Reminders to Physicians from an Introspective Computer Medical Record

A Two-Year Randomized Trial

CLEMENT J. McDONALD, M.D.; SIU L. HUI, Ph.D.; DAVID M. SMITH, M.D.; WILLIAM M. TIERNEY, M.D.; STUART J. COHEN, Ed.D.; MORRIS WEINBERGER, Ph.D.; and GEORGE P. McCABE, Ph.D.; Indianapolis, Indiana; and Lafayette, Indiana.

We developed a computer-stored medical record system containing a limited set of the total clinical data base—primarily diagnostic studies and treatments. This system responds to its own content according to physician-authored reminder rules. To determine the effect of the reminder messages generated by 1490 rules on physician behavior, we randomly assigned practitioners in a general medicine clinic to study or control groups. The computer found indications for six different actions per patient in 12 467 patients during a 2-year study: 61 study group residents who received computer reminders responded to 49% of these indications; 54 control group residents, to only 29% ($p < 0.0001$). Preventive care (occult blood testing, mammographic screening, weight reduction diets, influenza and pneumococcal vaccines) was affected. The intentions of the study group to use a given action for an indication predicted their response to the indications ($p < 0.03$, $r^2 = 0.33$). The intentions of the control residents did not.

WE HAVE DEVELOPED a computer-stored medical record system that is able to analyze and respond to its own contents according to physician-authored reminder rules. Each reminder rule specifies a patient's state and the clinical action indicated for that state. The computer follows these rules as standing orders; before each patient visit it searches the patient's electronic record for the specified indications and reminds the physicians about actions needed.

In previous work, we saw that a limited version of this reminder system reduced physician oversights and concluded that physician errors are often due to "channel noise" rather than practitioner ignorance (1). Others have seen similar benefits of computerized reminder systems (2-5).

Our previous studies were not large enough or long enough to provide detailed information about differences in the effects of the reminders on individual physicians or on individual clinical actions, or to identify potential effects on patient outcomes. We report the results of a 2-year randomized controlled trial of the long-term effects of a more comprehensive reminder system on 126 physicians. The purpose of this study was to determine the effect of the reminder system on physician's use of individual clinical actions, to identify and explain the variability in the response rate of physicians to indications for these actions, and to find improvements in patient outcomes that may result from the computer reminders.

► From the Regenstrief Institute for Health Care and the Department of Medicine, Indiana University School of Medicine, Indianapolis, Indiana; and the Department of Statistics, Purdue University, Lafayette, Indiana.

Methods

Since 1977, any patient seen by the General Medicine Service in the emergency room, clinics, or hospital wards has been registered with the Regenstrief computerized medical record system. After a patient was registered, the computer record included most of the patient's diagnostic studies, treatments, discharge and visit diagnoses, and clinic vital signs and test orders. Laboratory and medication data were entered automatically from computerized laboratory and pharmacy systems, respectively. The remaining information was hand-coded and entered into the computer by research technicians. We made no special effort to obtain any information about medical care provided before a patient's computer registration. As of the start of the study, more than 30 000 patients have been registered.

The computerized medical record plays an active role in the care of patients by reminding physicians of patient conditions that need attention. The computer's "judgment" in these matters is controlled by physician-authored rules, written in an English-like language called CARE(6).

For use in this study, a committee of three general medicine faculty members and appropriate consulting subspecialists developed a total of 1491 CARE rules that could generate 751 different reminder messages. (These numbers are not comparable with the numbers of rules or messages described in our previous studies because our current version of the CARE language is more succinct than the previous versions.) This set of rules included reminders about preventive care, such as the use of occult blood testing, mammography, cervical pap testing, purified protein derivative skin testing, adult immunization, and weight reduction, and about obtaining tests needed to complete the initial data base or to identify the cause of existing abnormalities (for example, the use of iron and iron-binding to diagnose microcytic anemia). There were rules about prophylactic treatment, such as the use of beta blockers to prevent sudden death in patients after myocardial infarction, the use of calcium supplements to prevent osteoporosis in menopausal women, and the use of long-acting nitrates to prevent angina. Finally, there were rules about treatment of active problems (such as the treatment of congestive heart failure). We have previ-

```
        THESE SUGGESTIONS ARE BASED ON PARTIAL DATA.
     YOUR CLINICAL JUDGMENT SHOULD ALWAYS TAKE PRECEDENCE.
CIRCLE "A" IF YOU AGREE WITH THE SUGGESTION, "D" IF YOU DISAGREE.

SAMPLE.PATIENT    #9999999-7    AGE: 55    SEX: F    RACE:    PHONE: 000-0000

Scheduled to  DEDALUS, STEPHEN on 19-DEC-83 at 01:30 PM (1)

Recent skull x-ray or bone scan reports "sinusitis" or sinus clouding.
clouding.  Antibiotics (R:1363) and/or decongestants may be requested.
SINUS X-RAY, 11-DEC-79, SINUS CLOUDING

The HIP Study (R:2083) showed that, in patients between the ages of
50 and 70, mammography & physical exam reduced the breast cancer
death rate by 1/2.  IU Oncology recommends breast exam every year &
"mammogram" every 1 to 3 years.  An alternative strategy would be
to reserve mammography screening for older women with obese
breasts.  Obesity is a risk factor for breast cancer, reduces the
sensitivity of physical exam, and improves the sensitivity of
mammography, since fat is an ideal contrast medium for breast
tissue.  R:2408, R:2386
AGE, 66.7
```

Figure 1. A typical reminder report produced during the study with fictitious patient and physician identification.

ously published the full set of protocols and their medical rationale (6). The computer had access to a limited subset of diagnostic studies, diagnoses, and treatments that were part of the patients' total clinical data. But within these constraints and those of the CARE language, the rules represent the practice ideals of the committee.

Operationally, the computer used these rules to review each patient's electronic medical record the day before each visit. For each patient, when it found conditions satisfying the reminder rule, a reminder was stored on a disk file. The computer gathered these messages as a printed report (Figure 1) that was attached to the charts of patients scheduled to be seen by the physicians.

The reminder messages included citations to the relevant medical literature (the numbers after "R:"). In this study, we asked the physicians to initial or notate each reminder so we could estimate the percentage of reports the physician actually received and read.

The provision of the reports was the experimental intervention in the study. The study was done in the general medicine clinic where physicians practiced in teams consisting of a single staff physician, three or four residents, and a nurse-clinician or a nurse. Because physicians on a given team sometimes saw one another's patients, we chose the practice team as the unit of randomization for the study. There were 27 such teams, each of which met one half day per week. Each team was randomized as to study or control. Physicians on a given team (whether faculty members, residents, or nurses) obtained reminder reports on patients assigned to that team. Physicians on control teams did not receive reminders, but the computer executed the reminder logic and kept records of the indications for actions just as it did for the study teams.

Residents are members of only one team and follow a group of patients one half day per week for the 3-year duration of their training. Faculty members and nurse clinicians practice on more than one team. The randomization was constrained to assure that individual faculty members and nurse-clinicians practiced on both study and control teams and thus could serve as their own controls.

For the analysis of patient outcomes, we aggregated patient results according to whether their assigned physician was in a study or control group and treated each patient as an independent observation.

We measured the effect of the computer reminders on physician behavior in terms of the physician's rate of reponse to the clinical indications that stimulated the computer reminders. The first step in developing this measure was to determine the percentage of the computer's indications for a given action to which the physician responded appropriately. To obtain this percentage, we created another set of CARE rules called compliance rules. The computer used these rules to judge whether the physician responded to the indications it discovered at each visit (the large number of patients prohibited manual chart review).

By necessity, the computer judgments were limited to consideration of actions that were usually recorded in the computer record (medication changes, referral to other clinics, and orders for diagnostic studies). The computer

Table 1. Computer-Recommended Action and Percent of Patients Eligible

Action	Patients Eligible	Major Eligibility Criteria
	%	
Stool occult blood	64	Yearly screening if age between 45 and 85
Cervical smear	45	Canadian Medical Association intervals for women with intact uterus
Hematocrit/hemoglobin	42	Initial baseline study and follow-up for decrease of less than −5
Chest roentgenogram	41	Initial screening of patients and follow-up of initial infiltrates and congestive heart failure
Pneumococcal vaccine	38	U.S. Public Health Service criteria except age over 65 instead of over 50
Tuberculosis skin test	27	If age is under 35, if no past history of positive test or tuberculosis (inner-city population)
Serum potassium	25	Yearly in patients on diuretic agents or potassium supplements
Mammogram	23	Every 18 months in women between 50 and 70
Influenza vaccine	23	U.S. Public Health Service criteria if patient visits clinic in fall
Diet	19	If weight greater than 130% of ideal weight
Reticulocyte count	14	Once for undiagnosed anemia
Total iron binding	11	Once in patients with undiagnosed anemia and a median cell volume under 90
Digitalis	10	Trial if patient has chest roentgenogram or echocardiagram showing congestive failure and unexplained dyspnea
Liver enzyme levels	10	Once a year if on hepatotoxic drugs and baselines for new patients
Antacids	9	Prophylaxis for high-dose aspirin, corticosteroids, or nonsteroidal anti-inflammatory agents

could not take into consideration actions such as counseling, reassurance, patient history, or physical examination, which were not reflected in its records. However, such actions would not be appropriate substitutes for many of the actions suggested by the computer (such as occult blood testing to detect colon cancer). Moreover, these limitations applied equally to physicians in the study and control groups and did not bias our results.

The full set of these compliance rules (55 pages of computer print-out) can be obtained from the author on request. In general, the compliance rule accepted the action suggested by the computer or a stronger action as an appropriate response to the detected indication. When a reminder message was triggered by an abnormal or significantly changed laboratory test, the computer generally accepted a repeat of the test in question as an appropriate response.

The response rate per visit contained redundant information about the need for the suggested action in an individual patient because it counted each visit at which the computer identified an indication in a patient as a separate event. Moreover, the per visit measure magnified the real difference between the study and control physicians because an indication for an action would be more likely to be overlooked by control physicians and therefore counted a greater number of times in the same patient. To eliminate this bias against the control physicians, we converted the physicians' per visit response rate into a per patient response rate for each study action. This measure counted an indication for an action only once per patient, no matter how many times it was found during the course of the study. We judged the physicians's response to this indication as positive if there was a response and negative if there was no response at a scheduled visit during the 2-year study. For the common actions, this measure approximates the physician's use of a given action per eligible patient.

To assure stable estimates of the individual physician's response rates, we excluded physicians with fewer than 100 reminder messages during the period of the study. This step eliminated interns who started their training in July 1980, 1 month before the end of the study and nine other short-term physicians. These exclusions eliminated fewer than 2% of the total reminder messages produced during the study and had no effect on the pooled average response rate. A total of 115 residents, 11 faculty members, and 4 nurse-clinicians were available for analysis.

In the separate analysis of physicians' response to suggestions about each of the 15 commonest actions, we eliminated physicians with five or fewer patients eligible for a given action to assure stable estimates of the response rate. Because physicians were less likely to have five patients eligible for the less frequently suggested actions, fewer physicians were available for the analysis of the effects on these actions.

In order to determine whether physicians' acceptance of the computer's indications for a given action could predict their response rate to the computer reminders, we asked departing residents in June 1980 and 1981 to complete a questionnaire. We chose to obtain this information at the end rather than the beginning of the patient's participation in the study to avoid confounding the effects of the questionnaire with the effects of the reminders. The disadvantage of this choice is that the physicians' opinions may have been influenced by the study intervention. For each of 12 different actions suggested by the computer, physicians were asked to answer on a scale from one to five whether they never, rarely, occasionally, usually, or always, take a given action when a given indication occurs in their own patients. The questionnaire asked for the intentions about multiple indications for a given action. For each action and each physician, we computed an overall intentions score by weighting the physicians' reported intentions for each indication by the frequency with which that indication occurred during the study and averaging this weighted in-

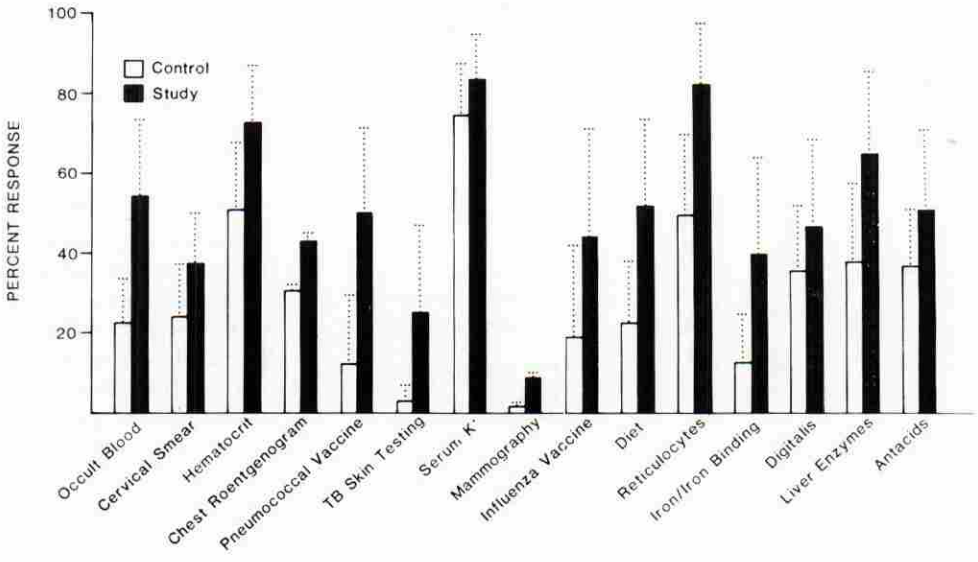

Figure 2. Residents' average response rate to indications for the 15 most frequently suggested clinical actions broken down by study and control. The dotted line arising from the top of each bar graph shows one standard deviation from the mean ($p = 0.015$ for digitalis results; $p < 0.0005$ for each of the others). Except for digitalis, these differences are significant after correcting the significance level to 0.0033 (Bonferroni correction).

tention over all the indications for a given action. We then used this overall intention score to predict the physician's response rate to suggestions about a given action.

In order to identify other factors that may predict the physician's rate of response to computer reminders, we obtained information about the physicians' number of years of training, the medical school they attended, their clinical ability as judged by faculty members, and their attitudes about the computer system and the reminder messages. Information about these items was available for overlapping subsets of the physicians studied.

To determine the effect of our experimental intervention on patient outcomes, we compared the average number of hospitalizations, emergency room visits, and the average values for each patient of blood urea nitrogen, hematocrit, aspartate transaminase, serum potassium, cholesterol, diastolic blood pressure, and body weight of patients in the study group with corresponding measurements made in the control group. The physiologic test variables we used were those that were ordered frequently enough by physicians to insure that most patients would have an observation on record and had the possibility of showing effects of corrective actions stimulated by the computer reminders. Information about emergency room visits and hospitalizations was obtained from the admission logs in emergency room visits at Wishard Memorial Hospital (there was no special interviewing at the end of the study). The laboratory measures were obtained during the course of routine clinical care. When a patient had more than one measurement of a given variable, we used a time-weighted average of these measures over the period of the study to represent that variable.

For the analysis of the effect of the reminder messages on faculty members (who served on both study and control teams), we used the individual as the unit of analysis and used paired *t*-tests to compare the differences in the response rate under study and control conditions. The results of nurse-clinicians were analyzed in the same fashion with paired *t*-tests and also with Wilcoxon signed rank tests (because of the small sample).

Residents served on a single team as either study or control subjects together with other residents of the same study status. We analyzed the effect of the study intervention on the residents' response rate by means of analysis of variance but used a hierarchical model with individual practice teams nested within study and control status to take into account possible effects of the physician's team. Within this model, we also estimated the interclass correlation of residents' response rates within a team (7). The analysis was done with the general linear model procedure in the Statistical Analysis System (8). Using the same approach, we did separate analyses on the residents' response rate to each of the 15 most commonly suggested actions and the response rate to all other actions combined. To identify differences among the three provider groups, we used analysis of variance procedures with the Student-Newman-Keuls multiple range tests.

We used regression analysis to determine which of the various resident factors was associated with the residents' overall response rate. Because the team on which the residents practiced had no significant effect on the overall response rate, we used the individual resident as the unit in these analyses. We did the regressions with and without arc sine transforms of the dependent variable and with and without weighting by the number of patients under the physician's care. Because the results of all four analyses were practically the same, we present the results only for the untransformed, unweighted analyses. All *p* values reported are two-tailed probabilities with a 0.05 significance level.

Results

OVERALL EFFECT OF COMPUTER REMINDERS

The patients cared for by study and control providers were not significantly different ($p > 0.2$) in sex or race. Nearly 65% of the patients were black, 65% women, and 60% over 50 years of age in both groups. The computer found indications for at least one clinical action in over 12 467 scheduled patients (more than 90% of all scheduled patients) who had computer review of their record during the 2-year course of study. The computer found indications for an average of six different clinical actions

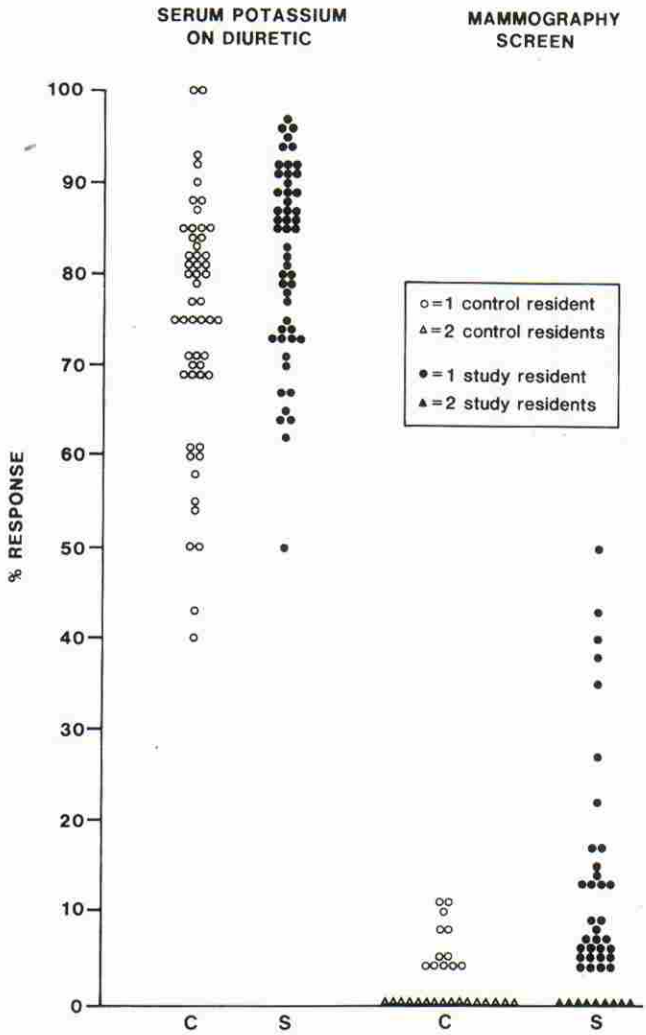

Figure 3. Distribution of residents response rates to suggestions about ordering serum potassium and screening mammography. Each circle represents the response rate of one resident. Each triangle represents the response rate of two residents.

per patient during this time.

The mean per patient response rate of 61 residents in the study group to the indications for the study actions was 49% and that of the 54 residents in the control group was 29%. The effect of the computer reminder messages on the residents' response rate was significant ($p < 0.0001$). The effect of the resident's team was not ($p = 0.1$, intraclass correlation $= 0.1$). The response rate for the 11 faculty members who served as their own controls was 44% and 29% in the study and control states respectively ($p < 0.01$). Curiously, the response rates of residents and faculty members in control sessions were identical, even though faculty members received reminders during study sessions that may have contaminated their control response rate.

Only four nurse-clinicians were included in the study. Their average response was 50% and 36% in the study and control state, respectively ($p < 0.03$ by paired t-test, $p < 0.06$ by Wilcoxon signed rank test). There were no statistically significant differences among the three provider classes in either study or control state.

Because the residents were the largest group of providers and were the only ones with numbers sufficient for more detailed analysis, most of the following results are limited to the effect of the computer reminder messages on residents.

EFFECT ON REMINDERS BY STUDY ACTION

During the period of the study, the computer found indications for more than 150 different clinical actions (diagnostic tests or treatments). We examined the response of the residents to indications for the 15 most frequently suggested clinical actions and to the responses to all of the remaining actions (in the aggregate).

For each of these actions, Table 1 shows the percentage of the patient population that was eligible for the action and the indications most frequently used by the computer to justify the action. Together, reminder messages about these 15 actions constituted 69.8% of these messages during the study. Figure 2 shows the average study and control resident's response rate to indications for each of these actions. The p value was 0.15 for digitalis usage; 0.0005 for cervical pap testing, chest roentgenogram, serum potassium, mammography, and antacid usage; and $p < 0.0001$ for the eight remaining actions in Table 1. All but the first of these is highly significant, after correcting the significance level to 0.0033 (using the Bonferroni correction required by the multiple statistical tests we did). The effect of the reminders on preventive care, such as cancer detection and immunization, was especially large. Residents in study groups were two to four times more likely to apply preventive care to their eligible patients than were control group physicians.

The aggregate effect of reminder messages about actions suggested less frequently than those in Table 1 was also significant ($p < 0.0001$). Among the less frequently suggested actions, the computer reminders had large numerical effects on serum amylase for abdominal pain, colon roentgenograms for hemoglobin-positive stools, urine cultures for pyurea, serum fluorescent treponemal antibody tests to follow-up positive VDRL tests, median cell volumes to detect anemia, metronidazole to treat trichomonas, multivitamins for alcoholic patients, vitamin K for unexplained prothrombin time elevations, prothrombin time after coumadin treatment, and T4 index to work up findings suspicious of hypo- or hyperthyroidism.

PHYSICIAN CHARACTERISTICS THAT PREDICT THEIR RESPONSE RATES

The response rate of both study and control residents differed substantially both across and within individual actions. Figure 3 shows the distribution of response rates for serum potassium and screening mammographic examination, respectively. Although potassium was one of the most, and mammography one of the least, "popular" actions, the more than 50% point spread between the extremes in physician response rate is typical of the other frequently suggested actions. In order to determine the cause of this variability, we examined a number of physician characteristics. Data about many of these character-

Table 2. Average Intentions Score and Response Rate for 12 Study Actions*

	Study Residents		Control Residents	
	Intentions Score	Response Rate	Intentions Score	Response Rate
Occult blood	4.6	0.55	4.2	0.22
Cervical smear	4.5	0.38	4.6	0.23
Chest roentgenogram	4.4	0.43	4.5	0.30
Pneumococcal vaccine	4.4	0.51	4.0	0.14
Tuberculosis skin test	3.5	0.26	3.2	0.03
Serum potassium	4.3	0.84	4.1	0.75
Mammography	3.1	0.08	2.3	0.02
Influenza vaccine	4.2	0.46	4.0	0.20
Diet	4.2	0.49	3.2	0.37
Digitalis	3.9	0.49	3.2	0.37
Antacids	2.7	0.50	2.6	0.37
Beta blockers	1.8	0.15	1.6	0.14

* Each intention score is based on the mean response of the residents (12 to 25) who were asked about their indications for a given action.

istics were available only for subsets of the total physician population. Therefore, we analyzed the relationship between these characteristics and the response rate univariately.

In these analyses, neither the number of years of training nor a faculty assessment of residents' clinical ability predicted the resident response rates. The attitude of the residents in the study groups about the computer system in general and the reminder messages in particular predicted their response rate, accounting for 15% of the variance ($p < 0.001$). The degree to which residents read the reports (as shown by their initials) predicted their response to a similar degree, explaining 15% ($p < 0.001$) of the variance. These two predictive variables were correlated ($r = 0.42$, $p < 0.001$); physicians who were positive about the computer were more likely to read the reports and vice versa.

DEGREE TO WHICH PHYSICIANS ACCEPTED COMPUTER INDICATIONS FOR THE STUDY ACTION

Table 2 shows the percentage of patients eligible for each of 12 actions, including 9 of the 10 most frequently suggested ones. The table also shows the average response rate and the average intentions score of study residents and control residents for each of these actions. Using this data, we examined the degree to which, as a group, the residents' overall intentions (as measured by the group mean intention score) predicted their overall behavior (as measured by the group mean response to an indication for a clinical action) across various actions. A similar analysis was done for control physicians. Among study residents, the physicians' intentions predicted their behavior, explaining 33% of the variance in response rate across the various actions ($p < 0.03$, $r^2 = 0.33$). Among control residents, a comparable analysis showed no significant relationship between intentions and actions ($p = 0.26$, $r^2 = 0.12$).

CAUSES OF NONRESPONSE TO COMPUTER REMINDERS

Gaps in the computer record existed at the time of computer review for three reasons. First, there was lack of systematic collection of patients' preregistration history. Second, because the computer review was done before the visit, there was no access to laboratory results or measurements obtained at the visit. Third, there were random delays in the delivery of some of the manually processed reports (for example, cervical pap smears). To determine the degree to which inappropriate reminder messages due to gaps in the medical record could account for the nonresponse rate, we reviewed charts of random samples of study patients in whom physicians did not respond to reminders about five frequently suggested clinical actions. For each such action, 100 charts were reviewed. We found important gaps in the computer's record in 37%, 50%, 10%, 17%, and 6% of the records reviewed respectively for occult blood tests, cervical smears, tuberculosis skin tests, pneumococcal vaccine, and screening mammographic examination. The gaps were the greatest with respect to cervical smears. This gap was due in large part to the many cervical pap smears recorded in the patient's chart from Obstretrics and Gynecology before their first medicine visit. (There was no systematic abstraction of the hospital chart at the time of registry in the Medicine Clinic computer.) Using a weighted average of the above figures, we estimate that computer record gaps could account for 29% of our study physicians' nonresponse.

Due to failures in delivery of the reports or lack of physician interest, study physicians did not initial the reminder report at 30% of the encounters. If we assume the physicians did not actually see these reminders, we could account for another 10% to 20% of physician nonresponse rate. Disagreement between physician and the computer regarding the indication for a given action accounts for another proportion of the study physician nonresponse. We are unable to quantitate the contribution of this factor to the overall nonresponse rate, but know that the physicians disagreed with the computer's indication for at least some of the suggested actions such as mammography and use of beta blockers.

Limitations of the computer reminder logic could also explain much of the nonresponse. The computer reminder rules were approximations of the actual logic used by physicians, and the physician was aware of circumstances the computer did not even consider. For example, the computer did not consider the possibility that a patient may have had bilateral mastectomies before suggesting screening mammography. One study patient had such surgery. Appropriately, the study physician did not

respond to that reminder.

PATIENT OUTCOMES

There were no significant differences between patients cared for by study physicians and those cared for by control physicians with respect to their overall number of hospitalizations and emergency room or clinic visits during the period of the study. Nor did they differ with respect to their time-averaged values of diastolic or systolic blood pressure, weight, serum glucose, serum hemoglobin, serum potassium, or blood urea nitrogen. In one subgroup—patients eligible for pneumococcal or influenza vaccine, there was a statistically significant effect. Patients from this subgroup who were cared for by study physicians had fewer winter hospitalizations and emergency room visits than those cared for by control physicians ($p < 0.02$) in the years influenza epidemics occurred.

Discussion

The computer reminder messages had a strong and persistent effect on patient care. There were significant effects on each class of provider and on physician usage of each of the 15 most commonly suggested actions. The usage of preventive care was at least twofold greater among physicians in the study group than among control group physicians.

There are two possible explanations for this observation. First, the difference between study and control response rates could be due to errors of commission by study physicians, assuming that the computer could pressure the physician into accepting suggestions about tests or treatments with which they disagree. This explanation is highly unlikely. On the average, study physicians ignored suggestions in about 50% of the actions suggested; they ignored the reminders about mammographic examination in 92% of patients. Physicians seem well able to resist suggestions with which they disagree.

The alternative explanation, and the one we believe is correct, is that the difference between study and control physicians is due to errors of omission by control physicians, that is, computer reminders improve the fidelity between a physician's actions and his intentions. This explanation is compatible with the relation seen between physicians' intentions and their actions— that is, by eliminating the obscuring errors of omission in the study group, we could discern a relationship between actions and intentions that was not apparent in the control group.

Gaps in the records considerably reduced the specificity of the reminder messages for some clinical actions. Cervical smears were most strongly affected. Half the physicians' nonresponse to these reminders were appropriate responses to a false-positive reminder message. However, we doubt this nonspecificity had an effect on the relative response rate of study physicians compared to controls because the missing data problem applied equally to both.

Other studies have shown that physicians often fail to live up to their collective (9) or personal (10-11) practice ideals. Hulka (11) reported that physicians fail to complete an average of 39%, 45%, and 49% of the items that they personally specify as necessary for the management of diabetes, hypertension, and dysuria, respectively. Among our control physicians, the short-fall between their stated ideals and their practice was especially notable with respect to preventive care. Physicians at large seem to have a similar difficulty. The percentage of patients eligible for influenza vaccine who actually receive the vaccination in a given year—20% (12)—is close to the percent vaccinated by our control physicians. Extrapolating from the four million doses of pneumococcal vaccine manufactured by 1980 (13) and the 24.7 million Americans eligible for the vaccine by virtue of their age, physicians at large vaccinated fewer than 16% of their eligible patients, a rate of vaccination comparable to that achieved by our control physicians. In one report less than half the patients with auto- or surgical splenectomy and fewer than 5% of patients with chronic obstructive lung disease or advanced age had been vaccinated within the first year of the vaccine's release (14). Finally, from a national survey we learn that 30% of black women and 13% of white women between the age of 45 and 64 living in metropolitan areas have not yet received their first cervical pap test (15).

We have shown that computer reminder messages reduce the disparity between the actions and intentions of our residents. Investigators from diverse clinical environments have shown that computer reminder systems of various types produce similar benefits (2-5). For selected problems the use of paper checklists also achieve such benefits. In one study, checklists increased the usage of occult blood testing, cervical pap testing, and mammography from control levels to degrees comparable to those we achieved (16). As simple memory prods, checklists may also improve diagnostic accuracy (17-18).

These observations lead to one conclusion. Busy physicians cannot perform according to their ideals without some form of mechanical aid. Other fields have long recognized the frailty of the human mind and provide memory aids (for example, the pilot's preflight checklist). If physicians are serious about achieving their stated ideals, they should do likewise. Today, the only practical aid for most physicians is the checklist. Checklists are most suitable for reminding about policies that apply to a large portion of one's population or require interventions at intervals, such as immunizations and screening. Checklists have the disadvantage of limited scope and burdensome processing.

Computer systems with capabilities similar to those described in this paper are now available, but cost puts them out of the reach of most practitioners. However, the cost of microcomputers is rapidly declining and will soon cease to be an obstacle. We predict that office practitioners will eventually acquire and use such systems, and as a result we would expect a greater response to reminders than seen in this study because some causes of nonresponse would be eliminated. For example, disagreement between the computer and practitioner's indications would be less likely because the office practitioner would

be free to define the rules in his practice according to his own ideals. In contrast, our study physicians had to accept one set of rules defined by a small committee. The occurrence of gaps in data needed for the computer reminders would also be less likely to occur. Our residents spend most of their time on the wards, not in the clinic, and were not consulted about the development of the rules. Thus, the residents had neither the time, resources, nor incentives to maintain computer records with the complete information needed by the reminder rules. The opposite is likely to be true in a private office practice.

The accuracy and scope of the computer reminder rules were severely limited by the availability of only a small portion of the total data base compared to that available to the physician at the patient encounter. The computer did not have access to most patient history and physical examination data, or any of the information collected at the visit. The computer contained only a coded precis of narrative reports. Reminder systems that can obtain information from the physician—through a dialogue—as well as from the electronic record are under development here and at Latter Day Saints Hospital in Salt Lake City. Such systems could escape from the current limits. The techniques of artificial intelligence provide additional potential to future reminder systems (19-20).

The computer reminder messages had no overall effect on our measures of patient outcomes. This fact does not diminish the importance of the large effect of the reminders on physician behavior because a number of factors limited the capability of our study to detect outcome differences. Firstly, there was no interviewing after the study to determine the patient's precise status at the end of the study. Our outcome measures depended on incomplete data, obtained in the routine care process. Second, our sample size was too small to find differences due to the clinical actions on which we had the greatest influence. For example, a current trial of occult blood testing estimated a need for 48 000 patients to be followed for 5 years and the Health Insurance Plan study required 60 000 patients to be followed 7 years to find outcome differences (21-22). We followed 8000 patients eligible for occult blood testing and 2000 patients eligible for mammography for only 2 years. In addition, the potential outcome effect of reminder messages was diluted by care provided during hospitalizations, and non-medicine clinic and emergency room visits where computer reminder messages were not available. Finally, we did see a reduction in the winter morbidity in one subgroup of patients—those eligible for pneumococcal and influenza vaccine.

Our results emphasize the slow pace of acceptance of new clinical practices. Mammographic screening, the only cancer screening technique with proven life-saving benefits (22) and supported by both the American Cancer Society and our local oncologists, had very limited acceptance by our physicians, despite a barrage of computer reminder messages. The computer reminder messages had no effect on the physicians' low use of beta blockers in patients after myocardial infarction (23) or calcium supplements in postmenopausal women (24-25) even though we cited clinical trials that supported use of these drugs. However, the lack of potency of citations to the medical literature should not be surprising. Almost 25 years ago, Coleman and associates (26) saw the weak effect of medical literature on the acceptance of new practices. We conclude that although computer reminder messages are potent activators of existing physicians intentions, they have little influence on the acceptance of new practices.

Coleman and associates saw that personal interactions strongly determined whether physicians would accept a new practice. We suspect that the size of the promotion—whether waged by professional societies or commercial interests—is also important. As in other pursuits, the truth may be that which "is loudly proclaimed and stoutly maintained."

ACKNOWLEDGMENTS: The authors thank Doris Paradise and the nurses and physicians of the General Medicine Clinic for their cooperation; Jeff Haas, Lonnie Blevins, and Dave Chamness who developed and managed the programs used in the study; the faculty members of the Indiana School of Medicine who provided advice about the study and the content of the CARE rules; and Ms. Shelly Glazener for her skills in developing this manuscript.

Grant support: in part by grants DHHS 02485 and DHHS 04090 from the National Center for Health Services Research, and grant PHS P60 AM 20542 from the National Institute for Arthritis, Metabolism and Digestive Diseases.

▶ Requests for reprints should be addressed to Clement J. McDonald, M.D.; Regenstrief Institute, 1001 West Tenth Street; Indianapolis, IN 46202.

References

1. McDonald CJ. Protocol-based computer reminders, the quality of care and the non-perfectability of man. *N Engl J Med.* 1976;**295:**1351-5.
2. Barnett GO, Winickoff R, Dorsey JL, Morgan MM, Lurie RS. Quality assurance through automated monitoring system and concurrent feedback using a computer-based medical information system. *Med Care.* 1978;**16:**962-70.
3. Rogers JL, Haring OM. The impact of a computerized medical record summary system on incidence and length of hospitalization. *Med Care.* 1979;**17:**618-30.
4. Wirtschafter D, Carpenter JT, Mesel E. A consultant-extender system for breast cancer adjuvant chemotherapy. *Ann Intern Med.* 1979;**90:**396-401.
5. Warner HR. *Computer Assisted Medical Decision Making.* New York: Academic Press; 1979.
6. McDonald CJ. *Action-Oriented Decisions in Ambulatory Medicine.* Year Book Medical Publishers; 1981.
7. Cochran WG, Snedecor GW. *Statistical Methods.* 6th Ed. Ames, Iowa: The Iowa State University Press; 1967:291.
8. *SAS User's Guide, 1979 Edition.* Cary, North Carolina: SAS Institute, Inc.; 1979:221-36.
9. Payne BC, Lyons TF, Dwarshius L, Kolton M, Morris W. *The Quality of Medical Care: Evaluation and Improvement.* Chicago: Hospital Research and Educational Trust; 1976.
10. Hare RL, Barnoon S. Medical care appraisal and quality assurance in the office practice of internal medicine. Special Report of the American Society of Internal Medicine, Washington, D.C., 1973.
11. Hulka BS. Quality of ambulatory care: an exploration of the discrepancy between explicit process criteria and performance. Hyattsville, Maryland: Department of Health, Education, and Welfare, 1980; DHEW publication no. (PHS) 80-3244.
12. Center for Disease Control. United States Immunization Survey. Atlanta: Department of Health, Education, and Welfare, 1979; DHEW publication no. (CDC) 79-8221.
13. Hilleman MR, Carlson AJ Jr., McLean AA, Vella PP, Weibel RE, Woodhour AF. Streptococcus pneumoniae polysaccharide vaccine: age and dose responses, safety, persistence of antibody, revaccination, and simultaneous administration of pneumococcal and influenza vaccines. *Rev Infect Dis.* 1981;**3:**S31-42.
14. Pantell RH, Stewart TJ. The pneumococcal vaccine: immunization at a crossroad. *JAMA.* 1979;**241:**2272-4.
15. Kleinman JC, Kopstein A. Who is being screened for cervical cancer? *Am J Public Health.* 1981;**71:**73-6.
16. Cohen DI, Littenberg B, Wetzel C, Neuhauser D. Improving

physician compliance with preventive medicine guidelines. *Med Care.* 1982;**20:**1040-5.
17. CHRISTENSEN-SZALANSKI JJ, BUSHYHEAD JB. Physicians' use of probabilistic information in a real clinic setting. *J Exp Psych [Hum Percept].* 1981;**7:**928-35.
18. DE DOMBAL FT, LEAPER DJ, HORROCKS JC, STANILAND JR, MCCANN AP. Human and computer-aided diagnosis of abdominal pain: further report with emphasis on performance of clinicians. *Br Med J.* 1974;**1:**376-80.
19. MILLER RA, POPLE HE JR, MYERS JD. Internist-I, an experimental computer-based diagnostic consultant for general internal medicine. *N Engl J Med.* 1982;**307:**468-76.
20. SHORTLIFFE EH. The computer as clinical consultant. *Arch Intern Med.* 1980;**140:**313-4.
21. GILBERTSEN VA, MCHUGH R, SCHUMAN L, WILLIAMS SE. The earlier detection of colorectal cancers: a preliminary report of the results of the Occult Blood Study. *Cancer.* 1980;**45:**2899-2901.
22. SHAPIRO S. Screening for early detection of cancer and heart disease. *Bull NY Acad Med.* 1975;**51:**80-95.
23. GREEN KG, PARK A, CHAMBERLAIN JA, et al. Improvement in prognosis of myocardial infarction by long-term beta andro-receptor blockade using practolol. *Br Med J.* 1975;**3:**735-40.
24. HORSMAN A, GALLAGHER JC, SIMPSON M, NORDIN BEC. Prospective trial of oestrogen and calcium in postmenopausal women. *Br Med J.* 1977;**2:**789-92.
25. RECKER RR, SAVILLE PD, HEANEY RP. The effect of estrogens and calcium carbonate on bone loss in postmenopausal women. *Ann Intern Med.* 1977;**87:**649-55.
26. COLEMAN J, MENZEL H, KATZ E. Social processes in physicians' adoption of a new drug. *J Chronic Dis.* 1959;**9:**1-19.