Critical Review Form

Clinical Prediction or Decision Rule

Likelihood Ratios for a Sleep Apnea Clinical Prediction Rule
Am J Respir Crit Care Med 1994;150:1279-1285

Objective: **"To identify independent predictors of disease severity using several statistical models". (p. 1279)**

Methods: **Consecutive patients referred to the Alberta Lung Association Sleep Clinic between February and September 1990 was approached for consent. Exclusion criteria included failure to agree to initial clinic visit, prior evaluation or treatment for sleep apnea, referral for a non sleep apnea sleep disorder, current tranquilizer or anti-depressant use, severe lung disease (Pao_2 < 55mm Hg or FEV_1 <50% predicted), neuromuscular or cerebrovascular disease, or recent MI or CHF or life-threatening dysrhythmia.**

HTN was defined as systolic BP > 160 or diastolic BP > 95mm Hg. All patients under went overnight polysomnography including electroencephalography, electrooculography, and electromyography. Apnea was defined as cessation of airflow for > 10 seconds while hypopnea was defined as airflow reduction with ≥ 3% decrease in oxygen saturation. The number of apnea and hypopnea events per hour was recorded with an "apnea-hypopnea index" ≥ 10 defining sleep apnea.

Experienced sleep clinic physicians were asked to assess the probability of sleep apnea for each patient after their clinical evaluation, but before the criterion, standard polysomnography testing. Additionally, patient and family reported features like snoring and sleep-related choking were collected with Likert-scale questionnaires.

After transforming the apnea – hypopnea score to a logarithmic scale. The authors developed a variety of linear models to identify significant predictors of sleep apnea. They also developed a logistic regression model and compared the various models with an ROC curve.

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 attempted – simply a derivation trial.
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
1b	Were all important predictors present in significant proportion of the study population?	Unknown since the authors fail to report the prevalence of candidate variables in their derivation cohort.
1c	Does the rule make clinical sense?	Not really (see below). As with most linear regression models, the end result is a complicated algebraic equation not readily computed at the bedside. Table 4 (reproduced below) can help quickly calculate the "Sleep Apnea Clinical Score", but a simple logistic model point-scoring system would be preferable.
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)?	No validation, so this CDR remains a Level IV tool.
3	*How well did the validation study meet the following criteria?*	
3a	Did the patients represent a wide spectrum of severity of disease?	For EM, probably not. These represent a self-selected cohort referred to one Canadian Sleep Clinic.
3b	Was there a blinded assessment of the gold standard?	Yes. "Persons scoring the sleep studies were blinded to the results of the questionnaire and anthropomorphic measurements". (p. 1280).
3c	Was there an explicit and accurate interpretation of the predictor variables & the actual rule without knowledge of the outcome?	Yes. "The inter-rater reliability coefficients for height, weight, neck circumference, and waist circumference measurements were 0.99; and for chest, hip, and thigh circumference measurements were 0.97, 0.94, and 0.80, respectively". (p. 1281)

3d	Did the results of the assessment of the variables or of the rule influence the decision to perform the gold standard?	No. By design all subjects had polysomnography.
4	How powerful is the rule (in terms of sensitivity & specificity; likelihood ratios; proportions with alternative outcomes; or relative risks or absolute outcome rates)?	• Among 263 clinic referrals, 200 were eligible and 82 (45.5%) were diagnosed with sleep apnea. • Linear regression identified a 3-item rule by which to risk stratify patients: neck circumference (cm), hypertension, and presence of snoring or nocturnal choking reported by bed-partner. To calculate the "Sleep Apnea Clinical Score" (SACS) see below. • Although they display multiple ROC curves (Fig 1 – 4, p.1282) the authors fail to report any AUC. However, the 4-variable linear model appears similar to a more complicated 16-item model as well as the logistic model and prior models by Viner and Crocker. • Physician gestalt as inferior to any of these models. • The LR^- for the SACS $\leq 5 = 0.25$ (95% CI 0.15-0.42) and for SACS > 15 $LR^+ = 5.17$ (95% CI 2.54-10.51).
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)?	Selection bias is quite likely which may compromise the model stability if ever validation is attempted.
2	What was the impact on clinician behavior and patient-important outcomes?	No clinician or patient importance outcomes were assessed. If this CDR is validated, future trials will need to demonstrate that they reduce unnecessary diagnostic test and/or reduce OSA-related morbidity.

Limitations

1) **Limited demographic reporting on the patient population to enhance understanding of the results' external validity. In particular, the prevalence of candidate variables in the model should have been reported.**

2) **No reporting of AUC.**

3) **Limited details provided on the logistic regression model assumptions or derivation.**

4) **No *a priori* power calculation, although with 82 diagnoses the 4-item rule ought to be appropriately powered by the 10 outcomes/predictor variable rule-of-thumb.**

5) **No attempt to perform recursive partitioning analysis which is a superior statistical method to maximize sensitivity at the expense of specificity (or vice-versa).**

6) **The linear regression equation difficult to operationalize for daily bedside use.**

7) **Derivation with no attempt to validate during this trial or in the 14-years since. Therefore, this instrument remains a Level IV CDR not ready for prime-time use.**

8) **No patient-important outcomes measured.**

Bottom Line

This single-center derived, 3-item CDR is superior to the clinical gestalt of experienced sleep physicians. If validated, this tool could augment risk-stratification of potential sleep apnea patients and allow focused testing for expedited referral to the finite criterion standard (polysomnography) for definitive testing.

SLEEP APNEA CLINICAL SCORE FOR THE LINEAR MODEL OF NECK CIRCUMFERENCE, HYPERTENSION, HABITUAL SNORING, AND PARTNER REPORTS OF NOCTURNAL CHOKING/GASPING

	Not Hypertensive Historical Features*			Hypertensive Historical Features*		
	None	One	Both	None	One	Both
Neck circumference (cm)						
28	0	0	1	0	1	2
30	0	0	1	1	2	4
32	0	1	2	1	3	5
34	1	2	3	2	4	8
36	1	3	5	4	6	11
38	2	4	7	5	9	16
40	3	6	10	8	13	22
42	5	8	14	11	18	30
44	7	12	20	15	25	42
46	10	16	28	21	35	58
48	14	23	38	29	48	80
50	19	32	53	40	66	110

* Historical features are habitual snoring and partner reports of nocturnal choking or gasping.