Investigating Developmental Delays Study: Comparison of SCQ and PDDST

Craig J. Newschaffer; Li-Ching Lee; Angeline David; Nora Lee

Center for Autism and Developmental Disabilities Epidemiology, Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD

Background

The Social Communication Questionnaire (SCQ) was developed as a screening instrument to identify children that should undergo additional assessment for autism spectrum disorders (ASD). Validation analyses focused on its ability to discriminate ASD cases from children with other diagnoses. The Pervasive Developmental Disorders Screening Test (PDDST) Stage I was developed to screen for ASDs among children referred to developmental disorders clinics and has been validated in a similar fashion against groups of children with other diagnoses. The authors of the SCQ report a sensitivity of 85% and specificity of 75% for differentiating diagnoses of PDDs (including autism) from other diagnoses (Berument et al., 1999). The author of the PDDST reported a sensitivity of 69% and a specificity of 63% (Siegal, 1996).

Although considered intrinsic measures of test performance, initial sensitivity and specificity estimates may depend greatly on the samples studied and, consequently, differences in apparent performance of the two tools could be related to differences in samples studies

In this analysis we report on the sensitivity, specificity, positive and negative predictive value of the SCQ and PDDST-II in a single sample of 284 children receiving preschool special education services, ages 3 to 5, in two school districts.

Methods

Surveys containing SCQ and PDDST-II questions were mailed to 1,013 parents of 3 to 5 year old children receiving special education services in Howard County, MD and Christiana School District, DE. Parents or guardians completed surveys and returned them by mail. 284 or (28.0%) surveys were received.

We used the "current" (not lifetime) version of the SCQ and the 1996 version of the PDDST-II with the order randomized.

For the full sample, only an imperfect gold standard of ASD status was available - parent self-report of a clinical ASD diagnosis and/or a documented autism special education classification. For a subsample of 41 subjects, ADI-R data were also available from follow-up clinic visits.

For each instrument we calculated estimates of areas under the ROC curve as well as sensitivity, specificity, predictive value positive, and predictive value negative using standard cutpoints and two alternative "optimal" cutpoints (Youlden's J, Index of Validity).

Figure 1. ROC Curves for SCQ and PDDST-II based parent report of ASD diagnosis and/or autism special education classification

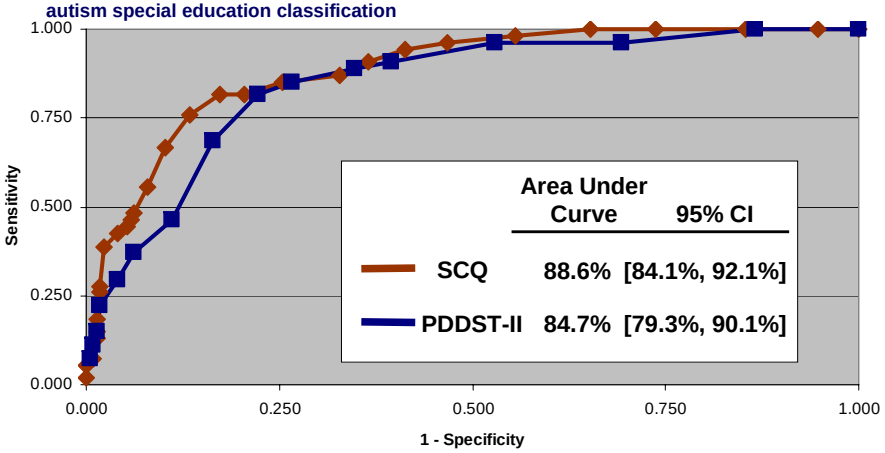


Table 1. SCQ and PDDST-II test performance measures at different cutpoints

Instrument	Cutpoint	Sn [95% CI]	Sp [95% CI]	PPV [95% CI]	NPV [95% CI]
SCQ	Youden's J >=12	81.5 [76.2, 86.8]	82.7 [80.1,85.2]	53.0 [47.5, 58.5]	94.9 [93.3, 96.5]
	Index of validity >=20	38.9 [32.3, 45.5]	97.8 [96.8, 98.9]	80.8 [73.0, 88.5]	87.0 [84.8, 89.1]
	Standard >=15	55.6 [48.8, 62.3]	92.0 [90.3, 93.0]	62.5 [58.5, 69.5]	89.6 [87.6, 91.6]
PDDST-II	Youden's J >=7	81.5 [76.2, 86.8]	77.8 [75.0,80.5]	46.8 [41.7, 52.0]	94.6 [92.9, 96.3]
	Index of validity >=12	22.2 [16.6, 27.9]	98.2 [97.3, 99.1]	75.0 [64.2, 85.8]	84.0 [81.8, 86.3]
	Standard >=4	90.7 [86.8, 94.7]	60.4 [57.2, 63.7]	35.5 [31.4. 39.6]	96.5 [94.9, 98]

Figure 2. ADI-R result by SCQ score

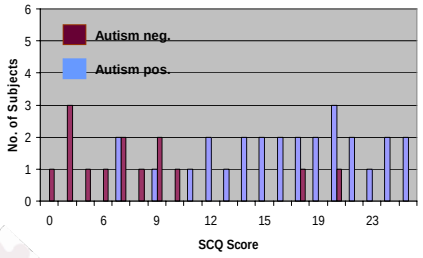
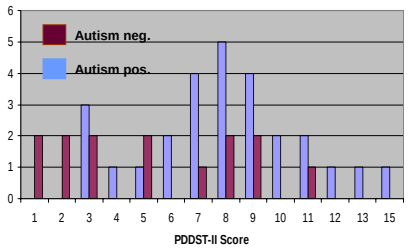


Figure 3. ADI-R result by PDDST-II score



Results

Figure 1 shows the ROC curves, with areas under the curves, for the SCQ and the PDDST-II using the imperfect gold standard of parental self-report of ASD diagnosis and/or autism special education classification. The discriminating ability of both tools are similar in this sample, with the SCQ performing only slightly better. The estimates in Table 2 suggest that, based on analyses in this common sample, at recommended cutpoints the SCQ is the more specific, and the PDDST-II the more sensitive screening tool. The data from the subsample with ADI-R scores (Figure 2 and Figure 3) supports this. An optimal cutpoint based on Youlden's J seeks to maximize both sensitivity and specificity (irrespective of prevalence in the sample) while the Index of validity maximizes the number correctly classified (will favor specificity when prevalence is low). In this sample, the optimal cutpoints for both instruments perform similarly (not unexpected, given the similar ROC curves).

Discussion

In the single sample used in this study, the SCQ and the PDDST-II had very different performance characteristics based on their recommended cutpoints. However, ROC curve and optimal cutpoint analyses suggest that the tools actually perform quite similarly. The difference in performance of the tools when used as recommended is clearly a function of the cutpoints. The selection of a recommended cutpoint reflects both the preference of the developer and the characteristics of the populations used in initial validation studies. Developers select cutpoints based on the setting in which they intend the tool to be used, judgments about the relative costs associated with false negatives and positives, and empirical data available from initial validation studies on test performance. Although reports on the development of a tool rarely give all the specifics regarding these decisions, it appears that the SCQ and the PDDST-II had similar general objectives and where initially validated using similar study designs. Our analysis shows that they will, in fact, perform similarly when the same criteria are used to select cutpoints and a common validation sample is used.

This study was supported by a cooperative agreement from the National Center on Birth Defects and Developmental Disabilities (Cooperative Agreement No. U10/CCU320408-03).