Title
Reliability of Clinical Assessments in Older Adults With Syncope or Near Syncope

Journal
Academic Emergency Medicine, 23(9)

ISSN
1069-6563

Authors
Nishijima, Daniel K
Laurie, Amber L
Weiss, Robert E
et al.

Publication Date
2016-09-01

DOI
10.1111/acem.12977

Peer reviewed

Published in final edited form as:

Acad Emerg Med. 2016 September ; 23(9): 1014–1021. doi:10.1111/acem.12977.

Reliability of clinical assessments in older adults with syncope or near syncope

Daniel K. Nishijima, MD, MAS[1], **Amber L. Laurie, MS**[2], **Robert E. Weiss, PhD**[3], **Annick N. Yagapen, MPH, CCRP**[2], **Susan E. Malveau, MSBE**[2], **David H. Adler, MD, MPH**[4], **Aveh Bastani, MD, FACEP**[5], **Christopher W. Baugh, MD, MBA**[6], **Jeffrey M. Caterino, MD, MPH, FACEP**[7], **Carol L. Clark, MD, MBA, FACEP**[8], **Deborah B. Diercks, MD, MPH**[9], **Judd E. Hollander, MD**[10], **Bret A. Nicks, MD, MHA**[11], **Manish N. Shah, MD, MPH**[12], **Kirk A. Stiffler, MD**[13], **Alan B. Storrow, MD**[14], **Scott T. Wilber, MD**[13], and **Benjamin C. Sun, MD, MPP**[2]

[1]Department of Emergency Medicine, UC Davis School of Medicine, Sacramento, CA

[2]Center for Policy and Research in Emergency Medicine, Department of Emergency Medicine, Oregon Heath & Science University, Portland, OR

[3]Department of Biostatistics, University of California, Los Angeles, CA

[4]Department of Emergency Medicine, University of Rochester, NY

[5]Department of Emergency Medicine, William Beaumont Hospital-Troy, Troy, MI

[6]Department of Emergency Medicine, Brigham & Women's Hospital, Boston, MA

[7]Department of Emergency Medicine, The Ohio State University Wexner Medical Center, Columbus, OH

[8]Department of Emergency Medicine, William Beaumont Hospital-Royal Oak, Royal Oak, MI

[9]Department of Emergency Medicine, University of Texas-Southwestern, Dallas, TX

[10]Department of Emergency Medicine, Thomas Jefferson University Hospital, Philadelphia, PA

[11]Department of Emergency Medicine, Wake Forest School of Medicine, Winston Salem, NC

[12]Department of Emergency Medicine, University of Wisconsin-Madison, Madison, WI

[13]Department of Emergency Medicine, Summa Health System, Akron, OH

[14]Department of Emergency Medicine, Vanderbilt University, Nashville, TN

Abstract

Objectives—Clinical prediction models for risk stratification of older adults with syncope or near syncope may improve resource utilization and management. Predictors considered for inclusion into such models must be reliable. Our primary objective was to evaluate the interrater agreement of historical, physical examination, and electrocardiogram (ECG) findings in older adults undergoing ED evaluation for syncope or near syncope. Our secondary objective was to

Corresponding Author: Daniel Nishijima, MD, MAS, University of California, Davis Medical Center, 4150 V. St. PSSB, 2100, Sacramento, CA 95817, TEL (916) 734.1376, FAX (916) 734.7950, dnishijima@ucdavis.edu.

assess the level of agreement between clinicians on the patient's overall risk for death or serious cardiac outcomes.

Methods—We conducted a cross-sectional study at 11 EDs in adults 60 years of age or older who presented with unexplained syncope or near syncope. We excluded patients with a presumptive cause of syncope (e.g., seizure), or if they were unable or unwilling to follow-up. Evaluations of the patient's past medical history and current medication use were completed by treating provider and trained research associate pairs. Evaluations of the patient's physical examination and ECG interpretation were completed by attending/resident, attending/advanced practice provider, or attending/attending pairs. All evaluations were blinded to the responses from the other rater. We calculated the percent agreement and kappa statistic for binary variables. Interrater agreement was considered acceptable if the kappa statistic was 0.6 or higher.

Results—We obtained paired observations from 255 patients; mean age was 73 years (SD 9 years), 137 (54%) were male and 204 (80%) were admitted to the hospital. Acceptable agreement was achieved in 18 of the 21 (86%) past medical history and current medication findings, none of the 10 physical examination variables, and 3 of the 13 (23%) ECG interpretation variables. There was moderate agreement (Spearman correlation coefficient, r=0.40) between clinicians on the patient's probability of 30-day death or serious cardiac outcome though, as the probability increased, there was less agreement.

Conclusions—Acceptable agreement between raters was more commonly achieved with historical rather than physical examination or ECG interpretation variables. Clinicians had moderate agreement in assessing the patient's overall risk for a serious outcome at 30 days. Future development of clinical prediction models in older adults with syncope should account for variability of assessments between raters and consider the use of objective clinical variables.

Introduction

Syncope is the transient loss of consciousness followed by spontaneous and complete recovery.[1] Syncope accounts for 740,000 emergency department (ED) visits and 250,000 hospital admissions in the US annually.[2] Differentiation between life-threatening etiologies such as arrhythmias or structural heart disease and benign etiologies such as vasovagal syncope is often difficult during an ED evaluation. This clinical dilemma is particularly pertinent to older adults (60 years or older) who have more co-morbidities and a higher prevalence of cognitive deficits than younger patients. Older adults with syncope also have a relatively high incidence of adverse outcomes – 6% of older adults with undifferentiated etiology of syncope in the ED experience death or serious cardiac outcome within 30 days.[3]

These factors contribute to a 85% hospitalization rate for older adults with syncope.[4] Hospitalization however, often does not ultimately lead to a diagnosis of the etiology of the syncopal event or to any therapeutic benefit. Up to 50% of admitted patients with syncope are discharged from the hospital without any clear etiology of the event and 60% receive no specific treatments during hospitalization.[5] Furthermore, admission has not been shown to improve one-year mortality in high-risk patients with syncope.[6]

The development of a clinical prediction model that accurately risk-stratifies older adults with syncope across a broad population has the potential to improve resource utilization and

management of these patients.[7] Prior clinical prediction models to risk stratify patients with syncope have been developed.[8] However, study design flaws (e.g., inclusion of young patients with clear vasovagal syncope or inclusion of subjects with serious conditions identified during ED evaluation) and the failure of external validation studies to replicate derivation study test characteristics have limited implementation of these instruments. The interpretation of high-risk clinical findings, in particular abnormal electrocardiogram (ECG) findings, has been suggested as one potential reason for the failure of external validation studies.[8]

Clinical findings considered for inclusion into clinical prediction models must by both reproducible and reliable.[9] Nonreproducible findings should not be included into clinical prediction instruments, as they will likely impair the performance of the instruments in clinical practice. Thus, the primary objective of our study was to evaluate the interrater agreement of clinical findings in older adults undergoing ED evaluation for syncope. Our secondary objective was to evaluate the level of agreement between clinicians on the patient's overall risk for death or serious cardiac outcomes.

Methods

Study Design

We conducted a multicenter cross-sectional study that was part of a larger prospective cohort study to derive and validate a novel risk prediction model for 30-day death or serious cardiac outcomes in older adults with unexplained syncope (ClinicalTrials.gov Identifier: NCT01802398). The study was approved by the Institutional Review Boards at all sites and written, informed consent was obtained from all participating subjects.

Study Setting and Population

We conducted the study at 11 academic EDs. This was a convenience sample of patients 60 years or older who presented to the ED with syncope or near-syncope. Patients with a presumptive cause of loss of consciousness due to seizure, stroke or transient ischemic attack, or hypoglycemia were excluded. Patients who were intoxicated from alcohol or other drugs, required medical or electrical intervention to restore consciousness, or who were unable or unwilling to provide informed consent and follow-up information were also excluded. We required at least two ED providers to evaluate the patient and rate items; acceptable provider pairs included: attending/resident, attending/advanced practice provider (e.g., nurse practitioner or physician's assistant), or any treating physician/site principal investigator.

Study Protocol

All patients underwent standardized history, physical examination, laboratory testing, and 12-lead ECG testing. Research assistants queried about symptoms associated with syncope directly from patients; therefore, we did not include items about symptoms in this assessment of interrater reliability. We collected data from initial treating providers on the patient's past medical history, medications, physical examination findings, and ECG interpretation. We also collected data on the probability that the patient will experience 30-

day cardiac death or a serious cardiac event as assessed by the treating provider. A serious cardiac event was defined as a significant arrhythmia, myocardial infarction, a new diagnosis of clinically significant structural heart disease, or a major cardiac intervention (pacemaker, implantable defibrillator, open-heart surgery, or angioplasty). Initial treating providers included attending or resident physicians or advanced practice providers. A second provider, blinded to the responses from the first treating provider, performed an independent evaluation during the patient's ED evaluation (eAppendix). Second providers included attending or resident physicians, advanced practice providers, or the site principal investigators who are all emergency physicians. Data collected from the second provider included ECG interpretation, physical examination findings, and the probability of the patient having a 30-day cardiac death or a serious cardiac event. Site research assistants completed second evaluations of the patients' past medical history and current medications based on review of the electronic medical record. Patients in whom the index ED visit was the first visit to the study site did not have second evaluations of past medical histories and current medications. Also, data abstracted by the site research assistants did not include any new findings identified after the index ED visit (i.e., included only data available to the treating provider during the ED visit). Finally, a third physician rater (a board certified cardiologist) who was blinded to all clinical data centrally reviewed all ECGs.

Measurements

Data variables collected were consistent with reporting guidelines for ED based syncope research.[10] We collected data on 17 comorbid factors such as a history of congestive heart failure or prior stroke with responses marked as present or not present/unknown. Data on current medications were organized by class of drug and included diuretics, beta-blockers, alpha blockers, nitrates, other antiarrhythmic agents (e.g., amiodarone), and calcium channel blockers. Physical exam findings included cardiac findings (e.g., carotid bruit, heart murmur), neurological findings (e.g., visual disturbances, speech abnormalities, focal weakness), and gastrointestinal findings (positive fecal occult blood test) and were marked as present, absent, or not assessed. ECG interpretations were based on the first ECG obtained in the ED and were categorized into three mutually exclusive categories: normal, isolated nonspecific ST segment/T wave abnormalities, or abnormal. Normal ECG interpretations included sinus tachycardias (>100 beats per minute [bpm]), sinus bradycardias (>40 bpm), isolated premature atrial or ventricular contractions, and incomplete right bundle branch blocks. Abnormal ECG interpretations included non-sinus rhythms (included paced rhythms), multiple premature ventricular complexes, sinus bradycardias (≤ 40 bpm), ventricular hypertrophies, short PR segment intervals (<10 milliseconds [ms]), axis deviations, atrioventricular node blocks, complete bundle branch blocks, Brugada patterns, Wolff-Parkinson-White patterns, abnormal QRS duration (>120 ms) or abnormal QTc prolongations (>450 ms), and evidence of acute or chronic ischemia. Estimation of the probability that the patient will experience 30-day cardiac death or serious cardiac event was assessed as a percentage from 0 to 100%. A serious cardiac event was defined as a significant arrhythmia (ventricular fibrillation, symptomatic ventricular tachycardia longer than 30 seconds, sick sinus syndrome, sinus pause longer than 30 seconds, Mobitz II heart block, complete heart block, symptomatic supraventricular tachycardia, or symptomatic bradycardia <40 beats per minute), myocardial infarction, a

new diagnosis of clinically significant structural heart disease (aortic stenosis <1cm^2, hypertrophic cardiomyopathy with outflow tract obstruction, severe pulmonary hypertension [mean pulmonary artery pressure >30 mmHg], left atrial myxoma/thrombus with outflow tract obstruction), or a major cardiac intervention (pacemaker or implantable defibrillator placement, open-heart surgery, or angioplasty).

Data Analysis

Data analyses were performed using SAS 9.4 (SAS Institute Inc. Cary, NC, USA). For each item the percent present per initial provider, raw agreement, the percent specific agreement within each response option, and the kappa statistic (with 95% confidence intervals [CI]) were calculated using normal approximation methods.[11] ECG findings were reviewed by three raters -- the initial provider, a second provider evaluator, and a central reader. Agreement statistics were calculated between the initial and second provider and for each provider versus the central reader. To compare agreement across the three raters, Fleiss's kappa and 95% CIs were also calculated.[12] We evaluated interrater reliability acceptability thresholds 0.60 for the kappa statistic.[13] To adjust for the prevalence and bias of clinical findings, we also calculated the prevalence adjusted bias adjusted kappa (PABAK).[14] Scatter plots of percent agreement and the kappa statistic were generated and grouped by historical, physical examination, and ECG findings. To evaluate the correlation between the two raters' estimated probability of 30-day cardiac death or serious cardiac event (non-normal continuous data), we calculated the Spearman correlation coefficient and generated scatter plots. A Bland-Altman plot was created to graphically illustrate discrepancies between raters by plotting the differences in the raters' probability of a cardiac event by their average probability of a cardiac event.

To evaluate if interrater agreements were similar by level of training (95% CIs for kappa overlapped), we stratified agreements based level of training. For historical variables we compared agreements if the initial evaluator was an attending physician or a resident or advanced practice provider. For physical examinations and ECG interpretations, we compared the agreements between attending/attending evaluations and attending/resident or advanced practice provider evaluations. We also generated scatter plots grouped by historical, physical examination, and ECG findings comparing attending/attending pairs and attending/resident or advanced practice provider pairs. For the estimated probability of 30-day cardiac death or serious cardiac outcome, we grouped probabilities into low (0 to <3%), intermediate (3 to 10%), and high (>10%) risk categories and compared the percent agreement for these groups between all pairs, attending/attending pairs, and attending/resident or advanced practice provider pairs. Grouping of risk was skewed to differentiate between low-risk patients.[15] We had an a priori target of 250 paired assessments that was based primarily on feasibility rather than power considerations.

Results

Characteristics of the Subjects

We collected paired independent observations by two clinicians from 255 subjects from Sept 2014 to Oct 2015 (10.1% of the 2,524 patients enrolled during this period). Subjects had a

mean age of 72.7 years (SD 8.8 years), 137 (53.5%) were male, and 204 (80.0%) were admitted to the hospital. There were 111 (43.9%) attending/resident pairs, 100 attending/advanced practice provider pairs (39.5%), and 42 (16.6%) attending/attending pairs. Descriptions of the pairs by site are represented in eTable 1. Demographic information and proportion of patients admitted to the hospital were similar between the enrolled subjects included in the interobserver agreement cohort and those that were not included (eTable 2).

Main Results

Historical Findings—A history of hypertension requiring medications (66.0%), coronary artery disease (26.0%), arrhythmia (23.3%), and diabetes requiring medications (22.8%) were the most common historical findings present per initial providers (Table 1). Beta-blockers (32.1%), diuretics (25.6%), and calcium channel blockers (19.5%) were the most common medications per initial providers. Percent agreements for patients' past medical history and current medications ranged from 83.7% to 99.5% while kappa statistics ranged from 0.34 to 0.94. Eighteen of the 21 (86%) of the historical findings had acceptable kappa statistics (0.60 or higher). The three historical findings that did not meet this threshold were history of peripheral vascular disease, alpha blocker use, and other antiarrhythmic agent use.

Physical Examination Findings—The most common physical exam findings present were heart murmur (10.6%) and positive fecal occult blood test (4.3%) (Table 2). Percent agreements for physical exam findings ranged from 60.4 to 95.7% and the kappa statistic ranged from -0.06 to 0.26. None of the 10 physical exam findings had kappa statistics 0.60 or higher.

ECG Findings—There were 248 ECG evaluations reviewed by all three raters of which 118 (47.6%) were reported as normal by the initial provider (Table 3). The most common abnormal ECG findings present per the initial provider were non-sinus rhythms (8.9%), acute or chronic ischemic changes (7.7%), and prolonged QTc intervals (7.3%). Percent agreement of specific ECG interpretations between the initial and the second rater ranged from 91.9 to 98.8% and the kappa statistic ranged from 0.18 to 0.79. Percent agreement and the kappa statistic between the initial and the second rater for an abnormal ECG were 82.3% and 0.65 (95% CI 0.55 to 0.74). Three of the 13 specific ECG findings had acceptable agreement (kappa 0.60 or higher) between raters: complete LBBB, multiple PVCs, and first degree heart block. Agreement statistics between the initial and the central rater, the second and the central rater, and across all three raters were similar to agreement statistics between the initial and second rater (eTable 3).

Clinical Impression for 30-day Outcome—There were 250 paired evaluations of the probability that the patient will experience 30-day cardiac death or a serious cardiac event. The median probabilities estimated from the initial provider and second providers were 5% (IQR 2 to 10%) and 5% (IQR 2 to 10%). The Spearman correlation coefficient between the first and second providers was 0.40 (see Figure 1, Panel A). When the average estimated rating of the probability of 30-day cardiac death or serious outcome was small, there was good agreement; however as the average estimate probability increased, the agreement decreased (Figure 1, Panel B).

Stratified Analyses by Level of Training—Stratified analyses of interrater agreement for historical, physical examination, and ECG variables are described in eTables 4 to 6 and the eFigure. For the estimated probability of 30-day cardiac death or serious cardiac outcome the overall percent agreement was higher between attending/attending pairs (43 pairs, total percent agreement 81.4% [95% CI 66.6 to 92.6%]) than attending/non-attending pairs (207 pairs, total percent agreement 47.8% [95% CI 40.9 to 54.9%]) (eTable 7).

Discussion

Our study is the most comprehensive assessment of interrater agreement in the ED assessment of older adults with syncope to date. The results of our 11-site study demonstrated a few patterns that describe interrater agreement in this patient population. First, it appears historical findings, specifically past medical history and medication use, had an overall higher level of interrater agreement compared to physical examination or specific ECG findings (Figure 2). The global assessment of whether an ECG was “abnormal” had acceptable interrater reliability. Second, physical examination findings were very rare and when present, demonstrated very poor interrater agreement as assessed via the kappa statistic; however, percent agreement was high (>80%) for 6 of the 10 physical findings. Specific ECG findings also did not have overall acceptable interrater agreement with only 3 of 13 ECG findings having kappa statistics ≥0.60; however, again percent agreement was high (>80%) for all of the specific ECG findings. Third, we found moderate agreement between clinicians on predicting the probability of 30-day death or serious cardiac outcome.

There were a number of findings in which the percent agreement was high (>80%) but the kappa statistic was fair or poor (<0.40) due to what is known as the “kappa paradox”.[16] In situations where the prevalence is very low or very high, the resulting kappa statistic may not fully reflect the reliability of the measure, necessitating the use of other measures such as percent agreement.[17] In particular, specific physical examination findings were infrequent. Eight of the 10 findings physical examination findings were coded as present by the initial rater less than 5 times. Our threshold of a kappa statistic ≥0.60 for acceptable interrater agreement is based on precedent.[18,19] Different thresholds of acceptability (e.g., kappa statistic ≥ 0.50) or use of the lower bound of the 95% confidence interval of the kappa statistic (≥ 0.40) have also been used.[20-22]

Only one other study prospectively assessed interrater agreement in the ED evaluation of syncope patients.[21] In the derivation of the San Francisco Syncope Rule (SFSR), to predict patients at risk for short-term serious outcomes, interrater agreement at a single site was assessed for physical examination and ECG findings. All six of the physical examination findings evaluated for interrater agreement (new neurological deficits, rales, abnormal heart sounds, carotid bruits, systolic murmurs, and diastolic murmurs), had a kappa statistic less than 0.60 (range 0.01 to 0.56).[21] Two ECG/rhythm findings were evaluated in the derivation of the SFSR -- abnormal ECG (new changes) had a kappa statistic of 0.69 (95% CI 0.61 to 0.77) and abnormal rhythm (non-sinus) had a kappa statistic of 0.56 (95% CI 0.45 to 0.67).

The lack of acceptable agreement of physical examination and ECG findings suggests that including subjective variables into clinical prediction models to risk stratify ED patients with

syncope will ultimately lead to models with unstable test characteristics in external validations studies. A systematic review by Serrano et al evaluated the accuracy and quality of clinical prediction models for syncope in the ED.[8] This review identified nine different syncope clinical prediction models.[3,21,23-29] Three of these models included physical examinations findings as one of the high-risk variables including: the presence of rales,[25] valvular heart disease on exam,[26] and positive fecal occult blood test.[26,28] None of these studies conducted an assessment of interrater agreement or have been externally validated.

All nine of the clinical prediction models included abnormal ECG findings as a high-risk clinical variable. This is not surprising as the ECG is widely considered the most important diagnostic test in the evaluation of ED syncope.[30,31] The definition of an “abnormal” ECG varies among published prediction models. Our findings suggest that a global assessment for “abnormal” findings had greater reliability than identification of specific ECG findings. In addition, it may be possible to improve accuracy of interval (e.g. QRS, QTc) and axis based abnormalities through direct abstraction of computer calculated values.

Future syncope prediction models may preferentially include highly objective variables such as age and diagnostic testing such as hematocrit, brain natriuretic peptide, and troponin testing.[32] Many of these diagnostic tests were not included in early clinical prediction models and more recent studies have suggested these tests may have a role in risk stratification of ED patients with syncope.[33] Incorporation of clinician impression into clinical prediction models may also improve risk stratification of ED patients.[9,15,34] However, our study suggests that clinicians did not have good agreement in assessing the patient's overall risk for a serious outcome at 30 days, further supporting the need for the development of an objective risk score.

Limitations

Our results should be interpreted in the context of some limitations. First, the generalizability of our study is limited as all sites were academic, teaching hospitals; therefore our results may not be applicable to smaller community hospitals with different patient populations or staffing models. Institutions solely staffed by attending physicians may produce higher levels of interrater agreement than those achieved within our cohort. Second, residents or advanced practice providers who may not have the clinical expertise as attending physicians did many of the evaluations. We did stratified analyses based on level of training to further evaluate differences in agreement by level of training. Third, some of the findings (e.g., carotid bruit or double vision) were rarely present or frequently not assessed, thus resulting in wide confidence intervals for measurements of reliability. In addition, variables collected for this study were based on prior studies and expert consensus;[10] however there may be other important clinical findings that we did not measure. Fourth, our study was a convenience sample of patients and thus sampling bias may occur, and estimates of kappa may not be comparable at other institutions where the prevalence of these findings is significantly different. However, an analysis of enrolled versus not enrolled patients demonstrated similar patient characteristics. Fifth, the primary outcome of the study was 30-day cardiac death or serious cardiac event. Other clinically

important etiologies of syncope such as sepsis or dehydration were not predicted in this study.

Conclusions

Acceptable agreement between raters was more commonly achieved with historical variables than with physical examination or ECG interpretation variables. Clinicians had only moderate agreement in assessing the patient's overall risk for a serious outcome at 30 days. Future development of clinical prediction models in older adults with syncope should account for variability of assessments between raters and consider the use of objective clinical variables.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Sources and Support: Research reported in this publication was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number R01HL111033. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Moya A, Sutton R, Ammirati F, et al. Guidelines for the diagnosis and management of syncope (version 2009). Euro Heart J. 2009; 30:2631–71.
2. Sun BC, Emond JA, Camargo CA Jr. Characteristics and admission patterns of patients presenting with syncope to U.S. emergency departments, 1992-2000. Acad Emerg Med. 2004; 11:1029–14. [PubMed: 15466144]
3. Sun BC, Derose SF, Liang LJ, et al. Predictors of 30-day serious events in older patients with syncope. Ann Emerg Med. 2009; 54:769–78. [PubMed: 19766355]
4. Birnbaum A, Esses D, Bijur P, Wollowitz A, Gallagher EJ. Failure to validate the San Francisco Syncope Rule in an independent emergency department population. Ann Emerg Med. 2008; 52:151–9. [PubMed: 18282636]
5. Linzer M, Yang EH, Estes NA 3rd, Wang P, Vorperian VR, Kapoor WN. Diagnosing syncope Part 2: Unexplained syncope. Clinical Efficacy Assessment Project of the American College of Physicians. Ann Intern Med. 1997; 127:76–86. [PubMed: 9214258]
6. Getchell WS, Larsen GC, Morris CD, McAnulty JH. Epidemiology of syncope in hospitalized patients. J Gen Intern Med. 1999; 14:677–87. [PubMed: 10571716]
7. Sun BC, Costantino G, Barbic F, et al. Priorities for emergency department syncope research. Ann Emerg Med. 2014; 64:649–55. [PubMed: 24882667]
8. Serrano LA, Hess EP, Bellolio MF, et al. Accuracy and quality of clinical decision rules for syncope in the emergency department: a systematic review and meta-analysis. Ann Emerg Med. 2010; 56:362–73. [PubMed: 20868906]
9. Green SM, Schriger DL, Yealy DM. Methodologic standards for interpreting clinical decision rules in emergency medicine: 2014 update. Ann Emerg Med. 2014; 64:286–91. [PubMed: 24530108]
10. Sun BC, Thiruganasambandamoorthy V, Cruz JD. Consortium to Standardize EDSRSR. Standardized reporting guidelines for emergency department syncope risk-stratification research. Acad Emerg Med. 2012; 19:694–702. [PubMed: 22687184]
11. Graham P, Bull B. Approximate standard errors and confidence intervals for indices of positive and negative agreement. J Clin Epidemiol. 1998; 51:763–71. [PubMed: 9731925]

12. Fleiss, JL.; Levin, B.; Paik, MC. The Measurement of Interrater Agreement Statistical Methods for Rates and Proportions. Third. New York, NY: John Wiley & Sons, Inc; 2003.
13. McHugh ML. Interrater reliability: the kappa statistic. Biochemia medica. 2012; 22:276–82. [PubMed: 23092060]
14. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther. 2005; 85:257–68. [PubMed: 15733050]
15. Nishijima DK, Sena M, Galante JM, et al. Derivation of a clinical decision instrument to identify adult patients with mild traumatic intracranial hemorrhage at low risk for requiring ICU admission. Ann Emerg Med. 2014; 63:448–56. [PubMed: 24314900]
16. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol. 1990; 43:543–9. [PubMed: 2348207]
17. Thompson WD, Walter SD. A reappraisal of the kappa coefficient. J Clin Epidemiolo. 1988; 41:949–58.
18. Stiell IG, McKnight RD, Greenberg GH, Nair RC, McDowell I, Wallace GJ. Interobserver agreement in the examination of acute ankle injury patients. Am J Emerg Med. 1992; 10:14–7. [PubMed: 1736906]
19. Stiell IG, Wells GA. Methodologic standards for the development of clinical decision rules in emergency medicine. Ann Emerg Med. 1999; 33:437–47. [PubMed: 10092723]
20. Kline JA, Mitchell AM, Kabrhel C, Richman PB, Courtney DM. Clinical criteria to prevent unnecessary diagnostic testing in emergency department patients with suspected pulmonary embolism. J Thromb Haemost. 2004; 2:1247–55. [PubMed: 15304025]
21. Quinn JV, Stiell IG, McDermott DA, Sellers KL, Kohn MA, Wells GA. Derivation of the San Francisco Syncope Rule to predict patients with short-term serious outcomes. Ann Emerg Med. 2004; 43:224–32. [PubMed: 14747812]
22. Gorelick MH, Atabaki SM, Hoyle J, et al. Interobserver agreement in assessment of clinical variables in children with blunt head trauma. Acad Emerg Med. 2008; 15:812–8. [PubMed: 19244631]
23. Colivicchi F, Ammirati F, Melina D, et al. Development and prospective validation of a risk stratification system for patients with syncope in the emergency department: the OESIL risk score. Euro Heart J. 2003; 24:811–9.
24. Costantino G, Perego F, Dipaola F, et al. Short- and long-term prognosis of syncope, risk factors, and role of hospital admission: results from the STePS (Short-Term Prognosis of Syncope) study. J Am Coll Cardiol. 2008; 51:276–83. [PubMed: 18206736]
25. Georgeson S, Linzer M, Griffith JL, Weld L, Selker HP. Acute cardiac ischemia in patients with syncope: importance of the initial electrocardiogram. J Gen Intern Med. 1992; 7:379–86. [PubMed: 1506942]
26. Grossman SA, Fischer C, Lipsitz LA, et al. Predicting adverse outcomes in syncope. J Emerg Med. 2007; 33:233–9. [PubMed: 17976548]
27. Martin TP, Hanusa BH, Kapoor WN. Risk stratification of patients with syncope. Ann Emerg Med. 1997; 29:459–66. [PubMed: 9095005]
28. Reed MJ, Newby DE, Coull AJ, Jacques KG, Prescott RJ, Gray AJ. The Risk stratification Of Syncope in the Emergency department (ROSE) pilot study: a comparison of existing syncope guidelines. Emerg Med J. 2007; 24:270–5. [PubMed: 17384381]
29. Sarasin FP, Hanusa BH, Perneger T, Louis-Simonet M, Rajeswaran A, Kapoor WN. A risk score to predict arrhythmias in patients with unexplained syncope. Acad Emerg Med. 2003; 10:1312–7. [PubMed: 14644781]
30. Brignole M, Alboni P, Benditt DG, et al. Guidelines on management (diagnosis and treatment) of syncope-update 2004. Executive Summary. Euro Heart J. 2004; 25:2054–72.
31. Huff JS, Decker WW, Quinn JV, et al. Clinical policy: critical issues in the evaluation and management of adult patients presenting to the emergency department with syncope. Ann Emerg Med. 2007; 49:431–44. [PubMed: 17371707]
32. Thiruganasambandamoorthy V, Ramaekers R, Rahman MO, et al. Prognostic value of cardiac biomarkers in the risk stratification of syncope: a systematic review. Intern Emerg Med. 2015; 10:1003–14. [PubMed: 26498335]

33. Reed MJ, Newby DE, Coull AJ, Prescott RJ, Jacques KG, Gray AJ. The ROSE (risk stratification of syncope in the emergency department) study. J Am Coll Cardiol. 2010; 55:713–21. [PubMed: 20170806]

34. Costantino G, Casazza G, Reed M, et al. Syncope risk stratification tools vs clinical judgment: an individual patient data meta-analysis. Am J Med. 2014; 127:1126 e1113–25.

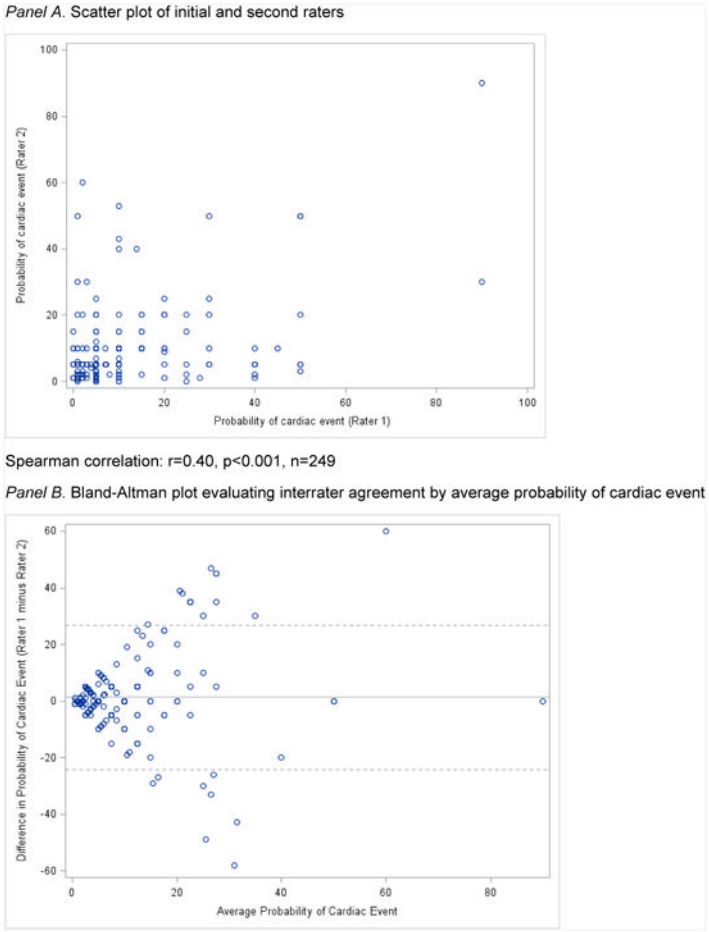

Figure 1. Interrater agreement of probability of 30-day serious cardiac event

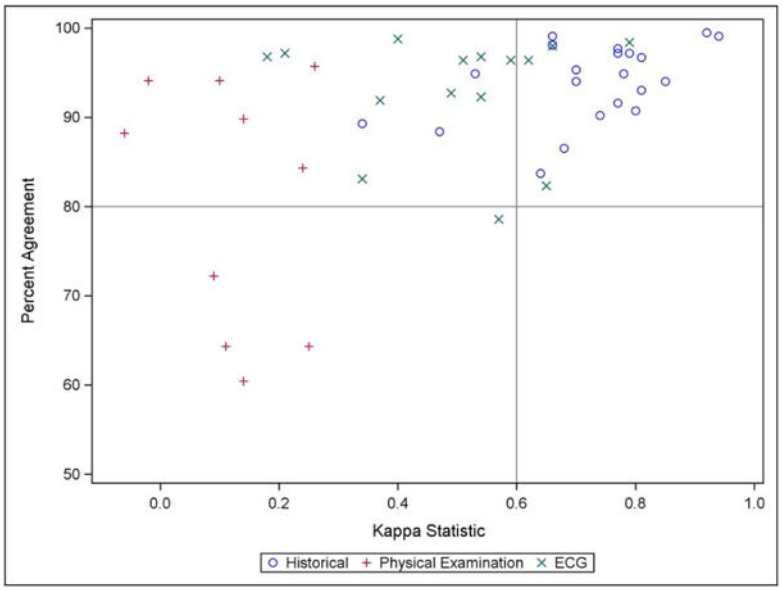

Figure 2. Percent agreement and kappa statistic of clinical findings

Table 1

Interrater agreement for historical findings, n=215 paired evaluations [a]

Finding [b]	n (%) per initial provider	% agreement	% positive agreement	% negative agreement	PABAK [c]	Kappa (95% CI)
Past Medical History						
Baseline cognitive impairment or dementia	14 (6.5)	97.2	78.6	98.5	0.94	0.77 (0.59 to 0.95)
Past stroke or transient ischemic attack	29 (13.5)	94.9	80.7	97.1	0.90	0.78 (0.65 to 0.90)
Congenital heart disease	2 (0.9)	99.1	66.7	99.5	0.98	0.66 (0.22 to 1.00)
Congestive heart failure	27 (12.6)	94.0	73.5	96.6	0.88	0.70 (0.55 to 0.85)
Ejection fraction <40% [d]	6 (2.8)	99.5	92.3	99.8	0.99	0.92 (0.77 to 1.00)
Peripheral vascular disease	13 (6.0)	94.9	56.0	97.3	0.90	0.53 (0.29 to 0.78)
Implanted permanent pacemaker	18 (8.4)	99.1	94.4	99.5	0.98	0.94 (0.86 to 1.00)
Implanted defibrillator	13 (6.0)	97.7	78.3	98.8	0.95	0.77 (0.58 to 0.96)
Coronary artery disease [e]	56 (26.0)	94.0	88.9	95.8	0.88	0.85 (0.77 to 0.93)
Structural heart disease [f]	22 (10.2)	96.7	82.9	98.2	0.94	0.81 (0.68 to 0.95)
Arrhythmia [g]	50 (23.3)	91.6	82.0	94.5	0.83	0.77 (0.66 to 0.87)
Seizure disorder	4 (1.9)	98.1	66.7	99.0	0.96	0.66 (0.35 to 0.97)
Diabetes requiring medication	49 (22.8)	93.0	85.4	95.4	0.86	0.81 (0.72 to 0.90)
Hypertension requiring medication	142 (66.0)	83.7	87.5	76.8	0.67	0.64 (0.54 to 0.75)
Chronic renal insufficiency [h]	18 (8.4)	95.3	72.2	97.5	0.91	0.70 (0.52 to 0.87)
Cancer requiring current active treatment	17 (7.9)	97.2	80.0	98.5	0.94	0.79 (0.62 to 0.95)
Current Medications						
Diuretics	55 (25.6)	86.5	80.7	91.5	0.73	0.68 (0.58 to 0.78)
Beta-blockers	69 (32.1)	90.7	92.1	93.2	0.81	0.80 (0.72 to 0.88)
Alpha blockers	19 (8.8)	88.4	66.7	94.2	0.77	0.47 (0.30 to 0.63)
Other antiarrhythmic agents [i]	14 (6.5)	89.3	45.5	94.7	0.79	0.34 (0.15 to 0.53)
Calcium channel blockers	42 (19.5)	90.2	87.6	94.2	0.81	0.74 (0.64 to 0.84)

[a] The second evaluation was completed by a research assistant based on review of the patient's electronic chart. Patients in whom the index ED visits was the first visit to the study site, did not have a second evaluation of past medical history and current medications. Data abstracted by the site research assistants did not include any new findings identified after the index ED visit (i.e., only included data available to the treating provider during the ED visit).

[b] A history of premature death in siblings or parents was not included as there were no patients with this per initial provider

[c] Prevalence adjusted bias adjusted kappa

[d] Based on most recent testing within 1 year

[e] Includes angina, myocardial infarction, positive stress test, history of coronary artery bypass grafting, percutaneous transluminal coronary angioplasty

[f] Aortic stenosis, pulmonary hypertention, cardiomyopathy, valvular heart disease, valve disease, idiopathic hypertrophic subaortic stenosis

[g] Ventricular arrhythmia/sudden death, supraventricular tachycardia (paroxysmal atrial tachycardia, paroxysmal supraventricular tachycardia, atrial fibrillation, atrial flutter), sick sinus syndrome, Mobitz II or complete heart block, junctional rhythm

[h] Creatinine 1.5 mg/gL for at least 3 months

[i] Unknown agreement was 14.3%

Table 2

Interrater agreement for physical examination variables, n=255 paired evaluations

Finding	n (%) per initial provider	% agreement	% positive agreement	% negative agreement	% unknown agreement	PABAK [a]	Kappa (95% CI)
Carotid bruit	0 (0.0)	64.3	0.0	74.9	36.0	0.29	0.11 (-0.01 to 0.24)
Heart murmur	27 (10.6)	84.3	40.9	91.2	0.0	0.69	0.24 (0.08 to 0.40)
Double vision	1 (0.4)	89.8	66.7	94.6	13.8	0.80	0.14 (-0.05 to 0.33)
Speech disturbance	4 (1.6)	94.1	0.0	97.0	0.0	0.88	-0.02 (-0.03 to -0.01)
Change in facial sensation	1 (0.4)	88.2	0.0	93.8	0.0	0.77	-0.06 (-0.08 to -0.04)
Unilateral weakness	4 (1.6)	94.1	20.0	97.0	0.0	0.88	0.10 (-0.11 to 0.31)
Facial drooping	1 (0.4)	95.7	50.0	97.8	18.2	0.91	0.26 (-0.04 to 0.55)
Abnormal cerebellar testing	2 (0.8)	72.2	40.0	82.9	24.4	0.44	0.09 (-0.04 to 0.23)
Abnormal gait	3 (1.2)	60.4	16.7	71.9	38.5	0.21	0.14 (0.02 to 0.25)
Positive fecal occult blood	11 (4.3)	64.3	90.9	40.5	73.4	0.29	0.25 (0.13 to 0.38)

[a] Prevalence adjusted bias adjusted kappa

Table 3

Interrater agreement (initial vs. second rater) for ECG interpretation, n=248 paired evaluations[a]

Finding [b, c]	n (%) per initial provider	% agreement[c]	% positive agreement	% negative agreement	PABAK [d]	Kappa (95% CI)
Normal	118 (47.6)	78.6	77.6	79.5	0.57	0.57 (0.47 to 0.67)
Isolated, nonspecific ST/T abnormalities	29 (11.7)	83.1	43.2	90.0	0.66	0.34 (0.18 to 0.49)
Abnormal	114 (46)	82.3	81.8	82.7	0.65	0.65 (0.55 to 0.74)
Non-sinus rhythms	22 (8.9)	91.9	41.2	95.7	0.84	0.37 (0.16 to 0.59)
Multiple premature ventricular contractions	7 (2.8)	98.0	66.7	99.0	0.96	0.66 (0.37 to 0.94)
Sinus bradycardia	5 (2)	97.2	22.2	98.6	0.94	0.21 (-0.16 to 0.57)
Left ventricular hypertrophy	11 (4.4)	96.8	55.6	98.3	0.94	0.54 (0.26 to 0.82)
Left axis deviation	10 (4)	96.4	52.6	98.1	0.93	0.51 (0.23 to 0.79)
Right axis deviation	4 (1.6)	98.8	40.0	99.4	0.98	0.40 (-0.15 to 0.94)
First degree heart block	13 (5.2)	96.4	64.0	98.1	0.93	0.62 (0.39 to 0.85)
Complete left bundle branch block	12 (4.8)	98.4	80.0	99.2	0.97	0.79 (0.59 to 0.99)
Complete right bundle branch block	9 (3.6)	96.4	60.9	98.1	0.93	0.59 (0.35 to 0.83)
Prolonged QRS	5 (2)	96.8	20.0	98.4	0.94	0.18 (-0.15 to 0.52)
Prolonged QTc	18 (7.3)	92.7	52.6	96.1	0.86	0.49 (0.28 to 0.69)
Acute or chronic ischemic changes	19 (7.7)	92.3	57.8	95.8	0.85	0.54 (0.35 to 0.72)

[a] initial provider included 62 attending, 95 midlevel and 91 residents and second providers included 228 attending, 2 midlevel and 18 residents.

[b] ECG findings of right ventricular hypertrophy, short PR interval, Brugada pattern and delta wave were not included as they were not present in either assessment

[c] Definitions of specific ECG findings are described in the text

[d] Prevalence adjusted bias adjusted kappa

Acad Emerg Med. Author manuscript; available in PMC 2017 September 06.