Failure to Validate the San Francisco Syncope Rule in an Independent Emergency Department Population

Adrienne Birnbaum, MD, MS
David Esses, MD
Polly Bijur, PhD
Andrew Wollowitz, MD
E. John Gallagher, MD

From the Department of Emergency, Albert Einstein College of Medicine, Bronx, NY.

Study objective: We conduct a prospective independent validation of the San Francisco Syncope Rule to identify emergency department (ED) syncope patients with short-term serious outcomes.

Methods: This was a prospective observational cohort study of adult patients presenting to a university hospital ED with acute syncope or near syncope. Patients meeting inclusion criteria as defined in the San Francisco Syncope Rule derivation were evaluated for 5 previously derived predictor variables: abnormal ECG result, shortness of breath, hematocrit level less than 30%, triage systolic blood pressure less than 90 mm Hg, and history of congestive heart failure. Hospital admission occurred at the discretion of the emergency physician, independent of the decision rule. Follow-up occurred through contact with the inpatient attending physician for admitted patients and by telephone contact with patients not hospitalized or those hospitalized and discharged before day 7. Predetermined outcome measures as defined by the San Francisco Syncope Rule were death, myocardial infarction, arrhythmia, pulmonary embolism, stroke, subarachnoid hemorrhage, significant hemorrhage, or any condition causing or likely to cause a return ED visit and hospitalization for a related event.

Results: Complete predictor and follow-up data were available for 713 of 743 (96%) enrolled patients. Sixty-one of 713 (9%) patients met predetermined criteria for serious outcome. Sixteen of 61 (26%; 95% confidence interval [CI] 16% to 39%) patients with a serious outcome were not identified as high risk by the rule. Rule performance to predict serious outcomes was sensitivity 74% (95% CI 61% to 84%), specificity 57% (95% CI 53% to 61%); negative likelihood ratio 0.5 (95% CI 0.3 to 0.7) and positive likelihood ratio 1.7 (95% CI 1.4 to 2.0).

Conclusion: In this independent validation study, sensitivity and negative likelihood ratio of the San Francisco Syncope Rule were substantially lower than reported in the original studies and suggest that the rule has limited generalizability. [Ann Emerg Med. 2008;52:151-159.]

0196-0644/$-see front matter

doi:10.1016/j.annemergmed.2007.12.007

INTRODUCTION

Background and Importance

Syncope, defined as transient loss of consciousness, is estimated to account for 1% to 3% of emergency department (ED) visits.[1] A recent analysis of the ED portion of the National Hospital Ambulatory Medical Care Survey estimated that 6.7 million syncope-related US ED visits (0.77% of total ED visits) occurred from 1991 to 2000 and that syncope admissions accounted for 2% of total ED admissions to the hospital.[2] Total annual costs for syncope-related hospitalizations in the United States approximate $2.4 billion.[3]

Usually of benign origin, syncope is occasionally a harbinger of significant morbidity and mortality. Hospitalization of many ED patients with syncope is driven by concern about potentially life-threatening causes despite recognized "low yield" for the majority of admissions.[4,5] Recent efforts have focused on prospective identification of ED patients with syncope who are at risk for early serious outcome in an attempt to hospitalize those patients most likely to benefit from admission.

The San Francisco Syncope Rule is a clinical decision rule derived for identification of ED patients with syncope or near syncope who are at low risk for short-term (7-day) serious outcome.[6] In the derivation cohort, this rule was reported to

Editor's Capsule Summary

What is already known on this topic

In validation studies, developers of the San Francisco Syncope Rule reported sensitivity above 96%, whereas an independent group found it to be 89%.

What question this study addressed

Sensitivity of the San Francisco Syncope Rule was determined in 713 syncope patients from a single center with a more racially and ethnically diverse patient population than that of previous validation studies.

What this study adds to our knowledge

The San Francisco Syncope Rule identified only 45 of 61 (74%) patients with serious outcomes and 17 of 25 (68%) of those in whom the serious outcome was not already evident during the emergency department visit. Most of the missed serious outcomes were arrhythmias.

How this might change clinical practice

The San Francisco Syncope Rule may be insufficiently sensitive to safely augment a strategy of cautious clinical judgment with high admission rates.

have 96% (95% CI 92% to 100%) sensitivity and 62% (95% CI 58% to 66%) specificity for the detection of serious 7-day outcomes.[6]

Attempts to validate this prediction rule have yielded discordant results.[7,8] Validation in a cohort of patients in the same ED setting in which the rule was derived reported high sensitivity (98% [95% confidence interval (CI) 89% to 100%]) similar to that reported in the derivation set.[7] In contrast, more recent efforts in an independent validation cohort yielded lower sensitivity (89% [95% CI 81% to 97%]).[8] We sought to provide additional independent validation of the rule.

Goals of This Investigation

The objective of this study was to validate the San Francisco Syncope Rule in an independent patient population. To accomplish this, we attempted to prospectively identify ED syncope patients with serious short-term outcomes, using the same definitions of predictor variables and serious outcomes used in the original derivation study.

MATERIALS AND METHODS

Study Design

The study had a single-setting, prospective, observational cohort design.

Setting

The study was conducted in the ED of Montefiore Medical Center, an urban academic center with approximately 80,000 adult visits per year, during January 5, 2005, to December 27, 2006. Data collection was performed by trained research associates present in the ED 24 hours per day, 7 days per week.

Selection of Participants

Adult patients aged 21 years and older, presenting to the ED with acute syncope (defined as transient loss of consciousness) or near syncope (sensation of impending but not actual loss of consciousness) as a reason for the ED visit, were eligible for inclusion. Patients with altered mental status, alcohol or illicit drug-related loss of consciousness, definite seizure, and transient loss of consciousness caused by head trauma were excluded.

Research associates prospectively screened and enrolled patients according to a predetermined protocol and predetermined screening criteria. Research associates identified potentially eligible patients through surveillance of the ED tracking system and query of nurses and physicians caring for patients. Research associates monitored these sources for patients presenting to the ED with complaints of syncope, loss of consciousness, fall, collapse, seizure, lightheadedness, tachycardia, bradycardia, shortness of breath, or chest pain. Patients were enrolled if the emergency physician determined that the reason for the ED visit was syncope or near syncope and none of the exclusion criteria were present. Written informed consent was obtained from all subjects. The study was approved by the Committee on Clinical Investigation of the Albert Einstein College of Medicine and the institutional review board of Montefiore Medical Center.

Interventions

Attending emergency physicians were asked by the research associates to complete a structured data collection instrument at the ED visit that dichotomously recorded 4 of the 5 predictor variables that compose the San Francisco Syncope Rule: complaint of shortness of breath, hematocrit level less than 30%, triage systolic blood pressure less than 90 mm Hg, or history of congestive heart failure. Physicians were permitted to use information provided by the patient, as well as information obtained from the medical record. The fifth predictor, abnormal ECG result, defined as not sinus rhythm or new changes compared with previous ECG, was determined by subsequent review of all ECGs by 2 senior physicians, one board certified in emergency medicine and one board certified in emergency medicine and internal medicine, with access to the hospital ECG database and blinded to the presence or absence of outcome or other predictor variables. Interobserver disagreements were discussed and consensus was obtained. Age, sex, race, and ethnicity were self-reported by patients.

The decision to admit or discharge enrolled patients from the ED was determined solely by the emergency physician independent of the decision rule. Follow-up through day 7 after the index ED visit was performed by trained research associates through structured interview with the admitting inpatient attending physician. Follow-up of patients discharged from the ED or not hospitalized through day 7 was also performed by telephone contact with the patients by the research associates

Table 1. Serious outcome definitions.

Outcome Measure	Definition
Death	Confirmed with findings in the medical record
Myocardial infarction	Any increase of troponin level or ECG change with an accompanying diagnosis of myocardial infarction on the discharge diagnosis and confirmed by the cardiology service involved in the care
Arrhythmia	Arrhythmia captured on monitoring and thought to have had a temporal relationship to the syncopal or near-syncopal event
Pulmonary embolism	Determined by high-probability ventilation-perfusion scan, CT of the chest, or angiography; confirmation on discharge diagnosis; and patient received treatment for pulmonary embolism or confirmed on autopsy
Stroke	Determined by discharge diagnosis, chart review to determine whether symptoms were temporally related to the admission, and confirmation that the admitting attending physician believed that the findings were thought to have been related or to have been a cause of the syncopal event
Subarachnoid hemorrhage	Same as stroke
Significant hemorrhage	Any episode of syncope or near syncope associated with a source of bleeding that required transfusion
Any condition causing or likely to cause a return ED visit and hospitalization for a related event	Any patients discharged from the ED or hospital after a syncopal event and then readmitted for the same or similar symptoms related to the initial syncopal event
	Patients admitted who required an acute intervention during their stay that would have caused them to return if they were discharged

using a structured data collection instrument. Contact was attempted as close to 7 days post-ED visit as possible. Home address, home telephone number, other telephone number (eg, work or cell), and e-mail address, as well as an address and 2 telephone numbers for another contact person, were collected by the research associates for all patients. Contact was initially generally attempted in the late morning. When an attempt was unsuccessful, subsequent attempts were made at other times of the day.

Methods of Measurement

Occurrence of a serious outcome (death, myocardial infarction, arrhythmia, pulmonary embolism, stroke, subarachnoid hemorrhage, significant hemorrhage, and any condition causing or likely to cause a return ED visit and hospitalization for a related event) was determined with predefined explicit criteria, as defined by the original investigators who derived the San Francisco Syncope Rule (Table 1).[6]

Data Collection and Processing

Data were entered into a database by SPSS Data Entry 4.0 (SPSS, Inc., Chicago, IL). Stata version 8.2 (StataCorp, College Station, TX) was used for data analysis.

Outcome Measures

The primary outcome measure was serious outcome within 7 days of the index ED visit. Research associates ascertained the presence or absence of serious outcome by using explicitly defined criteria (Table 1). Outcomes were reviewed by 2 primary study investigators to verify accurate categorization of serious outcomes. Research associates and the study investigators were blinded to the presence or absence of predictor variables when making determination of serious outcomes. Interobserver disagreements were resolved by discussion and consensus.

Primary Data Analysis

Descriptive analysis of patient characteristics stratified by occurrence of a serious outcome is reported as means with standard deviations and proportions with exact binomial 95% CIs.

Sensitivity, specificity, negative and positive predictive values, and negative and positive likelihood ratios for performance of the rule to predict serious outcomes are reported with 95% CIs.

Missing data were handled in the following way: For primary analysis, patients missing 1 or more predictor variables were excluded from analysis unless another predictor variable was positive (because patients with any of the 5 predictor variables identified as positive would be classified as "at risk" regardless of whether other predictor variables were positive or negative). Patients missing 7-day outcome (either because they were discharged from the ED on the index visit and were subsequently unavailable for follow-up or because they were admitted to the hospital, discharged in fewer than 7 days, and unavailable for follow-up after hospital discharge) were excluded from analysis (Figure). Patients with an inhospital serious outcome were included in the analysis even if they did not have full 7-day follow-up because these patients would be classified as having had a serious outcome irrespective of presence or absence of 7-day follow-up.

Both the κ statistic and simple proportionate agreement were used to assess interrater reliability of ECG interpretation.

Sensitivity Analyses

A planned secondary analysis was performed to assess the potential impact of missing predictor and outcome data on rule sensitivity to predict serious outcomes. The most optimistic estimate of rule sensitivity was calculated by assuming all missing predictor data were positive. Similarly, to address the potential impact of missing outcome data on sensitivity, the following assumptions were made: patients with at least 1 high-risk predictor were assumed to have a positive (serious) outcome; patients with negative and nonmissing predictor data were assumed to have a negative (nonserious) outcome; patients with 1 or more missing predictor variables and no positive

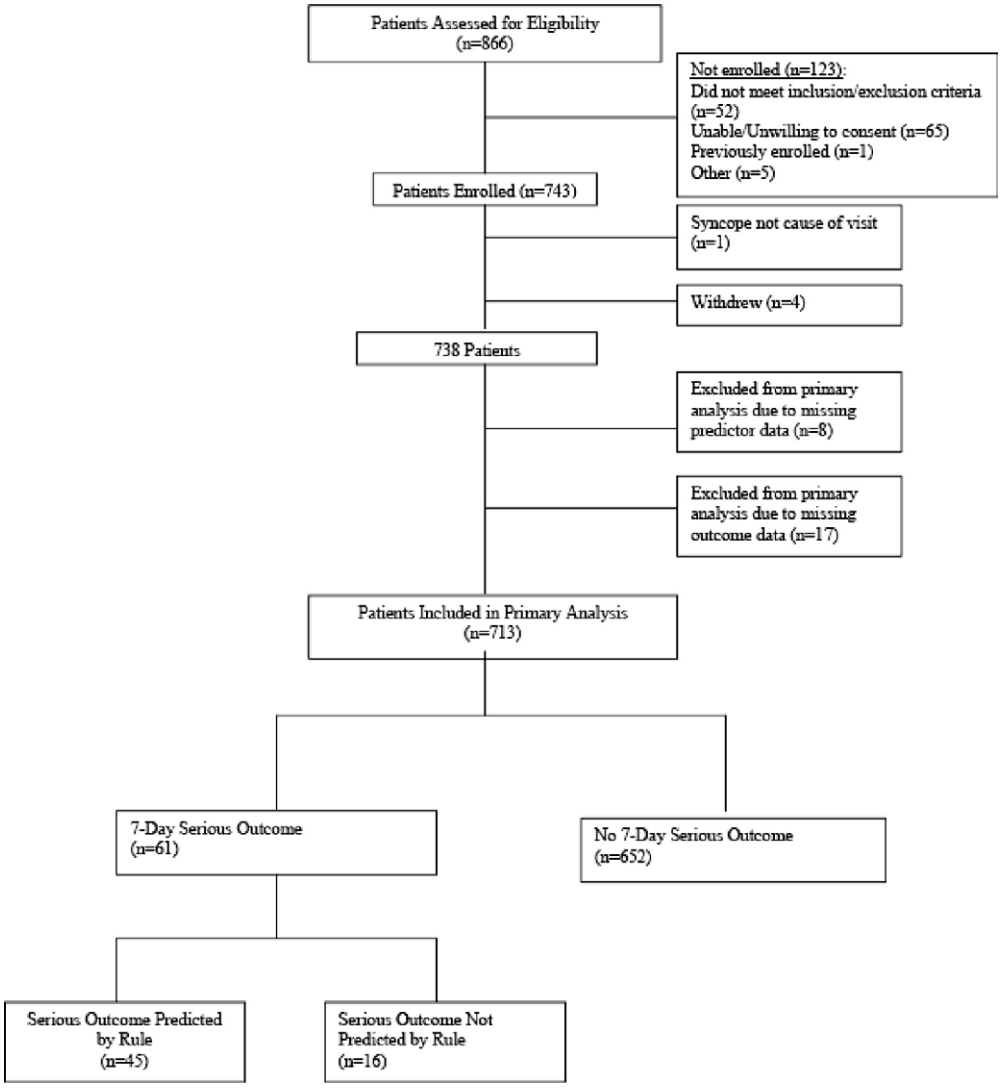

Figure. Patient flow diagram.

predictor were assumed to have a negative outcome. No patient was missing both predictor and outcome data.

A post hoc sensitivity analysis was performed to address rule performance for identification of serious outcomes that were not evident in the ED. This analysis was performed by ED medical record review of all identified serious outcomes performed by a senior board certified emergency physician to determine whether a predefined serious outcome was evident in the ED. Cases with serious outcomes that were evident in the ED were then excluded from analysis, and sensitivity of the rule to detect 7-day serious outcomes not evident in the ED was calculated.

The inclusion criterion for our study, acute syncope or near syncope as a reason for the ED visit, did not specifically require return to nonfocal neurologic status. Post hoc ED chart review of cases of neurologic serious outcome (subarachnoid hemorrhage and stroke) was performed to determine whether ED neurologic examinations and computed tomographic (CT) scan results were normal in these cases.

Reanalysis of rule sensitivity for detection of patients with serious outcomes was performed with ECG readings performed by emergency attending physicians at point of care to assess whether rule performance was likely to have been affected by the method of ECG interpretation.

A sample size of 730 patients was selected a priori according to the following assumptions and desired precision of the estimate: rule sensitivity of 95% with a lower limit of 95% CI extending no lower than 90% and a 7-day serious outcome rate of 10%. These assumptions were based on data reported by the original San Francisco Syncope Rule authors.[6]

RESULTS

Characteristics of Study Subjects

Characteristics of study subjects, stratified by outcome, are reported in Table 2. Information on flow of all participants in the trial from eligibility assessment to outcome ascertainment is presented in the Figure. Of 866 patients assessed for eligibility,

Table 2. Characteristics of study subjects.

Characteristic	All Patients (n=738)	No Serious Event Identified (n=677), 92%	7-Day Serious Event Identified (n=61), 8%	Difference Between Groups*
Age, y, mean (SD), range	61 (19) (21 to 101)	60 (19) (21 to 100)	66 (16) (22 to 101)	−6 (−11 to −1) 13 (7 to 19)
Age distribution, %				
21-40	17 (126)	18 (123)	5 (3)	2 (10 to 14)
41-60	30 (221)	30 (204)	28 (17)	−15 (−28 to −2)
61-80	37 (274)	36 (243)	51 (31)	−1 (−10 to 9)
81-101	16 (117)	16 (107)	16 (10)	
Sex, % (No.)				
Male	38 (282)	36 (246)	59 (36)	−23 (−36% to −10%)
Race/ethnicity, % (No.)				
Hispanic	39 (287)	39 (268)	31 (19)	8 (4 to 20)
Black	38 (279)	38 (256)	38 (23)	0 (−13 to 13)
White	17 (127)	17 (113)	23 (14)	−6 (−17 to 5)
Other	6 (45)	6 (40)	8 (5)	−2 (−9 to 5)
Disposition, % (No.)				
Admitted	83 (611)	81 (551)	98 (60)	−17 (−22 to −12)
Discharged	14 (107)	16 (107)	2 (1)	16 (13 to 19)
AMA†	3 (19)	3 (19)		3 (2 to 4)
Died in ED	(1)			−2 (−1 to 2)
SFSR predictors, % (No.)				
Abnormal ECG†	31 (225)	29 (194)	51 (31)	−22 (−35 to −9)
Shortness of breath	12 (86)	11 (75)	18 (11)	−7 (−17 to 3)
Hematocrit, %				
<30‡	5 (38)	4 (27)	18 (11)	−14 (−24 to −4)
SBP <90 mm Hg§	2 (15)	1 (10)	8 (5)	−7 (−14 to 0)
History of CHF	8 (61)	8 (53)	13 (8)	−5 (−14 to 4)
Any SFSR predictor	44 (328)	42 (283)	74 (45)	−32 (−44 to −21)

AMA, Against medical advice; *SFSR,* San Francisco Syncope Rule; *SBP,* systolic blood pressure; *CHF,* congestive heart failure.
*Difference between groups (no serious event identified)−(serious event identified).
†ECG not performed in 6 patients.
‡Hematocrit not performed in 8 patients.
§Triage blood pressure not performed in 1 patient.

743 patients were enrolled. One patient, a 33-year-old woman with vaginal bleeding who experienced syncope while in the ED, was excluded from analysis because syncope or near syncope was not the primary reason for the ED visit. Four additional patients withdrew from the study and were excluded, leaving 738 patients for analysis. Seven-day follow-up was available for 718 of 738 (97%) patients. The average number of contact attempts per patient was 3. Four patients who were discharged from the ED at the index visit were unavailable for follow-up despite multiple contact attempts. Median time to successful patient contact was 10 days (interquartile range 8 to 20 days) from the ED visit. Sixteen patients who were initially admitted to the hospital and discharged in fewer than 7 days were also unreachable for 7-day follow-up. Of these, 3 experienced a serious outcome while hospitalized.

Complete predictor variable data were available for 726 of 738 (98%) of patients. Exceptions were 6 patients for whom ECG was not performed, 8 patients for whom hematocrit level was not measured, and 1 patient for whom triage systolic blood pressure measurement was not documented. Of these, 3 were missing both ECG and hematocrit data. Four patients with missing predictor variable data had 1 or more other positive predictors. When missing data were handled as described in the "Primary Data Analysis" section above, 713 patients were available for analysis.

Main Results

Predefined serious outcomes were identified in 61 of 713 (9%; 95% CI 7% to 11%) patients. Rule performance characteristics are presented in Table 3. Sixteen of 61 (26%; 95% CI 16% to 39%) patients who experienced 1 or more serious outcomes were not identified by the prediction rule. Serious outcomes not identified by the rule consisted of 1 death, 8 arrhythmias, 3 strokes, and 1 subarachnoid hemorrhage, all thought to be causally related to the syncope. One patient required a blood transfusion for acute bleeding, and 2 patients returned to the ED within 7 days and were admitted for related medical problems. Arrhythmias in patients not identified by the prediction rule and determined to be related to the episode of syncope were sinus pause requiring pacemaker placement, 2 cases of Mobitz II second-degree atrioventricular block requiring pacemakers, junctional bradycardia with pulse rate of

Table 3. San Francisco Syncope Rule performance primary analysis.

	Serious Outcome		
Rule	**Yes, Number of Patients**	**No, Number of Patients**	**Total**
Rule positive	45	278	323
Rule negative	16	374	390
Total	61	652	713

Sensitivity=45/61=74% (95% CI 61% to 84%). Specificity=374/652=57% (95% CI 53% to 61%). Negative predictive value=374/390=96% (95% CI 93% to 98%). Positive predictive value=45/323=14% (95% CI 10% to 18%). Negative likelihood ratio=0.5 (95% CI 0.3 to 0.7). Positive likelihood ratio=1.7 (95% CI 1.4 to 2.0).

Table 4. San Francisco Syncope Rule Performance using assumptions for missing predictor and outcome data to maximize rule sensitivity.

	Serious Outcome		
Rule	**Yes, Number of Patients**	**No, Number of Patients**	**Total**
Rule positive	48	288	336
Rule negative	16	386	402
Total	64	674	738

Sensitivity=48/64=75% (95% CI 63% to 85%). Specificity=386/674=57% (95% CI 53% to 61%). Negative predictive value=386/402=96% (95% CI 94% to 98%). Positive predictive value=48/336=14% (95% CI 11 to 18%). Negative likelihood ratio=0.4 (95% CI 0.3 to 0.6). Positive likelihood ratio=1.8 (95% CI 1.5 to 2.0).

30 beats/min treated with a pacemaker, 2 cases of bradycardia requiring medication adjustment; 1 case of nonsustained ventricular tachycardia, and 1 case of alternating periods of junctional arrhythmia and slow atrial fibrillation with pauses treated with a pacemaker. The one death not predicted by the rule resulted from ventricular fibrillation/cardiac arrest, which occurred in a pharmacy soon after hospital discharge.

Physician judgment without application of the prediction rule resulted in admission of all 61 (100%; 95% CI 94% to 100%) patients with serious outcomes. The rule classified 45% of total patients as high risk, potentially decreasing overall admissions by 41% (from 86% to 45%). However, the rule failed to classify 16 of 61 (26%; 95% CI 16% to 39%) patients who experienced serious outcome as high risk.

The results of reanalysis of data performed by replacing missing data, as previously described, in such a way as to obtain the most optimistic estimate of rule sensitivity possible are reported in Table 4. There was no clinically significant difference between the estimate of rule sensitivity performed by excluding patients with missing data and that obtained by replacing missing data in such a way as to maximize sensitivity.

Percentage agreement for ECG reading by 2 independent raters was 82% (95% CI 79% to 85%). Interobserver agreement, as measured by kappa, was 0.53 (95% CI 0.46 to 0.60).

Post hoc sensitivity analysis performed to assess rule performance for identification of 7-day serious outcomes not identified in the ED resulted in identification of 25 cases of serious outcome that were not evident in the ED. Of these, 8 (32%; 95% CI 15% to 54%) were not identified by the rule, yielding a rule sensitivity of 68%; 95% CI 46% to 84%). Table 5 presents the number and type of serious outcome for all patients, those patients with serious outcomes not predicted by the rule, and those patients with serious outcomes not predicted by the rule and not identified in the ED.

Post hoc ED chart review of the 4 cases of neurologic serious outcomes demonstrated that ED neurologic examination results were normal or at baseline in all 4 cases, as documented by ED staff and neurology or neurosurgery consultants. However, ED CT scan results were positive in the case of subarachnoid hemorrhage and had new or possibly new findings consistent with infarct in 2 of the 3 cases of stroke. Exclusion of the 3 cases with abnormal CT scan results from analysis would have resulted in a rule sensitivity for detection of serious outcomes of 78% (95% CI 65% to 87%).

ECG interpretations performed by emergency physicians at the ED visit were available for 630 of 713 (88%) patients included in the primary analysis. Reanalysis of rule sensitivity for detection of patients with serious outcomes using ECG readings performed by emergency physicians at point of care resulted in rule sensitivity of 74% (95% CI 61% to 84%).

LIMITATIONS

Our study was designed to validate the findings of rule performance when applied as reported in the derivation study conducted by the initial investigators.[6] Our design did not therefore include a planned analysis to differentiate between rule performance to identify serious outcomes that were obviously identifiable in the ED and "occult" cases that were identified only after hospital admission, as has subsequently been reported in the internally and externally conducted validation studies.[7,8] Overall clinical utility of the rule to improve clinical decisionmaking may be better represented by ability of the rule to identify patients with serious outcomes that are not clinically apparent in the ED. To address this issue, we conducted a post hoc analysis of rule sensitivity for identification of the subset of patients with a 7-day serious outcome not identified in the ED. This analysis demonstrated a rule sensitivity of 68% (95% CI 46% to 84%). These results are similar to those reported in another recent attempt at external validation of the San Francisco Syncope Rule conducted by Sun et al.[8] These authors reported rule sensitivity for the subset of patients with a 7-day serious outcome diagnosed only after the index ED visit to be lower (69%; 95% CI 46% to 92%) than that for all 7-day serious outcomes (89%; 95% CI 81% to 97%).

We did not make an attempt to assess the ability of the rule to predict serious outcomes beyond 7 days. We chose 7-day outcome to match the methodology of the initial derivation study as closely as possible.[6] Despite use of 7-day outcome as the primary endpoint in the San Francisco Syncope Rule

Table 5. Serious outcomes.

Outcome	All Patients, n=61*	Not Predicted by Rule, n=16	Not Predicted by Rule and Serious Outcome Not Diagnosed in ED, n=8
Death	4	1	1†
Myocardial infarction	3		
Arrhythmia	33	8	5‡
Pulmonary embolism	3		
Stroke	7	3	
Subarachnoid hemorrhage	1	1	
Significant hemorrhage	9	1	
Return/hospitalization	5	2	2§

*Sixty-five events occurred in 61 patients.
†Patient experienced witnessed cardiac arrest in pharmacy. Initial cardiac rhythm was ventricular fibrillation.
‡Arrhythmias were sinus pause treated with pacemaker, Mobitz II second-degree atrioventricular block treated with pacemaker, junctional bradycardia with pulse rate of 30 beats/min treated with pacemaker, nonsustained ventricular tachycardia, and periods of junctional rhythm and slow atrial fibrillation with pauses treated with pacemaker.
§One patient returned to the ED for recurrence of near syncope. The other returned to another ED and was admitted; further details were unavailable.

derivation study, Quinn et al[7] chose to use 30-day serious outcome not determined during the ED evaluation in their validation study, which limits direct comparison of data from our validation study, as well as that of Sun et al,[8] to the validation performed by Quinn et al.[7] However, Quinn[9] has subsequently observed that reporting of 30-day outcomes instead of 7-day outcomes did not significantly change the results or conclusions.

The frequency of syncope cases as a percentage of total ED visits during the study period (0.5%) was substantially lower, and the hospital admission rate (86%) higher, than reported in previous validation studies.[7,8] Demographic characteristics of our patient population may partially account for these findings. Nonwhite and Hispanic groups have been reported to have lower incidences of ED presentation for syncope than reference groups,[2] and admission decisions may be influenced by the presence of a relatively indigent patient population for whom access to primary care, outpatient testing, and medical specialty follow-up is often unavailable in a timely fashion. However, it is also possible that despite methodology that attempted to maximize enrollment by using data collectors that were present in the ED 24 hours per day, 7 days per week, not all patients presenting to the ED with syncope were screened for enrollment, introducing the potential for selection of a sample with different characteristics than those of previous validation studies.

Complete predictor data and complete follow-up data were not available for 2% and 3% of patients, respectively. However, the magnitude of the proportion of missing data was small and therefore unlikely to have a meaningful effect on our findings. Nonetheless, we took a conservative approach to missing data by removing all patients with incomplete data from the primary analysis. A secondary analysis of rule sensitivity was then conducted to assess the impact of missing data. This analysis was performed by replacing missing predictor and outcome data with values that resulted in the optimal estimate of rule sensitivity. As shown in Tables 3 and 4, estimates of rule sensitivity were minimally affected by missing data.

ECG interpretation, performed in our study by 2 senior physicians, one board certified in emergency medicine and the other board certified in emergency medicine and internal medicine, with access to the hospital ECG database containing old ECG results on all patients previously known to the health care system, varied somewhat from the methodology described in the derivation study, which used physician reading of ECGs at point of care. We chose the former method in an attempt to optimize performance of this predictor variable by providing standardized interpretation with unlimited access and time for comparison to previous ECGs. Reanalysis of rule sensitivity for detection of patients with serious outcomes using ECG readings performed by emergency physicians at point of care did not affect rule sensitivity and did not support the likelihood that this aspect of the methodology had a substantial effect on determination of rule performance.

A limitation of ECG interpretation in our study was the finding of only moderate interobserver agreement when measured using the κ statistic. However, simple interobserver agreement was 82%, and κ has been shown to substantially underestimate true concordance when unbalanced marginal totals in the standard 4-fold κ table are present, as was the case in our data set.[10] Quinn et al[7] reported κ values of 0.55 for abnormal rhythm (not sinus) and 0.68 for abnormal ECG (new changes). Sun et al[8] reported a κ of 0.5 for abnormal ECG results.

Several methodologic aspects of our study may limit the generalizability of results. Study inclusion was limited to adult patients 21 years of age and older. The results cannot therefore be extrapolated to the pediatric age group. This was a single-site study conducted at an urban ED with a large, indigent minority population. Although the possibility exists that the rule might perform differently in a different patient population, there is no apparent reason to suspect a significant directional effect.

DISCUSSION

The San Francisco Syncope Rule, a clinical decision rule derived to predict ED syncope patients at risk for short-term

serious outcomes,[6] meets several methodologic standards that have been set for the development of such rules.[11] First, the problem is potentially well suited to development of a clinical decision rule because prevalence of the clinical problem is relatively high and current admission practice is recognized as being costly and inefficient.[2-5] Second, the San Francisco Syncope Rule was derived according to rigorous methodologic standards.[12] Finally, sensitivity of the rule to predict serious outcome has been demonstrated to be high in the original derivation study and validation study performed at the same site (96%, 95% CI 92% to 100%; and 98%, 95% CI 89% to 100%, respectively).[6,7]

It is essential that clinical decision rules be prospectively validated in new patient populations before application to clinical practice because many statistically derived rules fail to perform well when tested in new populations.[11] An attempt at independent validation of the San Francisco Syncope Rule at a different clinical site[8] resulted in report of a somewhat lower point estimate of sensitivity (89%; 95% CI 81% to 97%). We report the results of a prospective independent validation study, which demonstrate a rule sensitivity of 74% (95% CI 61% to 84%) to predict serious 7-day outcome, which is substantially lower than that of either previously published validation study. In our cohort, the San Francisco Syncope Rule would have classified 45% of the patients as high risk, potentially decreasing overall admissions by 41% (from 86% to 45%). However, the rule failed to classify 16 of 61 (26%; 95% CI 16% to 39%) patients who experienced serious outcome as high risk.

Potential reasons for differences in observed performance of clinical decision rules include factors such as instability in the model, differences in prevalence of disease, and differences in application of the rule.[13-16]

Differences in patient population and differences in application of the rule may result in inconsistent performance. Although mean age was nearly identical in the studies reported by Quinn et al,[7] Sun et al,[8] and us, the initial derivation and internal validation studies included children, whereas the 2 external validation studies enrolled only adults. The possibility exists that younger age may be associated with superior rule performance. We chose to limit our sample to adult patients according to the assumption that a clinical decision rule would have the greatest potential to decrease hospitalization in the nonpediatric age group. Our patient population is skewed strongly toward black and Hispanic patients. Race and ethnicity of patients was not specifically reported by Quinn et al.[7] The cohort described by Sun et al[8] was more than three-quarters white. The prevalence of serious outcome in our study (9%; 95% CI 7% to 11%) was similar to that reported in the derivation set (12%; 95% CI 9% to 14%), making this another unlikely explanation for differences in performance.

With regard to rule application, we attempted to adhere as closely as possible to that described in the derivation study. Physicians providing predictor information were given verbatim definitions by research associates of predictors as reported in the derivation study to use in completing the data collection instrument. Outcome measures were similarly adapted from the original work.

The lower estimate of sensitivity reported in our study compared to that reported by Sun et al[8] may result in part from differences in methodology used for ECG classification. Because limited access to previous ECGs, Sun et al[8] relied exclusively on predefined criteria of ECG abnormality, without reference to whether abnormalities were new or old. In contrast, our methodology used the definition supplied by the original authors that considered ECG results to be abnormal only if they were non-sinus rhythm or if abnormalities were not known to be old and provided the physician interpreters with full access to the hospital-wide ECG database. It is possible that this aspect of the methodology resulted in a lower rate of abnormal ECG classification.

The principal goal of a clinical decision rule is revision of disease probability.[17] The negative likelihood ratio, defined as the likelihood that a negative test result would be found in a patient experiencing a serious outcome, compared with the likelihood of a negative test result occurring in a patient without a serious outcome, gives an estimate of the ability of the rule to revise pretest disease probability to effectively rule out the target disorder of interest. The negative likelihood ratio of 0.5 found in this study, which is consistent with our reported sensitivity of 74%, suggests that application of the San Francisco Syncope Rule to ED syncope patients exerts only a weak influence on the reduction of pretest probability of serious outcome derived from undifferentiated clinical judgment.

It is possible that syncope represents a symptom for which the high degree of complexity of plausible causal associations linking risk factors and outcomes may limit the clinical usefulness of a lock-step algorithmic solution derived from traditional statistical techniques such as recursive partitioning. Other techniques such as neural networks have been advocated as decisionmaking aids for clinical problems with high levels of complexity. However, issues such as transportability, clinical interpretation of models, and clinical acceptability of neural networks have been cited as issues limiting widespread application to medical decisionmaking.[18]

Although conclusions drawn from post hoc review of small numbers of outcomes should be interpreted with caution, in our sample serious outcomes attributable to neurologic causes and hemorrhage were all identified in the ED, whereas the one death associated with ventricular fibrillation/cardiac arrest and 5 cases of arrhythmia were not. Future focus on identification or refinement of predictor variables with the ability to improve on prospective identification of patients at risk for arrhythmia-related events may have the potential to enhance risk-stratification tools.

In summary, the results of our effort to independently validate the ability of the San Francisco Syncope Rule to identify 7-day serious outcomes in ED patients with syncope suggest that this risk-stratification tool is less predictive than previously reported in both previous validation studies. Our

data do not support use of the San Francisco Syncope Rule to safely improve on clinical judgment in our population.

Supervising editor: Robert Silbergleit, MD

Author contributions: AB, DE, PB, and EJG conceived the study and designed the trial. DE and AW supervised the conduct of the trial and data collection. AB, DE, and AW managed the data, including quality control. PB provided statistical advice on study design and data analysis. AB analyzed the data. AB drafted the article, and all authors contributed substantially to its revision. AB takes responsibility for the paper as a whole.

Funding and support: By *Annals* policy, all authors are required to disclose any and all commercial, financial, and other relationships in any way related to the subject of this article, that might create any potential conflict of interest. The authors have stated that no such relationships exist. See the Manuscript Submission Agreement in this issue for examples of specific conflicts covered by this statement.

Publication dates: Received for publication June 20, 2007. Revisions received September 24, 2007, and October 30, 2007. Accepted for publication December 3, 2007. Available online February 20, 2008.

Presented as an abstract at the Society for Academic Emergency Medicine annual meeting, May 2007, Chicago, IL.

Reprints not available from the authors.

Address for correspondence: Adrienne Birnbaum, MD, MS, Emergency Medicine, 1B21 Building #6, Jacobi Medical Center, Bronx, NY 10461; 718-918-5815, fax 718-918-7459; E-mail adrienne.birnbaum@gmail.com.

REFERENCES

1. American College of Emergency Physicians. Clinical policy: critical issues in the evaluation and management of patients presenting with syncope. *Ann Emerg Med.* 2001;37:771-776.
2. Sun BC, Emond JA, Camargo CA. Characteristics and admission patterns of patients presenting with syncope to US emergency departments, 1992-2000. *Acad Emerg Med.* 2004;11:1029-1034.
3. Sun BC, Emond JA, Camargo CA. Direct medical costs of syncope-related hospitalizations in the United States. *Am J Cardiol.* 2005; 95:668-671.
4. Gallagher EJ. Hospitalization for fainting: high stakes, low yield. *Ann Emerg Med.* 1997;29:540-542.
5. Kapoor WN. Syncope. *N Engl J Med.* 2000;343:1856-1862.
6. Quinn JV, Stiell IG, McDermott DA, et al. Derivation of the San Francisco Syncope Rule to predict patients with short-term serious outcomes. *Ann Emerg Med.* 2004;43:224-232.
7. Quinn J, McDermott D, Stiell I, et al. Prospective validation of the San Francisco Syncope Rule to predict patients with serious outcomes. *Ann Emerg Med.* 2006;47:448-454.
8. Sun BC, Mangione CM, Merchant G, et al. External validation of the San Francisco Syncope Rule. *Ann Emerg Med.* 2007;49:420-427.
9. Quinn JV. In reply. Thirty-day versus 7-day outcomes in the San Francisco Syncope Rule. *Ann Emerg Med.* 2006;48:760.
10. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol.* 1990;43:543-549.
11. Stiell IG, Wells GA. Methodologic standards for the development of clinical decision rules in emergency medicine. *Ann Emerg Med.* 1999;33:437-447.
12. Gallagher EJ. Shooting an elephant. *Ann Emerg Med.* 2004;43: 233-237.
13. Wasson JH, Sox HC, Neff RK, et al. Clinical prediction rules: application and methodologic standards. *N Engl J Med.* 1985; 313:793-799.
14. Charlson ME, Ales KL, Simon R, et al. Why predictive indexes perform less well in validation studies. *Arch Intern Med.* 1987; 147:2155-2161.
15. Poses RM, Cebul RD, Collins M, et al. The importance of disease prevalence in transporting clinical prediction rules: the case of streptococcal pharyngitis. *Ann Intern Med.* 1986;105:586-591.
16. Wigton RS, Connor JL, Centor RM. Transportability of a decision rule for the diagnosis of streptococcal pharyngitis. *Arch Intern Med.* 1986;146:81-83.
17. Gallagher EJ. Clinical utility of likelihood ratios. *Ann Emerg Med.* 1998;31:391-397.
18. Griffith J. Artificial neural networks: are they ready for use as clinical decision aids? *Med Decis Making.* 2000;20:243.