

Improving the Definition of Developmental Delay

Marjolaine M. Limbos, PhD,* David P. Joyce, MD, CCFP†

AQ: A

ABSTRACT: ●●●.

(*J Dev Behav Pediatr* 33:000–000, 2012) **Index terms:** developmental delay, screening, primary care, well-child visit.

AQ: B

We thank Dr. Glascoe and her colleagues for their useful comments regarding our study on the accuracy of screening for developmental delay using the PEDS and ASQ.¹ The sensitivity and specificity of a screening test is determined by comparing it with the best available gold standard for the diagnosis of the condition in question.² In the case of developmental delay and disability, assessment using a number of professionally administered, well-established, reliable, validated, and comprehensive psychological tests is generally accepted as the gold standard.³ Such testing is used widely in clinical practice, was used in our recent study, and has been used in previous studies of screening for developmental delay.^{1,4} As Dr. Glascoe and colleagues point out, psychological tests for developmental delay do not come with a simple yes-or-no result, because of the spectrum of developmental delays, ranging from being simply at-risk to having a disability. Often an arbitrary cutoff is chosen to distinguish between those with developmental delay and those with normal development. In the case of developmental and intellectual disability, the definition is more widely agreed upon—individuals must score 2 or more SDs below the mean and have concurrent significant deficits in adaptive functioning.^{5,6} On the other hand, defining developmental delay is more problematic, but scoring below 1.0 to 1.5 SDs from the mean is often used in research and clinical settings.³ In our study, we found the sensitivity and specificity of the ASQ and PEDS to remain relatively stable whether the 10th percentile (approximately 1.3 SDs) or 1.5 SDs below the mean was used as the cutoff for classifying developmental delay.

Glascoe and colleagues point out that in some settings inclusion of children scoring below the 25th percentile might be prudent, as this would ensure that at-risk children who qualify for special education and other early interventions are identified and referred for services.

Although we agree with the principle of identifying as many at-risk children as possible, we have some reservations about this practice. First, we believe that it is unlikely that brief screening tests, such as the ASQ and PEDS, would have enough discriminative power to correctly classify children using this broader definition of developmental delay. Second, many standardized screening tests, such as the ASQ, set their threshold for a positive screen rather high, so that only those children who are performing significantly below their age-equivalent peers screen positive.⁷ For example, 2 SDs below the mean on any domain of the ASQ is set as the cutoff for a positive screen based on the standardization sample for that test. It follows that if a broader definition of developmental delay is used (i.e., the 25th percentile), many children identified as screen negative on the ASQ would be considered falsely negative, resulting in apparent decrease in the sensitivity of the test. Third, we believe that the risks of identifying too many children with developmental delay may outweigh the potential benefits. It can unnecessarily raise parent concern and in some cases reduce expectations, leading to a self-fulfilling prophecy. This latter point has long been a criticism of developmental screening in young children and is a reason why we decided to use more stringent criteria for developmental delay. Finally, the decision of what cutoff to use can vary widely in studies but should reflect the expected prevalence of developmental delay in the population studied, which is generally considered to be 10% to 15% of children.^{3,8}

We repeated the analysis of sensitivity and specificity, classifying all children who scored below the 25th percentile on the criterion measures as having a developmental delay. As expected, the sensitivity fell significantly for both the PEDS (57%) and the ASQ (48%), although specificity remained in the moderate range for both tests. Changing the cutoff on the criterion from the 10th to the 25th percentile effectively moves children who were originally classified as having normal development into the category of having a developmental delay. Those with developmental delay can be grouped as either true positives (TP) or false negatives (FN), so the number of one or both of these categories may increase with this change of definition. Since the number of false

AQ: C From the *British Columbia Children's Hospital, Vancouver, BC, Canada; †Department of Family Practice, University of British Columbia, Vancouver, BC, Canada.

Disclosure: The authors declare no conflict of interest.

AQ: D Address for reprints: Marjolaine Limbos, PhD, ●●●; e-mail: david.joyce@utoronto.ca.

Copyright © 2011 Lippincott Williams & Wilkins

negatives is in the denominator of the equation for sensitivity ($TP/TP+FN$), a rise in false negatives leads to a decrease in sensitivity.

In summary, we agree that inclusion of all children who may benefit from early intervention programs is an important goal of developmental screening. However, depending on where the “goal posts” are set in the definition of developmental delay, the observed sensitivity and specificity of the screening test will change. In the case of the PEDS and ASQ, it seems that accurately classifying this broader cohort of at-risk children may be difficult. Guided by the findings of our study, and by current clinical practice, we believe that currently available broad-band screening tests are most able to identify those children who meet more traditional criteria for developmental delay, specifically those at or below the 10th percentile on psychological tests.

REFERENCES

1. Limbos MM, Joyce DP. Comparison of the ASQ and PEDS in screening for developmental delay in children presenting for primary care. *J Dev Behav Pediatr*. 2011;32:499-511.
2. Montori VM, Wyer P, Newman TB, Keitz S, Guyatt G; Evidence-Based Medicine Teaching Tips Working Group. Tips for learners of evidence-based medicine: 5. The effect of spectrum of disease on the performance of diagnostic tests. *CMAJ*. 2005;173:385-390.
3. Eriksson M, Westerlund M, Miniscalco C. Problems and limitations in studies on screening for language delay. *Res Dev Disabil*. 2010;31:943-950.
4. Glascoe FP, Byrne KE, Ashford LG, Johnson KL, Chang B, Strickland B. Accuracy of the Denver II in developmental screening. *Pediatrics*. 1992;89:1221-1225.
5. The AAIDD Ad Hoc Committee on Terminology and Classification. *Intellectual Disability: Definition, Classification, and Systems of Supports*. 11th ed. Washington, DC: American Association on Intellectual and Developmental Disabilities; 2010.
6. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. text rev. Washington, DC: American Psychiatric Association; 2000.
7. Squires J, Potter L, Bricker D. *The ASQ User's Guide for the Ages & Stages Questionnaires: A Parent-Completed, Child-Monitoring System*. 2nd ed. Baltimore, MD: Paul H. Brookes Publishing Co; 1999.
8. Boyle CA, Decoufle P, Yeargin-Allsopp M. Prevalence and health impact of developmental disabilities in US children. *Pediatrics*. 1994;93:399-403.