CAD-MDD: Not Diagnostic, Lacks Screening Data

To the Editor: In the recent article by Gibbons and colleagues,[1] their Computerized Adaptive Diagnostic Test for Major Depressive Disorder (CAD-MDD) was presented as a "diagnostic screening tool." However, the authors themselves stated, "Screening measures, like the CAD-MDD, are not diagnostic measures…."[1(p670)] Consistency is needed. The term *diagnostic screening tool* is self-contradictory, and calling this test "diagnostic" is misleading. Screening tests do not make diagnoses—they assess the likelihood of undeclared disease. Diagnostic procedures (tests or interviews) then follow positive screens. Despite its claimed efficiency, CAD-MDD cannot eliminate the need for diagnostic interviews.

The true screening performance of CAD-MDD was not tested. The stated positive predictive value (PPV) of 0.66 would not apply in primary care and epidemiology[1] because these settings do not match the derivation sample. MDD prevalence was 20%, but in primary care it approximates 5% in the general adult population.[2–4] The authors acknowledged this issue but did not compute how this prevalence confound[5] would compromise CAD-MDD performance. With 5% prevalence, sensitivity of 0.94, and specificity of 0.82 (cross-validated results; see Figure 2 and p 672 in the article[1]), the PPV would be 0.22, not 0.66. Negative predictive value would change little, from 0.98 to 0.996. Moreover, the specificity of 0.87 (before cross-validation) highlighted in the abstract[1] is unrealistic because of broad psychiatric exclusions.[1] In populations that have not been "scrubbed" in that way, PPV would be even lower than 0.22.

Data reporting and analyses were suboptimal. Descriptions of and results for the 2 subsamples were not reported separately before aggregation. Specificity was probably no better than 0.50 in the clinical subsample and close to 1.0 in the "scrubbed" control subsample, but these values can only be estimated because of incomplete data reporting. Test-retest reliability—a standard requirement—was not reported. Receiver operating characteristic curve areas were not reported. The confidence of positive/negative screen results (Table 2 in the article[1]) was not reported for the 68 false positive, 127 true positive, 7 false negative, and 454 true negative cases of MDD. If the decision tree iterated to equally high casewise confidence statements for false positive cases as for true positive cases, and the same for the negative cases, that would challenge the value of the computer-generated confidence statements.

Overall, this report features misleading labeling and unjustified claims for major diagnostic screening applications of CAD-MDD, but it lacks real screening data. It also lacks clear logic, transparent data presentation, and essential analyses. The claim that "We now have the ability to efficiently screen large populations for MDD"[1(p674)] is misleading. The test could screen out MDD in general populations with high confidence (0.996), but that is very different from screening "for MDD." A positive screen with CAD-MDD in primary care would move the likelihood estimate from 0.05 to approximately 0.20 or less. Thus, it is not an alternative approach to lengthy diagnostic assessment. CAD-MDD is untested and not ready for research "in primary care, psychiatric epidemiology, molecular genetics, and global health,"[1(p669)] much less for commercial launch.[6] At best, it is a prototype approaching readiness for field testing of its true screening performance, which certainly will be worse than is depicted here.[1] Caveat emptor.

REFERENCES

1. Gibbons RD, Hooker G, Finkelman MD, et al. The computerized adaptive diagnostic test for major depressive disorder (CAD-MDD): a screening tool for depression. *J Clin Psychiatry*. 2013;74(7):669–674.
2. Williams JW Jr, Kerber CA, Mulrow CD, et al. Depressive disorders in primary care: prevalence, functional disability, and identification. *J Gen Intern Med*. 1995;10(1):7–12.
3. O'Connor EA, Whitlock EP, Beil TL, et al. Screening for depression in adult patients in primary care settings: a systematic evidence review. *Ann Intern Med*. 2009;151(11):793–803.
4. Arroll B, Goodyear-Smith F, Kerse N, et al. Effect of the addition of a "help" question to two screening questions on specificity for diagnosis of depression in general practice: diagnostic validity study. *BMJ*. 2005;331(7521):884–887.
5. Galen RS, Gambino SR. *Beyond Normality: The Predictive Value and Efficiency of Medical Diagnoses*. New York, NY: Wiley; 1975.
6. Psychiatric Assessments Inc. DBA Adaptive Testing Technologies corporate Website. http://www.adaptivetestingtechnologies.com/. Accessed September 20, 2013.

Bernard J. Carroll, MBBS, PhD, FRCPsych
bcarroll40@comcast.net

Author affiliations: Pacific Behavioral Research Foundation, Carmel, California.

Potential conflicts of interest: **Dr Carroll** receives royalties from licensing the Brief Carroll Depression Scale and the Carroll Depression Scale-Revised to Multi-Health Systems, Inc (www.mhs.com).
Funding/support: None reported.
J Clin Psychiatry 2014;75(1):85 (doi:10.4088/JCP.13lr08758).