Reliability and Validity of a Two-Question Alcohol Screen in the Pediatric Emergency Department

Anthony Spirito, PhD,^a Julie R. Bromberg, MPH,^{b,c} T. Charles Casper, PhD,^d Thomas H. Chun, MD, MPH,^{b,c} Michael J. Mello, MD, MPH,^{b,c} J. Michael Dean, MD, MBA,^d James G. Linakis, PhD, MD,^{b,c} for the Pediatric Emergency Care Applied Research Network

abstract

BACKGROUND AND OBJECTIVE: A multisite study was conducted to determine the psychometric properties of the National Institute of Alcohol Abuse and Alcoholism (NIAAA) 2-question alcohol screen within pediatric emergency departments (PEDs).

METHODS: Participants ($N = 4838$) included 12- to 17-year-old subjects treated in 1 of the 16 participating PEDs across the United States. A criterion assessment battery (including the NIAAA 2-question alcohol screen and other measures of alcohol, drug use, and risk behaviors) was self-administered on a tablet computer. A subsample ($n = 186$) was re-administered the NIAAA 2-question screen 1 week later to assess test-retest reliability.

RESULTS: Moderate to good test-retest reliability was demonstrated. A classification of moderate risk or higher on the screen had the best combined sensitivity and specificity for determining a diagnosis of alcohol use disorder (AUD) for all students. Any past year drinking among middle school students increased the odds of a diagnosis of an AUD according to *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* criteria, whereas the optimal cutoff for high school ages was ≥ 3 drinking days in the past year. The optimal cutoff for drinking days determining a positive Alcohol Use Disorders Identification Test score among middle school subjects was ≥ 1 drinking day, whereas the optimal cutoff for high school subjects was ≥ 2 drinking days.

CONCLUSIONS: The NIAAA 2-question screen is a brief, valid approach for alcohol screening in PEDs. A positive screen suggests that referral for further evaluation is indicated to determine if an adolescent has an AUD.

Departments of ^aPsychiatry & Human Behavior and ^bEmergency Medicine, The Warren Alpert Medical School of Brown University, Providence, Rhode Island; ^cDepartment of Emergency Medicine, Rhode Island Hospital, Providence, Rhode Island; and ^dDepartment of Pediatrics & PECARN Data Coordinating Center, University of Utah, Salt Lake City, Utah

Dr Spirito contributed to the study design, formulated the manuscript concept, and drafted the initial manuscript; Ms Bromberg, Dr Chun, and Dr Mello contributed to the design of the study, and critically reviewed and edited the manuscript; Dr Casper contributed to the design of the study, supervised the analyses, and reviewed and revised the manuscript; Dr Dean contributed to the design of the study, and reviewed and edited the manuscript; Dr Linakis contributed to the design of the study, formulated the manuscript concept, and critically reviewed and edited the manuscript; and all authors approved the final manuscript as submitted.

DOI: 10.1542/peds.2016-0691

Accepted for publication Sep 22, 2016

Address correspondence to James G. Linakis, PhD, MD, Rhode Island Hospital, Department of Emergency Medicine, 55 Claverick St, 2nd Floor, Providence, RI 02903. E-mail: james_linakis_phd@brown.edu

WHAT'S KNOWN ON THIS SUBJECT: Early identification of youth alcohol problems is strongly recommended, yet there is no consensus regarding the best alcohol screening tool for adolescents. Preliminary evidence identified the National Institute of Alcohol Abuse and Alcoholism 2-question screen as a potential tool for pediatric emergency department clinicians.

WHAT THIS STUDY ADDS: This study determined the psychometric properties of the National Institute of Alcohol Abuse and Alcoholism 2-question alcohol screen in a large, diverse pediatric emergency department sample. The screen was found to have adequate reliability and concurrent/convergent validity.

To cite: Spirito A, Bromberg JR, Casper TC, et al. Reliability and Validity of a Two-Question Alcohol Screen in the Pediatric Emergency Department. *Pediatrics*. 2016;138(6):e20160691

TABLE 1 NIAAA Two-Question Screen Risk Assessment Based on the Number of Drinking Days in the Past Year

Middle school:

1. Do you have any friends who drank beer, wine or any drink containing alcohol in the past year?
2. How about you—in the past year, on how many days have you had more than a few sips of beer, wine or any drink containing alcohol?

High school:

1. In the past year, on how many days have you had more than a few sips of beer, wine or any drink containing alcohol?
2. If your friends drink, how many drinks do they usually drink on an occasion?

Age	1–5 d	6–11 d	12–23 d	24–51 d	≥52 d
12–15 y	Moderate	High	High	High	High
16 y	Low	Moderate	High	High	High
17 y	Low	Moderate	High	High	High

The earlier that youth initiate alcohol use, the more likely they are to use other drugs and engage in other problem behaviors, such as sex without contraception, delinquency, and school dropout.^{1,2} For these reasons, medical^{3,4} and federal^{5–7} organizations recommend alcohol screening and intervention (when appropriate) for adolescents within pediatric emergency departments (PEDs) and other health care settings. Previous studies in primary care^{8,9} and in the emergency department^{10–13} have found that although a large majority of physicians have favorable attitudes toward alcohol disorder screening, such services are underutilized. A substantial portion of adolescents use PEDs as their only source of medical care.^{14,15} These individuals are more likely to report substance use and mental health problems, highlighting a need for PED-based alcohol screening.^{16,17} Although the PED is an ideal venue for alcohol screening, screening instruments must involve minimal training and implementation time to be feasible.

In 2011, the National Institute of Alcohol Abuse and Alcoholism (NIAAA) developed an alcohol screening tool for youth that asks about the patient's drinking frequency and friends' drinking to determine alcohol risk. This tool's items and risk levels have been operationally defined by the NIAAA⁶ and are summarized in Table 1. Due in part to its brevity, this screen is ideal for PEDs and pediatric primary settings. Initial analyses of the NIAAA

2-question screen indicated that it may be an effective predictor of current and future alcohol problems,^{18,19} although, to date, the screen has not been rigorously tested.^{19–21} The objective of the present study was to determine the test-retest reliability and concurrent and convergent validity of the NIAAA 2-question screen when delivered in the PED setting.

METHODS

Youth treated in 1 of the participating PEDs in the Pediatric Emergency Care Applied Research Network (PECARN) were enrolled in the study. Established in 2001, PECARN was the first pediatric emergency care research network and currently consists of 18 PEDs located across the country and a data coordinating center. Sixteen of the sites participated in this study (as noted in the Acknowledgments).

All sites received institutional review board approval before enrolling participants. Due to the potential legal implications of adolescent high-risk behavior (eg, illicit alcohol or drug use), a Certificate of Confidentiality was obtained from the US Department of Health and Human Services. Inclusion criteria were as follows: (1) 12 to 17 years of age; (2) seen in the PED for a non-life-threatening injury, illness, or mental health condition; and (3) in the opinion of the clinical staff, were medically, cognitively, and behaviorally stable. Youth were excluded if they were: (1) in severe

acute emotional distress (eg, suicidal, suspected by the clinical staff of being a victim of child abuse); (2) in the opinion of the clinical staff, cognitively impaired and unable to provide informed assent; (3) unaccompanied by an adult qualified to give written permission for the youth's participation in research; (4) unable to read and speak English or Spanish; (5) parents unable to read and speak English or Spanish; (6) without a telephone or an address of residence; or (7) previously enrolled in this study. Adolescents who met inclusion/exclusion criteria and their parent(s) were approached by study staff and asked to provide written assent and written parental permission, respectively.

After enrollment, a criterion assessment battery, including the NIAAA 2-question screen and other measures of alcohol, drug use, and risk behavior was self-administered on a tablet computer. In accordance with the NIAAA guidelines,⁶ the screen was used to group participants into 4 categories: nondrinkers and those with low, moderate, and high risk. These risk classifications are determined based on the number of drinking days in the past year (Table 1). Also of note, when making decisions about referral for further evaluation, clinicians were asked to consider whether patients have friends who drink (middle school, ages 11–14 years) or binge drink (high school, ages 14–18 years). Both the middle school (which asks about peer alcohol use first) and the high

TABLE 2 NIAAA 2-Question Screen Risk Assessment at Baseline and 1-Week Follow-up

Variable	1-Week Follow-up NIAAA Risk Assessment				Total
	Nondrinker	Lower Risk	Moderate Risk	High Risk	
Baseline NIAAA risk assessment					
Nondrinker	126 (92%)	6 (4%)	5 (4%)	0	137
Lower risk	3 (15%)	13 (65%)	3 (15%)	1 (5%)	20
Moderate risk	11 (55%)	0	9 (45%)	0	20
High risk	0	1 (11%)	3 (33%)	5 (56%)	9
Total	140 (75%)	20 (11%)	20 (11%)	6 (3%)	186

Cases above the diagonal ($n = 14$) represent youth endorsing higher risk on retest, and cases below the diagonal ($n = 18$) represent youth endorsing lower risk on retest.

school (which asks first about the individual's own use) versions were administered as appropriate. The definition of binge drinking varies by age and sex; thus, for the purposes of this analysis, we assumed that the participant's friends were the same sex and in the same age category as the participant.

Test-Retest Reliability

A random sample of enrolled participants was contacted by telephone and e-mail 7 to 14 days after the PED visit to repeat the NIAAA 2-question screen to measure test-retest reliability.

Concurrent and Convergent Validity

Concurrent validity, the degree to which the results of a test are comparable to those of an established gold standard measure of the same construct, was assessed with the Alcohol- and Substance-Use Disorder module of the Diagnostic Interview Schedule for Children (DISC).²² The DISC, the most widely used and studied mental health interview, has been tested in both clinical and community populations²³ ages 9 to 17 years and has been used in a number of emergency department screening studies.^{24,25} The DISC has been shown to have high sensitivity (0.73–1.0 for psychiatric disorders, including substance use disorder).²² The DISC was used as the criterion measure for *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5), diagnoses based on participant responses. A question about craving substances was added to the DISC so that the DSM-5

diagnosis for an alcohol use disorder (AUD) could be derived.

Convergent validity, the degree to which 2 tests designed to assess the same construct are related, was measured by using the Alcohol Use Disorders Identification Test (AUDIT),²⁶ the most widely used screen for adolescent alcohol misuse. The AUDIT is a 10-question screen focusing on the quantity and frequency of alcohol use, alcohol dependence, and alcohol-related consequences.²⁷ It has adequate internal consistency ($\alpha = 0.85$ [consumption] and 0.61 [consequences]).²⁴

Statistical Analysis

Because there are 2 different versions of the NIAAA 2-question screen, analyses were performed for the sample as a whole, as well as for middle school and high school ages separately (Table 1). Test-retest reliability was calculated by using an intraclass correlation coefficient (ICC)²⁸ for the overall NIAAA 2-question screen score and for the individual question regarding number of days drinking in the past year. A Fleiss-Cohen weighted κ was also calculated based on categorization of number of drinking days in the past year as none versus any. To assess the relationship between responses to the 2 questions, for middle school, a summary of the distribution of question 2 responses (yes/no, participant drank in the past year) was examined against responses to question 1 (yes/no, any friends

who drank in past year) by using a κ coefficient. These responses were dichotomized due to a high number of zeros. Because there was more variability in the numeric responses for high school participants, a Pearson's correlation coefficient was calculated for this group.

Concurrent validity was examined by using a logistic regression model comparing the odds of a DISC diagnosis (yes/no) against risk categories of the NIAAA 2-question screen. Differences between levels representing a single change in categorization (nondrinker versus low risk, low versus moderate risk, and moderate versus high risk) were tested with the Wald test. The Cochran-Armitage test was used to examine the trend to receive a DISC diagnosis across all of the screen categories. A receiver-operating characteristic curve (ROC) analysis was used to investigate possible cutpoints on the NIAAA 2-question screen score for detecting a DISC diagnosis. The optimal cutpoint was defined as the point at which the sum of sensitivity and specificity was maximized. Test characteristics were calculated at each potential cutpoint, and the area under the curve was used to provide an assessment of the overall accuracy of the screen in predicting DISC diagnosis.

Convergent validity was examined by comparing the AUDIT scores between risk categories of the NIAAA 2-question screen and testing the differences between levels representing single changes in categorization (nondrinker versus

low risk, low versus moderate risk, and moderate versus high risk) with the Wilcoxon rank-sum test, followed by a test of independence. An analysis of variance was used to examine whether AUDIT scores differed across all of the screen categories. AUDIT scores were also compared between participants classified as drinkers and nondrinkers on the NIAAA 2-question screen for each age group by using the Wilcoxon rank-sum test. An ROC analysis was used to investigate possible cutpoints on the NIAAA screen score for detecting an AUDIT score ≥ 4 , which has been used as the clinical cutoff in previous studies with adolescents.^{24,26,29}

Sensitivity was used as the basis for our sample size requirements. We assumed a target sensitivity of 90%. For the 95% confidence interval (CI) around sensitivity to be within $\pm 2.5\%$, ~5000 participants would be needed. We determined that ~200 participants with 1-week follow-up would provide a stable estimate of test-retest reliability.

Multiple imputation was used to handle nonresponse in analyses involving the AUDIT and DISC surveys. We generated 5 imputations by fully conditional specification,³⁰ using backward selection to choose sufficiently predictive models for each variable given the others. The models were constrained to include at least 1 item from the NIAAA 2-question screen to preserve any association between NIAAA outcomes and the other variables. In analyzing the imputed data, Kruskal-Wallis tests were replaced by global tests on the coefficients in a proportional ordinal logistic regression model and χ^2 tests were replaced by Wald tests on a coefficient in a logistic regression model. All analyses were performed by using SAS version 9.4 (SAS Institute, Inc, Cary, NC).

RESULTS

Preliminary Analyses

The analyses include results from the 4838 participants who completed baseline activities in the PED. Approximately the same number of participants was recruited from each of the sites. Participants were equally distributed across sex and age. Forty-six percent of participants identified as white and 26% identified as black; 26% identified as Hispanic. Overall, of the 4838 participants who completed baseline activities, 4.1% of the AUDIT total scores were missing and therefore imputed. Missing data were due to a participant responding “I prefer not to answer,” an item that precluded calculating a total score. In 6.5% of the cases, an AUDIT-positive participant was not diagnosed with an AUD on the DISC, and in 1.9% of cases, a participant who received a diagnosis on the DISC did not reach the AUDIT cutoff score of 4.

Test-Retest Analyses

A total of 186 (68%) of the 274 participants assigned to the test-retest follow-up group completed their 1-week follow-up assessment (average completion date was 9.8 days from enrollment). There were no differences in age, sex, or any of the baseline alcohol use variables between those who completed or did not complete the retest. Of those who completed the retest, 44% completed it online and 56% completed it over the telephone.

On retesting, 14 youth reported a higher (7.5%) and 18 reported a lower (9.7%) NIAAA risk category. The ICC for the 4 NIAAA categories as a score was 0.67 for the entire sample (95% CI, 0.58–0.74), 0.67 for the middle school sample (95% CI, 0.51–0.78), and 0.65 for the high school sample (95% CI, 0.54–0.75) (Table 2). Weighted κ coefficients were as follows: entire sample, $\kappa = 0.63$ (95% CI, 0.51–0.75); middle school sample, $\kappa = 0.57$ (95% CI, 0.27–0.87); and

high school sample, $\kappa = 0.61$ (95% CI, 0.48–0.75). These results suggest moderate agreement.³¹

When responses were dichotomized according to whether participants reported no drinking at all or any drinking days, κ coefficients were as follows: entire sample, $\kappa = 0.65$ (95% CI, 0.52–0.77); middle school, $\kappa = 0.58$ (95% CI, 0.25–0.91); and high school, $\kappa = 0.63$ (95% CI, 0.48–0.78).

The ICC for the number of days the participant reported drinking in the past year for the total sample was 0.32 (95% CI, 0.18–0.44), indicating a fair level of agreement. The ICC was higher for the middle school group (0.50 [95% CI, 0.30–0.66]; $n = 66$) than for the high school group (0.30 [95% CI, 0.13–0.46]; $n = 120$).

Relationship Between Self-report and Friends' Drinking Questions

For the whole sample, the age group-adjusted risk of self-reported drinking was higher among those whose friends drank than among those whose friends did not drink (relative risk [RR], 3.4 [95% CI, 3.1–3.7]). For the middle school group, the risk of drinking was 8-fold higher among those whose friends drank (RR, 8.1 [95% CI, 6.1–11.1]). For the high school group, the risk of drinking was 3-fold higher among those whose friends drank (RR, 2.9 [95% CI, 2.7–3.2]). For high school students, the Pearson correlation coefficient between the 2 questions on the screen was $r = 0.29$ (95% CI, 0.26–0.33; $P < .01$).

Concurrent Validity

Table 3 summarizes whether a participant received a DSM-5 diagnosis of an AUD on the DISC according to categories of the NIAAA 2-question screen. Each change in risk category on the NIAAA 2-question screen leads to a significant difference in DISC diagnosis of an AUD. Table 4 indicates that a classification of moderate risk or higher on the

NIAAA 2-question screen had the best combined sensitivity (89% [95% CI, 69–100] for middle school and 88% [(95% CI, 81–96] for high school) and specificity (91% [95% CI, 90–92] for middle school and 81% [95% CI, 80–82] for high school) for determining an AUD on the DISC.

Figure 1 indicates that for middle school students, a DSM-5 diagnosis of AUD was predicted best by any self-reported drinking in the past year (which is identical to a classification of moderate risk or higher) on the NIAAA 2-question screen. The table accompanying Fig 1 indicates that the optimal cutoff for high school ages, however, was ≥ 3 drinking days in the past year for predicting DISC alcohol use diagnosis according to the DSM-5, with a sensitivity of 93% (95% CI, 87–99) and a specificity of 81% (95% CI, 79–82).

Convergent Validity

Table 5 presents the AUDIT scores according to categories of the NIAAA 2-question screen. With the exception of the low-risk category compared with the moderate-risk category, each change in the screen categories led to a significant difference in AUDIT scores. The overall test comparing NIAAA risk categories and the analysis of variance post hoc test for trends were statistically significant. The Wilcoxon rank-sum test also showed significant differences in the distribution of

TABLE 3 Distribution of DSM-5 Diagnosis of AUD and NIAAA 2-Question Screen Risk Assessment at Baseline

NIAAA Risk	DSM-5 Diagnosis of AUD		Total	P
	No	Yes		
Nondrinker	3633 (100%)	3	3636	—
Lower risk	403 (98%)	6 (2%)	409	<.01
Moderate risk	566 (95%)	30 (5%)	596	.01
High risk	156 (79%)	41 (21%)	197	<.01

The displayed *P* values are based on a logistic regression model comparing the odds of DSM-V diagnosis between those of a given NIAAA risk assessment versus those with the next lowest risk assessment. The Wald test that all of the coefficients associated with NIAAA risk assessment in a logistic regression model of the odds of DSM-V diagnosis yields a *P* value <.01. The Cochran-Armitage test for trend yields a *P* value <.01.

AUDIT scores between drinkers and nondrinkers. Of the participants classified as high risk on the screen, the majority were also categorized as high risk on the AUDIT (ie, 64% had an AUDIT score ≥ 4).

For the clinical cutoff of 4 on the AUDIT, a cutoff of high risk on the NIAAA 2-question screen provided the highest combined sensitivity (78% [95% CI, 63–94]) and specificity (92% [95% CI, 90–93]) for middle school students. A cutoff of lower risk or greater provided the highest sensitivity (95% [95% CI, 93–97]) and specificity (74% [95% CI, 72–75]) for high school students (Table 6). ROC analyses based on the number of self-reported drinking days are shown in Fig 2. The table accompanying Fig 2 indicates that the AUDIT clinical cutoff for middle school students is best predicted by using a cutoff of ≥ 1 drinking day on the NIAAA 2-question screen, with 78% sensitivity (95% CI, 63–94) and 92% specificity (95% CI, 90–93). For high school ages, however, ≥ 2

drinking days best predicted the AUDIT clinical cutoff, with 90% sensitivity (95% CI, 87–93) and 82% specificity (95% CI, 80–83).

DISCUSSION

This article presents psychometric data on a brief measure recommended by the NIAAA to screen for youth alcohol risk. Data were collected from a large, ethnically, racially, and geographically diverse sample from 16 PEDs within the PECARN network. DSM-5 diagnoses were found for 2% of the sample, which is consistent with data from the National Survey on Drug Use and Health.³²

Moderate to good test-retest reliability was found.³³ Test-retest reliability was comparable across the middle and high school samples using both ICC and κ approaches. When responses were dichotomized into drinks versus does not drink, agreement was good.³⁴ Approximately 17% of the sample

TABLE 4 ROC for the NIAAA 2-Question Screen When Predicting DSM-5 Diagnoses by Middle and High School Participants

Group	Predict DSM-5 if NIAAA risk is:	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC
Middle school age group	$\geq$ Nondrinker	1.00 (1.00–1.00)	0.00 (0.00–0.00)	0.00 (0.00–0.01)	—	0.91
	$\geq$ Moderate risk	0.89 (0.69–1.09)	0.91 (0.90–0.92)	0.05 (0.01–0.08)	1.00 (1.00–1.00)	—
	$\geq$ High risk	0.33 (0.02–0.63)	0.99 (0.98–0.99)	0.13 (0.00–0.27)	1.00 (0.99–1.00)	—
	> High risk	0.00 (0.00–0.00)	1.00 (1.00–1.00)	—	1.00 (0.99–1.00)	—
High school age group	$\geq$ Nondrinker	1.00 (1.00–1.00)	0.00 (0.00–0.00)	0.02 (0.02–0.03)	—	0.90
	$\geq$ Lower risk	0.97 (0.93–1.01)	0.67 (0.65–0.69)	0.07 (0.05–0.08)	1.00 (1.00–1.00)	—
	$\geq$ Moderate risk	0.88 (0.81–0.96)	0.81 (0.80–0.82)	0.10 (0.08–0.13)	1.00 (0.99–1.00)	—
	$\geq$ High risk	0.53 (0.41–0.65)	0.95 (0.95–0.96)	0.22 (0.16–0.28)	0.99 (0.98–0.99)	—
	> High risk	0.00 (0.00–0.00)	1.00 (1.00–1.00)	—	0.98 (0.97–0.98)	—

AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value.

TABLE 5 Distribution of AUDIT Overall Scores and NIAAA 2-Question Screen Risk Assessment at Baseline

Variable	Calculated AUDIT Score, (%)				Total	P
	0	1–3	4	> 4		
Baseline NIAAA risk assessment						
Nondrinker	3494 (96)	118 (3)	6 (0)	18 (1)	3636	—
Lower risk	126 (31)	198 (48)	32 (8)	53 (13)	409	<.01
Moderate risk	233 (39)	227 (38)	37 (6)	99 (17)	596	.28
High risk	13 (7)	45 (23)	13 (6)	127 (64)	197	<.01

The displayed *P* values are based on the Wilcoxon rank-sum test comparing participants with a given NIAAA risk assessment versus those with the next lowest risk assessment. A test of independence between AUDIT scores and NIAAA risk assessment groups yielded a *P* value <.01. The analysis of variance test for trend yielded a *P* value <.01.

changed answers on retesting, with comparable rates reporting higher or lower risk categories. This inconsistency might have been related to mode of assessment; all baseline data were collected online while more than one-half of the retest questions were completed over

the telephone with an interviewer. Overall, reliability of the NIAAA 2-question screen was adequate or better given that reliability statistics are affected by the number of items in a scale.³³ The 1 exception was the “drinking days” item for which reliability was fair for middle school

students but poor for high school students. The lower coefficients on the continuous variable of drinking days, compared with the categorical and dichotomous classifications, suggest recall problems when a specific number of drinks is asked of a respondent.

FIGURE 1

ROC analyses predicting DSM-5 diagnosis of AUD by using the NIAAA 2-question screen self-reported number of drinking days for middle school and high school subjects. AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value.

Concurrent validity was examined by using ROC analyses and revealed that categorizing youth as low versus moderate or higher risk on the NIAAA 2-question screen had the best combined sensitivity and specificity for determining a DSM-5 diagnosis of an AUD of any severity (mild, moderate, or severe). This outcome was true regardless of whether the youth was in middle or high school. Similarly, for middle school students, a DSM-5 diagnosis of an AUD was associated with any drinking in the past year as self-reported on the screen. However, the optimal cutoff for high school ages was ≥ 3 drinking days for predicting an AUD. Our finding in high school students that 3 days is an optimal cutoff is consistent with a recent study by Clark et al³⁵ of rural youth attending an outpatient primary care appointment. This study found that 3 drinks in the past year was also the best predictor for middle school students, but we found any drinking to be the best predictor. The difference between studies with middle school students may have been due to the greater percentage of rural middle school students in the sample by Clark et al. We chose to explore test characteristics using a cutpoint that maximizes the sum of sensitivity and specificity. As with

TABLE 6 ROC for the NIAAA 2-Question Screen When Predicting AUDIT Clinical Cutoff Score by Middle and High School Participants

Group	Predict AUDIT ≥ 4 if NIAAA risk is:	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC
Middle school age group	$\geq$ Nondrinker	1.00 (1.00–1.00)	0.00 (0.00–0.00)	0.01 (0.01–0.02)	—	0.86
	$\geq$ Moderate risk	0.78 (0.63–0.94)	0.92 (0.90–0.93)	0.12 (0.07–0.17)	1.00 (0.99–1.00)	—
	$\geq$ High risk	0.34 (0.16–0.52)	0.99 (0.99–1.00)	0.41 (0.20–0.62)	0.99 (0.99–0.99)	—
	> High risk	0.00 (0.00–0.00)	1.00 (1.00–1.00)	—	0.99 (0.98–0.99)	—
High school age group	$\geq$ Nondrinker	1.00 (1.00–1.00)	0.00 (0.00–0.00)	0.12 (0.11–0.13)	—	0.89
	$\geq$ Lower risk	0.95 (0.93–0.97)	0.74 (0.72–0.75)	0.33 (0.30–0.36)	0.99 (0.99–1.00)	—
	$\geq$ Moderate risk	0.71 (0.66–0.76)	0.86 (0.85–0.87)	0.41 (0.37–0.45)	0.96 (0.95–0.96)	—
	$\geq$ High risk	0.36 (0.31–0.41)	0.98 (0.98–0.99)	0.74 (0.68–0.81)	0.92 (0.91–0.93)	—
	> High risk	0.00 (0.00–0.00)	1.00 (1.00–1.00)	—	0.88 (0.87–0.89)	—

AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value.

any screen, a different cutpoint may be used, depending on the tradeoff between sensitivity and false-positive findings.

With respect to convergent validity, a simple classification using 1 item (drinker versus nondrinker) on the NIAAA 2-question screen had the best combined sensitivity and specificity with respect to a clinical cutoff of 4 on the AUDIT for middle school students. The optimal cutoff for high school ages, however, was ≥ 2 drinking days on the screen to predict a clinical cutoff score of 4 on the AUDIT. These cutoff scores, for both AUDs and the AUDIT clinical cutoff, err on the side of overclassification.

There are some limitations to this study that should be considered. First, although the sample was large and diverse, it is not representative of the general population. The study was limited to adolescents being treated in a PED; thus, generalization to other populations may be limited. Second, the order of administration of the criterion measures was varied, but the NIAAA 2-question screen was always administered first, which may have had an effect on the outcomes. Third, correlations of the criterion instruments with the NIAAA 2-question screen might have been affected somewhat because both the AUDIT and DISC ask about frequency of alcohol use but use different response formats than the free choice item on the NIAAA 2-question screen. In addition, the correlations are also

FIGURE 2

ROC analyses predicting AUDIT score by using the NIAAA 2-question screen self-reported number of drinking days for middle school and high school subjects. AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value.

affected by the reliability and validity of each criterion measure. Fourth, the measures were all self-administered, and participants were informed that responses would not be shared with clinical staff; therefore, we cannot

comment on how the screen would perform when the questions are asked by a health care provider. Fifth, test-retest reliability may have been affected by the fact that only about two-thirds of the designated sample

completed the retest, although there were no differences between completers and noncompleters. In addition, most of the retest sample were nondrinkers.

CONCLUSIONS

The NIAAA 2-question screen, which categorizes youth risk level according to frequency of alcohol use, is a valid, rapid, and simple approach for PED-based alcohol screening that is briefer than other comparable screens. Self-administration may be a useful way to screen in a busy clinical practice and has the potential advantage of eliciting more accurate responses from youth.³⁶ However, the NIAAA screen maximizes sensitivity in identifying youth who may be at risk for alcohol use problems. Therefore, either more conservative cutoff scores could be used or additional questioning will be necessary to determine if an adolescent should be referred for further evaluation. Future research should examine the predictive validity of the NIAAA 2-question screen in detecting AUDs at later time periods as well as examining if cutoff scores differ by specific age groups.

ACKNOWLEDGMENTS

The authors acknowledge PECARN and the participating PECARN

sites, including: Baylor College of Medicine/Texas Children's Hospital (R. Shenoi); Boston Children's Hospital (M. Monuteaux); Children's Hospital of Colorado (L. Bajaj); The Children's Hospital of Philadelphia (J. Fein); Children's National Medical Center (K. Brown); Cincinnati Children's Hospital Medical Center (J. Grupp-Phelan); Columbia University/Children's Hospital of New York–Presbyterian (L. Chernick); Hasbro Children's Hospital (A. Spirito); Lurie Children's Hospital of Chicago (E. Powell); Medical College of Wisconsin (M. Levas); Nationwide Children's Hospital (D. Cohen); Nemours/Alfred I. duPont Children's Hospital (C. Mull); St Louis Children's Hospital/Washington University (F. Ahmad); University of California, Davis (T. Horeczko and C. Vance); University of Michigan (A. Rogers); and University of Pittsburgh (B. McAninch and B. Suffoletto). Our efforts would not have been possible without the commitment of the investigators and research coordinators from these sites.

The authors also thank the PECARN Steering Committee members: R. Stanley (chair), B. Bonsu, C. Macias, D. Brousseau, D. Jaffe, D. Nelson, E. Alpern, E. Powell, J. Chamberlain, J. Bennett, J.M. Dean, L. Bajaj, L. Nigrovic, N. Kuppermann, P. Dayan,

P. Mahajan, R. Ruddy, and R. Hickey. A special thanks to the staff at the Data Coordinating Center, including H. Gramse, S. Zuspan, J. Wang, J. M. Dean, M. Ringwood, and T. Simmons, for their dedication and assistance throughout the study. Lastly, the authors thank the subjects and their parents for participating in this study.

ABBREVIATIONS

AUD: alcohol use disorder
AUDIT: Alcohol Use Disorders Identification Test
CI: confidence interval
DISC: Diagnostic Interview Schedule for Children
DSM-5: *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*
ICC: intraclass correlation coefficient
PECARN: Pediatric Emergency Care Applied Research Network
PED: pediatric emergency department
NIAAA: National Institute of Alcohol Abuse and Alcoholism
ROC: receiver-operating characteristic curve
RR: relative risk

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

Copyright © 2016 by the American Academy of Pediatrics

FINANCIAL DISCLOSURE: The authors have indicated they have no financial relationships relevant to this article to disclose.

FUNDING: All phases of this study were supported in part by the National Institute of Alcohol Abuse and Alcoholism (1R01AA021900 to Drs Spirito and Linakis). This project is supported in part by the Health Resources and Services Administration, Maternal and Child Health Bureau, Emergency Medical Services for Children Network Development Demonstration Program, under cooperative agreements U03MC00008 and U03MC00001, U03MC00003, U03MC00006, U03MC00007, U03MC22684, and U03MC22685. This information or content and conclusions are those of the authors and should not be construed as the official position or policy of, nor should any endorsements be inferred by, the Health Resources and Services Administration, the US Department of Health and Human Services, or the US Government. Funded by the National Institutes of Health (NIH).

POTENTIAL CONFLICT OF INTEREST: The authors have indicated they have no potential conflicts of interest to disclose.

REFERENCES

1. Hays RD, Ellickson PL. Associations between drug use and deviant behavior in teenagers. *Addict Behav*. 1996;21(3):291–302
2. Weinberg NZ, Glantz MD. Child psychopathology risk factors for drug abuse: overview. *J Clin Child Psychol*. 1999;28(3):290–297
3. Kokotailo PK; Committee on Substance Abuse. Alcohol use by youth and adolescents: a pediatric concern. *Pediatrics*. 2010;125(5):1078–1087

4. American College of Emergency Physicians. Alcohol screening in the emergency department policy statement. Available at: www.acep.org/Content.aspx?id=29074. Accessed February 11, 2012
5. Higgins-Biddle J, Dilonardo J. *Alcohol and Highway Safety: Screening and Brief Intervention for Alcohol Problems as a Community Approach to Improving Traffic Safety*. DOT HS 811. Washington, DC: National Highway Traffic Safety Administration; 2015
6. National Institute on Alcohol Abuse and Alcoholism. Alcohol screening and brief intervention for youth: a practitioner's guide. Available at: www.niaaa.nih.gov/YouthGuide. Accessed December 1, 2011
7. Substance Abuse and Mental Health Services Administration. Screening, brief intervention and referral to treatment. Available at: www.samhsa.gov/sbirt. Accessed October 13, 2016
8. Harris BR, Shaw BA, Sherman BR, Lawson HA. Screening, brief intervention and referral to treatment for adolescents: attitudes, perceptions and practice of New York school-based health center providers. *Subst Abus*. 2016; 37(1):161–167
9. Yuma-Guerrero PJ, Lawson KA, Velasquez MM, von Sternberg K, Maxson T, Garcia N. Screening, brief intervention, and referral for alcohol use in adolescents: a systematic review. *Pediatrics*. 2012;130(1):115–122
10. Chun TH, Spirito A, Rakowski W, D'Onofrio G, Woolard RH. Beliefs and practices of pediatric emergency physicians and nurses regarding counseling alcohol-using adolescents: can counseling practice be predicted? *Pediatr Emerg Care*. 2011;27(9):812–825
11. D'Onofrio G, Nadel ES, Degutis LC, et al. Improving emergency medicine residents' approach to patients with alcohol problems: a controlled educational trial. *Ann Emerg Med*. 2002;40(1):50–62
12. Graham DM, Maio RF, Blow FC, Hill EM. Emergency physician attitudes concerning intervention for alcohol abuse/dependence delivered in the emergency department: a brief report. *J Addict Dis*. 2000;19(1):45–53
13. O'Rourke M, Richardson LD, Wilets I, D'Onofrio G. Alcohol-related problems: emergency physicians' current practice and attitudes. *J Emerg Med*. 2006;30(3):263–268
14. Oster A, Bindman AB. Emergency department visits for ambulatory care sensitive conditions: insights into preventable hospitalizations. *Med Care*. 2003;41(2):198–207
15. Wilson KM, Klein JD. Adolescents who use the emergency department as their usual source of care. *Arch Pediatr Adolesc Med*. 2000;154(4):361–365
16. Cunningham R, Knox L, Fein J, et al. Before and after the trauma bay: the prevention of violent injury among youth. *Ann Emerg Med*. 2009;53(4):490–500
17. Rhodes KV, Gordon JA, Lowe RA. Preventive care in the emergency department, part I: clinical preventive services—are they relevant to emergency medicine? Society for Academic Emergency Medicine Public Health and Education Task Force Preventive Services Work Group. *Acad Emerg Med*. 2000;7(9):1036–1041
18. Chung T, Smith GT, Donovan JE, et al. Drinking frequency as a brief screen for adolescent alcohol problems. *Pediatrics*. 2012;129(2):205–212
19. Brown S, Donovan J, McGue M, Shulenberg J, Zucker R, Goldman M. Youth alcohol screening workgroup II: determining optimal secondary screening questions. *Alcohol Clin Exp Res*. 2010;34(suppl 2):267A
20. Smith GT, Chung T, Martin CS, Donovan JE, Windle M. Youth alcohol screening workgroup I: measuring consumption of alcohol as a screener in children and adolescents. *Alcohol Clin Exp Res*. 2010;34(suppl 2):267A
21. Kelly SM, Gryczynski J, Mitchell SG, Kirk A, O'Grady KE, Schwartz RP. Validity of brief screening instrument for adolescent tobacco, alcohol, and drug use. *Pediatrics*. 2014;133(5):819–826
22. Fisher PW, Shaffer D, Piacentini JC, et al. *Sensitivity of the Diagnostic Interview Schedule for Children*, 2nd edition (DISC-2.1) for specific diagnoses of children and adolescents. *J Am Acad Child Adolesc Psychiatry*. 1993;32(3):666–673
23. Shaffer D, Fisher P, Lucas CP, Dulcan MK, Schwab-Stone ME. NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. *J Am Acad Child Adolesc Psychiatry*. 2000;39(1):28–38
24. Chung T, Colby SM, Barnett NP, Monti PM. Alcohol use disorders identification test: factor structure in an adolescent emergency department sample. *Alcohol Clin Exp Res*. 2002;26(2):223–231
25. Chung T, Colby SM, O'Leary TA, Barnett NP, Monti PM. Screening for cannabis use disorders in an adolescent emergency department sample. *Drug Alcohol Depend*. 2003;70(2):177–186
26. Fairlie AM, Sindelar HA, Eaton CA, Spirito A. Utility of the AUDIT for screening adolescents for problematic alcohol use in the emergency department. *Int J Adolesc Med Health*. 2006;18(1):115–122
27. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption—II. *Addiction*. 1993;88(6):791–804
28. McGraw K, Wong S. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. 1996;1:30–46
29. Chung T, Colby SM, Barnett NP, Rohsenow DJ, Spirito A, Monti PM. Screening adolescents for problem drinking: performance of brief screens against DSM-IV alcohol diagnoses. *J Stud Alcohol*. 2000;61(4):579–587
30. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res*. 2007;16(3):219–242
31. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37–46
32. Substance Abuse and Mental Health Services Administration. Behavioral health trends in the United States: results from the 2014 National Survey on Drug Use and Health, 2015. HHS Publication No. SMA 15-4927, NSDUH Series H-50. Available at: <http://www.samhsa.gov/data/sites/default/files/NSDUH-FRR1-2014/NSDUH-FRR1-2014.pdf>. Accessed October 12, 2016

33. Fleiss J. *Statistical Methods for Rates and Proportions*. 2nd ed. New York, NY: Wiley; 1981
34. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174
35. Clark DB, Martin CS, Chung T, et al. *Screening for Underage Drinking and Diagnostic and Statistical Manual of Mental Disorders*. 5th Edition Alcohol Use Disorder in rural primary care practice. *J Pediatr*. 2016;173:214–220
36. Dolezal C, Marhefka SL, Santamaria EK, Leu CS, Brackis-Cott E, Mellins CA. A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. *Arch Sex Behav*. 2012;41(2):401–410