Convergent Validity and Interrater Reliability of Estimating the ABCD² Score From Medical Records

Koto Ishida, MD; Jonathan M. Raser-Schramm, MD, PhD; Christina A. Wilson, MD, PhD; Scott E. Kasner, MD; Michael T. Mullen, MD; Steven R. Messe, MD; Brett Cucchiara, MD

Background and Purpose—The ABCD² score is increasingly used for risk stratification of transient ischemic attack patients. We sought to determine the reliability and convergent validity of retrospective ABCD² score estimation from medical records.

Methods—We compared ABCD² scores that were prospectively determined by a vascular neurology attending to scores determined retrospectively from medical record review. Emergency department records and neurology consult notes for patients with acute transient ischemic attack were abstracted with explicit ABCD² scoring redacted. Scores were estimated by 2 independent raters using these records. Estimated ABCD² component scores, total scores, and risk category were compared both between retrospective raters and with prospectively obtained scores. Reliability was assessed using unweighted κ statistics.

Results—Interrater reliability was substantial with 72% exact agreement in total score between retrospective raters (κ=0.64) and nearly perfect with 82% agreement for ABCD² category (κ=0.71). Interrater agreement was best for age and diabetes mellitus and poorest for clinical features and duration. Agreement between the retrospective raters and prospectively obtained score was >90% for age, blood pressure, and diabetes mellitus, but only ≈70% for clinical features and duration. Retrospectively, estimated total ABCD² score exactly matched the prospective score in 58% of patients for rater 1 and 44% of patients for rater 2. Retrospectively, estimated ABCD² category matched the prospectively scored category in 67% of patients for rater 1 and 71% of patients for rater 2.

Conclusions—The ABCD² score can be abstracted from medical records with substantial interrater reliability but limited convergent validity. This may lead to misclassification of risk category in more than one third of patients. (Stroke. 2013;44:XXX-XXX.)

Key Words: ABCD² score ■ medical records ■ risk stratification ■ transient ischemic attack

Accurate risk stratification of patients with transient ischemic attack (TIA) is critically important to facilitate efficient evaluation and management of these patients. To this end, clinical risk scores, such as the ABCD² score, have been developed to help identify patients at highest risk of stroke during the early period after TIA. The ABCD² score is based on age, blood pressure, clinical features, duration of symptoms, and presence of diabetes mellitus. It has been validated in observational studies, both by the initial group developing the score and by independent groups, with mixed results. Although prospective ABCD² scoring is used to aid clinical decision making, retrospective abstraction is frequently used both for research purposes and score validation. Several validation studies have either used retrospective extraction of ABCD² scoring from medical records or failed to explicitly state how ABCD² scoring was obtained.1-6 However, the accuracy of retrospective abstraction of the score from medical records has never been established. The aim of the current study was to assess the convergent validity (accuracy compared with a prospectively assigned score) and interrater reliability of retrospective estimation of the ABCD² score from medical records.

Methods

Patients with suspected TIA who presented to the emergency department at our hospital between December 2010 and March 2012 and who were seen by a board-certified vascular neurology attending were included. TIA was defined in the traditional sense, as acute onset of focal cerebral or monocular symptoms lasting <24 hours and presumed to be due to a vascular cause. Patients were included only if there was complete resolution of symptoms at the time of emergency department presentation or hospital admission. All patients for whom there was sufficient clinical suspicion to justify diagnostic testing for a neurovascular cause were eligible for inclusion. This study was approved by the University of Pennsylvania Institutional Review Board. At our institution, standard attending initial evaluation templates for all patients with TIA include ABCD² scoring, and attending physicians were explicitly instructed to perform ABCD²
scoring themselves and not rely on housestaff reports. The ABCD² score prospectively calculated by the vascular neurology attending was considered the gold standard.

Medical records from the emergency department as well as the initial neurology resident consultation note were redacted for patient identifiers as well as any explicit ABCD² score documentation. Records were photocopied and distributed to 2 neurologists, who independently determined each component of the ABCD² score and calculated the total score. Agreement between the 2 retrospective raters was assessed, as was agreement between retrospective and prospective raters, for both the total score as well as the individual component items of the score.

Analysis of interrater agreement for the retrospective scorers used weighted and unweighted \( \kappa \) scores assessing individual ABCD² items, total ABCD² score, and ABCD² category (0–3, 4–5, 6–7). Comparison of each retrospective scorer with the prospective attending score (gold standard) was done using percent agreement.

### Results

Raters retrospectively reviewed medical records for 102 TIA patients. Mean patient age was 62 years; 53% were female. Based on prospective neurovascular attending scores, 33% of patients were categorized as low risk (ABCD² score, 0–3); 58% as moderate risk (4–5), and 9% as high risk (6–7).

Interrater reliability of retrospective raters was substantial with 72% exact agreement in total score between raters (\( \kappa =0.64 \)) and nearly perfect with 82% agreement for ABCD² category (\( \kappa =0.71 \)). Retrospectively, estimated total ABCD² score exactly matched the prospective attending score in 58% of patients for rater 1 and 44% of patients for rater 2; retrospective ABCD² category matched the prospectively scored category in 67% of patients for rater 1 and 71% of patients for rater 2. The Table summarizes this data and presents agreement between raters for the individual ABCD² components.

### Discussion

In our cohort, the ABCD² score was abstracted from medical records with substantial interrater reliability but limited convergent validity. Clinical features (C) and duration (D) display the greatest variability among all ABCD² components in retrospective evaluation. These are also the only 2 components that rely on subjective history and patient self-report. Together, they comprise 4 of the 7 possible points of the ABCD² score. In risk stratification of TIA patients in clinical practice, the ABCD² category rather than the exact score is often used to determine short-term risk of ischemic stroke, and division of the score into 3 categories corresponding to low (0–3), moderate (4–5), and high (6–7) risk. We found that roughly one third of patients were misclassified for ABCD² category based on retrospective scoring.

Estimation of clinical scores based on medical record review is widespread, but the accuracy of this approach needs to be confirmed for each individual clinical score. Formal evaluation of the National Institutes of Health stroke scale, for example, has demonstrated that it can be abstracted from medical records with a high degree of reliability and validity.

Our results on the validity of retrospective abstraction of the ABCD² score from medical records serve as a cautionary note. Our study has several limitations. The patient cohort came from a single tertiary care center, only patients presenting to the emergency department were included, and the vascular neurology attending was assumed to be the gold standard. Medical records reviewed by raters included neurology resident consult notes, which, given considerable exposure to the ABCD² score during training, might have been biased to include more relevant historical detail than a typical generalist physician or community-based neurologist might include in their documentation. This suggest that, if anything, our results might reflect an overly optimistic assessment of the validity of the ABCD² score. Finally, we did not assess the actual predictive value of ABCD² scoring for short-term stroke risk (criterion validity), so the clinical significance of our findings remain uncertain. It is possible that more accurate ABCD² scoring obtained prospectively might result in better performance of the ABCD² score at predicting risk.

### References


Convergent Validity and Interrater Reliability of Estimating the ABCD² Score From Medical Records
Koto Ishida, Jonathan M. Raser-Schramm, Christina A. Wilson, Scott E. Kasner, Michael T. Mullen, Steven R. Messe and Brett Cucchiara

Stroke. published online January 22, 2013;
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2013 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/early/2013/01/22/STROKEAHA.112.675611

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org//subscriptions/