Critical Review Form
Clinical Prediction or Decision Rule

External Validation of the San Francisco Syncope Rule
Annals of Emergency Medicine 2007; 49:420-427

Objective: **"To evaluate the accuracy of the San Francisco Syncope Rule to identify "low-risk" patients in an independent prospective validation sample." (p 421)**

Methods:

From April 2005 – April 2006 at West Los Angeles Veterans Affairs Medical Center adult patients with a complaint of syncope or near-syncope were recruited 7-days/week from 8A-10P. Research assistants queried the charge nurse and physician staff, as well as reviewing the ED log to identify eligible patients. Exclusion criteria included witnessed seizure, head trauma related LOC, ongoing confusion, intoxication, age <18 years, DNR/DNI orders or lack of follow-up information. All elements of the SFSR were collected prospectively, although none of the components were mandated.

Unlike the SFSR definition of abnormal ECG, investigators labeled an abnormal ECG as treating physician note of any non-sinus rhythm, bundle branch block, left axis deviation, ventricular hypertrophy, 2° or 3° AV block, Q/ST/T ischemic changes or non-specific ST/T abnormality. The clinical evaluation forms were completed by 2nd – 4th year residents and whenever possible independently by an attending physician. When an ECG was not ordered by the treating physician a study ECG was obtained and hidden from the treating physician.

Outcomes included death, MI, arrhythmias, PE, CVA/TIA, SAH, aortic dissection, newly diagnosed structural heart disease, or significant hemorrhage/anemia requiring transfusion as well as any syncope-related acute intervention (pacemaker, angioplasty, valve surgery, balloon pump, vasopressors, or surgery for AAA, ruptured spleen or ectopic pregnancy). Outcomes were assessed by 14-day medical record review or telephone follow-up reviewed independently by two EM physicians. Potential serious outcomes were adjudicated by a three EM physician panel. The study was powered for 98% sensitivity with 10% CI if 50 primary outcomes occurred.

The authors conducted a sensitivity analysis gauged to optimize SFSR performance by assuming missing variable data was positive (sensitivity maximized) or absent (specificity maximized). They did the same with missing outcome data.

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, this is prospective validation on a population distinct from that used to derive or validate the SFSR.
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, in the original derivation 50 candidate predictor variables from the syncope literature were evaluated.
1b	Were all important predictors present in significant proportion of the study population?	Yes, as demonstrated in Table 1 (p 423) SFSR variables were present in 5% (SBP <90) to 37% (abnormal ECG) of the study population.
1c	Does the rule make clinical sense?	Yes, the rule has content validity based upon decades of syncope research. Additionally, the CHESS mnemonic is easy to remember: "Don't faint while playing CHESS with your grandfather" helps clinicians remember the population most-often missed by the SFSR – the elderly.
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)?	This is a Level III CDR, validated on a second, geographically proximate population from the original SFSR derivation and validation.
3	*How well did the validation study meet the following criteria?*	
3a	Did the patients represent a wide spectrum of severity of disease?	Yes, according to Table 1, 58% were admitted or transferred which seems representative of other US syncope papers.
3b	Was there a blinded assessment of the gold standard?	A panel of three emergency physicians were "blinded to the structured data forms completed by treating physicians" as they determined occurrence and timing of serious clinical events.
3c	Was there an explicit and accurate interpretation of the predictor variables & the actual rule without knowledge of the outcome?	Table 3 displays the Kappa for SOB (0.5), abnormal ECG (0.5) and any high-risk predictor (0.6). The timing of data collection relative to outcomes was not explicitly stated but was presumably

		before knowledge of outcomes for the majority.
3d	Did the results of the assessment of the variables or of the rule influence the decision to perform the gold standard?	No, "physicians were instructed to treat and admit patients in their usual manner, without any specific study intervention or testing".
4	How powerful is the rule (in terms of sensitivity & specificity; likelihood ratios; proportions with alternative outcomes; or relative risks or absolute outcome rates)?	• 81% of 709 screened subjects were consented with 91% telephone and 6% medical record review follow-up (3% lost to follow-up). • 58% were admitted or transferred. • 32% did not obtain hematocrit and 7% lacked an ECG. • 11.7% had a 7-day serious event (56/477) and only 3.4% (16/477) were not identified in the ED. **ALL CASES with missing predictors "present"** 7-day serious outcome Yes No SFSR+ 50 295 SFSR- 6 126 56 421 477 Sen 89 Spec 30 LR+ 1.3 (1.1 – 1.4) LR- 0.36 (0.17 – 0.77) 7-day serious outcome Yes No Gestalt + 56 223 Gestalt - 0 198 56 421 Sen 100 Spec 47 LR+ 1.9 (1.7 – 2.0) LR - 0 (0 – 0.3) **Delayed 7-day serious outcome not obvious in ED (N=437) missing variables set to present** Yes No SFSR+ 11 295 SFSR- 5 126 16 421 Sen 69 Spec 30 LR+ 0.98 (0.7-1.4) LR- 1.0 (0.5-2.2)

		Delayed 7-day serious outcome not obvious in ED. Missing variables set to absent. Yes No SFSR+ 11 173 SFSR- 5 248 16 421 Sen 69 Spec 59 LR+ 1.7 (1.2-2.4) LR- 0.5 (0.3-1.1) • Estimates of sensitivity were minimally affected by missing predictor data. • Sensitivity and specificity were lower when attending physician data collection forms were used.
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 performed, but the rule does not appear to improve upon clinical gestalt with these physicians on this population.
2	What was the impact on clinician behavior and patient-important outcomes?	No impact analysis performed.

Limitations

1) **Potential *selection bias* with non-consecutive subject enrollment but no biologically plausible reason to suspect those presenting between 10P-8A would differ from those presenting at other times.**

2) **Hematocrit not obtained on 32% of subjects but sensitivity analysis does not suggest these results would alter the SFSR prognostic test characteristics.**

3) **Very small number of serious outcome events. One formula for power calculation for diagnostic tests is:**

$$TP + FN = Z^2 \times \frac{[SN\,(1 - SN)]}{W^2} \qquad \text{and} \qquad N\,(SN) = \frac{TP + FN}{P}$$

Where Z = Area under curve (2 SD= 1.96)
TP = True positives
FN = False negatives
SN = Sensitivity
W = Width of confidence internal
P = Prevalence of disorder

In the current study
Z = 1.96
SN = 0.98
W = 0.05
P = 0.034

TP + FN = 30 N (SN) = 30/0.34 = 886 subjects

Therefore, the current study may have been under-powered.

Bottom Line:

The SFSR does not validate well in an LA VA ED population. Sensitivity analysis for missing predictor variables and outcome results do not suggest that missing data explains these discrepant results. Before widespread use of the SFSR external validation still needs to occur. Alternatively larger trials focused on the ED population without clinician recognition of serious outcome during index visit may enhance the prognostic characteristics of the SRSR, particularly if adequately powered and focused on dysrhythmia recognition (the most frequently missed serious outcome).