

Research Paper ■

Implementation of Clinical Guidelines via a Computer Charting System:

Effect on the Care of Febrile Children Less than Three Years of Age

DAVID L. SCHRIGER, MD, MPH, LARRY J. BARAFF, MD, KELLY BULLER, MD, MPH, MANALI AYATCHIT SHENDRIKAR, MPH, SAMEER NAGDA, EDWARD J. LIN, VLADISLAV J. MIKULICH, SHAN CRETIN, PHD, MPH

Abstract **Objective:** The authors have shown that clinical guidelines embedded in an electronic medical record improved the quality, while lowering the cost, of care for health care workers who incurred occupational exposures to body fluid. They seek to determine whether this system has similar effects on the emergency department care of young children with febrile illness.

Design: Off-on-off, interrupted time series with intent-to-treat analysis.

Setting: University hospital emergency department.

Subjects: 830 febrile children less than 3 years of age and the physicians who treated them.

Interventions: Implementation of an electronic medical record that provides real-time advice regarding the content of the history and physical examination and recommendations regarding laboratory testing, treatment, diagnosis, and disposition.

Measurements: Documentation of essential items in the medical record and after-care instructions; compliance with guidelines regarding testing, treatment, and diagnosis; charges.

Results: The computer was used in 64 percent of eligible cases. Mean percentage documentation of 21 essential history and physical examination items increased from 80 percent during the baseline period to 92 percent in the intervention phase (13 percent increase; 95 percent CI, 10–15 percent). Mean percentage documentation of ten items in the after-care instructions increased from 48 percent at baseline to 81 percent during the intervention phase (33 percent increase; 95 percent confidence interval, 28–38 percent). All documentation decreased to baseline when the computer system was removed. There were no demonstrable improvements in appropriateness of care, nor was there evidence that appropriateness worsened. Mean charges were not changed by the intervention.

Conclusion: The intervention markedly improved documentation, had little effect on the appropriateness of the process of care, and had no effect on charges. Results for the febrile child module differ from those for the module for occupational blood and body fluid exposure (a more focused and straightforward medical condition), underscoring the need for implementation methods to be tailored to specific clinical complaints.

■ JAMIA. 2000;7:186–195.

Affiliation of the authors: University of California–Los Angeles School of Medicine, Los Angeles, California.

This work was supported in part by grant HSO6284 from the Agency for Health Care Policy and Research and by an unrestricted grant to Dr. Schriger for health services research from the MedAmerica Corporation.

Correspondence and reprints: David Schriger, MD, MPH, 924 Westwood Boulevard, Suite 300, Los Angeles, CA 90024-2924; e-mail: <schriger@ucla.edu>.

Received for publication: 7/14/99; accepted for publication: 11/2/99.

It has been repeatedly documented that guidelines must be actively implemented if they are to modify patient care.¹ Computers have been established as one way to improve physician performance, yet results are mixed.^{2,3} In contrast to the numerous studies that have evaluated the utility of computers in the execution of a single task (e.g., using the laboratory, taking preventive health care measures), few studies have evaluated the computer implementation of a comprehensive guideline. The difficulties in producing implementation software from broad guideline documents has been noted.⁴

We previously reported that clinical guidelines embedded in an electronic medical record (EMR) markedly improved documentation and appropriateness of testing and treatment decisions while reducing charges by 28 percent for emergency department patients treated for occupational exposure to blood or body fluids.⁵ We now report on the use of the same system in the emergency department care of febrile children less than 3 years of age.

This presenting complaint is one of five modules developed for the EDECS (Emergency Department Expert Charting System) project. We developed each module with the goal of improving the quality and cost-effectiveness of care. We developed a module for the care of the febrile children because: 1) improved quality of care and medical record documentation were highly desirable goals in their own right, and this clinical problem is the second most common cause of malpractice complaints in emergency medicine, 2) we wanted to standardize the management of children with fever without a source, and 3) we wanted to include a pediatric condition in our evaluation of EDECS to determine whether any unique concerns arose when the system was used in the care of children.

The fundamental principle of the EDECS system is that if guidelines are to have impact, they must be integrated into functions that physicians find useful in the routine care of patients. The febrile child module was designed to provide evidence-informed advice regarding the content of the history and physical examination, use of laboratory testing, interpretation of laboratory results, administration of medications, diagnosis, disposition, and content of after-care instructions for patients sent home. We hypothesized that this guideline-embedded EMR would produce at least a 20 percent absolute improvement in documentation in the medical record and after-care instructions, and similar improvements in the appropriateness of testing, treatment, and selected diagnoses.

Methods

Setting and Patients

The study was conducted at the UCLA Emergency Medicine Department from 1992 to 1995. Of the 37,000 patients seen in the UCLA Emergency Medicine Department each year, 5,500 (15 percent) are children, and 1,400 of these children are less than 3 years of age and present with a febrile illness. Children less than 3 years of age were eligible for the study if they presented with febrile illness or were noted to have a temperature of 38°C or higher at triage. Patients were excluded if they had an underlying condition or disease likely to alter the management of their acute febrile illness (e.g., organ transplant, immunosuppression, congenital heart disease, renal failure); if respiratory distress (acute bronchospasm or stridor) was the primary reason for the visit; or if the patient presented for a scheduled recheck. In our department, about 1,100 children meet these eligibility requirements each year. The majority of children are cared for by 1 of 36 emergency medicine residents. Before 1993, residents in pediatrics were assigned to the emergency department during certain hours of the day. From 1993 to 1995, pediatric residents were available more sporadically. All house staff were supervised by attending physicians who were board-certified in emergency medicine. The attending physician typically functioned in a supervisory role, permitting the resident to operate independently so long as there was no substantive disagreement about the proper course of action. While there was no formal attempt to ensure that the attending faculty agreed and complied with the guideline, the faculty were well aware of our publications in this area and did not express specific disagreement.

Parents of patients were not asked to provide informed consent, since the intervention sought only to improve documentation and compliance with existing departmental practice regarding the management of febrile children. The exemption from informed consent was deemed acceptable by the study section of the Agency for Health Care Policy and Research that funded the project, and exemption forms were filed with the UCLA Institutional Review Board.

Guideline Development

In anticipation of this project, we reviewed the literature and compiled existing guidelines addressing the management of children with fever. Following standard guideline development methods,⁶⁻⁸ we created guidelines for various aspects of the management of

```

:Assign      N=1
and :Test    fever >=38 & age < 90
and :Assign  r$="Temperature >= 38 and age < 90 days"
and :Fail

Or :Test    age > 90
and :Test   Start("B", patientRace$) & hx sickle<>"No"
and :Test   fever > 39
and :Test   fws$="Yes"|diag{15}, desc$<>"
and :Assign r$=:r$+"Black pt. with SSD/?SSD at risk for occult bacteremia"
and :Fail

[... more rules like these]

Or :Test    r$<>"
and :Assign labs{N}.recom$="14"
and :Assign labs{N}.rationale$=:r$

Or :Succeed

```

Figure 1 Section of the rule for ordering a complete blood count (lab #1). The "Fail" forces the software to consider each indication. The rationale is captured in the "r\$" variable, and the screen color for CBC is set to red (value 14) in the "labs{}recom" array.

febrile children and published baseline analyses and a clinical guideline for the care of children with fever without a source.⁹⁻¹² Rules for other aspects of the care of pediatric patients were derived from standard textbooks.

Software Development and Operation

We programmed the EDECS in the rule-based expert-system shell Applications Manager (Intelligent Environments, Tewksbury, Massachusetts) using Database Manager (the database packaged with OS-2) as the database and OS-2 as the operating system (both IBM). The Applications Manager program is an expert system shell/client-server application development environment that uses the proprietary Universal Rules Language for defining the medical logic. Separate rule modules were made for each of the program's functions (checking adequacy of documentation, selecting tests, selecting treatments, choosing diagnoses, selecting diagnoses, and preparing after-care instructions) (Figure 1). All modules are run every time the physician updates a screen. Details of programming goals and methods have been described elsewhere.^{5,13}

The EDECS program begins with separate screens for history of present illness, past medical history, and physical examination. Each screen presents color-coded items, with essential items in red. Essential items must be answered before the history and physical are deemed complete; however, each item has an "Unknown" option so that clinicians are not forced to answer incorrectly when information is unavailable. Data entry is accomplished primarily with the mouse: the user selects from pull-down boxes and pick lists. After completing the history, the physician reviews screens for laboratory testing and treatment. Color-

coding indicates whether each commonly used laboratory test or treatment is recommended, optional, or not recommended for that patient. While all tests and treatments available at UCLA can be ordered through EDECS, only those commonly indicated in the evaluation and treatment of febrile children are visible on the main screens.

When the physician selects the "Rationale" button, the reasoning for each recommendation is displayed. When a physician deviates from a recommendation, EDECS requests an explanation. Once initial testing and treatment are completed, further advice regarding the need for additional testing and treatment, the need for re-evaluation, and the disposition is provided. For patients who are to be sent home, the program prepares a proposed set of after-care instructions that the physician may modify. The program then prints any prescriptions, the after-care instructions, and the medical record and stores selected variables in the permanent database.

The Experiment

This prospective off-on-off interrupted time series experiment was conducted in three phases. During the baseline phase (Phase 1, May 1992 to Dec 1993), care proceeded in the usual way with handwritten medical records. During the intervention phase (Phase 2, Nov 1994 to Mar 1995), physicians were asked to use EDECS when treating febrile children. During the postintervention control phase (Phase 3, Jul 1995 to Dec 1995), the EMR was removed and handwritten charts were again used exclusively.

During Phase 2, each physician on rotation in the emergency department was informed of the experiment, consented (all agreed), and was given a 15-minute orientation, solely regarding technical issues related to operating the software. Eligible patients were identified by the triage nurse who flagged the chart and attached a "Febrile child data collection form." This form provided the physician with a list of the history and physical examination items that were required for all patients. The form was optional, was not part of the medical record, and was included so that in the absence of bedside computing the physician would not have to go back to the bedside every time the computer asked a question for which the data had not been collected. After completing the history and physical examination, the physician would leave the bedside and, using the EDECS computer in the charting area of the emergency department, enter the history and physical examination results, and order tests and treatments.

Dependent Variables

The primary outcome measures for this study were the quality of documentation of the medical record and after-care instructions, the appropriateness of testing and treatment decisions and diagnoses, the percentage of testing and treatment charges attributable to indicated activities, and the per-patient charges for each visit. We characterized the quality of documentation of the medical record by determining (for each of the 21 essential items) what percentage of charts had the item, and calculated an overall documentation score by averaging the percentage documentation across the 21 items, giving each item equal weight. Essential items are those elements needed to negotiate the process-of-care rules specified in the guidelines for all patients.

Appropriateness of testing and treatment were calculated as the number of appropriate decisions divided by the total number of decisions.⁵ For example, each child was scored as needing or not needing ceftriaxone, on the basis of explicit rules established before the onset of the experiment. The actual ceftriaxone decision was deemed appropriate when the action matched the indication (indicated–given or not indicated–not given). We determined the appropriateness of decisions regarding the diagnosis of otitis media in a similar manner by determining whether the documented findings on physical examination met explicit criteria warranting this diagnosis and comparing this to whether the diagnosis was made. It should be noted that the “appropriateness of otitis media diagnosis” is affected by both the quality of documentation and the quality of the decision making and does not include reference to an independent gold standard. Charge data were adjusted so that each item retained its initial charge throughout the experiment, regardless of whether a price change occurred. Charges were classified as physician/facility, laboratory, or treatment.

Data Collection, Power Calculations, and Statistical Methods

Throughout the experiment the emergency department log was checked every other day to identify eligible subjects. Handwritten charts for all subjects not seen with EDECS in Phase 2 (the intervention phase), a 50 percent random sample of charts from Phase 1, and a 20 percent random sample of charts from Phase 3 were abstracted onto a 175-item form. We sampled for two reasons—to extend each study period for a longer time, thereby limiting vulnerability to seasonal variation; and to minimize the labor required for chart

abstraction. We included all Phase 2 charts, since the majority were already in electronic format and did not require abstraction.

Abstractors were trained and tested on standardized cases until they demonstrated an error rate consistently less than 2 percent, and inter-rater reliability was checked periodically. Data on laboratory results and charges were downloaded from the hospital mainframe computer and combined with data from the EMR and data from the chart abstractions. As a further control for seasonal variation, case complexity was established using the patient’s age, the triage nurse’s intake note, and the first sentence of the physician note, to categorize the patient into one of three categories (simple, intermediate, or complex).*

From pilot studies, we expected 50 percent compliance with documentation rules and testing and treatment guidelines in the control phase, and designed the experiment to have at least 90 percent power to detect an increase to 80 percent, assuming independence among observations.¹⁴ We used the patient as the unit of analysis, because most physicians participated in only one phase of the study. We accounted for the potential nonindependence of patient observations resulting from clustering on physician by using Huber adjustments to the logit and regression procedures. For categoric variables (percentage documentation, percentage compliance, appropriateness of diagnosis), we examined hypotheses regarding among-phase differences using the logit procedure (at times including variables for potential confounders), and calculated 95 percent confidence intervals (CIs) for between-phase differences using linear regression.¹¹ Linear regression was used for statistical testing of charges using raw and log-transformed charges. Summary statistics were created by averaging the individual items in each scale and performing regressions on these averages.

Results

We abstracted 352 charts from the 718 eligible subjects in the initial control period (Phase 1) and 104 charts from the 538 eligible subjects in the postintervention control period (Phase 3) (Table 1). By definition, all of these charts were handwritten. The EDECS system was used for 239 (64 percent) of the 374 patients cared for during the intervention period (Phase 2); the other 135 charts were handwritten. One hundred eighty-five different physicians each had primary responsibility

*A copy of the algorithm is available from the author.

Table 1 ■

Demographic and Clinical Features of Events, Patients, and Physicians, by Phase of Study

| Type | Phase 1 | Phase 2 | Phase 3 |
|--|-------------|--------------|-------------|
| | Control | Intervention | Control |
| No. of patients | 718 | 374 | 538 |
| No. analyzed* | 352 | 374 | 104 |
| Patients: | | | |
| Mean age (SD) in years | 1.12 (0.70) | 1.31 (0.69) | 1.31 (0.78) |
| Female | 47% | 48% | 45% |
| Complexity: | | | |
| Simple (1) | 61% | 55% | 57% |
| Intermediate (2) | 32% | 37% | 22% |
| Complex (3) | 7% | 8% | 21% |
| Average mean (SD) | 1.46 (0.63) | 1.52 (0.63) | 1.64 (0.81) |
| Physicians (N = 185): | | | |
| No. (%) of cases seen solely by attending | 6 (2) | 8 (2) | 2 (2) |
| Mean (SD) postgraduate year of resident physicians | 1.91 (0.81) | 2.73 (0.99) | 2.79 (0.98) |
| Specialty of doctors who saw at least one patient in the phase | | | |
| Emergency physician | 39% | 57% | 69% |
| Other | 61% | 43% | 31% |
| Patients seen by: | | | |
| Emergency physicians | 39% | 83% | 85% |
| All other physicians | 61% | 17% | 15% |

*These numbers represent 50 percent random sample of Phase 1 and a 20 percent random sample of Phase 3.

for at least one of the 830 patients in the study group. One hundred forty-one physicians participated in one phase of the study, 35 physicians in the intervention phase and one of the two control phases, and 9 physicians in all three phases. Fifty percent of physicians saw one or two study patients; the average was 4.5 patients per physician, and the range was 1 to 32. Emergency medicine residents were the primary providers in 62 percent of all cases, pediatric residents in 30 percent, family medicine residents in 5 percent, and emergency department faculty in 3 percent. Pediatric interns and residents saw many more patients in Phase 1 (55 percent of all Phase 1 cases) than in the other phases, when emergency department residents saw 85 percent of patients. The average postgraduate year (PGY) increased throughout the study, because emergency medicine trainees are PGY 2–4, whereas pediatric trainees are PGY 1–3. Patient age and gender distributions were similar among phases; however, Phase 3 cases were of higher complexity than those of Phases 1 and 2.

Seventy-six physicians cared for at least one of the 374 eligible patients during the intervention phase; the majority saw fewer than three patients (Figure 2). Thirty-two percent of these 76 physicians never used

EDECS, and 26 percent used it in every case (Figure 3). The cases of the 25 percent of physicians who each saw more than five cases accounted for 64 percent of all cases seen with EDECS. This observation and the U-shaped distribution of the results suggest that EDECS use was not random; some physicians gravitated toward it while others avoided it. Emergency physicians used EDECS in 68 percent of eligible cases, whereas others used EDECS in 41 percent of eligible cases (difference, 27 percent; 95 percent CI, 14–40 per-

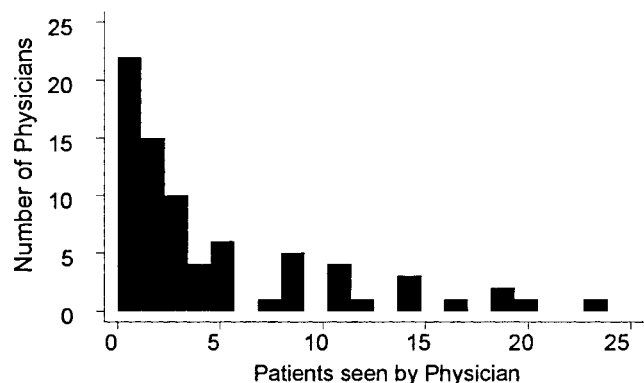


Figure 2 Number of patients seen by each physician during the intervention phase (Phase 2). The total number of patients was 374.

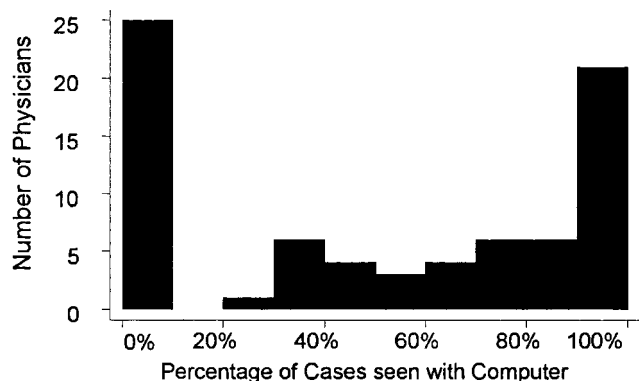


Figure 3 Number of physicians using EDECS in each indicated percentage of cases.

cent). Seven of 44 (16 percent) emergency physicians and 17 of 32 (53 percent) other physicians (difference, 37 percent; 95 percent CI, 17–58 percent) never used EDECS.

Documentation was higher in Phase 2 than in Phase 1 for 20 of 21 essential history and physical examination items, and was higher in Phase 2 than in Phase 3 for all 21 items (Table 2). Overall documentation of

the 21 items was 80 and 74 percent for control phases 1 and 3, respectively, and 92 percent for the intervention phase (Phase 2–3 difference, 13 percent; 95 percent CI, 10–15 percent). Documentation of handwritten charts was fairly constant across phases (80, 78, and 74 percent, respectively), and the combination of 78 percent documentation on the 135 handwritten Phase 2 charts and 100 percent documentation on the 239 EDECS charts produced the overall documentation percentage (92 percent). Thus, the use of EDECS in just over two thirds of eligible cases was sufficient to produce significant improvements in overall documentation for Phase 2.

While EDECS stimulated the doctors to create more complete charts, it did not always change behavior in ways that reliably produced a more complete physical examination. For example, the percentage of ear examinations that included insufflation (blowing a puff or air toward the ear drum to see if it moves) was 15 percent in Phase 1, 22 percent in Phase 2 (10 percent on handwritten charts, 30 percent on EDECS charts), and 5 percent in Phase 3 (Phase 1–2 difference, 8 percent; 95 percent CI, 2–14 percent). Since EDECS en-

Table 2 ■

Documentation of Medical History and Physical Examination Items

| | % of Patients | | | Difference in % (95% CI) | |
|------------------------------|----------------------|----------------------|----------------------|--------------------------|---------------------|
| | Phase 1 (N = 352) | Phase 2 (N = 374) | Phase 3 (N = 104) | Phases 1 vs. 2 | Phases 1 vs. 3 |
| Medical history: | | | | | |
| Allergies | 94 | 100 | 91 | 6 (4 to 8) | -3 (-8 to 3) |
| Days ill | 93 | 98 | 91 | 5 (2 to 8) | -2 (-7 to 3) |
| Highest temperature | 66 | 83 | 64 | 17 (9 to 24) | -2 (-14 to 10) |
| Behavior | 32 | 77 | 32 | 45 (35 to 54) | -1 (-13 to 12) |
| Fluid intake | 64 | 85 | 44 | 21 (13 to 28) | -20 (-31 to -10) |
| Cough | 72 | 91 | 72 | 19 (13 to 25) | 0 (-11 to 11) |
| Vomiting | 84 | 90 | 81 | 7 (1 to 12) | -3 (-12 to 7) |
| Diarrhea | 79 | 90 | 74 | 10 (4 to 17) | -5 (-15 to 4) |
| Antipyretic use | 96 | 98 | 92 | 2 (-1 to 5) | -4 (-9 to 1) |
| Has primary care MD | 54 | 85 | 38 | 31 (24 to 38) | -15 (-28 to -2) |
| Physical examination: | | | | | |
| Temperature | 100 | 98 | 92 | -1 (-3 to 0) | -7 (-13 to -2) |
| Pulse | 95 | 99 | 92 | 4 (2 to 6) | -3 (-9 to 3) |
| Respirations | 85 | 96 | 80 | 11 (7 to 15) | -5 (-14 to 3) |
| Weight | 93 | 98 | 97 | 5 (2 to 8) | 4 (0 to 8) |
| Hydration status | 42 | 81 | 28 | 39 (31 to 47) | -14 (-27 to -2) |
| Behavior | 89 | 95 | 89 | 7 (2 to 11) | 1 (-8 to 9) |
| Ears | 84 | 88 | 67 | 4 (-2 to 10) | -16 (-27 to 5) |
| Throat/oropharynx | 87 | 92 | 85 | 5 (0 to 11) | -2 (-11 to 6) |
| Neck | 55 | 86 | 49 | 30 (22 to 39) | -6 (-19 to 7) |
| Lungs | 98 | 99 | 97 | 1 (-1 to 3) | -1 (-4 to 3) |
| Abdomen | 97 | 98 | 94 | 1 (-1 to 4) | -3 (-7 to 2) |
| Overall (95% CI) | 80 (78 to 81) | 92 (90 to 94) | 74 (72 to 77) | 13 (10 to 15) | -5 (-8 to 2) |

Table 3 ■

Documentation of After-care Instructions

| | % of Patients | | | Difference in % (95% CI) | |
|--|---------------|---------------|---------------|--------------------------|------------------|
| | Phase 1 | Phase 2 | Phase 3 | Phases 1 vs. 2 | Phases 1 vs. 3 |
| Antipyretics* | 66 | 85 | 58 | 20 (11 to 28) | -8 (-20 to 5) |
| Antipyretic dosing* | 44 | 78 | 20 | 35 (26 to 44) | -24 (-34 to -13) |
| Lethargy* | 28 | 77 | 37 | 48 (39 to 58) | 9 (-5 to 22) |
| Not eating/drinking* | 38 | 78 | 67 | 41 (30 to 51) | 30 (16 to 43) |
| No urine output* | 8 | 69 | 7 | 61 (53 to 70) | 0 (-6 to 6) |
| Fever past 2 days* | 52 | 82 | 57 | 30 (21 to 38) | 5 (-9 to 18) |
| Antibiotic name† | 81 | 89 | 78 | 9 (0 to 18) | -3 (-19 to 13) |
| Antibiotic dosing† | 54 | 80 | 31 | 26 (14 to 37) | -24 (-43 to -5) |
| Antibiotic reasoning/hints/side effects† | 4 | 65 | 6 | 61 (51 to 71) | 2 (-6 to 10) |
| Oral rehydration‡ | 58 | 93 | 72 | 35 (26 to 44) | 14 (-3 to 30) |
| Overall % (95% CI) | 48 (45 to 51) | 81 (76 to 86) | 50 (48 to 52) | 33 (28 to 38) | 2 (-2 to 5) |

*For Phase 1, $n = 344$; Phase 2, $n = 358$; Phase 3, $n = 95$.

†For Phase 1, $n = 225$; Phase 2, $n = 199$; Phase 3, $n = 36$.

‡For Phase 1, $n = 184$; Phase 2, $n = 311$; Phase 3, $n = 53$.

couraged (but did not mandate) the use of this technique, the three-fold improvement in compliance is notable; it demonstrates the power of real-time computer reminder systems. The failure to approach 100% compliance, however, highlights the difficulty of achieving universal, voluntary implementation of a practice guideline.

Percentage documentation of each desirable item in the after-care instructions was significantly higher in Phase 2 than in the other phases (Table 3). Overall documentation for the ten items was 48 percent in Phase 1, 81 percent in Phase 2, and 50 percent in Phase 3 (Phase 1–2 difference, 33 percent; 95 percent CI, 28–38 percent). The percentage documentation in handwritten after-care instructions remained constant across the three phases (48, 48, and 50 percent, respectively), and the 81 percent documentation in the intervention phase results from a blending of 100 percent documentation on 229 computer after-care instructions with 48 percent documentation on 129 handwritten charts from this phase. Covariates used in the logistic regressions included physician specialty, years of training, case complexity, and a dummy variable indicating whether the treating physician contributed six or more cases to the study. The inclusion of any combination of these did not change the magnitude and significance of findings regarding documentation of history, physical examination, and after-care instructions.

Otitis media was the sole diagnosis in 24 percent of cases; viral syndrome was diagnosed in 52 percent of cases. Sepsis, fever without source, and all other di-

agnoses each accounted for less than 5 percent of cases, precluding any meaningful analysis of the effect of EDECS on these conditions. The percentage of cases diagnosed as otitis media declined from Phase 1 to Phase 2 and from Phase 2 to Phase 3. Patients seen with EDECS were less likely to be given this diagnosis (odds ratio, 0.38; 95 percent CI, 0.19–0.75), even after adjustments were made for phase of the study (secular trend), physician specialty, case complexity, and number of cases seen by physician (Figure 4). The guideline and the EDECS program required physicians to document at least two abnormalities (of four separate ear examination variables) to justify the di-

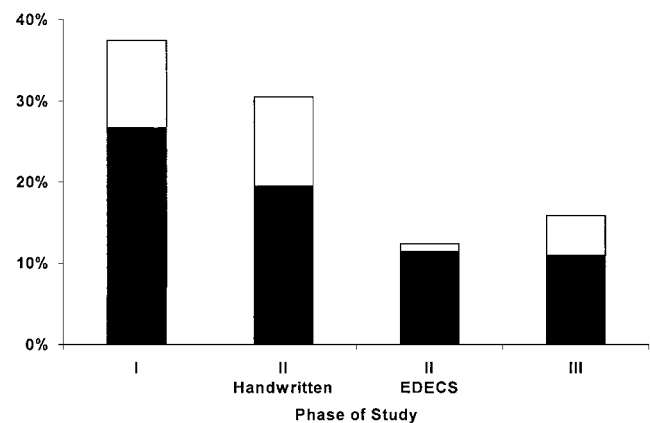


Figure 4 Percentage of cases diagnosed as otitis media by physical examination findings. The black portion of each column indicates diagnoses supported by physical findings; the white portion indicates diagnoses not supported by physical findings.

agnosis of otitis media. The reduction in otitis media diagnoses when the EMR was used was predominantly due to a decrease in the number of cases of unsubstantiated otitis media diagnoses. In only two cases did physicians override the EDECS suggestion and make the diagnosis of otitis media in the absence of two findings. Use of EDECS increased the frequency of viral syndrome diagnoses; the magnitude of the increase mirrored the decrease in the frequency of otitis media diagnoses.

Across all patients, the use of diagnostic tests was deemed appropriate in the following percentage of cases: complete blood count, 84 percent; chest radiography, 70 percent; blood culture, 83 percent; urinalysis, 81 percent; urine culture, 84 percent; and lumbar puncture, 86 percent. There were no important differences in utilization or appropriateness of diagnostic tests among phases or between EDECS and handwritten charts in Phase 2. The decision to use or withhold oral antibiotics was made appropriately in 96 percent of all cases, assuming the final diagnosis was correct. Fifty-eight percent of children were discharged home on oral antibiotics.

Intramuscular ceftriaxone, which at the time of the study was indicated only in fever without a source, was given to 10 percent of the patients discharged home. In only 35 (44 percent) of the 80 administrations was there evidence of this diagnosis. There was, however, no difference in rate or appropriateness of use among phases or between handwritten and EDECS charts.

Median total charges were similar in the three phases (\$216, \$216, and \$222, respectively). Mean phase 3 charges (\$635) were higher than charges in the other phases (Phase 1, \$357; Phase 2, \$387); however, this difference disappeared when charges were adjusted for patient complexity. The percentage of charges attributed to physician, facility, laboratory, and treatment remained consistent across phases, as did the appropriateness of these charges regardless of whether charges or log-charges were used.

Discussion

The EDECS febrile child module improved documentation of the medical record and after-care instructions, reaffirming our experience with the EDECS occupational exposure to blood and body fluids (OEBBF) and low-back pain modules.^{5,16} This is not surprising. Physicians using a system that prompts them in real time would be expected to outperform physicians relying solely on their memory.¹⁷ Feedback

has also been successfully used to improve documentation of the care of febrile children, but we believe that prompting in real time is more efficient and is more likely to achieve near-100 percent compliance.¹⁸ Logic suggests that better documentation of essential information is a prerequisite for purposeful improvement of the quality of care. Without this documentation, it is impossible to identify areas in need of improvement, and evidence suggests that those who produce better documentation provide better care.¹⁹

In contrast to the OEBBF module, however, the EDECS febrile child module did not produce substantial or significant changes in physician test ordering or treatment decisions. The guideline development effort focused on the testing and treatment of children with fever without a source, but only 36 children (4.3 percent of the study group) were in this category. The analysis of testing and treatment decisions was further complicated by poor documentation in the handwritten phases, uncertainty regarding the validity of physical findings (Was the "red" ear really red?), and uncertainty regarding the validity of diagnoses (Did the child diagnosed with "otitis media" really have this condition?). We underestimated the magnitude of these problems, and investigators in this area would be wise to heed them in planning future efforts. It is likely that multicenter evaluations will be needed to garner the number of subjects needed to evaluate the less frequently used rules in the program. It is also likely that confirmation of findings and diagnoses (either through second examinations or gold-standard tests) will be needed to prove that the cases in the intervention and control phases are truly equivalent.

The failure of our intervention to modify test and treatment ordering behavior should not be attributed solely to problems with diagnostic classification and low power. It is likely that the willingness of physicians to follow the guidelines is inversely proportional to the degree of controversy.^{1,20} The fever-without-a-source section of the EDECS febrile child module was based on the published results of evidence-informed expert panel processes, and other EDECS rules drew on published guidelines and recommendations.¹² Nevertheless, many disagree with these guidelines,²¹⁻²⁴ and it is likely that not all EDECS users agreed with them. These factors may also contribute to patterns regarding the uses of EDECS. The EMR was used in 96 percent of eligible OEBBF cases, 79 percent of eligible low-back pain cases, and 64 percent of eligible febrile child cases, suggesting that users perceived that they were less in need of assistance in the treatment of febrile children or did not like the advice that was provided.^{5,17}

There was evidence, however, that this module did have positive effects on care. Otitis media was diagnosed less often and viral syndrome more often when EDECS was used. This translates into a lower use of antibiotics, which many would consider desirable. Furthermore, when physicians used EDECS to diagnose otitis media, the documented physical examination findings justified the diagnoses in more than 92 percent of cases, an improvement over the 70 percent justified cases in the handwritten charts. At best, this finding represents an improvement in care; at worst, an improvement in documentation that facilitates quality improvement.

Average charges were unchanged by the intervention, an expected result since there was no demonstrable change in ordering patterns and consideration of cost was not part of the guideline development process. We are currently building a Web site that will make the OEBOF program available to all medical providers, so that all OEBOF patients can be treated according to continuously updated Centers for Disease Control and Prevention guidelines. Before a similar undertaking can occur for febrile children, we will need to further build consensus regarding the best approach to fever without a source and gain more understanding of what is required to get physicians to comply with guidelines for its treatment.

This study's internal validity could have been affected by temporal confounding, bias from differential documentation among phases, and errors in abstracting the handwritten charts. It could also be argued that the correct unit of analysis should have been either the physician (since the target of the intervention is the physician's behavior) or the site (if the intervention is viewed as an attempt to fix "the system"). Using the site as the unit of analysis is impractical and seldom done. Furthermore, since 76 percent of physicians participated in only one phase of the study, it is difficult to deem the physician the proper unit of analysis. Instead, we used the patient as the unit of analysis but performed robust Huber regression to adjust for any clustering on physician. The return to baseline in the second control phase, and the similarity of these results to those in the OEBOF experiment, provide further evidence that observed differences were not random and were due to the intervention.

It might also be argued that the active ingredient in the intervention was the paper prompt that physicians carried into the room rather than the software. The investigation by Adams et al.²⁵ of the "key" elements in de Dombal's abdominal pain diagnostic system revealed that an unaided accuracy rate of 47 percent rose to 59 percent with structured forms and to 72

percent with structured forms and real-time computer assistance. Our OEBOF study confirmed this finding, whereas the current study neither supports nor refutes it.⁵ External validity is affected by the use of a single university hospital as the setting and the use of house staff, who may feel pressured by the faculty-investigators to use the computer, as the primary users.

These data show that computer-assisted medical care does not provide homogeneous results across heterogeneous complaints. Changing physician behavior is a complex process, and specific circumstances of practice setting, clinical knowledge and prevalent belief, pace of practice, and characteristics of providers must be acknowledged and accommodated if an intervention is to successfully change behavior.²⁶ We cannot assume that a single computer intervention will effectively handle all complaints in all settings, and local modification will be an important part of any implementation. Although there is strong evidence that computer-assisted medical decision making is an effective way of changing behavior, many computerized guidelines will need to be formally evaluated before we can prospectively identify the optimal implementation strategy for a given problem.

The authors thank Mostafa Hassanvand (EDECS programming), William H. Rogers, PhD (statistical consulting), and Phil Arce (charge data analysis) and Sripha Ouk, Linh Hoang, Bogdan Alexandrescu, Sumeet Shendrikar, Reina Rodriguez, Reza Danesh, Joanna Cheng, Ed Nguyen, Kristal Liu, Andrew Liu, Patrick Gibbons, and Ryan Narasaki for help with chart abstraction, data entry, and data cleaning.

References ■

1. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet*. 1993;342:1317-22.
2. Hunt DL, Haynes BH, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. *JAMA*. 1998;270(15):1339-46.
3. Balas EA, Austin SM, Mitchell JA, Ewigman BG, Bopp KD, Brown GD. The clinical value of computerized information services: a review of 98 randomized clinical trials. *Arch Fam Med*. 1996;5(5):271-8.
4. Tierney WM, Overhage JM, Takesue BY, et al. Computerizing guidelines to improve care and patient outcomes: the example of heart failure. *J Am Med Inform Assoc*. 1995;2(5):316-22.
5. Schriger DL, Baraff LJ, Rogers WH, Cretin S. Implementation of clinical guidelines using a computer charting system: effect on the initial care of health care workers exposed to body fluids. *JAMA*. 1997;278(19):1585-90.
6. Eddy DM. Analyzing indirect evidence. In: *Methodology Perspectives*. Rockville, Md: U.S. Department of Health Services, Public Health Service, Agency for Health Care Policy

- and Research, 1994:5–14.
7. Woolf SH. An organized analytic framework for practice guideline development: using the analytic logic as a guide for reviewing evidence, developing recommendations, and explaining the rationale. In: *Methodology Perspectives*. Rockville, Md: U.S. Department of Health Services, Public Health Service, Agency for Health Care Policy and Research, 1994:105–14.
 8. Schriger DL. Training guideline panels. In: *Methodology in Clinical Practice Guideline Development: Methodology Perspectives*. Rockville, Md: U.S. Department of Health Services, Public Health Service, Agency for Health Care Policy and Research, 1994:115–21.
 9. Baraff LJ, Oslund SA, Schriger DL, Stephen ML. Probability of bacterial infections in febrile infants less than three months of age: a meta-analysis. *Pediatr Infect Dis J*. 1992; 11(4):257–64.
 10. Baraff LJ, Lee SI, Schriger DL. Outcomes of bacterial meningitis in children: a meta-analysis. *Pediatr Infect Dis J*. 1993;12:389–94.
 11. Baraff LJ, Oslund S, Prather M. Effect of antibiotic therapy and etiologic microorganism on the risk of bacterial meningitis in children with occult bacteremia. *Pediatrics*. 1993; 92:140–3.
 12. Baraff LJ, Bass JW, Fleisher GR, et al. Practice guideline for the management of infants and children with fever without source 0–36 months of age. *Pediatrics*. 1993;92:1–12; and *Ann Emerg Med*. 1993;22:1198–210.
 13. Schriger DL, Baraff LJ, Wilton R, McCoy JM. EDECS: the emergency department expert charting system. *Proc 13th Annu Symp Comput Appl Med Care*. 1989:611–5.
 14. Baraff LJ, Schriger DL, Starkman S. Compliance with a standard for the emergency department management of epileptics who present after an uncomplicated convulsion. *Ann Emerg Med*. 1990;19:45–50.
 15. StataCorp. *Stata Statistical Software: Release 6.0*. College Station, Tex: Stata Corporation, 1999.
 16. Day F, Hoang LP, Ouk S, Nagda S, Schriger DL. The impact of a guideline-driven computer charting system on the emergency care of patients with acute low back pain. *Proc 19th Annu Symp Comput Appl Med Care*. 1995:576–80.
 17. McDonald CJ. Protocol-based computer reminders, the quality of care, and the non-perfectibility of man. *N Engl J Med*. 1976;295:1351–5.
 18. Holbrook J, Aghababian R. A computerized audit of 15,009 emergency department records. *Ann Emerg Med*. 1990;19: 139–44.
 19. Kahn KL, Rogers WH, Rubenstein LV, et al. Measuring quality of care with explicit process criteria before and after implementation of the DRG-based prospective payment system. *JAMA*. 1990;264:1969–73.
 20. Bero L, Grilli R, Grimshaw JM, Oxman AD (eds). *The Cochrane Effective Practice and Organisation of Care Review Group. The Cochrane Database of Systematic Reviews. Issue 3*. Oxford, England: Update Software, 1998.
 21. Kramer MS, Shapiro ED. Management of the young febrile child: a commentary on recent practice guidelines. *Pediatrics*. 1997;100:128–34.
 22. Schriger DL. Management of the young febrile child: clinical guidelines in the setting of incomplete evidence. *Pediatrics*. 1997;100:136.
 23. Baraff LJ, Schriger DL, Bass JW, et al. Management of the young febrile child: commentary on practice guidelines. *Pediatrics*. 1997;100:134–6.
 24. Wittler RR, Cain KK, Bass JW. A survey about management of febrile children without source by primary care physicians. *Pediatr Infect Dis J*. 1998;17:271–7.
 25. Adams ID, Chan M, Clifford PC, et al. Computer-aided diagnosis of acute abdominal pain: a multicentre study. *BMJ*. 1986;293:800–4.
 26. U.S. Congress, Office of Technology Assessment. *Identifying Health Technologies That Work: Searching for Evidence*, Washington, DC: U.S. Government Printing Office, Sept 1994:177–98. Publication OTA-H-608.