Critical Review Form

Therapy

A Computerized Handheld Decision-Support System to Improve Pulmonary Embolism Diagnosis, *Ann Intern Med* 2009; 151:677-686

Objective: To test "the hypothesis that a hand held computer-based clinical decision-support system (CDSS) intended to guide diagnostic testing for pulmonary embolism could improve diagnostic decision making more than paper-based educational materials." (p. 677)

Methods: Cluster randomized trial of 20 French emergency departments that a) volunteered to participate and b) had access to D-dimer testing, spiral CT, and venous ultrasonography. Each center prospectively enrolled patients with clinically suspected PE presenting to the ED. Exclusion criteria included previously excluded or confirmed venous thromboembolism (VTE), referral from an inpatient setting, failure to obtain patient consent, absence of any diagnostic testing, death before testing was complete or emergency physician failure to document any final diagnostic impression.

During the pre-intervention period (June 2005-November 2005) all 20 ED's were provided handheld computers to accustom physicians to real-time data entry about patient characteristics, PE probability and diagnostic testing performed. Based upon the author's defined "appropriateness criteria" (appropriate = testing until probability <5% to exclude PE or >85% to diagnose PE) ED's were stratified into low and high proportions of "appropriate work-up" for cluster randomization into two groups of 10 via a random number table.

In the intervention period (December 2005-June 2006) the intervention group had a CDSS activated in their hand held device. The CDSS asked clinicians to sequentially a) enter the variables required to generate a PE-risk using the revised Geneva clinical decision rule; b) enter an estimate of PE probability which could be The Geneva CDR-derived estimate or the clinician's clinical gestalt; c) review all pertinent tests labeled as "appropriate" or "inappropriate" based upon their meta-analysis derived diagnostic test characteristics as level of invasiveness. The results of each test were then imputed and the CDSS informed the physician to consider PE excluded, diagnosed or to keep testing. On the contrary, the control group received posters and pocket cards that demonstrated the revised Geneva score, a diagnostic

algorithm based on clinical probability and D-dimer level + ultrasonography and spiral CT. Hand held devices remained in these control group ED's for real-time data collection, but the CDSS was not turned on.

The primary outcome was appropriateness of the diagnostic work up for PE. Secondary outcomes included adherence to recommendations, number of tests performed in patients with appropriate diagnostic work up, and clinical outcomes at 3 months assessed by 2 investigators blinded to group assignment.

Based upon a projected 15% absolute difference improvement between the two groups, a sample of 665 patients in each group would provide 90% power at 5% significance. Accounting for exclusions and lost to follow up investigators planned to recruit 1500 patients. In analyzing the data, investigators assumed worst case scenarios pretest probability if none was provided by the physician before testing. Specifically, if PE was ruled out then the pretest probability was assumed to have been 85%, and if PE was ruled in then the pretest probability was assumed to have been 5%. Subgroup analysis of patients with real-time data entry was planned *a priori*. Mixed logistic regression models were used to adjust effect size of the intervention strategy for study period, patient risk factors (age, known heart disease, chronic lung disease, current anticoagulation, and previous VTE).

Guide		Comments
I.	**Are the results valid?**	
A.	**Did experimental and control groups begin the study with a similar prognosis (answer the questions posed below)?**	
1.	Were patients randomized?	Yes. "We used a random-number table to assign emergency departments to use either computer-based or paper guidelines. We stratified randomization to include 5 centers each that had a low and high proportion of "appropriate work-up during the pre-intervention period in each group." (p. 678)
2.	Was randomization concealed (blinded)?	Yes. "Two researchers who were blinded to center identities allocated trial assignments." (p. 678)
3.	Were patients analyzed in the groups to which they were randomized?	There is no clear statement of intention-to-treat analysis, but that appears to be what the investigators report.

4.	Were patients in the treatment and control groups similar with respect to known prognostic factors?	No. Table 2 (p. 681) demonstrates that in the intervention phase the paper-guidelines group had more chronic respiratory disease (13.1% vs. 18.0%, p = 0.008) less chest pain (65.4% vs. 57.8%, p = 0.002), and more altered mental status (1.4% vs. 4.4%, p < 0.001).
B.	**Did experimental and control groups retain a similar prognosis after the study started (answer the questions posed below)?**	
1.	Were patients aware of group allocation?	Uncertain. No clear statement of patient blinding.
2.	Were clinicians aware of group allocation?	ED's were aware of whether the CDSS was turned on or not. It is unclear whether ED's or individual clinicians were aware of the study's objectives.
3.	Were outcome assessors aware of group allocation?	Uncertain. "Investigators reviewed patient charts and collected information on tests performed within 72 hours of ED evaluation", but whether or not they were blinded to patient's study site group allocation is not stated.
4.	Was follow-up complete?	No. As demonstrated in Figure 4 (p. 684) a total of 283/1645 (17.2%) declined or were lost to follow up including 27% of CDSS group with PE excluded (vs. 11% of paper guidelines group with PE excluded).
II.	**What are the results (answer the questions posed below)?**	

1.	How large was the treatment effect?	• Enrolling hospitals were similar in number of physicians (11.7 CDSS vs. 12/3 in paper guidelines) and volume (mean admissions 40500 in CDSS vs. 32500 in paper guidelines); 61 patients (median) per center were analyzed. • 992 analyzable patients were enrolled pre-intervention and 1645 analyzable patients post intervention. • Data was entered real-time pre-intervention in 41% of CDSS vs. 80% post-intervention; significantly more clinicians provided a pre-test estimate in the post intervention CDSS group than in the paper guidelines group: 83.3% vs. 37% $p < 0.001$. • CDSS group performed fewer mean tests per patient with appropriate work up 1.76 vs. 2.25 $p < 0.001$. • The CDSS group had greater improvement in appropriate testing with absolute change 30.2% vs. 10.9%, difference in change 19.3% (95% CI 2.9-35.6, $p = 0.023$). This improvement occurred mainly in patients with PE ruled out. • Definite thromboembolic events were observed in 2/375 (0.5%) of CDSS patients at 3-months vs. 6/589 (1%) of paper guideline patients.
	How precise was the estimate of the treatment effect?	See 95% CI above.

III.	How can I apply the results to patient care (answer the questions posed below)?	
1.	Were the study patients similar to my patient?	Yes, adult ED patients with suspected PE. The proportions of low, intermediate or high risk stratifications are not provided, but are probably similar to our institution.
2.	Were all clinically important outcomes considered?	Yes, including 3-month VTE-rates. "However, our study was not designed to detect" differences in these secondary outcomes. (p. 684) It would also have been interesting to assess clinician acceptance and impact on ED thoroughfare.
3.	Are the likely treatment benefits worth the potential harm and costs?	Yes, cheap and relatively straight forward CDSS on a hand held device increase the appropriateness of PE diagnostic work up with negatively impacting patient outcomes come with simple guideline dissemination.

Limitations

1) **Unequal distribution of patients between groups.**
2) **Significant proportion lost to follow up, especially in CDSS subset with PE excluded.**
3) **Under powered to detect 3-month outcome differences between groups.**
4) **No assessment of clinician acceptability or impact on ED thoroughfare.**

Bottom Line

Paper-based guideline dissemination that does not integrate decision support tools into physician work flow at the time and location of decision making have limited effect on moving emergency clinicians towards more Bayesian-appropriate PE diagnostic strategies.