Critical Review Form
Clinical Prediction or Decision Rule

Failure to Validate the San Francisco Syncope Rule in an Independent Emergency Department Population, *Annals of EM* 2008 (in press)

Objective: **"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". (p 2)**

Methods: **Trained research assistants staffed Montefiore Medical Center 24-hours per day, 7-days per week from Jan 2005 – Dec 2006 enrolling adults >21 years old with acute syncope (transient loss of consciousness) or near-syncope (sensation of impending, but not actually loss of consciousness). Exclusion criteria included altered mental status, alcohol or drug-related loss of consciousness (LOC), definite seizure or head-trauma related LOC.**

At the point-of-care EM physicians recorded 4/5 elements of the SFSR, while the fifth predictor (ECG changes) was interpreted at a later time by two senior EM physicians. Independent of the decision rule, treating clinicians made disposition decisions.

The primary outcome was 7-day serious outcome (death, MI, arrhythmia, PE, stroke, SAH, significant hemorrhage, ED recidivism for a related condition with admission or emergent intervention) as assessed by structured patient interview in-hospital or by phone.

To optimize the SFSR prognostic test characteristics, the investigators conducted a sensitivity analysis assuming missing data were positive and missing outcomes were assumed present if one variable was positive. Additionally, a post-hoc analysis of events not readily apparent at the index ED evaluation was used to evaluate the SFSR on this subset. Finally, since the inclusion criteria didn't exclude those without a return to baseline neurological status, the authors assessed neurologic serious outcomes separately.

The study was powered to sensitivity 95% with 5% CI and 10% serious outcome 7-day prevalence with 730 subjects.

Guide		Comments
I.	***Is this a newly derived instrument* (Level IV)?**	
A.	Was validation restricted to the retrospective use of statistical techniques on the original database? (If so, this is a Level IV rule & is not ready for clinical application).	No, validation was on new patient populations separate from the derivation trials so this is nota Level IV CDR.
II.	**Has the instrument been validated? (Level II or III). If so, consider the following:**	
1a	Were all important predictors included in the derivation process?	Yes, the SFSR was derived from 50 predictor variables assessed for accuracy and reliability using accepted CDR standards.
1b	Were all important predictors present in significant proportion of the study population?	Yes, as illustrated by Table 2 (p 5) predictors were present in 2% (systolic BP <90) to 31% (abnormal ECG) of the sample.
1c	Does the rule make clinical sense?	Yes, content validity based upon decades old literature and the CHESS mnemonic is easy to remember.
2	Did validation include prospective studies on several different populations from that used to derive it (II) or was it restricted to a single population (III)?	Single urban New York hospital. Broad demographics of this population, in comparison to San Francisco derivation and validation sites, provide enhanced external validity.
3	*How well did the validation study meet the following criteria?*	
3a	Did the patients represent a wide spectrum of severity of disease?	No, as reported in Table 5, the majority of serous outcomes were arrhythmia with no AAA rupture, spleen rupture, or ectopic rupture.
3b	Was there a blinded assessment of the gold standard?	Yes, “research associates and the study investigators were blinded to the presence or absence of predictor variables when making determination of serious outcomes”. (p 3)
3c	Was there an explicit and accurate interpretation of the predictor variables & the actual rule without knowledge of the outcome?	Probably, “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 SFSR". Investigators did not assess accuracy or reliability of this data acquisition.

3d	Did the results of the assessment of the variables or of the rule influence the decision to perform the gold standard?	No, "the decision to admit or discharge enrolled patients from the ED was determined solely by the emergency physician independent of the decision rule" (p 2)
4	How powerful is the rule (in terms of sensitivity & specificity; likelihood ratios; proportions with alternative outcomes; or relative risks or absolute outcome rates)?	• 86% (743/866) of eligible subjects were enrolled but 5 were dropped and 25 had incomplete data recorded. • 97% had 7-day follow-up and 98% had complete predictor variables leaving 713 for analysis. • 9% (61/713) had serious outcome of which 26% (16/61) were <u>not</u> identified by the SFSR including 1 death, 8 arrhythmias, 3 strokes and 1 SAH. • 83% of subjects were admitted and clinician judgment resulted in admission of all 61 serious outcomes by identifying 45% of the cohort as high-risk the rule would have decreased admission from 83% to 45% (p 6) • Expert read of ECG had κ=0.53 • Post-hoc analysis of cases not identified in the ED did not improve sensitivity (68%). Nor did sensitivity analysis with optimal data input or exclusion of abnormal CT/neuro cases or POC ECG interpretation (sensitivity 74%).

		Serious Outcome **Yes** **No** **SFSR+** **45** **278** **SFSR-** **16** **374** Sen 74% Spec 57% LR+ 1.1 (95% CI 1.4-2.0) LR- 0.5 (95% CI 0.3-0.7)
III.	**Has an impact analysis demonstrated change in clinical behavior or patient outcomes as a result of using the instrument? (Level I). If so, consider the following:**	

1	How well did the study guard against bias in terms of differences at the start (concealed randomization, adjustment in analysis) or as the study proceeded (blinding, co-intervention, loss to follow-up)?	No impact analysis was performed.
2	What was the impact on clinician behavior and patient-important outcomes?	Since the rule did not validate with impressive prognostic test characteristics, the authors did not contemplate cost-savings and other benefits.

Limitations

1. **Single center validation. SFSR may not work as well on the demographically diverse populations of the Bronx compared with the more homogenous San Francisco derivation/original validation populations. This seems unlikely as a rationale that the SFSR will validate elsewhere based on the disappointing LA results (see PGY III paper).**
2. **Pediatric patients excluded, so not applicable to them.**

Bottom Line

The SFSR did not validate on a demographically diverse patient population and should not be used outside of San Francisco. Potential reasons why the rule failed to validate in one-setting versus another include model instability, variable disease prevalence or differences in application of the rule.