Accuracy of a computerized clinical decision-support system for asthma assessment and management

Laura J Hoeksema,1 Alia Bazzy-Asaad,2 Edwin A Lomotan,3 Diana E Edmonds,4 Gabriela Ramírez-Garnica,5 Richard N Shiffman,4 Leora I Horwitz6,7

ABSTRACT
Objective To evaluate the accuracy of a computerized clinical decision-support system (CDSS) designed to support assessment and management of pediatric asthma in a subspecialty clinic.
Design Cohort study of all asthma visits to pediatric pulmonology from January to December, 2009
Measurements CDSS and physician assessments of asthma severity, control, and treatment step.
Results Both the clinician and the computerized CDSS generated assessments of asthma control in 767/1032 (74.3%) return patients, assessments of asthma severity in 100/167 (59.9%) new patients, and recommendations for treatment step in 66/167 (39.5%) new patients. Clinicians agreed with the CDSS in 543/767 (70.8%) of control assessments, 37/100 (37%) of severity assessments, and 19/66 (29%) of step recommendations. External review classified 72% of control disagreements, 56% of severity disagreements, and 76% of step disagreements as CDSS errors. The remaining disagreements resulted from pulmonologist error or ambiguous guidelines. Many CDSS flaws, such as attributing all ‘cough’ to asthma, were easily remediable. Pediatric pulmonologists failed to follow guidelines in 8% of return visits and 18% of new visits.
Limitations The authors relied on chart notes to determine clinical reasoning. Physicians may have changed their assessments after seeing CDSS recommendations.
Conclusions A computerized CDSS performed relatively accurately compared to clinicians for assessment of asthma control but was inaccurate for treatment. Pediatric pulmonologists failed to follow guideline-based care in a small proportion of patients.

INTRODUCTION
Guidelines for the treatment of pediatric asthma were created in 1991 by the National Asthma Education and Prevention Program1 and were updated in 1997,2 2002,3 and 2007.4 Thus, for at least two decades, there has been nationally disseminated expert guidance regarding the appropriate treatment of pediatric asthma. Adherence to these guidelines reduces asthma emergency-department visits and hospitalizations.5 6 Nonetheless, only a quarter of patients with persistent asthma symptoms are taking anti-inflammatory medication as recommended by the guidelines.7 Consequently, there has been little discernible improvement in pediatric asthma outcomes in the USA since the development of the asthma guidelines.8 For example, the rate of emergency department visits for asthma decreased only slightly from 1992 to 2006.8

Clinical decision support through an electronic health record has been proposed as a promising approach to improving guideline-based care.9 10 Several such systems have been implemented for asthma management,11–15 but overall results have been mixed.17 18 While some studies have reported improvements in documentation,13 processes,11 15 or outcomes,14 others have not been successful.12 18

It is possible that one barrier to successful practice change is inaccuracy of the systems themselves. Guidelines are often written with vague or underspecified language, complicating the translation into computer algorithms.19 Furthermore, tailoring therapy to symptoms is a substantially more complicated decision process and requires correspondingly more complex decision support than simple reminders for preventive care or alerts for medication interventions.20 21 Adding another layer of complexity, the newest guidelines for asthma management require the assessment of severity (for new patients) or control (for return patients) to be based both on current impairment and on future risk.4 Yet, few studies have reported the accuracy and validity of the advice produced by clinical decision-support systems (CDSS) for asthma care.22–25 One found 91% accuracy in distinguishing between mild and severe asthma,23 one found a weighted $\kappa$ of 0.69–0.72 for agreement of asthma severity between computerized CDSS and clinical experts,24 and one study of an asthma control tool (not computer-based) found a correlation of $R^2=0.54–0.59$ with expert opinion.25 None of the computer-based studies assessed the accuracy of the CDSS in ‘real life’ use by practicing clinicians.

We developed a computerized CDSS for asthma management, based on the most recent 2007 asthma guidelines.4 The system automatically provided assessments of impairment, risk, control, and severity, and generated treatment recommendations for new patients. The system was designed in collaboration with pediatric pulmonologists and was widely used in a pediatric pulmonary clinic. In this study, we explore, in detail, the cases in which practicing pediatric pulmonologist and CDSS assessments were at variance in order to better understand computerized CDSS and clinician diagnostic capabilities.

METHODS
Setting
This study was conducted in the pediatric pulmonology clinic at Yale-New Haven Children's...
Hospital. This clinic is staffed by nine providers including five pediatric pulmonology attendings, three pediatric pulmonology fellows, and a nurse practitioner. Each year, there are approximately 1200 clinic visits for asthma in addition to visits for a variety of other respiratory diseases.

System development
A CDSS for pediatric asthma based on National Asthma Education Program Expert Panel Report 3 (EPR-3) was developed by tagging relevant sections of the guideline text as elements of the Guideline Elements Model, and then developing rules using EXTRACTOR transforms. A set of forms were then designed to be visually similar to the figures contained in EPR-3. The system was implemented in January 2009. Two pediatric pulmonologists were part of the CDSS design team and assisted with design, implementation planning, and launch of the computerized CDSS.

The EPR-3 guidelines recommend categorizing asthma severity (intermittent, mild persistent, moderate persistent, or severe persistent) for new patients and asthma control (well controlled, not well controlled, or very poorly controlled) for returning patients. Despite these distinct terminologies, however, both are categorized based on a similar assessment of level of impairment (symptoms, pulmonary function tests, and use of short-acting β agonists) and risk (exacerbations requiring oral steroids, hospitalizations, and acute/ER visits). In both cases, a particular level of severity or control is accompanied by a recommended intensity (or ‘step’) of treatment.

We designed a set of forms within the ambulatory electronic medical record (Centricity EMR/formerly ‘Logician,’ General Electric, Fairfield, CT) to capture key historical and clinical data related to impairment and risk for new patients (figure 1), and modified it slightly to capture similar data for returning patients (figure 2). Clinicians could only utilize the new patient (severity assessment) form for new patients, even if the patient already carried a diagnosis of asthma. Each form contained check boxes or radio buttons for clinicians to record patient information in a structured fashion. These data served as inputs to the computerized CDSS, which encoded the data, calculated the patient’s level of severity or control, and posted it to assist clinicians (figure 3). For new patients, a treatment step was then recommended once the asthma severity had been determined (figure 4). If the clinicians selected a different treatment step than that calculated by the computerized CDSS, they were alerted by red text showing the CDSS recommendation on the screen and were encouraged to enter the reason for the variance.

![Figure 1](https://example.com/figure1.png)
in a free text box. The CDSS did not suggest a treatment step for return visits because it was unable to translate the medication list into an existing level of treatment. Consequently, it could not suggest a revised treatment step.

**Study sample and design**

We designed a prospective cohort study of agreement with the computerized CDSS in the first year of implementation. All pediatric asthma visits occurring from January 6, 2009 to December 31, 2009 were eligible for inclusion. Visits were classified as asthma visits if patients were referred to the pediatric pulmonology clinic for evaluation or treatment of asthma. Visits in which the clinician did not enter enough information to activate the computerized CDSS were excluded from the disagreement analysis.

Cases in which there was a disagreement between the pediatric pulmonologist and the computerized CDSS relating to asthma control (for follow-up visits), severity (for initial visits), or treatment step (for initial visits) were reviewed by a primary care physician (LJH or LIH) and by the chief of pediatric pulmonology (AB-A). We reviewed all disagreements occurring during the entire year for new visits. Due to the higher volume of control disagreements, we reviewed all disagreements occurring in the first 5 months for return visits and randomly selected an additional 25 control disagreements from the second half of the year to analyze. We determined the primary reason for disagreement by reviewing the full visit record, including both free text and structured data entry. We then jointly developed a qualitative taxonomy of reasons for disagreement. The study was approved by the Yale Human Investigation Committee, which granted a waiver of signed informed consent and an Health Insurance Portability and Accountability Act (HIPAA) waiver to review the charts.

**Main measures**

Outcome measures included rate of agreement by providers with the system for asthma control (return visits only), asthma severity (new visits only), and asthma treatment (new visits only). We hypothesized that more experienced providers would be more likely to provide guideline-adherent care, and therefore recorded provider experience. Since racial disparities in asthma treatment are well documented, we also collected data on patient race.

**Statistical analysis**

We used descriptive statistics to describe the rate of record completion, the rate of agreement with CDSS, and the number of disagreements attributable to each reason for disagreement. We qualitatively described reasons for disagreement. We used paired $t$ tests to determine whether clinicians systematically over- or under-rated severity or control compared to the CDSS. Finally, we used $\chi^2$ tests to determine whether provider experience level or patient race was associated with agreement on asthma control. There were an insufficient number of new patients to perform the same analyses for asthma severity.

**RESULTS**

From January 6, 2009 to December 31, 2009, there were a total of 1199 visits for asthma care, including 1032 return visits and 167 new patient visits. A total of 568 (47.4%) of visits were seen by attending physicians, 276 (23.0%) were seen by pulmonology...
fellows, and 355 (29.6%) were seen by the nurse practitioner. A total of 573 (47.8%) of patients were white, 269 (22.4%) were black, and 269 (22.4%) were Hispanic. Female patients accounted for 471 (39.3%) visits, and the mean age of all patients was 8.1 years (SD 5.2).

Asthma control
Clinicians recorded a control assessment for 880/1032 return patients (85.3%). In 767 of these 880 visits (87.2%), they also entered enough impairment and risk information to trigger a CDSS severity assessment. Of these 767 visits, the providers agreed with the computerized CDSS 70.8% of the time (543 visits). Clinicians rated their patients as significantly better controlled than the computerized CDSS did (p < 0.0001, paired t test). This can be seen visually in table 1 which illustrates that in 188 of the 225 (83.6%) cases in which there was a disagreement, physicians assessed their patients as being better controlled than the computerized CDSS assessment.

There was no significant difference in frequency of disagreements between provider and CDSS assessments in the first and second 6 months of implementation (27.9% vs 30.6%, p = 0.41). We reviewed, in detail, all 94 disagreements that occurred during the first 5 months of the study (table 2), and 25 randomly selected charts from the second half of the study period. Nearly one-third reflected providers’ failure to follow guideline-based care: 33 cases (28%). For example, a provider categorized a patient as being well controlled, even though she noted that the patient had some limitation in normal activity, which, according to the guideline, should move the patient to the ‘not well controlled’ category. In 28/33 (85%) of these cases, there was only one factor such as cough or short acting β agonist use which led to the discrepancy between the provider and the CDSS; however, there was no consistent culprit factor. A total of 11/224 (4.9%) disagreements were accompanied by explanatory comments by the provider (ie, ‘symptoms only during viral illness’).

There were significant differences in rate of agreement with CDSS control assessment by attendings, fellows, and nurse practitioners. Those with most clinical experience (attendings) agreed with the CDSS assessment most often (78.2%), while those with least clinical experience (pulmonary fellows) agreed with the CDSS assessment least often (63.2%). The nurse practitioner’s agreement level was intermediate (66.0%); p = 0.0004. Considering only cases in which providers failed to follow guidelines, the trend was similar and still statistically significant: attending guideline deviation 1.3%, NP guideline deviation 4.7%.
deviation 3.3%, and fellow guideline deviation 4.4% (p = 0.03).
The N was too small to perform similar analyses for asthma-severity assessments or step selection. Patients’ race was not significantly associated with failure to follow guidelines for asthma control assessment (p = 0.08).

**Asthma severity**

Clinicians recorded a severity assessment for 131/167 new patients (78.4%). In 100 of these 131 visits (76.3%), they also entered enough information to trigger a computerized CDSS severity assessment. In 37 of the 100 visits with both a provider and computerized CDSS assessment (37%), providers agreed with the CDSS. On average, the computerized CDSS rated patients as having more severe asthma than providers did (p = 0.02, paired t test). Of the 63 charts in which there were disagreements, 41 clinicians assessed their patients as having less severe asthma than the CDSS did, and 22 clinicians assessed their patients as having more severe asthma than the CDSS did (table 3). There was no significant difference in rate of disagreement with the computerized CDSS in the first 6 months and second 6 months of the CDSS implementation (61.5% vs 59.1%, p = 0.47).

We reviewed all charts in which there were disagreements during the first year of implementation (N = 63). We found that disagreements were attributable to three main causes: guideline deficiencies, provider errors, and computerized CDSS inadequacies (table 2). A total of 3/63 (4.8%) disagreements were accompanied by explanatory comments by the provider (ie, ‘these are due to albuterol’).

Guideline deficiencies accounted for six cases (10% of all variances). In these cases, the guidelines themselves were sufficiently vague or absent that both CDSS and provider interpretations of severity were plausible.

Variance rates were driven by providers in 54% (21/63) of cases. In 14 cases, providers did not adhere to guideline recommendations. For instance, a provider might document a severe symptom of impairment, yet categorize the patient as having only mild or moderate asthma severity without any comment. These were categorized as non-adherence. In four cases, providers explicitly disagreed with the guidelines. For example, in one case, the provider declined to categorize a 9-month-old infant as having severe asthma because it would require higher doses of

**Table 1** Provider assessment versus computerized clinical decision-support system assessment of asthma control

<table>
<thead>
<tr>
<th>Provider assessment</th>
<th>Computerized clinical decision-support system assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Well controlled</td>
</tr>
<tr>
<td>Well controlled</td>
<td>419</td>
</tr>
<tr>
<td>Not well controlled</td>
<td>32</td>
</tr>
<tr>
<td>Very poorly controlled</td>
<td>3</td>
</tr>
</tbody>
</table>

Bold text indicates agreement.
medication than the physician was comfortable providing to such a young child. In three cases, the provider made a documentation error and inadvertently selected the wrong checkbox, in contradiction to data documented elsewhere in the note.

The remaining variances (56%) were attributable to computerized CDSS inadequacies. These included the inability to appropriately categorize asthma severity in patients who had already been started on treatment (14 cases), the inability to recognize symptoms entered as free text (six cases), the inability to distinguish symptoms caused by asthma and those caused by other illnesses such as allergic rhinitis or reflux disease (eight cases), and computerized CDSS algorithm errors, such as assigning all patients with at least two exacerbations a year to moderately severe category, even though these patients may alternatively be categorized as mild or severe depending on other criteria (eight cases).

Table 3 Provider assessment versus computerized clinical decision-support system assessment of asthma severity

<table>
<thead>
<tr>
<th>Provider assessment</th>
<th>Computerized clinical decision-support system assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermittent</td>
<td>Intermittent</td>
</tr>
<tr>
<td>Intermittent</td>
<td>1</td>
</tr>
<tr>
<td>Mild persistent</td>
<td>9</td>
</tr>
<tr>
<td>Moderate persistent</td>
<td>6</td>
</tr>
<tr>
<td>Severe persistent</td>
<td>0</td>
</tr>
</tbody>
</table>

Bold text indicates agreement.

DISCUSSION

In this postimplementation study of a CDSS to enhance asthma management in a pediatric pulmonary clinic, we found that computerized CDSS assessments were accurate compared to expert clinician review in 80% of all control assessments, 66% of all severity assessments, and 59% of all step recommendations. Practicing pediatric pulmonologists failed to strictly follow guideline recommendations in 8% of return visits and 18% of new patients.

The reasons providers and the computerized CDSS disagreed were quite different for assessments of control versus severity. The majority of control variances were caused by providers attributing asthma-like symptoms such as cough to other conditions such as allergic rhinitis, gastrointestinal reflux, or acute upper-respiratory infection. Since the computerized CDSS was designed always to treat these as asthma-related, it tended to assess patients as being less well controlled than providers thought they were. By contrast, many disagreements about asthma severity were caused by the fact that many of the ‘new’ patients arriving for subspecialty consultation had in fact already been diagnosed and treated. According to the EPR-5 guidelines, severity assessment in these patients should take into consideration existing medications. This was not feasible for the CDSS and therefore led to additional errors.

Analysis of variances between computerized CDSS assessments and clinician assessments can provide insights into areas for improvement in decision-support design. Our taxonomy of
reasons for discrepancies between provider and computerized CDSS assessments identified some obvious gaps in the design of the CDSS that could be quickly remedied to improve its assessment capabilities. For instance, modifying impairment questions to ensure they are asthma-specific (ie, changing ‘any cough symptoms’ to ‘any asthma-related cough symptoms’) would immediately eliminate half the disagreements about control. It is notable that these flaws existed in the system, despite integral involvement by practicing pediatric pulmonologists from the start of the design phase, as recommended by informatics experts.28 Our study thus demonstrates the critical importance of carefully analyzing the reasons for practicing clinician disagreements with decision support in the post-implementation period in order to improve design and effectiveness. Unfortunately, although the value of postimplementation audits is well recognized,29 30 published reports of computerized CDSS interventions typically lack such information.

Our analysis of disagreements between the computerized CDSS and clinician assessments also gave us insight into clinical care provision. About a third of disagreements about severity or control were attributable to provider errors. Other studies have shown that computerized CDSSs may be more accurate than clinicians.23 24 Although guidelines may legitimately be disregarded in certain clinical contexts, it is notable that those who were most often at variance with the computerized CDSS were the least clinically experienced providers: pulmonary fellows. Consequently, we believe it likely that these events represent true deviations from appropriate care, demonstrating that even clinical experts may derive some benefit from computerized CDSSs. Furthermore, it is impossible to determine how many assessments were altered in real-time by providers who noted the computerized CDSS recommendations and may therefore have improved their care.

Lastly, analysis of disagreements between computerized CDSS and clinicians also gives us an insight into CDSS capabilities. Most computerized CDSSs provide one-step alerts or guidance based on relatively simple rules (ie, suggest pharmacotherapy for LDL above goal if diabetes is in the problem list).21 31 These are typically modestly effective.32 The computerized CDSS in this study, by contrast, was designed to perform much more cognitively rich work—to determine, based on a large variety of inputs, a patient’s acuity of illness, and then, based on that assessment, to suggest a tailored treatment regimen. Systematic reviews show that CDSSs are much less successful for this type of activity than for simpler activities such as preventive care or drug dosing.20 In this context, the relatively high accuracy of the computerized CDSS asthma control assessment (80%) is a notable achievement, although the system did not perform as well for new patients, and even 80% accuracy may not be sufficiently high for widespread use or acceptance.

The strengths of this study lie in our ability to determine precisely what information providers were using to determine asthma severity and control, and therefore to pinpoint reasons for disagreement with a computerized CDSS following guideline protocols. However, this study does have some limitations. We relied on chart notes to determine practitioners’ clinical reasoning. We do not know if providers chose to discount certain information, if they preferentially weighted certain information in a different fashion than the guidelines suggest, or if they inadvertently checked an incorrect box on the clinical decision support screen. Furthermore, in most cases we were not able to assess whether these clinical experts deliberately disagreed with the guidelines, or were simply not aware of the guideline-recommended practice. However, all cases in which a provider did not follow guidelines were reviewed by an expert clinician to determine whether the difference was appropriate (an ambiguous guideline or a reasonable disagreement) or a lapse in recommended care. We do not know how often providers changed their assessments to become concordant with the guidelines once they viewed the CDSS assessments. We did not assess patient outcomes. Finally, we performed this study in a specialty clinic, and the results may not be generalizable to a primary care setting. It is possible that the rate of disagreement would be lower in a setting in which the clinicians were not content experts.

We have used the results of this analysis to substantively alter the next iteration of the decision-support system, which is aimed at primary care physicians. We have changed the wording of the cough assessment to clarify that it refers only to asthma-related cough. We have added specific questions about adherence, inhaler technique, and environmental controls. Unless these are all recorded as appropriate, the computerized CDSS will not provide advice about treatment step. This eliminates the problem of the computerized CDSS recommending a higher step of therapy for patients who would probably respond to existing therapy if their inhaler technique were improved. For patients who are not well-controlled, we have added additional questions about alternative diagnoses or psychosocial factors. Finally, we have improved the computerized CDSS so that it can now identify existing treatments and recommend management for return patients as well as new patients. We expect these alterations to reduce computerized CDSS errors and increase the CDSS face validity and utility for practitioners.

In conclusion, we found that a CDSS designed to assess and manage pediatric asthma patients in a pediatric pulmonology practice performed well for return visits, with providers both entering data appropriately and agreeing with most of its assessments. We further found that 8% of return visits and 18% of new visits to an academic pediatric pulmonology practice did not conform to guideline-based practice, suggesting that even expert clinicians may benefit from clinical decision support. Finally, examining cases in which pulmonologists did not agree with the computerized CDSS proved to be a valuable method both of identifying guideline-deviant care and of improving the CDSS itself. This is an evaluative step that should be undertaken after implementation of complex decision-support systems.

Funding This work was supported by contract HHSAG290200810011 from the Agency for Healthcare Research and Quality and grant T15-LM07065 from the National Library of Medicine. At the time this study was conducted, LH was supported by the CTSA Grant UL1 RR024139 and KL2 RR024138 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health, and National Institutes of Health roadmap for Medical Research.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES


