Improving Preventive Care by Prompting Physicians

E. Andrew Balas, MD, PhD; Scott Weingarten, MD; Candace T. Garb, BS; David Blumenthal, MD; Suzanne Austin Boren, MHA; Gordon D. Brown, PhD

Objectives: To assess the impact of prompting physicians on health maintenance, answer questions regarding the mode of delivery, and identify opportunities and limitations of this information intervention.

Methods: Systematic electronic and manual searches (January 1, 1966, to December 31, 1996) were conducted to identify clinical trial reports on prompting clinicians. Three eligibility criteria were applied: (1) randomized controlled clinical trial, (2) clinician prompt, alert, or reminder in the study group and no similar intervention in the control group, and (3) measurement of the intervention effect on the frequency of preventive care procedures. Data were abstracted by independent reviewers using a standardized abstraction form, and quality of methodology was scored. A series of meta-analyses on triggering clinical actions was performed using the random-effects method. The statistical analyses included 33 eligible studies, which involved 1547 clinicians and 54,693 patients.

Results: Overall, prompting can significantly increase preventive care performance by 13.1% (95% confidence interval [CI], 10.5%-15.6%). However, the effect ranges from 5.8% (95% CI, 1.5%-10.1%) for Papanicolaou smear to 18.3% (95% CI, 11.6%-25.1%) for influenza vaccination. The effect is not cumulative, and the length of intervention period did not show correlation with effect size ($R = -0.015, P = .47$). Academic affiliation, ratio of residents, and technique of delivery did not have a significant impact on the clinical effect of prompting.

Conclusions: Dependable performance improvement in preventive care can be accomplished through prompting physicians. Vigorous application of this simple and effective information intervention could save thousands of lives annually. Health care organizations could effectively use prompts, alerts, or reminders to provide information to clinicians when patient care decisions are made.

Arch Intern Med. 2000;160:301-308

With the growing emphasis on health maintenance, preventive care, physicians are increasingly expected to perform tasks that are unrelated to the acute problem of the patient. The most common reason that women give for not undergoing screening for breast cancer and cervical cancer is that it was not offered or recommended by their physicians, especially by the youngest group of internists and family practitioners. Various quality scorecard systems encourage plans and clinicians to use every episode of patient care to promote health maintenance. Based on the latest recommendations of the American Academy of Pediatrics, a visit to a physician for a child’s broken arm becomes an opportunity to update immunization.

Controversies exist regarding whether physician prompting affects clinical outcomes. Some consider prompts to be very effective in bringing physicians’ attention to a necessary clinical action and to avoid medical errors due to information overload. Others note that prompts do not always produce the expected change, and some analyses indicated nil to moderate effect on clinical practices.

Unfortunately, the significant benefits of prompting preventive care have never been persuasively communicated to clinicians and health care plans. Researchers also continue placing people in control groups of prompting trials, raising serious ethical questions. Conventional, nonquantitative reviews have failed to highlight that thousands of lives could be saved by prompting.

Conventional approaches to research synthesis, such as reviews or odds ratio calculations, cannot specify the change that results from prompting preventive care. Furthermore, most health services research studies have been conducted in an academic environment, raising questions about the generalizability of results. This study highlights opportunities to elevate preventive care, use diverse and inexpensive technologies for prompting, and improve health plan scorecard performance. The systematic review of studies is supplemented with

From the School of Medicine, University of Missouri, Columbia (Drs Balas and Brown and Ms Garb and Boren); Health Services Research, Cedars-Sinai Medical Center, Los Angeles, Calif (Dr Weingarten); and Health Policy Research and Development Unit, Massachusetts General Hospital, Boston (Dr Blumenthal).
METHODS

Studies meeting the following criteria were eligible for inclusion in this series of meta-analyses: (1) randomized controlled clinical trial, (2) physician prompt in the study group and no similar intervention in the control group, and (3) measurement of the effect on the number of preventive care activities. Studies qualified for automatic exclusion by failing to (1) randomly assign subjects to intervention and control groups and (2) test the randomization by comparing the intervention and control groups at baseline. Preventive care actions can be administered by clinicians with a variety of specialties. Articles that involved clinicians with certain clinical specialties or that focused on unique preventive care actions (eg, alcohol abuse counseling) were not excluded.

COLLECTION OF REPORTS

Many articles eligible for this series of meta-analyses were already in place as part of the Columbia Registry. The Columbia Registry is a systematic collection of medical management trials at the University of Missouri, School of Medicine. Extensive searches were done to collect all additional relevant controlled clinical trial reports: bibliographic database retrievals (eg, MEDLINE, CINAHL, HEALTH), manual searches (monographs, unindexed publications, reference lists), and informal contacts. By mapping the eligibility criteria, the search strategy for the collection of trials on prompting included a combination of medical subject heading (MeSH) terms and textwords (TWs) for each eligibility criterion. The words are MeSH terms unless noted as TW. The criteria were (1) random (truncated TW), group (truncated TW), random allocation (MeSH and TW), randomized controlled trial (publication type), or clinical trial (publication type); (2) checklist (TW), encounter forms (TW), tags (TW), triggers (TW), reminder systems, alert (TW), reminder (TW), leaflets (TW), stickers (TW), messages (TW), or tailored messages (TW); and (3) preventive health services, immunization, vaccination, smoking, smoking cessation, mass screening, mammography, prenatal care, hypertension, blood pressure, diabetes mellitus, alcoholism, substance-related disorders, vaginal smears, hypercholesterolemia, glaucoma, or occult blood.

SCORING AND ABSTRACTION

Two research associates (C.T.G. and S.A.B.) checked eligibility and abstracted information from the reports using standardized and reproducible methods. The methodologic quality of each of the eligible trial reports was evaluated using a validated clinical trial scoring system tailored for health services research trials. The scoring system awards a score on an arbitrarily selected scale from 1 to 100. Aspects of the studies scored included sample definition, testing randomization, intervention, effect variable definition, blinding, numeric table of effect variables, ratio of withdrawals, and analysis of effect variables. Since more than half of the items are study characteristics, as opposed to reporting characteristics, the minimum required score was set at 50 in this study. Subsequently, raw data were derived from all eligible articles (eg, site, patient sample, clinician sample, intervention, and effects). These covariates were required for a study to be included. Authors were also contacted if additional information was needed.

STATISTICAL SYNTHESIS

The health maintenance rate was defined as the ratio of the number of preventive care actions to the number of eligible physician-patient encounters (opportunities). The number of eligible physician-patient encounters was used as the denominator, because the latest clinical recommendations indicate that virtually every visit regardless of reason is an opportunity to provide preventive care. Therefore, a missed opportunity is any physician-patient encounter that does not address preventive care. In addition, the higher the number of eligible studies, 3 trials were conducted in private offices recruited by the researchers, and 30 were conducted in university-affiliated clinics (19 studies) or public clinics (11 studies). Specialty of physicians included internal medicine (15 studies), family practice (general practice) (20 studies), and obstetrics (1 study). In all but one39 of the 33 studies, the patients were adults. The number of patients in the studies ranged from 57 patients33 to 7397 patients, with an average of 1657. The average ratio of patients to clinicians was 35.3.

RESULTS

Literature searches identified 101 pertinent clinical studies. During the filtering process, 68 studies were excluded for various reasons: same data as from another eligible study (3 studies), planned studies (2 studies), not a clinician reminder intervention but a patient reminder intervention (54 studies), other irrelevant reminder (7 studies), and no control group (2 studies). This left 33 eligible reports on randomized controlled clinical trials. No studies were excluded for only failing to test randomization by baseline comparison of intervention and control groups. All studies meeting the inclusion criteria and not qualifying for exclusion exceeded the quality scoring threshold. Using the replicable scoring technique, the average ± SD quality score was 69.7 ± 7.9. The characteristics of all eligible studies are listed in Table 1. All studies were parallel group trials, with the exception of the 2 studies that had a crossover design. In the pool of studies, 17 randomized patients directly, and 16 studies randomized through physicians. In health services research, patients are often randomly assigned to groups through their clinicians, because it is often the clinician who is directly targeted by the intervention. Overall, data on 1547 clinicians and 54693 patients were included in this meta-analysis. In the group of eligible studies, 3 trials were conducted in private offices recruited by the researchers, and 30 were conducted in university-affiliated clinics (19 studies) or public clinics (11 studies). Specialty of physicians included internal medicine (15 studies), family practice (general practice) (20 studies), and obstetrics (1 study). In all but one39 of the 33 studies, the patients were adults. The number of patients in the studies ranged from 57 patients33 to 7397 patients, with an average of 1657. The average ratio of patients to clinicians was 35.3.
of physician-patient encounters, the more opportunities for improvement exist. The clinical effect of prompts was estimated by the difference between the health maintenance rate in the intervention and control groups. The time during which health maintenance is observed is incorporated in the denominator. Point and interval (0.95) estimates of the prompting effect were calculated with models based on random-effects assumptions. These assumptions were raised by the diversity of clinical settings and groups of subjects analyzed. Homogeneity is not an assumption of the random-effects model used in this analysis.

To estimate the overall prompting effect, we used the (trimmed) modified DerSimonian-Laird estimator. The estimate was calculated by a weighted average of the individual rate differences. For each study, the weight was the reciprocal of the sum of an unbiased estimate of the within-study and the between-study variance so that studies having a large variance (and therefore less reliable) received a small weight (eg, in case of studies with few observations). In addition, we estimated the variance and the confidence interval (CI) of the estimated overall rate difference.

Furthermore, a random-effects regression model was applied to identify covariates with possible influence on prompting effect on health maintenance activities. Covariate analysis attempts to identify reasons for variation in the rates. Such factors are sources of heterogeneity in the efficacy of prompts among the studies. In this study, 9 variables were assumed to be associated with the effect of physician prompts based on preliminary literature review: trial quality, site, reimbursement method, characteristics, clinical specialty, size, delivery mode, targeted clinical action, and duration. Single-covariate and multiple-covariate regression models were used. The regression coefficients were estimated by the least-squares method. In addition, SEs and P values were calculated to determine the significance of the effects (covariates). The proportional reduction in variance among studies between conditional and unconditional models measured the degree to which a covariate accounted for the overall variance of the study.

The prompts were always delivered before a scheduled encounter. A few of the studies specified exactly when the prompt was delivered: once per year during the patient’s month of birth, following randomization, in 1 month in advance of due date for targeted action, the night before the scheduled visit, the morning of the scheduled visit, before the visit at the reception desk, or during routine intake procedures.

The prompts contained a variety of information for the clinicians. The most frequently included aspects indicated were patient name, patient diagnosis, notice of deviation from standard (overdue procedures or tests), criteria for procedures or tests, recommendation for treatment or action, indication of when upcoming procedures or tests are due, information on previous procedures or tests, and space for clinician to indicate that the procedure or test was offered or why it was not offered. Five of the studies provided generic prompts that were not patient specific.

The generic prompts were a checklist of guidelines applicable to any patient. Frequently, there was no room to record data of performed actions on the generic prompts. Some studies indicated that the prompts were repeated until the procedures or tests were completed.

In some studies, patients were reminded in addition to the clinicians. Several studies indicate that offering prompts to clinicians and patients amplifies the effect of the intervention. Patients were reminded by telephone, letter, postcard, or a copy of their health maintenance report.

One study, in addition to randomizing clinicians to groups who always received prompts and never received prompts, had a group of clinicians who sometimes received prompts. These clinicians were less likely than the always reminded clinicians and more likely than the never reminded clinicians to vaccinate their patients.

For most trials included in this study, the targeted actions were considered complete if the clinician made a note in the medical record. In addition to a medical record...
audit, one study also interviewed physicians following the encounters.40 Two studies indicated that a random selection of medical records was audited to obtain the rates of compliance with the recommended procedures.27,41 In other studies, patients reported the completion of the targeted action through postencounter telephone interviews.4,30

Only 7 of the 33 studies reviewed provided costs of the reminder intervention. None of the articles included the start-up costs of establishing the computerized reminder systems. Instead, the cost analyses focused on the operating expenses of the reminder systems. Studies showed that the computer-based reminder system cost $0.78 per patient per year to operate,26 the cost to maintain the patient’s medical record and produce reminders was estimated at $2 per patient visit; the cost per cervical cancer screening gained was $11.75; the cost per blood pressure screening gained was $1.7040; the computerized system cost $0.02 per record review42; with the use of the physician reminder system, improvement in preventive services was achieved for less than $5 per extra procedure completed48, and the cost per additional vaccination recorded was $0.43.49

IMPACT ON QUALITY SCORECARD PERFORMANCE

Prompts to clinicians resulted in a significant increase in the performance of all 16 preventive care procedures, in-
cluding cancer screening (fecal occult blood, mammography, Papanicolaou smear), immunization (influenza vaccination, pneumococcal vaccination [Pneumovax], tetanus vaccination), diabetes management, hemoglobin management, blood pressure management and follow-up, cardiac care, cholesterol management, smoking cessation, glaucoma screening, alcohol abuse counseling, prenatal care, and tuberculosis testing. Most of the studies addressed the clinical areas of cancer screening and prevention (20 studies), immunization (14 studies), and diabetes management (4 studies). Eleven of the studies addressed 2 of these clinical areas, whereas 6 of the studies focused on other unlisted areas. The rate difference and calculated tolerance (ie, the number of additional but unpublished negative studies that could reverse the conclusions of this study) for selected cancer screening and immunization procedures are presented in Table 2. Only 6 of the clinical procedures are presented in Table 2, because the remaining clinical procedures were tested in fewer than 6 studies. All calculated tolerance levels exceeded the corresponding threshold tolerance, and it is therefore unlikely that unpublished negative studies would overturn these results. Physician prompting for the 6 procedures presented in Table 2 (fecal occult blood test, influenza vaccination, mammogram, Papanicolaou smear, pneumococcal vaccination, and tetanus vaccination) could expect to save 8333 lives annually.

### OTHER EFFECTS OF PROMPTING

The procedure-specific analysis of covariates indicated that the number of clinicians subjected to prompting, the ratio of residents among those receiving prompts, or the academic affiliation of the participating clinics did not make a difference in the effect of prompting. There are only 4 significant impacts based on a P<.05 level. With Bonferroni adjustment for multiple testing (P<.001), only 1 impact would be considered statistically significant, ie, larger patient-clinician ratio predisposes to larger improvement in tetanus vaccination as a result of prompting (0.0017 ± 0.0004, P < .001). A similar effect was also observed for pneumococcal vaccination (0.0073 ± 0.0025, P = .004). The results of the multiple covariate analysis indicate that prompting has significantly less effect on Papanicolaou smear rates in capitated outpatient care and when prompting is computer generated (−0.097 ± 0.043 [P = .02] and −0.078 ± 0.039 [P = .05], respectively). The few significant regression coefficients indicate the resistance of prompting to various covariates.

Presentation of the prompt appeared to have no significant effect on clinical responses (Table 3). Twenty-six studies using prompts attached to the front of the patient’s medical record made a 14.0% (95% CI, 11.1%-16.9%) rate difference in the overall performance of prompts, whereas 7 studies using prompts placed elsewhere, such as tagged progress notes or computer monitor display, achieved a rate difference of 12.1% (95% CI, 5.4%-18.9%).

Generation of prompts also did not show a significant difference in the clinical effect of prompting (Table 3). Generation of the prompting message through computerized records occurred in 25 studies. Although computerization appeared to adversely change the impact of

<table>
<thead>
<tr>
<th>Targeted Procedure</th>
<th>No. of Studies</th>
<th>Rate Difference, % (95% CI)</th>
<th>Calculated Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fecal occult blood test</td>
<td>11</td>
<td>13.7 (4.7-22.8)</td>
<td>189</td>
</tr>
<tr>
<td>Mammogram</td>
<td>14</td>
<td>11.5 (7.1-16.0)</td>
<td>226</td>
</tr>
<tr>
<td>Papanicolaou smear</td>
<td>15</td>
<td>5.8 (1.5-10.1)</td>
<td>165</td>
</tr>
<tr>
<td>Influenza vaccination</td>
<td>9</td>
<td>18.3 (11.6-25.1)</td>
<td>151</td>
</tr>
<tr>
<td>Pneumococcal vaccination</td>
<td>8</td>
<td>17.2 (6.1-28.4)</td>
<td>132</td>
</tr>
<tr>
<td>Tetanus vaccination</td>
<td>8</td>
<td>11.1 (5.0-17.5)</td>
<td>152</td>
</tr>
</tbody>
</table>

*Some studies contained an analysis of more than 1 targeted intervention. CI indicates confidence interval.

### Table 2. Effect of Prompting on Selected Procedures*

<table>
<thead>
<tr>
<th>Targeted Procedure</th>
<th>No. of Studies</th>
<th>Rate Difference, % (95% CI)</th>
<th>Calculated Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fecal occult blood test</td>
<td>11</td>
<td>13.7 (4.7-22.8)</td>
<td>189</td>
</tr>
<tr>
<td>Mammogram</td>
<td>14</td>
<td>11.5 (7.1-16.0)</td>
<td>226</td>
</tr>
<tr>
<td>Papanicolaou smear</td>
<td>15</td>
<td>5.8 (1.5-10.1)</td>
<td>165</td>
</tr>
<tr>
<td>Influenza vaccination</td>
<td>9</td>
<td>18.3 (11.6-25.1)</td>
<td>151</td>
</tr>
<tr>
<td>Pneumococcal vaccination</td>
<td>8</td>
<td>17.2 (6.1-28.4)</td>
<td>132</td>
</tr>
<tr>
<td>Tetanus vaccination</td>
<td>8</td>
<td>11.1 (5.0-17.5)</td>
<td>152</td>
</tr>
</tbody>
</table>

*Some studies contained an analysis of more than 1 targeted intervention. CI indicates confidence interval.
prompting on Papanicolaou smear rates, this confounding effect could not be confirmed by the multivariate analysis (−0.091 ± 0.051, P = .07). Studies using computer-generated prompts achieved a rate difference of 13.6% (95% CI, 10.9%-16.3%) compared with 10.1% (95% CI, 1.3%-18.9%) for studies using non–computer-generated prompts. However, computerized record keeping and prompt generation was associated with larger studies. The average study size for the computerized studies was 2033 ± 2241 patients compared with noncomputerized studies, which averaged 577 ± 677 patients (P = .009).

Finally, the length of intervention period did not show a significant influence on the overall difference made by prompting. The studies ranged in duration from 5 to 161 weeks, with a total of 1596 study weeks. The average study length was 83 weeks. There was no association between the length of study and the health maintenance rate (R = −0.015, P = .47).

RATE DIFFERENCE

The rate differences and 95% CIs for the analyzed studies are presented in Figure 1. The resulting cumulative health maintenance rate difference was 13.1% (95% CI, 10.5%-15.6%). Only 7 of the studies, when analyzed individually, did not demonstrate significant improvement in the performance of preventive care items.29,32,33,39,43,46,47

CUMULATIVE RATE DIFFERENCE

The cumulative rate differences and 95% CIs for the analyzed studies are presented in Figure 2. The studies were cumulated chronologically based on publication year. Since additional studies were included in the health maintenance rate difference calculations, the CI narrowed. Since the publication of the first randomized prompting study by McDonald37 in 1976, cumulative effect has always indicated significant improvement in preventive care.

COMMENT

The results of this study document that prompting physicians can lead to a significant improvement in health maintenance. The observed and much needed increase in performance of preventive care efforts could reap substantial reductions in total mortality. The many prompting tools offer a wide selection of options that are equally effective and easily applicable in most health care organizations (eg, checklists attached to the patient chart, tagged notes, computer-generated encounter forms, prompting stickers, patient-carried prompting cards). The fact that the beneficial effect of prompting does not seem to endure highlights the need for a complex intervention package to improve health maintenance activities that incorporates but is not limited to prompts. Larger incremental changes in clinical practices probably require combination with other interventions (eg, education, feedback, patient involvement).39

The objective of this article was to analyze the human response to a trigger on the part of clinicians to prompting messages. The measure of the clinical procedure is a quantification of this response. There is no practical link between the odds ratio and the rate of change perceived by clinicians and patients to prompting. By using the rate difference, improvement in HEDIS (The Health Plan Employer Data and Information Set) scorecard performance can be forecast. For example, a managed care organization that has an inferior 34% diabetic eye examination rate could become an above average performer by the application of prompting (national average, 38.4%).35

The rationale of prompting is to promote well-established and effective clinical interventions. For example, it has been documented that uncomfortable and inconvenient preventive care procedures are performed at a much lower rate.36 Yet, even Papanicolaou smear rates that are known to be more resistant to change can still be influenced. Prompts also have been used successfully to remind clinicians to discuss sigmoidoscopy with their patients. Rather than assume that sigmoidoscopy is too uncomfortable for all patients to tolerate, clinicians should inform the patient of the potential benefits and risks and enable the patient to make an informed decision.37 In one study using prompts, screening sigmoidoscopy was offered to 58% of patients.38 Significantly
more patients in the intervention group completed the sigmoidoscopy (29% vs 2%) compared with the control group. The risks associated with the physician prompting intervention are low, whereas the potential benefits are significant in both statistical and practical terms. If a particular decision is supported by one prompt, then it is unlikely to be overwhelming even when several decisions are supported by several prompts. Local factors vary substantially among different health care settings and may determine the exact amount of improvement.

Physicians need reliable information about costs to implement reminder systems. Statements regarding cost without substantiating data are made habitually in reports of clinical trials. Most of the reported costs seem reasonable but still represent a significant operating expense for the practice. The cost of a system to support preventive care reminders could be higher for a practice without computers or a database.

The debate about priorities in cancer research highlights the need to compare the promises of prompting preventive care and the treatment of cancer. For example, a recent analysis estimated the number of patients who would have to be treated to save one life by mammography or by adding a combination product of cyclophosphamide, methotrexate, and fluorouracil to tamoxifen in the treatment of breast cancer (600 and 25, respectively). The annual number of mammographies is about 7.9 million and there are 24,000 newly detected cases of breast cancer in patients between the ages of 50 and 69 years. Therefore, an 11.5% improvement in the rate of breast cancer screening achieved through prompting could save 50% more lives than adding cyclophosphamide, methotrexate, and fluorouracil to tamoxifen. This calculation does not compare the human and financial costs of prevention and chemotherapy. Medical research has traditionally emphasized the discovery of new technologies, tests, devices, and medications while many scientific advances have not achieved their potential for improving patient care. Our study indicates that the benefits of quality improvement and better implementation may be comparable to the benefits of some technological advances.

Meta-analysis is informative when the same intervention has been tested repeatedly, and therefore the effect can be estimated quantitatively. To commit resources to a quality improvement intervention, clinicians and provider organizations need to know the magnitude of the expected benefits. In our study, quantitative analysis was necessary because of the apparent lack of awareness of the benefits of reminding, numerous repeated tests of the same intervention, and the ambiguous recommendations coming from nonquantitative reviews. Based on the significant effect of prompting, future experimentation and placement of randomly selected patients to control groups should raise ethical questions.

One potential problem with defining the health maintenance rate as the ratio of the number of preventive care actions to the number of visits is that it can be changed simply by altering the denominator. Thus, in situations in which the patient is seen often, the apparent health maintenance rate would go down, yet it would not necessarily be appropriate to engage in preventive actions at all visits.

Numerous factors have been identified as obstacles to preventive care, but the interest of patients lies in those that can be changed to improve quality. Lurie et al found that patients who had free care received more preventive services than patients who shared the cost. Fragmentation of care, patients seeing several physicians of different specialties or at different locations, also presents a barrier to preventive care. Among patients surveyed by Bindman et al, the most important factor associated with receipt of preventive care was having a regular source of care. Often patients and physicians have an expectation that the agenda will be limited to the immediate concern, with preventive care reserved for the checkup, or periodically scheduled physical examination, which may not take place in a timely fashion, if at all.

The American College of Physicians recommends that "the external oversight of care be restructured and that the primary locus of quality assurance be returned to the medical profession and to health institutions." Timely possession of pertinent information could empower physicians to direct medical resources and improve the quality of patient care. Such reallocation of information is more likely to lead to much needed changes that can actually prevent disease and benefit patients.

Accepted for publication October 14, 1999.

This research was supported by a grant from the Center for Health Management Research of the National Science Foundation and University of Washington, Seattle, and by grant LM05545 from the National Library of Medicine, Bethesda, Md.

We are indebted to the following colleagues for statistical advice and constructive assistance: Jane C. Johnson,