participants.

Information was also available in the Word documents provided to indicate where some participants were known university students. One-hundred and eighteen participants (48.2%) were identified as known university students and 127 (51.8%) unknown with regard to their university status. In order to provide greater clarity on how well the QuickScreen test is performing for potentially better compensated dyslexics, we agreed to repeat the analysis (i.e., calculation of the diagnostic accuracy measures) splitting the results by this university grouping.

Methods

The sensitivity of a diagnostic test indicates how good it is at finding people with the condition in question. It is the probability that someone who has the condition is identified as such by the test. Whereas the specificity of a diagnostic test indicates how good it is at identifying people who do not have the condition. It is the probability that someone who does not have the condition is identified as such by the test.

In this case, the QuickScreen test has five possible outcome indications. Therefore, we can calculate the sensitivity of each category in identifying people with dyslexia (treating each test category as a "test positive") and also the specificity of each category in identifying people without dyslexia (treating each category as a "test negative").

Another important set of accuracy measures are the predictive values of the test.. These are also termed the "post-test probabilities" and provide the probability of a positive or negative diagnosis given the test result. The predictive values therefore provide important information on the diagnostic accuracy of the test for a particular participant, answering the question "How likely is it that I have or don't have dyslexia given the test result that I have received?"

The predictive values depend on the prevalence of the condition in question in the population, i.e., the proportion of individuals who have dyslexia, as well as the sensitivity and specificity of the test. As the sample of data available are a selection of "cases" with a positive dyslexia diagnosis and "controls" with a negative dyslexia diagnosis from observational data, rather than a random sample from the population, the true prevalence is unknown. Therefore we can't reliably estimate the predictive values directly from the data available. Following discussion with Dr Dee Walker, it was agreed that we should use an estimated prevalence of 10% in calculating the predictive values, based on previous research studies and the figures quoted by dyslexia organisations. The observed prevalence in the data available was considerably higher than this (78.8%), indicating an oversampling of dyslexic participants. In screening situations, the prevalence is almost always small and the positive predictive value low, even for a fairly sensitive and specific test.

For each QuickScreen test category, we therefore estimate its sensitivity, specificity, positive and negative predictive value. We also provide 95% confidence intervals for each to capture our uncertainty in the estimates. If repeated studies were undertaken and the 95% confidence interval was calculated for each study, 95% of the intervals would contain the true value.

The standard estimation of binominal proportions, such as the sensitivity and specificity of a diagnostic test (i.e., taking the observed sample proportion), has been shown to be less than adequate, particularly when the sample size is relatively low. Applying a continuity correction can provide a better estimate and allow more accurate confidence intervals to be developed.

Using the logit transformation in calculating the confidence interval can also help to meet the assumptions of normality and avoid producing limits beyond the possible boundary values of 0 and 100%. Therefore, we provide diagnostic accuracy measure values using the continuity adjusted

estimates and continuity adjusted logit intervals (for further information and the formulae applied see: *D. N. Mercaldo, X-H Zhou, and K. F. Lau; 2005*[1]).

Alongside these diagnostic accuracy measures, we have carried out a statistical test to assess whether there is evidence of an association between the QuickScreen test outcome and the independent dyslexia diagnosis. This would be expected if the test is useful in discriminating between dyslexic and non-dyslexic individuals. Fisher's exact test is applied (rather than a large sample test such as the Chi-square test, for example) to account for the fact that we have relatively low sample sizes, which can bias the results in asymptotic tests (as the normal approximation of the multinomial distribution can fail).

As discussed in the Introduction section above, the analysis is also repeated, split by the known and unknown university status grouping.

Validity

It should be noted when interpreting the results of this analysis that their validity depends on the applicability of the sample participants to the population of interest. This includes the spectrum of severity of dyslexia in the sample. Where this might not reflect the target population, a study is sometimes said to suffer from "spectrum bias".

The potential for other biases such as classification bias, where misclassification of participants in their independent dyslexia diagnosis may have occurred, should also be considered.

A more formal, prospective cohort study may provide a more reliable assessment of the diagnostic test accuracy, by helping to eliminate potential sources of bias such as those described above.

Results

The results of the analysis outlined in the Methods section are presented below, first for the full 245 participants and then for the known university student group (n=118) and finally the unknown university status group (n=127).

All Participants

A Fisher's exact test (on the data in Table 1) finds strong statistical evidence (p-value < 0.0001) of an association between the independent dyslexia diagnosis and the QuickScreen test indication.

The proportion of participants without dyslexia who received each QuickScreen test result (i.e., sample specificity) and the proportion of participants with dyslexia who received each QuickScreen test result (i.e., sample sensitivity) are shown in Table 2.

[1] Mercaldo, Nathaniel David; Zhou, Xiao-Hua; and Lau, Kit F., "Confidence Intervals for Predictive Values Using Data from a Case Control Study" (December 2005). UW Biostatistics Working Paper Series. Working Paper 271. http://biostats.bepress.com/uwbiostat/paper271

	None	Borderline	Mild	Moderate	Strong	Total
Negative	55.8%	44.2%	0%	0%	0%	100%
Positive	5.7%	24.9%	33.7%	32.1%	3.6%	100%

Table 2: Raw sample specificity (Negative row) and sensitivity (Positive row) values for each QuickScreen test category, based on the results for the full 245 participants.

For example, 55.8% of participants without dyslexia received a QuickScreen indication of "None", and 32.1% of participants with dyslexia receive a QuickScreen indication of "Moderate".

The proportion of participants with and without dyslexia in each QuickScreen test category are shown in Table 3. These are the raw sample predictive values, based on the observed sample prevalence, and do not reflect estimates for the population.

	None	Borderline	Mild	Moderate	Strong
Negative	72.5%	32.4%	0%	0%	0%
Positive	27.5%	67.6%	100%	100%	100%
Total	100%	100%	100%	100%	100%

Table 3: Raw sample predictive values (Negative and Positive) for each QuickScreen test category, based on the results for the full 245 participants.

For example, 72.5% of those participants with a QuickScreen test result of "None" were non-dyslexics, and 100% of those participants with a QuickScreen test result of "Strong" were dyslexic.

The diagnostic accuracy measures for each QuickScreen test category, estimated using the adjusted method (with adjusted logit confidence intervals) and assuming a 10% prevalence of dyslexia are shown in Table 4.

QuickScreen Test Result	Diagnostic Measure	Estimate	95% Confidence Interval
None	Sensitivity	6.6%	(3.8%, 11.0%)
	PPV	1.3%	(0.7%, 2.3%)
	Specificity	55.4%	(42.3%, 67.8%)
	NPV	98.7%	(97.7%, 99.3%)
Borderline	Sensitivity	25.4%	(19.8%, 31.9%)
	PPV	5.9%	(4.1%, 8.4%)
	Specificity	44.6%	(32.2%, 57.7%)
	NPV	94.1%	(91.6%, 95.9%)
Mild	Sensitivity	34.0%	(27.7%, 40.9%)
	PPV	52.3%	(21.3%, 81.7%)
	Specificity	3.4%	(0.8%, 13.1%)
	NPV	47.7%	(18.3%, 78.7%)
Moderate	Sensitivity	32.5%	(26.3%, 39.3%)
	PPV	51.2%	(20.5%, 81.0%)
	Specificity	3.4%	(0.8%, 13.1%)
	NPV	48.8%	(19.0%, 79.5%)
Strong	Sensitivity	4.5%	(2.4%, 8.5%)
	PPV	12.8%	(3.1%, 40.3%)
	Specificity	3.4%	(0.8%, 13.1%)
	NPV	87.2%	(59.7%, 96.9%)
Mild, Moderate or Strong	Sensitivity	69.1%	(62.3%, 75.1%)
	PPV	69.0%	(35.7%, 90.0%)
	Specificity	3.4%	(0.8%, 13.1%)
	NPV	31.0%	(10.0%, 64.3%)
Borderline, Mild, Moderate or Strong	Sensitivity	93.4%	(89.0%, 96.2%)
	PPV	18.9%	(14.8%, 23.8%)
	Specificity	44.6%	(32.2%, 57.7%)
	NPV	81.1%	(76.2%, 85.2%)
None or Borderline	Sensitivity	30.9%	(24.9%, 37.7%)
	PPV	3.4%	(2.8%, 4.2%)
	Specificity	96.6%	(86.9%, 99.2%)
	NPV	96.6%	(95.8%, 97.2%)

Table 4: Estimates of the diagnostic accuracy measures for each QuickScreen test category using the adjusted logit method, for the full 245 participants (with 10% prevalence). PPV = Positive Predictive Value; NPV = Negative Predictive Value.

In addition to considering each category in isolation, the measures for some combinations of the QuickScreen test result are also provided. For example, we estimate that 96.6% (95% Confidence

Interval [CI] = 86.9%, 99.2%) of non-dyslexic individuals will receive a QuickScreen indication of "None or Borderline". An individual receiving a QuickScreen indication of "Mild, Moderate or Strong" is estimated to have a 69.0% (95% CI = 35.7%, 90.0%) probability of a positive dyslexia diagnosis.

University Group

The results of the analysis outlined in the Methods section for the known university student group are presented below.

Test results were available for 118 known university students with an independent dyslexia diagnosis; 77 (65.3%) had a positive diagnosis and 41 (34.7%) a negative diagnosis. Of these 118 participants, 28 (23.7%) received a QuickScreen indication of None; 41 (34.7%) an indication of Borderline; 33 (28.0%) Mild; 15 (12.7%) Moderate; and 1 (0.8%) Strong (as shown in the cross-tabulation in Table 5).

	None	Borderline	Mild	Moderate	Strong	Total
Negative	23	18	0	0	0	41 (34.7%)
Positive	5	23	33	15	1	77 (65.3%)
Total	28 (23.7%)	41 (34.7%)	33 (28.0%)	15 (12.7%)	1 (0.8%)	118 (100%)

Table 5: Cross-tabulation of the dyslexia diagnosis (Negative/Positive) versus the QuickScreen test result (None/Borderline/Mild/Moderate/Strong) for the 118 known university participants.

A Fisher's exact test on these data finds strong statistical evidence (p-value < 0.0001) of an association between the independent dyslexia diagnosis and the QuickScreen test indication.

The proportion of known university students without dyslexia who received each QuickScreen test result (i.e., sample specificity) and the proportion of known university students with dyslexia who received each QuickScreen test result (i.e., sample sensitivity) are shown in Table 6

	None	Borderline	Mild	Moderate	Strong	Total
Negative	56.1%	43.9%	0.0%	0.0%	0.0%	100%
Positive	6.5%	29.9%	42.9%	19.5%	1.3%	100%

Table 6: Raw sample specificity (Negative row) and sensitivity (Positive row) values for each QuickScreen test category, based on the results for the 118 known university participants.

The proportion of known university students with and without dyslexia in each QuickScreen test category are shown in Table 7. These are the raw sample predictive values, based on the observed sample prevalence, and do not reflect estimates for the population.

	None	Borderline	Mild	Moderate	Strong
Negative	82.1%	43.9%	0%	0%	0%
Positive	17.9%	56.1%	100%	100%	100%
Total	100%	100%	100%	100%	100%

Table 7: Raw sample predictive values (Negative and Positive) for each QuickScreen test category, based on the results for the 118 known university participants.

The diagnostic accuracy measures for each QuickScreen test category for the known university students are shown in Table 8. These are estimated using the adjusted method (with adjusted logit confidence intervals) and assuming a 10% prevalence of dyslexia.

QuickScreen Test Result	Diagnostic Measure	Estimate	95% Confidence Interval
None	Sensitivity	8.6%	(4.1%, 16.9%)
	PPV	1.7%	(0.8%, 3.5%)
	Specificity	55.6%	(41.0%, 69.3%)
	NPV	98.3%	(96.5%, 99.2%)
Borderline	Sensitivity	30.8%	(21.7%, 41.7%)
	PPV	7.2%	(4.6%, 10.9%)
	Specificity	44.4%	(30.7%, 59.0%)
	NPV	92.8%	(89.1%, 95.4%)
Mild	Sensitivity	43.2%	(32.9%, 54.1%)
	PPV	52.8%	(21.5%, 82.1%)
	Specificity	4.3%	(1.0%, 16.0%)
	NPV	47.2%	(17.9%, 78.5%)
Moderate	Sensitivity	20.9%	(13.4%, 31.1%)
	PPV	35.2%	(11.3%, 69.8%)
	Specificity	4.3%	(1.0%, 16.0%)
	NPV	64.8%	(30.2%, 88.7%)
Strong	Sensitivity	3.6%	(1.2%, 10.8%)
	PPV	8.6%	(1.6%, 35.8%)
	Specificity	4.3%	(1.0%, 16.0%)
	NPV	91.4%	(64.2%, 98.4%)
Mild, Moderate or Strong	Sensitivity	63.0%	(52.0%, 72.8%)
	PPV	62.0%	(28.9%, 86.8%)
	Specificity	4.3%	(1.0%, 16.0%)
	NPV	38.0%	(13.2%, 71.1%)
Borderline, Mild, Moderate or Strong	Sensitivity	91.4%	(83.1%, 95.9%)
	PPV	18.6%	(14.1%, 24.2%)
	Specificity	44.4%	(30.7%, 59.0%)
	NPV	81.4%	(75.8%, 85.9%)
None or Borderline	Sensitivity	37.0%	(27.2%, 48.0%)
	PPV	4.1%	(3.1%, 5.4%)
	Specificity	95.7%	(84.0%, 99.0%)
	NPV	95.9%	(94.6%, 96.9%)

Table 8: Estimates of the diagnostic accuracy measures for each QuickScreen test category using the adjusted logit method, for the 118 known university participants (with 10% prevalence). PPV = Positive Predictive Value; NPV = Negative Predictive Value.

Unknown University Status Group

The results of the analysis outlined in the Methods section for the unknown university status group are presented below.

Test results were available for 127 participants with unknown university status with an independent dyslexia diagnosis; 116 (91.3%) had a positive diagnosis and 11 (8.7%) a negative diagnosis. Of these 127 participants, 12 (9.4%) received a QuickScreen indication of None; 30 (23.6%) an indication of Borderline; 32 (25.2%) Mild; 47 (37.0%) Moderate; and 6 (4.7%) Strong (as shown in the cross-tabulation in Table 9).

	None	Borderline	Mild	Moderate	Strong	Total
Negative	6	5	0	0	0	11 (8.7%)
Positive	6	25	32	47	6	116 (91.3%)
Total	12 (9.4%)	30 (23.6%)	32 (25.2%)	47 (37.0%)	6 (4.7%)	127 (100%)

Table 9: Cross-tabulation of the dyslexia diagnosis (Negative/Positive) versus the QuickScreen test result (None/Borderline/Mild/Moderate/Strong) for the 127 unknown university status participants.

A Fisher's exact test on these data finds strong statistical evidence (p-value < 0.0001) of an association between the independent dyslexia diagnosis and the QuickScreen test indication.

The proportion of participants with unknown university status without dyslexia who received each QuickScreen test result (i.e., sample specificity) and the proportion of participants unknown university status with dyslexia who received each QuickScreen test result (i.e., sample sensitivity) are shown in Table 10.

	None	Borderline	Mild	Moderate	Strong	Total
Negative	54.5%	45.5%	0.0%	0.0%	0.0%	100%
Positive	5.2%	21.6%	27.6%	40.5%	5.2%	100%

Table 10: Raw sample specificity (Negative row) and sensitivity (Positive row) values for each QuickScreen test category, based on the results for the 127 unknown university status participants.

The proportion of participants with unknown university status with and without dyslexia in each QuickScreen test category are shown in Table 11. These are the raw sample predictive values, based on the observed sample prevalence, and do not reflect estimates for the population.

	None	Borderline	Mild	Moderate	Strong
Negative	50.0%	16.7%	0%	0%	0%
Positive	50.0%	83.3%	100%	100%	100%
Total	100%	100%	100%	100%	100%

Table 11: Raw sample predictive values (Negative and Positive) for each QuickScreen test category, based on the results for the 127 unknown university status participants.

Notably, the proportion of participants in the Borderline group with a positive diagnosis is somewhat higher in the unknown university status group compared with the known university student group (83.3% compared with 56.1%).

The diagnostic accuracy measures for each QuickScreen test category for the participants with unknown university status are shown in Table 12. These are estimated using the adjusted method (with adjusted logit confidence intervals) and assuming a 10% prevalence of dyslexia.

QuickScreen Test Result	Diagnostic Measure	Estimate	95% Confidence Interval
None	Sensitivity	6.6%	(3.3%, 12.7%)
	PPV	1.4%	(0.6%, 3.0%)
	Specificity	53.4%	(29.2%, 76.0%)
	NPV	98.6%	(97.0%, 99.4%)
Borderline	Sensitivity	22.5%	(15.9%, 30.8%)
	PPV	5.1%	(2.8%, 9.2%)
	Specificity	46.6%	(24.0%, 70.8%)
	NPV	94.9%	(90.8%, 97.2%)
Mild	Sensitivity	28.3%	(21.0%, 37.0%)
	PPV	19.6%	(5.9%, 48.4%)
	Specificity	12.9%	(3.2%, 40.4%)
	NPV	80.4%	(51.6%, 94.1%)
Moderate	Sensitivity	40.8%	(32.4%, 49.8%)
	PPV	26.0%	(8.4%, 57.2%)
	Specificity	12.9%	(3.2%, 40.4%)
	NPV	74.0%	(42.8%, 91.6%)
Strong	Sensitivity	6.6%	(3.3%, 12.7%)
	PPV	5.4%	(1.3%, 20.0%)
	Specificity	12.9%	(3.2%, 40.4%)
	NPV	94.6%	(80.0%, 98.7%)
Mild, Moderate or Strong	Sensitivity	72.5%	(63.9%, 79.8%)
	PPV	38.4%	(14.2%, 70.1%)
	Specificity	12.9%	(3.2%, 40.4%)
	NPV	61.6%	(29.9%, 85.8%)
Borderline, Mild, Moderate or Strong	Sensitivity	93.4%	(87.3%, 96.7%)
	PPV	18.2%	(11.4%, 27.8%)
	Specificity	46.6%	(24.0%, 70.8%)
	NPV	81.8%	(72.2%, 88.6%)
None or Borderline	Sensitivity	27.5%	(20.2%, 36.1%)
	PPV	3.4%	(2.4%, 4.7%)
	Specificity	87.1%	(59.6%, 96.8%)
	NPV	96.6%	(95.3%, 97.6%)

Table 12: Estimates of the diagnostic accuracy measures for each QuickScreen test category using the adjusted logit method, for the 127 unknown university status participants (with 10% prevalence). PPV = Positive Predictive Value; NPV = Negative Predictive Value.

Potential Further Work

The analysis presented in this report provides an initial assessment of the diagnostic accuracy of the QuickScreen dyslexia test. Further work could potentially be undertaken to expand on this initial analysis and to develop the test further.

In addition to the overall QuickScreen test indications, individual scores are available for various processes such as visual, verbal, memory, reading, comprehension, etc. By using the individual test scores and additional participant demographics we could potentially build a model to predict the probability of dyslexia. This model could then be used to possibly adjust the current QuickScreen indication category boundaries to optimise the resulting diagnostic accuracy measures. This piece of work may be particularly useful in helping to distinguish between individuals currently in the Borderline group by accounting for participants' university status.

Speed of Processing

Another area of potential further research, highlighted by Dr Dee Walker, is to explore how the QuickScreen speed of processing results vary between participants with and without dyslexia.

Table 13 below shows a cross-tabulation of the dyslexia diagnosis versus the speed of processing results available from the QuickScreen data.

	No Difficulties	Average	Difficulties	Total
Negative	27	23	2	52 (21.3%)
Positive	12	103	77	192 (78.7%)
Total	39 (16.0%)	126 (51.6%)	79 (32.4%)	244 (100%)

Table 13: Cross-tabulation of the dyslexia diagnosis (Negative/Positive) versus the QuickScreen speed of processing result (No Difficulties/Average/Difficulties) for the full 245 participants.

We observe that of those 52 participants with a negative dyslexia diagnosis 27 (51.9%), 23 (44.2%) and 2 (3%) have No Difficulties, Average and Difficulties speed of processing results, respectively. Whereas of those 192 with a positive dyslexia diagnosis 12 (6.3%), 103 (53.6%) and 77 (40.1%) have No Difficulties, Average and Difficulties speed of processing results, respectively.

Hence, there appears to be a clear association between speed of processing and dyslexia diagnosis. This supports the case for considering including speed of processing as an explanatory variable in a model for the probability of dyslexia. Furthermore, rather than using the categorical speed of processing result, the continuous score for speed of processing may further distinguish between those more or less likely to have dyslexia.