Efficacy of an Evidence-Based Clinical Decision Support in Primary Care Practices
A Randomized Clinical Trial

Thomas G. McGinn, MD, MPH; Lauren McCullagh, MPH; Joseph Kannry, MD; Megan Knaus, MPH; Anastasia Sofianou, MS; Juan P. Wisnivesky, MD, DrPH; Devin M. Mann, MD, MS

IMPORTANCE There is consensus that incorporating clinical decision support into electronic health records will improve quality of care, contain costs, and reduce overtreatment, but this potential has yet to be demonstrated in clinical trials.

OBJECTIVE To assess the influence of a customized evidence-based clinical decision support tool on the management of respiratory tract infections and on the effectiveness of integrating evidence at the point of care.

DESIGN, SETTING, AND PARTICIPANTS In a randomized clinical trial, we implemented 2 well-validated integrated clinical prediction rules, namely, the Walsh rule for streptococcal pharyngitis and the Heckerling rule for pneumonia.

INTERVENTIONS AND MAIN OUTCOMES AND MEASURES The intervention group had access to the integrated clinical prediction rule tool and chose whether to complete risk score calculators, order medications, and generate progress notes to assist with complex decision making at the point of care.

RESULTS The intervention group completed the integrated clinical prediction rule tool in 57.5% of visits. Providers in the intervention group were significantly less likely to order antibiotics than the control group (age-adjusted relative risk, 0.74; 95% CI, 0.60-0.92). The absolute risk of the intervention was 9.2%, and the number needed to treat was 10.8. The intervention group was significantly less likely to order rapid streptococcal tests compared with the control group (relative risk, 0.75; 95% CI, 0.58-0.97; \( P = .03 \)).

CONCLUSIONS AND RELEVANCE The integrated clinical prediction rule process for integrating complex evidence-based clinical decision report tools is of relevant importance for national initiatives, such as Meaningful Use.

TRIAL REGISTRATION clinicaltrials.gov Identifier: NCT01386047
Efficacy of a Clinical Decision Support System

Up to 30% of all medical care in the United States can be considered unnecessary.\(^1\)\(^-\)\(^3\) There is widespread consensus among health care experts, policymakers, and consumers that health information technology has great potential to help reduce this overtreatment, particularly electronic health records (EHRs) with clinical decision support (CDS). To this end, the federal government has distributed more than $6 billion in incentive payments to health care professionals and hospitals to implement CDS in conjunction with EHRs to improve performance on high-priority health conditions.\(^4\) Nevertheless, this potential has not been realized, highlighting the need to improve provider (including attending physicians, residents, fellows, and nurse practitioners) adoption and use of CDS.\(^5\)\(^-\)\(^8\)

Most forms of CDS involve input of 1 or 2 simple data points. For example, influenza vaccination reminders embedded in EHRs use the time of the year to trigger providers to act, and colon cancer screening reminders use patient age as the trigger. More complex forms of decision support that are customized and pull data from various data fields hold great potential to reduce waste. Successful examples of CDS have reduced prescribing of brand-name antibiotics,\(^9\) improved lipid management in patients receiving renal transplants,\(^10\) enhanced compliance with guidelines for treating human immunodeficiency virus,\(^11\)\(^-\)\(^13\) and decreased ordering of tests when costs were displayed,\(^14\) in addition to age-specific alerts that have minimized inappropriate prescribing in the elderly.\(^15\)\(^-\)\(^20\) Yet, large-scale adoption rates continue to be a challenge.

Integrated clinical prediction rules (CPRs) represent a form of CDS used at the point of care. They are well-validated forms of evidence-based tools that help providers accurately estimate probabilities and have demonstrated in research trials the ability to reduce waste, while simultaneously enhancing quality of care.\(^21\) However, to integrate CPRs into an EHR requires data from various components of the history, physical examination, and basic laboratory test results instantaneously, making them difficult to include in the EHR. Prior attempts to integrate such complex forms of evidence at the point of care to influence frontline decision making have failed because of a lack of adoption by providers.\(^22\)\(^-\)\(^23\)

We chose to examine the effect of new methods for evidence integration and to assess the influence this integration had on the management of upper respiratory tract infections. For this study, we examined 2 underused CPRs, the rule by Heckerling et al\(^24\) for pneumonia and the rule by Walsh et al\(^25\) for streptococcal pharyngitis (data not shown). Although there are other scores developed for these diseases, the Walsh and Heckerling rules have been validated in various settings and are considered the standard of care, yet they are not uniformly applied in clinical practice.\(^25\)\(^-\)\(^27\)

Our study sought to improve provider adoption of this complex CDS at the point of care in real time by seamlessly blending CPRs into provider work flow. Phase 1 of the study evaluated novel usability techniques that draw on low-cost technology and real ambulatory clinical settings and created a tool that includes CPRs in a widely used commercial EHR platform. This first phase of the development and testing is described in previous publications.\(^28\)\(^-\)\(^29\)

Methods

The CPR study consisted of the following 2 phases: (1) the development and usability testing of the CPR tool and (2) a randomized clinical trial in which we evaluated the effectiveness of the CPR for reducing antibiotic ordering rates among enrolled providers, as well as the uptake of the tool. Phase 1 and the usability testing that were central to the development are described in greater detail elsewhere.\(^28\)

Setting

The institutional review board at Mount Sinai School of Medicine approved the study protocols. The CPR randomized clinical trial was conducted within 2 large urban ambulatory primary care practices at Mount Sinai Medical Center in New York City with a mix of providers, including attending physicians, residents, fellows, and nurse practitioners. The patient population of the clinical practice was racially/ethnically diverse, with almost 56% of patients self-identifying as Hispanic, 35% as African American, 7% as white, and 2% as other races/ethnicities (data not shown).

Exclusion and Inclusion Criteria

Attendings, residents, fellows, and nurse practitioners working in the outpatient primary care clinic at the medical center were eligible to participate in the study. Providers were not enrolled if they were part of the study team.

Study Design

Primary care providers were randomized using a random number generator to a control group (usual care) or an intervention group (Figure 1). The intervention group participated in an in-person 1-hour training consisting of an overview of CPRs...
and the evidence supporting the Walsh and Heckerling CPRs in particular, as well as the study protocols, demonstration of the tool in the EHR, and a video of a simulated patient encounter in which the tool is used. Providers randomized to the control group received background information on the Walsh and Heckerling CPRs (Figure 2).

From November 1, 2010, through October 31, 2011, we activated the CPR tool for intervention providers by modifying their EHR user access settings. The CPR tool appeared on the provider’s screen during a clinical encounter when the provider entered 1 or more keywords in the fields for chief complaint, diagnosis, or orders (data not shown). Criteria for triggering included chief complaints and diagnoses associated with pharyngitis or pneumonia or a diagnosis and test order combination.28 Chief complaints relevant to pharyngitis included sore throat, throat discomfort, and streptococcal pharyngitis; for pneumonia, they included possible pneumonia and “chest hurts when breathing.”28 Providers were invited to complete an CPR risk score calculator and were given management recommendations based on the score. The recommendation for patients with a total score of -1 (unlikely to have streptococcal pharyngitis) was to provide supportive care. The recommendation for patients with a total score of 0 or 1 (intermediate likelihood of streptococcal pharyngitis) was to obtain a throat swab or symptom resolution before deciding about antibiotic therapy. The recommendation for patients with a total score of 2 or higher (high likelihood of streptococcal infection) was to start empirical antibiotic therapy, with no culture necessary. Similar recommendations, but specific to pneumonia, were given after completing the pneumonia risk score calculator. Patient encounters with the study providers were monitored, and data were collected from the EHR for the intervention period.

Outcomes
Primary Outcomes

We assessed changes in provider patterns of ordering antibiotics as the primary outcome. Antibiotic orders were evaluated by examining EHR data related to patient encounters in which the tool was triggered. The study team derived a list of antibiotics for pharyngitis and pneumonia used in the analysis, with consultation from other primary care physicians. We measured the frequency, rates, and type of antibiotics prescribed after using the CPR risk score calculator and bundled order sets. Additional outcomes measured were how often chest radiographs, rapid streptococcal tests, and pharyngitis throat cultures were ordered during eligible patient visits. We recorded the occurrence of 1 or more emergency department and outpatient clinic visits and antibiotics ordered 2 weeks after an CPR visit as a proxy for the appropriateness of care as a result of using the tool.

Process Outcomes

Step-by-step analysis of provider use of the CPR tool has been critical to understanding its uptake and to shaping its ongoing development. Therefore, we also assessed the uptake and use of the tool by measuring the rate of eligible visits in which (1) the provider opened the tool after it was triggered (based on the criteria in the Study Design subsection of the Methods section), (2) the provider completed the risk score calculator embedded in the tool, (3) the provider viewed the bundled order set associated with the score generated, and (4) the provider signed the associated bundled order set.

Provider demographics such as sex, age, and years of practice were collected in a self-reported survey and analyzed. Patient characteristics were selected for analysis on the basis of their potential to influence provider decisions. These included age, sex, smoking status, any allergies to antibiotics (penicillins, cephalosporins, quinolones, and macrolides), and relevant comorbidities (asthma, chronic obstructive pulmonary disease, diabetes mellitus, and congestive heart failure).

Statistical Analysis

Differences in the characteristics of patients seen by providers in the intervention and control groups were measured using t test and χ² test as appropriate. Provider characteristics were compared using similar methods (data not shown). Process, primary, and secondary outcomes were compared for all intervention and control visits and were stratified by pneumonia or pharyngitis CPR tool. Absolute risk (risk difference) and
relative risk (RR) ratios of antibiotic prescribing in the intervention group vs the control group were compared using a generalized estimating equation model (logit link) with robust standard errors to account for clustering of patient visits (encounters) within providers and subsequent correlated outcomes. Because the median age of patients differed in the intervention and control groups, we performed a secondary analysis adding age as a covariate in our final age-adjusted models. The same approach was applied for our secondary outcomes of how often chest radiographs, rapid streptococcal tests, and pharyngitis throat cultures were ordered. Similar generalized estimating equation analysis compared rates of antibiotic prescribing and secondary outcomes in the intervention group vs the control group 2 weeks after a visit. Analyses were conducted with statistical software (SAS, version 9.2; SAS Institute, Inc) using 2-sided P values.

### Results

From November 1, 2010, through October 31, 2011, a total of 168 primary care providers provided informed consent and were enrolled in the study. There were no statistically significant differences in demographics between providers in the intervention and control groups (data not shown).

In total, more than 40 003 patient visits occurred during the study period. The median ages of the patients seen at the study visits were statistically different for the intervention and control groups (P = .001). No statistically significant differences in sex or race/ethnicity were observed between the 2 groups. Similarly, no statistically significant differences were found in the patient health data on smoking status, relevant comorbidities, or relevant allergies to the antibiotics on the CPR list or other medication allergies (Table 1).

### Primary Outcomes

We found statistically significant differences in antibiotic prescriptions, rapid streptococcal test orders, and pharyngitis throat cultures between the control and intervention groups (Table 2). Providers in the intervention group were significantly less likely to order antibiotics (age-adjusted RR, 0.74; 95% CI, 0.60-0.92; P = .008). The absolute risk of the intervention was 9.2%, and the number needed to treat was 10.8.

When stratified by pharyngitis and pneumonia, the RR of ordering remained significant for pneumonia (age-adjusted RR, 0.79; 95% CI, 0.64-0.98; P = .03). The intervention group was significantly less likely to order rapid streptococcal tests in pharyngitis encounters (age-adjusted RR, 0.75; 95% CI, 0.58-0.97; P = .03). No statistically significant differences were observed in chest radiograph ordering among patients with possible pneumonia (age-adjusted RR, 0.98; 95% CI, 0.60-1.62; P = .95) or in pharyngitis throat culture orders (age-adjusted RR, 0.54; 95% CI, 0.18-1.64; P = .28).

The RR ratio for ordering quinolones (broad-spectrum antibacterial drugs) in the intervention group compared with the control group was 0.50 (95% CI, 0.29-0.88; P = .02) (Table 3). No significant differences were observed between the other antibiotics, including penicillins, cephalosporins, and macrolides.

No significant differences were found between the intervention and control groups in the proportions of visits resulting in a patient returning to the emergency department (P > .99) or outpatient clinic (P = .10) for follow-up treatment (Table 4). Likewise, no differences were observed in rates of antibiotic orders between the intervention and control groups 2 weeks after CPR visits (P = .45).

### Process Outcomes

The CPR tool was triggered 984 times. Of these, 398 patients were seen by control providers and 586 by intervention providers. The CPR tool was triggered 984 times. Of these, 398 patients were seen by control providers and 586 by intervention providers.
Overall, the tool had high adoption rates in the intervention group, with 62.8% having opened the tool and 57.5% of providers accepting the tool. Among the encounters seen by intervention providers, the pharyngitis tool was more commonly used than the pneumonia tool (Table 5). In total, 74.3% of 374 pharyngitis encounters resulted in an initial acceptance of the triggered CPR tool compared with 42.5% of 212 pneumonia encounters (RR, 0.65; 95% CI, 0.54-0.78; P < .001).

In addition, providers were significantly more likely to complete later steps for the pharyngitis tool than for the pneumonia tool. Examples include completion of the risk score calculator (RR, 0.69; 95% CI, 0.57-0.83) and completion of bundled order sets (RR, 0.60; 0.46-0.77) (P < .001 for both). For most patient visits, these scores categorized patients as being at low risk for pharyngitis (55.4%) and pneumonia (96.6%). Providers were more likely to open bundled order sets for pharyngitis visits than for pneumonia visits. For pharyngitis visits, 63.1% of bundled sets were opened, and 50.5% were complete.
Discussion

The CPR is a form of complex CDS that holds great promise for reducing overtreatment and testing. Although numerous well-validated CPRs exist, few studies30-33 have reported significant adoption of CPRs, and they have found little effect on provider behavior.

The CPR study brought CDS to the point of care in real time by embedding 2 well-validated CPRs in an ambulatory EHR. The intervention group had a high overall adoption rate (62.8%); orders for antibiotic therapy were significantly reduced (age-adjusted RR, 0.74; 95% CI, 0.60-0.92; P = 0.008), as were orders for rapid streptococcal tests (age-adjusted RR, 0.75; 95% CI, 0.58-0.97; P = 0.03). The antibiotics that were ordered more were appropriate (ie, narrow spectrum). When comparing pneumonia vs pharyngitis antibiotic ordering, a significant difference was observed only for pneumonia-related visits. The reason is unclear but is attributed to the CPR tool design or to the integration into provider work flow, and further investigation is required.

Prior investigations of CDS achieved adoption rates between 10% and 20%,34 In addition to our high overall adoption rate of 62.8%, we found that 57.5% of all relevant encounters resulted in the provider completing the risk score calculator and generating a risk score assessment. This exceeds the adoption rates for another CDS intervention that attempted to reduce antibiotic prescriptions for upper respiratory tract infections but resulted in a reduction of only 16%.35 We believe our results indicate that providers may have perceived the tool as being helpful with the clinical diagnosis of pharyngitis and pneumonia, enhancing clinical work flow and improving patient care, which are hallmarks of CDS interventions that are well received by physicians.23,36,37 We speculate that the high adoption rates were a result of the comprehensive user-centered development process we created for this study, including pilot testing with providers, usability testing, focused user training on CDS (rarely provided in other studies), and collaboration with a multidisciplinary team that included experts in usability and informatics and CDS specialists.29,38

The CPR tool does not guarantee a reduction in antibiotic use after implementation. If overuse of antibiotics was an issue at baseline, the tool should reduce use. Although some clinicians may have been triggered to order antibiotics, the overall rate of ordering antibiotics was reduced. This study was unusual in that it specifically addressed many of the challenges of integrated decision support identified in earlier studies (cited in the introduction). These included ensuring that the CDS tool used only off-the-shelf functionality from the vendor, minimizing the frequency of alerts and carefully balancing alerts in relation to workflow.28 In addition, the study team conducted usability testing and focus groups with providers, incorporated a supportive training program, and provided personalized on-call troubleshooting during the first weeks of implementation. Finally, this study was performed in the setting of a widely available commercial EHR rather than a proprietary system available to only a single institution. Resources to maintain the tool and implementation resources consisted of a part-time EHR programmer to assist with the design and troubleshoot the EHR program, significantly enhancing the potential for replication.39

Our study had several limitations. It was performed at 2 primary care practice sites within a large academic health center with significant numbers of trainees. Although off-the-shelf technology was used to facilitate exportation of the method to other sites, further dissemination studies are needed to identify the effect of clinical setting on CPR effectiveness. In addition, our study took place in an inner-city academic setting; therefore, its use in more diverse settings may offer different insight.

Table 5. Integrated Clinical Prediction Rule Tool Use in the Intervention Group

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total (N = 586)*</th>
<th>Pharyngitis (n = 374)</th>
<th>Pneumonia (n = 212)</th>
<th>Absolute Risk Difference</th>
<th>Relative Risk (95% CI)†</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tool opened</td>
<td>368 (62.8)</td>
<td>278 (74.3)</td>
<td>90 (42.5)</td>
<td>31.8</td>
<td>0.65 (0.54-0.78)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Risk score calculator opened</td>
<td>337 (57.5)</td>
<td>249 (66.6)</td>
<td>88 (41.5)</td>
<td>25.1</td>
<td>0.69 (0.57-0.83)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Pharyngitis scoreb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low, −1 to 0</td>
<td>...</td>
<td>138/249 (55.4)</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Medium, 1 to 2</td>
<td>...</td>
<td>93/249 (37.3)</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>High, 3 to 4</td>
<td>...</td>
<td>18/249 (7.2)</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Pneumonia score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low, −1 to 3</td>
<td>...</td>
<td>...</td>
<td>85/88 (96.6)</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>High, 4 to 5</td>
<td>...</td>
<td>...</td>
<td>3/88 (3.4)</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Bundled order sets opened</td>
<td>319 (54.4)</td>
<td>236 (63.1)</td>
<td>83 (39.2)</td>
<td>23.9</td>
<td>0.68 (0.56-0.83)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Bundled order sets completed</td>
<td>246 (42.0)</td>
<td>189 (50.5)</td>
<td>57 (26.9)</td>
<td>23.6</td>
<td>0.60 (0.46-0.77)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Abbreviation: ellipsis, not applicable.

* Total of all visits in the intervention group.
† Risk ratios, CIs, and P values are calculated from a generalized estimating equation log-binomial model adjusting for clustering of orders or visit by provider and using robust standard error estimators.
In conclusion, we demonstrated high adoption rates (>50%) of a complex decision support tool that integrates evidence-based diagnostic tools at the point of care. Successful integration and adoption were related to its user-centered design and implementation process, which resulted in a reduction in antibiotic ordering and in point-of-care testing and throat cultures. High adoption rates of CDS interventions that lead to changes in outcomes need to be further researched. Our study showed one such approach to successful CDS interventions in a commercial EHR.

ARTICLE INFORMATION

Author Contributions: Study concept and design: McGinn, Kanny, Wisnivesky, Mann. Acquisition of data: Kanny, Mann. Analysis and interpretation of data: McGinn, McCullah, Knaus, Sofianou, Wisnivesky, Mann. Drafting of the manuscript: McGinn, McCullah, Kanny, Knaus, Sofianou, Mann. Critical revision of the manuscript for important intellectual content: McGinn, Kanny, Wisnivesky, Mann. Obtained funding: McGinn. Administrative, technical, and material support: McCullah, Kanny, Knaus. Study supervision: McGinn, Kanny.

Conflict of Interest Disclosures: Dr Wisnivesky is a member of the research board of EHE International; has received consulting fees from Merck, IMS Health, and UBS; and is the recipient of a research grant from GlaxoSmithKline.

Funding/Support: This study was supported by grant 5 R1B HS 018491 from the Agency for Health and Quality Research (Dr McGinn).

Role of the Sponsor: The Agency for Health and Quality Research participated in the design and conduct of the study and in the collection, management, analysis, and interpretation of the data.

Additional Contributions: Jaccquy Arciniegas, MPH, Diego Chileuska, BA, Dan Edonyabo, MS, Alice Li, MD, and Lucas Romero, BA, assisted with the study.

REFERENCES
If a clinical prediction rule falls off the shelf of a physician's office, does it make a sound? I think not.

Many clinical prediction rules to improve the quality and efficiency of medical care have been developed and published, but they are not widely used. There are many complicated reasons for their underuse. First, not all clinical prediction rules have been validated on independent samples in diverse clinical settings; therefore, physicians have reason to question the reproducibility of the results. Second, even when validated clinical prediction rules are available, physicians tend to favor their own judgment when caring for a specific patient. Perhaps most important, physicians do not use clinical prediction rules because most of them are time-consuming and cumbersome, requiring entering data into a calculator or computer, interrupting the flow of our clinical encounter and thought processes.

Electronic health records could potentially promote the use of clinical prediction rules by integrating the decision rule into the clinical work flow. The authors of “Efficacy of an Evidence-Based Clinical Decision Support in Primary Care Practices” demonstrate the value of this approach by embedding 2 underused clinical prediction rules—the Walsh rule for pharyngitis and the Heckerling rule for pneumonia—into the electronic health record used by 2 primary care practices in a large academic medical center.

Physicians randomized to the intervention were prompted when they typed in keywords, such as sore throat or possible pneumonia, to open an electronic tool in the margin. They could then enter the necessary information into the calculator embedded in the electronic tool, and the result would produce a bundled order set that they could sign. The control physicians simply received information about the 2 rules.

By making it easier for physicians to complete the clinical prediction rules and take the appropriate actions, the authors demonstrated an impressive decrease in the ordering of antibiotics and a 50% decrease in the ordering of the broad-spectrum quinolones. Of note, the tool was developed with focus group input from physicians and usability testing, and supportive training was provided to physicians. Here too is a lesson for all health systems developing electronic health records: physician input, customization, and training are critical for success. Otherwise, we will have electronic health records, accessible and legible, full of cut-and-pasted material, and our care will be little improved.