Critical Review Form
Diagnostic Test

Comparison of the Unstructured Clinician Estimate of Pretest Probability for Pulmonary Embolism to the Canadian Score and the Charlotte Rule: A Prospective Observational Study, *Acad EM* 2005; 12: 587-593

Objectives: **"To perform a head-to-head comparison of the unstructured clinical estimate of pretest probability for PE with a structured decision rule." (p 587)**

Methods:

Prospective observational study at Carolina's Medical Center from October 2001 – July 2004 with 142 health care professionals (36 EM attendings, 99 residents, 7 PA's, and one research assistant) who used a standardized data collection instrument to assign unstructured pre-test probability of PE (<15%, 15-40%, >40%) after the history/physical exam. A previously validated stepwise protocol using a whole blood D-dimer and alveolar dead space measurement was used to evaluate all suspected PE patients with abnormal results further with VQ scan or CT angiogram. All subjects had 45-day follow-up (medical record review, postal mailings, telephone calls, Social Security Death Index inspection and the North Carolina Chief Medical Examiner's record review). Interobserver agreement between gestalt, Canadian Clinical Decision Rule (CDR) and the Charlotte rule were assessed by independent review of 154 patients by 62 different clinicians.

Guide		Comments
I.	**Are the results valid?**	
A.	**Did clinicians face diagnostic uncertainty?**	Yes, all assessments were performed and recorded *before* objective testing.
B.	**Was there a blind comparison with an independent gold standard applied similarly to the treatment group and to the control group? (Confirmation Bias)**	Non – low risk patients (abnormal D-dimer or alveolar dead space) had VQ or CT, but not all subjects had Gold standard testing performed. A confirmation bias may exist by these methods, so the authors attempted to capture all PE's not tested for by their detailed 45-day follow-up methods described above.

C.	**Did the results of the test being evaluated influence the decision to perform the gold standard? (Ascertainment Bias)**	Possibly since the Charlotte Rule was part of the pre-existing diagnostic algorithm.

II.	**What are the results?**	
A.	**What likelihood ratios were associated with the range of possible test results?**	• Clinicians were able to evaluate pre-test probability for 98.9% of subjects using Gestalt, 97.7% using Canadian and 95.2% using Charlotte. No difference in PE prevalence was noted between subjects with and without missing data. • The majority of subjects were low-risk (69% by gestalt, 73% by Canadian and 88% by Charlotte) with the corresponding PE prevalence 2.6%, 3.0%, and 4.2% (Fig. 1-2, p. 590) • <u>As training level increases the prevalence of PE increases.</u> • The Kappa (κ) for Canadian and Charlotte was 0.47 (0.33-0.61) and 0.85 (0.69-1.0) respectively. Gestalt had κ = 0.60 (0.46-0.74). Comparing Gestalt to Canadian κ= 0.31 and to Charlotte κ = 0.13. <u>Low-Risk Method</u> — <u>LR (-)</u> — <u>PE Prevalence</u> Gestalt — 0.45 (0.35-0.56) — 2.7% Canadian (score <2) — 0.48 (0.38-0.59) — 2.9% Charlotte (“safe”) — 0.72 (0.66-0.80) — 4.2%

III.	How can I apply the results to patient care?	
A.	**Will the reproducibility of the test result and its interpretation be satisfactory in my clinical setting?**	The Charlotte Rule demonstrated excellent inter-rater reproducibility, Gestalt moderate reproducibility and the Canadian rule less moderate reproducibility. Given the wide range of physicians tested in this study, one would expect similar results at any United States academic ED.
B.	**Are the results applicable to the patients in my practice?**	Yes, based upon Table 1 (p.588) age and symptom presentation appear similar to BJH PE pts. Note that tachypnea (33%), tachycardia (33%) and hypoxemia (10%) were noted in a minority of these patients. Radiology often contests PE in the differential diagnosis and these figures may be useful to quote in the future.
C.	**Will the results change my management strategy?**	Yes. I'll recognize Gestalt has an equal footing with CDR's for the reliable, reproducible exclusion of PE. None of these three strategies, however, has sufficiently low LR- to exclude PE to <1%, the accepted stop-testing threshold (*Annals EM* 2003; 42: 266-275).
D.	**Will patients be better off as a result of the test?**	Yes, if clinicians recognize more accurate pre-test probabilities and the limitations of Gestalt, Charlotte or Canadian CDR to produce posterior probability < 1%.

Limitations

1. **Single center study with ascertainment bias favoring the home grown Charlotte rule.**
2. **Alveolar dead space measurement not used outside of Charlotte.**
3. **Lower than previously reported prevalence suggests large group of no-risk subjects were tested.**

Bottom Line

None of these methods (gestalt, Charlotte, Well's) is sufficient to bring the post-test probability below 1% by themselves. Clinical gestalt ($LR^- = 0.45$) displays similar ability to identify low-risk PE patients as the Charlotte Rule ($LR^- = 0.72$), and the Canadian Well's PE CDR < 2 ($LR^- = 0.48$). Clinical gestalt had moderate reproducibility ($\kappa=0.60$) and compares well with Well's criteria ($\kappa=0.47$) although the Charlotte Rule's reproducibility is superior ($\kappa=0.85$). More experienced clinicians tend to have more low risk patients diagnosed with PE (a patient deemed low-risk by an experienced clinician is more likely to have a PE diagnosed than a patient labeled low-risk by a less experienced clinician).