

RESEARCH LETTER

Performance and Utilization of an Emergency Department Electronic Screening Tool for Pneumonia

Appropriate treatment of pneumonia begins with accurate diagnosis. However, clinicians have difficulty integrating data for clinical decision making.¹ Significant variability in pneumonia management exists in the emergency department (ED).² Decision support might decrease variability and improve care, but physician utilization is historically low.³ An alerting tool is needed for physicians to utilize computer-based pneumonia decision support.

We developed a real-time electronic screening tool that identifies patients with pneumonia by applying Bayesian probabilistic logic to the electronic medical record, and we implemented the tool in 4 EDs. A “P” appears

on the ED electronic tracker board when pneumonia likelihood reaches 40%. Clicking on the icon displays pneumonia likelihood and the relevant data (Figure). After confirming the diagnosis, the ED physician proceeds with a linked decision support tool that provides management recommendations.

See Invited Commentary at end of letter

We measured tool sensitivity, specificity, and predictive values against physician panel case review. We also measured physician acknowledgment of the screening tool alert and utilization of the linked pneumonia management tool.

Methods. In 1995, Intermountain Healthcare began deploying an electronic medical record in Salt Lake County, Utah EDs, currently staffed by 80 attending physicians and 24 emergency medicine residents. A paper-based

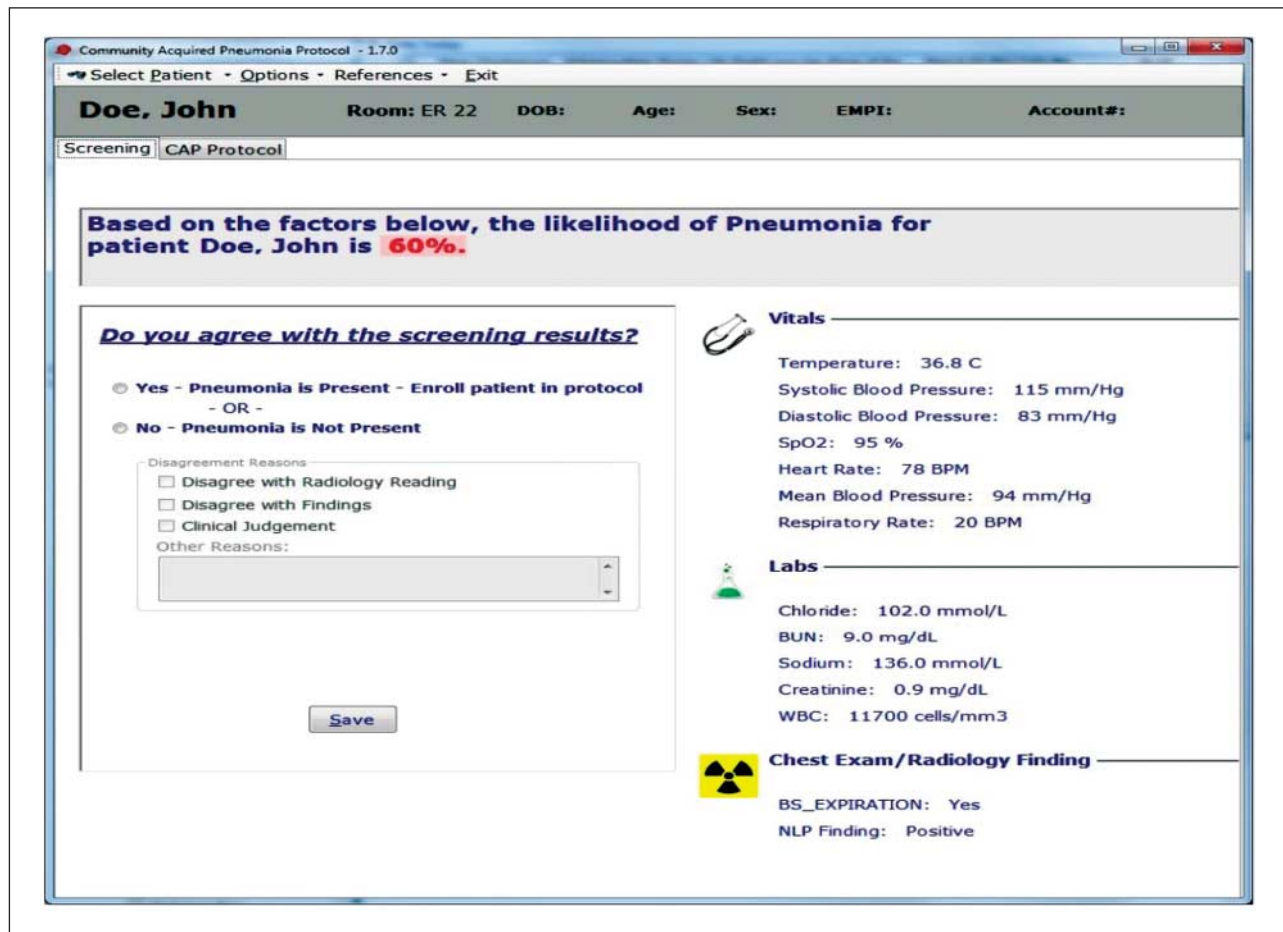


Figure. Screen brought up after clinician clicks the “P” alert on the electronic tracking board. Estimated likelihood of pneumonia as well as relevant clinical features are presented to the clinician, who may agree or disagree with the diagnosis.

pneumonia guideline was implemented with moderate success in 1996.⁴ Electronic screening tool development began in 1998,⁵ but implementation was impractical until dictated radiology reports became electronically available a median of 20 minutes after study completion.

The pneumonia screening tool is based on a medical screening and diagnosis framework within Intermountain Healthcare's computerized environment. The tool monitors ED patient data, extracts relevant clinical features, and calculates the pneumonia probability. The tool development reference data set included 48 449 Intermountain ED patient encounters from 2008 through 2010, including 2413 patients with pneumonia. Utilized data include 6 vital sign variables, 6 laboratory values, 25 nursing assessment variables, patient age, patient's chief complaint, and findings extracted from the chest imaging report using natural language processing. For example, complaints of fever or cough increase likelihood of pneumonia. Since pneumonia is "a constellation of suggestive clinical features with a demonstrable infiltrate,"^{6(p539)} the tool runs when any ED patient's dictated chest imaging report becomes available.

Study Population. The study population included all ED patients with chest imaging during two 2-month periods in 2011, immediately following implementation in May and then 6 months later. We thereby studied tool performance during early summer and early winter periods of pneumonia incidence and measured physician utilization over time. We randomly selected from each time period 300 plus 60 additional patients for whom the tool alerted for possible pneumonia. Three physician authors independently reviewed patient records to determine presence of pneumonia. Initial disagreement was resolved by discussion and consensus. The Intermountain Healthcare institutional review board approved tool development, implementation, and data collection as quality improvement.

Statistical Analysis. Random sample data allowed unbiased estimates and CIs using standard formulas. To correct for verification bias in the combined sample, sensitivity and specificity estimates were computed using the Begg and Greenes approach⁷ with 95% CIs by bias-corrected bootstrap. Tool utilization was analyzed by χ^2 analysis.

Results. Chest imaging was performed on 13 859 ED patients during the 2 study intervals. Pneumonia incidence was 8.2%. Among patients with pneumonia confirmed by physician panel review (n=109), mean (SD) age was 56.2 (20.7) years; 56% were women; and hospital admission rate was 48%.

The screening tool alerted in 81 of 109 true cases of pneumonia (74%) by physician review. The tool falsely alerted in 78 of 612 patients without pneumonia (13%). Sensitivity was 40.9% (95% CI, 32.0%-52.0%); specificity, 96.6% (95% CI, 95.9%-97.3%); positive predictive value, 50.9% (95% CI, 42.9%-58.9%); and negative predictive value, 95.0% (95% CI, 92.9%-96.7%). Imaging reports incorrectly read positive by the tool caused 59% of false-positive cases. Most false-negative cases were caused by lack of clinical pneumonia features other than chest imaging.

Among true-positive pneumonia cases, physicians' agreement with the tool alert increased over the first 6 months nonsignificantly from 37% (15 of 41) to 53% (21 of 40) ($P=.18$). Of true pneumonia cases, tool utilization increased from 12% (6 of 49) to 48% (29 of 60) ($P<.001$).

Comment. We developed and implemented an electronic screening tool for pneumonia with moderate accuracy compared with physician review. We have not found a similar tool reported previously.⁶ Our tool's greatest challenge has been natural language processing translation of dictated, unstructured radiology reports into discrete conclusions. Variability in radiology reports makes interpretation difficult both for computers and clinicians. Changing to structured reporting would improve tool performance (eg, "new parenchymal opacity consistent with pneumonia in the appropriate clinical setting: yes/no/possibly").

Choosing the alert threshold of 40% was another challenge. Screening tools usually maximize sensitivity over specificity, but we emphasized specificity to avoid "alert fatigue."⁸ Because lower sensitivity resulted, we allowed clinicians to launch the management tool via a desktop icon in a parallel process.

Accuracy and utilization will likely improve with iterative improvements. The next performance test will be to determine if the tool helps physicians diagnose pneumonia and improve clinical outcomes.⁶

Nathan C. Dean, MD
Barbara E. Jones, MD
Jeffrey P. Ferraro, MS
Caroline G. Vines, MD
Peter J. Haug, MD

Published Online: March 18, 2013. doi:10.1001/jamainternmed.2013.3299

Author Affiliations: Division of Pulmonary and Critical Care Medicine (Drs Dean and Jones) and Emergency Department (Dr Vines), Intermountain Medical Center, and Department of Internal Medicine, University of Utah (Drs Dean and Jones), Salt Lake City; and Homer Warner Center for Informatics Research, Salt Lake City, Utah (Mr Ferraro and Dr Haug).

Correspondence: Dr Dean, Intermountain Medical Center, 5121 S Cottonwood St, Murray, UT 84107 (nathan.dean@imail.org).

Author Contributions: *Study concept and design:* Dean, Jones, and Haug. *Acquisition of data:* Dean, Jones, Ferraro, and Vines. *Analysis and interpretation of data:* Dean, Jones, and Ferraro. *Drafting of the manuscript:* Dean and Jones. *Critical revision of the manuscript for important intellectual content:* Dean, Ferraro, Vines, and Haug. *Statistical analysis:* Jones. *Administrative, technical, and/or material support:* Ferraro. *Study supervision:* Haug. *Tool implementation:* Dean.

Conflict of Interest Disclosures: None reported.

Funding/Support: This study was funded in part by Intermountain Healthcare and the Intermountain Research and Medical Foundation Salt Lake City, Utah, and by Pfizer US from an unrestricted, investigator-initiated competitive grant.

Role of the Sponsors: The sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. None of the authors have relevant financial interests, activities, relationships, or affiliations.

Previous Presentation: This work was presented in abstract form at the European Respiratory Society International Conference; September 3, 2012; Vienna, Austria. **Additional Contributions:** The authors thank Al Jephson, BS, for data extraction; Gregory J. Stoddard, MS, for statistical analysis; Dominik Aronsky, MD, for his original work on the electronic tool and many suggestions since; Todd L. Allen, MD, for helping design and implement the tool in the ED; Herman B. Post, BS, Kumar Mynam, MS, and Darin Wilcox, BS, for their design and programming of the tool; and the ED physicians providing care at the Salt Lake County Intermountain Hospitals for their interest in quality improvement and helpful suggestions during development of the electronic decision support tool.

1. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science*. 1981;211(4481):453-458.
2. Dean NC, Jones JP, Aronsky D, et al. Hospital admission decision for patients with community-acquired pneumonia: variability among physicians in an emergency department. *Ann Emerg Med*. 2012;59(1):35-41.
3. Vines CJ, Dean NC. Technology implementation impacting the outcomes of patients with CAP. *Semin Respir Crit Care Med*. 2012;33(3):292-297.
4. Dean NC, Silver MP, Bateman KA, James B, Hadlock CJ, Hale D. Decreased mortality after implementation of a treatment guideline for community-acquired pneumonia. *Am J Med*. 2001;110(6):451-457.
5. Aronsky D, Haug PJ. Diagnosing community-acquired pneumonia with a Bayesian network. *Proc AMIA Symp*. 1998;11(3):632-636.
6. Mandell LA, Wunderink RG, Anzueto A, et al; Infectious Diseases Society of America; American Thoracic Society. Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. *Clin Infect Dis*. 2007;44(suppl 2):S27-S72.
7. Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. *Biometrics*. 1983;39(1):207-215.
8. Ash JS, Sittig DF, Campbell EM, Guappone KP, Dykstra RH. Some unintended consequences of clinical decision support systems. *AMIA Annu Symp Proc*. 2007;Oct 11:26-30.

INVITED COMMENTARY

Computers and the Diagnosis of Pneumonia

Pneumonia and influenza together ranked as the eighth leading cause of death in the United States and led to over 1 million hospital admissions in 2009, with pneumonia accounting for the majority of the deaths.¹ While most cases will resolve with treatment, patients admitted to the intensive care unit have mortality rates as high as 37%.² Prior studies show that timely initiation of appropriate antibiotic therapy is associated with decreased mortality.³ In addition to its clinical and public health importance, pneumonia is also one of several conditions targeted in “pay for performance” approaches, and hospitals’ performance on certain measures of pneumonia quality is publicly reported. Thus, establishing an accurate and timely diagnosis of this illness is imperative.

Over the past 30 years, several clinical prediction rules have been developed to aid clinicians in the diagnosis of

pneumonia. Despite the validation of such rules, which rely on the presence of signs and symptoms to estimate the probability of pneumonia, there is no combination of history and physical examination findings that can rule in or rule out pneumonia with certainty.⁴ Consequently, the presence of an infiltrate found on chest radiograph in the right clinical scenario is considered the reference standard for the diagnosis of pneumonia. Clinicians use a combination of symptoms, signs, and laboratory and radiographic findings to diagnose and treat patients with pneumonia and to prognosticate outcomes, and the advent of electronic medical records and computerized decision support means that each of these data points is now available for computers to do this work for us.

It is in this context that Dean et al⁵ report on the development of a screening tool to identify patients with pneumonia in the emergency department (ED) using data prospectively amassed within an electronic medical record.⁵ The tool monitors ED patients and calculates the probability of pneumonia based on their chief complaint, age, vital signs, laboratory findings, nursing assessment, and chest radiography reports interpreted using natural language processing. The screening tool alerts physicians when the probability of pneumonia reaches a 40% threshold. The authors focus on patients who had already undergone chest imaging to determine the diagnostic test characteristics of the tool, with adjustment for verification bias. To assess the tool, the gold standard for the diagnosis of pneumonia was not the chest radiograph (since these findings were inputs for the screening tool) but instead was an implicit determination of pneumonia by a panel of 3 physicians who reviewed the medical record. The tool alerted in 74% of true cases of pneumonia and in 13% of patients without pneumonia, with a reported sensitivity of 40.9% and a specificity of 96.6%.

For a diagnostic prediction rule to have benefit, it should have a few key qualities: (1) usability; (2) accuracy exceeding physician judgment alone; and (3) implications for patient treatment or the outcomes or costs of care.⁶ Undoubtedly, the tool described by Dean et al⁵ would fulfill the first criteria in health care settings that use similar electronic medical records—it runs in real time on every patient and alerts physicians when the probability of pneumonia exceeds a given threshold. This limits the need for a human to input information. However, it is less clear whether the tool fulfills the other 2 criteria. It is likely that pneumonia is already relatively high on the differential diagnosis of any patient presenting with signs and symptoms that lead the clinician to order a chest radiograph. Furthermore, as the authors state, the real question is whether this tool helps physicians make better decisions about whether to treat patients and whether treatment can improve the clinical outcomes that matter for pneumonia, such as mortality, medical complications, or recovery from symptoms and impaired functional status. Ultimately, physicians will need to be convinced that the tool works better than their own judgment and benefits patients.

Notably, the authors chose to increase the specificity of the tool over its sensitivity, to avoid a high false-