Original Investigations JAMIA

Research Paper ■

A Meta-analysis of 16 Randomized Controlled Trials to Evaluate Computer-based Clinical Reminder Systems for Preventive Care in the Ambulatory Setting

STEVEN SHEA, MD, WILLIAM DUMOUCHEL, PHD, LISA BAHAMONDE, BA

Abstract **Objective:** Computer-based reminder systems have the potential to change physician and patient behaviors and to improve patient outcomes. We performed a meta-analysis of published randomized controlled trials to assess the overall effectiveness of computer-based reminder systems in ambulatory settings directed at preventive care.

Design: Meta-analysis.

Search Strategy: Searches of the Medline (1966–1994), Nursing and Allied Health (1982–1994), and Health Planning and Administration (1975–1994) databases identified 16 randomized, controlled trials of computer-based reminder systems in ambulatory settings.

Statistical Methods: A weighted mixed effects model regression analysis was used to estimate intervention effects for computer and manual reminder systems for six classes of preventive practices.

Main Outcome Measure: Adjusted odds ratio for preventive practices.

Results: Computer reminders improved preventive practices compared with the control condition for vaccinations (adjusted odds ratio [OR] 3.09; 95% confidence interval [CI] 2.39–4.00), breast cancer screening (OR 1.88; 95% CI 1.44–2.45), colorectal cancer screening (OR 2.25; 95% CI 1.74–2.91), and cardiovascular risk reduction (OR 2.01; 95% CI 1.55–2.61) but not cervical cancer screening (OR 1.15; 95% CI 0.89–1.49) or other preventive care (OR 1.02; 95% CI 0.79–1.32). For all six classes of preventive practices combined the adjusted OR was 1.77 (95% CI 1.38–2.27).

Conclusion: Evidence from randomized controlled studies supports the effectiveness of data-driven computer-based reminder systems to improve prevention services in the ambulatory care setting.

■ **JAMIA.** 1996;3:399–409.

Affiliations of the authors: Division of General Medicine, Department of Medicine, College of Physicians and Surgeons, Columbia University, New York, NY (SS, LB); Division of Epidemiology, Columbia University, School of Public Health, New York, NY (SS); Department of Medical Informatics, College of Physicians and Surgeons, Columbia University, New York, NY (SS, WD); Division of Biostatistics, Columbia University School of Public Health, New York, NY (WD).

Supported by a grant from The Commonwealth Fund.

Correspondence and reprint requests to: Steven Shea, MD, Division of General Medicine, 622 West 168th Street, New York, NY, 10032.

Received for publication: 5/21/96; accepted for publication: 7/15/96.

The potential for computers to improve clinical care has been appreciated for many years, but the gradual development of effective technology has only recently made high-performance systems with rapid response times affordable and acceptable to potential users.[1,2] One strategy for using computers to improve clinical care has relied on user-driven programs that provide consultation to clinicians seeking assistance with complex diagnostic or management issues.[3-5] Users evoke these programs, input some or all of the necessary clinical data, and receive consultative output. A second strategy has been to develop data-driven programs that provide reminders in less complex but more frequently occurring clinical situations without requiring users to input new data or to evoke the program.[6-8] These passive systems are triggered by conditions in a clinical database that satisfy specified logic.[6-8] One data-driven reminder system, triggered by computer-based order entry, has been tested in the inpatient setting, where a controlled study found a reduction in length of stay, diagnostic test charges, pharmacy charges, and overall costs.[9] In the ambulatory setting, the same system, again triggered by computer-based order entry, reduced diagnostic test ordering and ancillary testing costs when physicians were shown prior test results,[10] computer-generated predictions of abnormal results,[11] or test prices.[12]

An important clinical application of data-driven systems in the ambulatory setting is the area of preventive care. Evidence-based standards for preventive services have been published[13,14] and are widely accepted, but extensive data indicate that the use of mammography, Pap smears, vaccinations, and other preventive services falls short of their full public health potential.[15] Interest in the potential of computer-based reminder systems to improve preventive services has led to a number of controlled trials testing the effects of these systems. These studies have been reviewed[1,2] but not summarized using meta-analysis. We performed a meta-analysis of these studies in order to assess the current state of the evidence concerning the effectiveness of these systems and to provide a summary estimate of the magnitude of the effect of computer-generated reminders across studies. Two specific issues were of particular interest. First, we used an extension of conventional meta-analytic methods to combine different categories of preventive services to provide a pooled estimate using all available data of the effect of computer-based reminder systems on physicians' actions in implementing preventive services. Second, we sought to identify sources of heterogeneity among these studies and to assess their effects on the findings.[16]

Methods

Search Strategy

The Medline (1966–December 1995), Nursing and Allied Health (1982–October 1995), and Health Planning and Administration (1975–November 1995) databases were searched using the key phrase "reminder systems." This strategy yielded 90 articles. Examination of these articles for references missed by the database search produced 28 more articles. Additional database searches were performed using the following key words: software, computers, ambulatory care, preventive health services, primary prevention, HMO, family practice, professional practice, attitude to computers, automatic data processing, primary health care, and decision support systems/management. No additional pertinent references were obtained from these searches.

The 118 articles were reviewed to determine relevance. Studies of noncomputer-based reminder systems—such as telephone calls, postcards, questionnaires, chart stamps, screening credit cards, medication reminders, and visual reminders—were excluded unless these types of reminders were studied in relation to computer-generated reminders. Criteria defining randomization and the use of controls, developed before this review, were then applied. We defined randomization as the selection of subjects for intervention and control conditions using a randomization process. We defined "controlled" as a study in which the intervention group was compared with a concurrent control group. We included only studies with a control group that received no intervention (e.g., studies comparing computer reminders with patient reminder cards but not with any true control group were excluded). Studies using *only* historical controls or before—after designs were excluded. Studies with concurrent controls that *also* reported comparisons with historical controls were included,[21,29,38] and both types of comparisons were included in the analysis. Multiple reports of single studies were combined, as were multiple preventive practices reported in a single publication, as described below. Sixteen separate randomized, controlled studies[17-38] of computer-based reminder systems for preventive services in ambulatory settings were identified (Table 1).

These 16 studies were heterogeneous with respect to the preventive services targeted by the reminder systems (Table 1). We therefore grouped these preventive services into six categories (Table 2). The studies were also heterogeneous with respect to how intervention

Table 1 ■

Characteristics of 16 Randomized Controlled Studies Included in the Meta-analysis

Author*	Dates Conducted	Setting	Intervention	Preventive Services
Barnett[17]	1975–77	Harvard Community Health Plan (HMO) (Boston, MA)	Patients	Hypertension follow-up
McDonald[18,19]	1977–79	University Hospital, General Medicine Practice (Indianapolis, IN)	Physicians	Influenza vaccine Pneumococcal vaccine Mammography Fecal occult blood test Pap smear PPD
Tierney[20]	1983–84	University Hospital, General Medicine Practice (Indianapolis, IN)	Physicians	Pneumococcal vaccine Mammography Fecal occult blood test Pap Smear PPD
Turner[21]	1984	University Hospital, General Medical Practice (Philadelphia, PA)	Physicians	Blood pressure check Influenza vaccine Tetanus vaccine Pneumococcal vaccine Fecal occult blood testing Rectal exam Glaucoma screening
McDowell[22] McDowell[23] Rosser[24] McDowell[25] Rosser[26]	1985–86	University Hospital, Family Medicine Practice (Ottawa, Ontario)	Physicians Patients‡	Blood pressure check Pap smear Tetanus vaccine Influenza vaccine Assess smoking status
Chambers[27]	1986–87	University Hospital, Family Practice Center (Philadelphia, PA)	Patients	Mammography
Becker[28]	1986–87	University Hospital, General Medicine Practice (Charlottesville, VA)	Physicians Patients‡	Blood pressure check Fecal occult blood test Tetanus vaccine Influenza vaccine Mammography Pap smear Glaucoma screening Dental screening Pneumococcal vaccine
Ornstein[29]	1988–89	University Hospital, Family Practice Center (Charleston, SC)	Physicians‡ Patients	Cholesterol screening Fecal occult blood test Mammography Pap smear Tetanus vaccine
Litzelman[30]	1989	University Hospital, General Medicine Practice (Indianapolis, IN)	Physicians	Mammography Fecal occult blood test Pap smear
McPhee[31] Fordham[32]	1989†	University Hospital, General Medicine Practice (San Francisco, CA)	Physicians	Clinical breast exam Mammography Fecal occult blood test Rectal exam Sigmoidoscopy Pelvic exam Pap smear
Szilagyi[33]	1989	University Hospital, Pediatric Clinic (Rochester, NY)	Patients	Influenza vaccine
Landis[34]	1990	Family Health Center (Ashville, NC)	Physicians	Mammography
Moran[35]	1990	Urban Health Center (Dorchester, MA)	Patients	Influenza vaccine
McPhee[36]	1991†	Community-based practices primary care	Physicians	Clinical breast exam Mammography Fecal occult blood test Rectal exam Sigmoidoscopy Pelvic exam Pap smear Smoking assessment Smoking counseling Dietary assessment Dietary counseling

Table 1 continued on page 402.

Table 1 ■ *(Continued)*

Author*	Dates Conducted	Setting	Intervention	Preventive Services
Burack[37]	1994†	Five separate clinical sites: 2 health department clinics 1 HMO practice 2 Outpatient practice sites of a private hospital (Detroit, MI)	Patients	Mammography
Frame[38]	1994†	Rural multiple-office family practice (Dansville, NY)	Families (Patients)	Hypertension screening Cholesterol screening Tetanus vaccine Fecal occult blood test Pap smear Mammography Clinical breast exam Teach self-examination Assess smoking status Weight Teach reporting of postmenopausal bleeding

*Superscript refers to number in reference list.
†Date of publication; dates study conducted not given.
‡Primary unit of randomization.

strategies involving computer reminders, manual reminders, or both were compared with each other and with a control group. We therefore considered four intervention conditions (computer reminders, manual reminders, both, and control) and two sampling designs (random allocation of physicians together with their patient panels, or of individual patients). In addition, whether the control group was concurrent or based on a before–after comparison was also considered in the analysis. Each of these aspects of heterogeneity was represented in our overall model in order to estimate the effect of each intervention strategy.

Each data point (Table 3) in the overall model consisted of two quantities: the number of subjects in a group and a proportion reflecting the success rate for a preventive outcome in that group. The 16 studies each contributed a variable number of data points to the analysis (Table 3, second column). For example, a group receiving computer reminders with a reported success rate for influenza vaccine contributed one data point. If success rates for that group were reported for multiple preventive practices (e.g., pneumococcal vaccine and mammography as well as influenza vaccine), then multiple data points were contributed for that group, even if the preventive practices were within the same category. If the study design compared baseline and follow-up time periods for a group of subjects, then that group contributed two data points for each preventive practice reported. The correlations (non-independence) of these data points was reflected in the covariance matrix of the overall model and considered in the statistical adjustment procedure. The control group in each study contributed its own data point.

The third column of Table 3 shows the median number of patients per data point for each of the 16 studies, regardless of whether the study randomized patients or physicians. This distinction was represented through specification of an additional variable in the overall model. Number of Intervention Conditions (fourth column) refers to computer reminders, manual

Table 2 ■

Categories of Preventive Services

Category	Specific Preventive Service
Vaccinations	Influenza vaccine Pneumococcal vaccine Tetanus vaccine
Breast cancer screening	Clinical breast exam Mammography
Cervical cancer screening	Pelvic exam Pap smear
Colorectal cancer screening	Rectal exam Fecal occult blood test Sigmoidoscopy
Cardiovascular risk reduction	Blood pressure check Hypertension follow-up Smoking assessment Smoking counseling Dietary assessment Dietary counseling Cholesterol screening
Other preventive services	Glaucoma screening TB skin test (PPD) Dental screening Teach self-examination (skin, testicular, and breast cancer) Weight Teach reporting of postmenopausal bleeding

Table 3 ■

Summary of the Data Taken From Each of the 16 Studies

Study	Number of Data Points Contributed by Each Study	Median Number of Patients Per Data Point	Number of Intervention Conditions	Number of Preventive Services*	Number of Patient Groups	Number of Time Periods Considered	Location in Cited Papers of Data: n = number of patients; p = success rate (rate of completion of preventive maneuver)
Barnett[17]	2	57	2	1 (1)	2	1	Table 1 (24 mo. data)
McDonald[18,19]	20	1773	2	10 (5)	2	1	Table 1 and Figure 2
Tierney[20]	20	409	4	5 (5)	4	1	Figures 3 and 4 (Hemoccult, pneumococcal vaccine, TB skin test, Pap smear, mammogram only; n assumed equal for all intervention arms.)
Turner[21]	30	67	4	5 (5)	3	2	n from Table 3, p based on physician averages
McDowell[22,23,24] Rosser[24,26]	25	1059	3	5 (3)	5	1	p from Table 2, n from text. Family data and efficacy analyses omitted.
Chambers[27]	2	571	2	1 (1)	2	1	Table 2
Becker[28]	24	77	3	8 (5)	3	1	Table 4
Ornstein[29]	40	1030	4	5 (5)	4	2	Table 4
Litzelman[30]	6	1460	2	3 (3)	2	1	Table 3 (all physicians)
McPhee[31] Fordham[32]	21	(197)**	3	7 (3)	3	1	Figure 2 (Control group), Table 4 (n, p imputed to match the effect sizes and significance levels of the reported analysis of covariance)
Szilagyi[33]	3	63	2	1 (1)	3	1	Table 1
Landis[34]	4	32	4	1 (1)	4	1	Table 1
Moran[35]	3	136	2	1 (1)	3	1	Table 1 (overall)
McPhee[36]	22	(5073)**	2	9 (5)	2	1	Values of n and p imputed to match the effect sizes and significance levels of the reported analysis of covariance
Burack[37]	2	1034	2	1 (1)	2	1	Table 3 (all sites)
Frame[38]	44	634	3	11 (6)	2	2	Table 4
TOTALS	268		44	(51)	46	19	

*Number in parentheses is the number of specific preventive services listed in Table 2 that were represented in the study.
**Based on the imputed data set rather than actual reported design.

reminders, both, or control. The Number of Preventive Practices (fifth column) refers to the specific preventive practices listed in the right-hand column of Table 2, and the number in parentheses refers to the number of groups of preventive practices included in the study (range one to six).

The number of patient groups (Table 3, sixth column) refers to the number of distinct groups of patients in a study. For example, one group may have received computer reminders and also served as its own historical control, thereby being counted as one group, although contributing two data points, (column 2) and two intervention conditions (column 4). The number of time periods considered (seventh column) was one if contemporaneous controls were used, two if historical controls were compared.

Statistical Methods

We used a weighted mixed effects model regression analysis (Table 4) to analyze the observed success rates, where success refers to completion of the preventive maneuver. The response variable was the logit of the success rate p. The residual variance for each point was assumed to be $1/[np(1 - p)]$, where n is the number of patients. The six categories of preventive services and the four intervention conditions (including control) were entered into the model as fixed effects. The interaction of preventive service category with intervention (24 levels), the main effect of the study (16 levels), patient groups within the study (46 levels), interventions within the study (44 levels), preventive service categories within the study (57 levels), and time periods within the study (19 levels) were entered into the model as random effects. Restricted

Table 4 ▪

The Mixed Effects Model Used to Combine the Data From the 16 Studies

Defining formula	$\gamma_{ijktm} = \alpha_m + \beta_k + \gamma_{mk} + a_i + b_{j(i)} + c_{k(i)} + d_{m(i)} + e_{t(i)} + \varepsilon_{ijktm}$
Factor levels	*Outcome m* = vaccinations, screen breast, screen cervix, screen colon, screen CV risk, Other prev. meas. *Intervention k* = control, computer reminder, manual reminder, both computer and manual reminders *Study i* = one of 16 studies *Cohort j* = separate groups of patients within a study, not always randomized (at most 5 cohorts per study) *Period t* = at most two periods of observation per *Cohort*, each with a different *Intervention* status
Variable/definitions	y_{ijktm} = logit(p_{ijktm}) = $\log_e[p_{ijktm}/(1 - p_{ijktm})]$ p_{ijktm} = observed proportion of eligible patients provided *Outcome m* in *Cohort j*, *Intervention k*, *Period t* of *Study i* n_{ijktm} = number of eligible patients used to determine p_{ijktm} α_m = base rate (logit scale) for *Outcome m* β_k = log odds ratio of effect of *Intervention k* for the "average" *Study* (= 0 for Control) and for the "average" *Outcome* γ_{mk} = deviation from average of *Intervention k* effect for *Outcome m* a_i = main effect of *Study i* on logit(p_{ijktm}) $b_{j(i)}$ = effect of *Cohort j* within *Study i* on logit(p_{ijktm}) $c_{k(i)}$ = study-specific effect of *Intervention k* within *Study i* on logit(p_{ijktm}) (for all outcomes) $d_{m(i)}$ = study-specific effect on *Outcome m* within *Study i* $e_{t(i)}$ = time *Period t* deviation in logit(p_{ijktm}) within *Study i* ε_{ijktm} = sampling error for logit(p_{ijktm}). Variance of ε_{ijktm} assumed equal to $\{n_{ijktm} \max[.01, p_{ijktm}(1 - p_{ijktm})]\}^{-1}$
REML variance estimates (all effects assumed independent normally distributed with mean 0)	$V(\gamma_{mk}) = 0.0585$ $V(a_i) = 0.1481$ $V(b_{j(i)}) = 0.0207$ $V(c_{k(i)}) = 0.0641$ $V(d_{m(i)}) = 0.6133$ $V(e_{t(i)}) = 0.0257$

maximum likelihood (REML) was used to estimate the six variance components using the PROC MIXED program (SAS Inc., 1992). Coefficients and standard errors from the model were used to estimate the odds ratio, its 95% confidence interval, and attained significance level for each combination of preventive service category and intervention (compared with control). We also used the model to compare intervention strategies averaged over all preventive service categories. For each study, we used the model to compute the estimated odds ratio comparing interventions with control averaged overall preventive service categories. To check for publication bias, we graphed the log odds ratio estimates for each study against their standard errors and computed their rank correlation coefficient. To estimate the effect, if any, of study design, we ran further analyses with additional fixed-effect terms in the model, identifying studies that randomized patients, studies that randomized providers, and studies that used more than one time period. We computed F-tests for whether these terms interacted with an intervention condition.

Results

The findings of the 16 studies were uncorrelated with the sample sizes in the studies. Specifically, the magnitude of the odds ratios was not significantly correlated with the standard errors of the natural logarithm of the odds ratios (Kendall's R = 0.16; N = 26; P = 0.24) (Fig 1).

Comparisons of Reminder Systems versus Control Group

Computer reminders increased preventive practices compared with a control group for four of the six groups of preventive practices, including vaccinations, breast cancer screening, colorectal cancer screening, and cardiovascular risk reduction, and for all six practices combined (OR 1.77; 95% CI 1.38–2.27) (Table 5). Manual reminders increased preventive practices compared with a control group for the same four groups of preventive practices and for all six practices combined (OR 1.57; 95% CI 1.20–2.06). Computer plus manual reminders increased preventive practices compared with a control group for all six groups of preventive practices and for all six combined (OR 2.23; 95% CI 1.67–2.98). Thus, both methods of generating reminders were effective overall. Both methods had the greatest effect on vaccinations, somewhat smaller effects on colorectal cancer screening and cardiovascular risk reduction, and lesser effects on breast and cervical cancer screening. Overall, and over the time periods studied, computer reminders increased preventive practices by 77% compared with a control group.

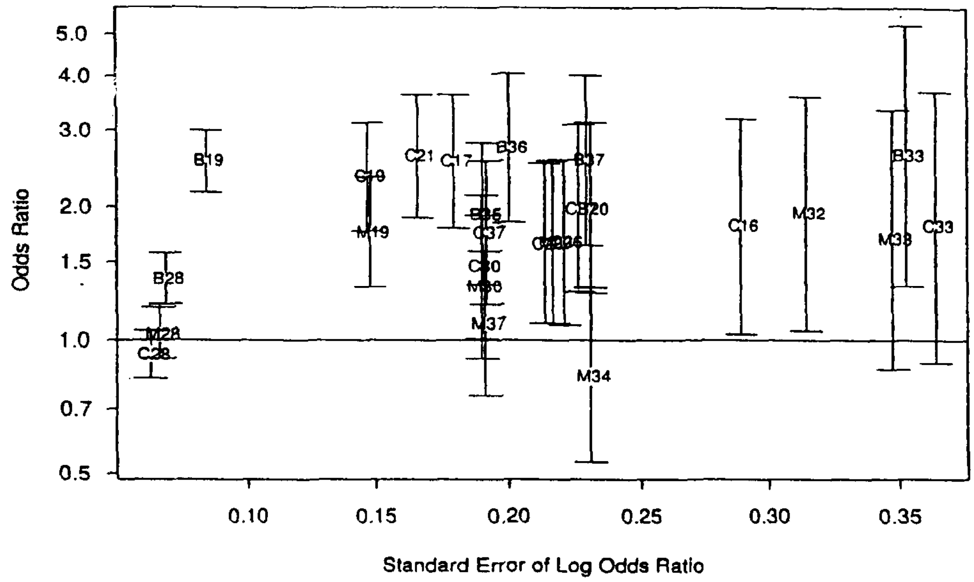

Figure 1 This figure shows the 95% non-simultaneous confidence intervals for odds ratio of intervention vs. control. Intervals are labeled with a letter denoting the intervention (C: Computer, M: Manual, B: Both) and the number of the study in the list of references. If there were a publication bias against nonsignificant studies, one would expect studies having larger standard errors to have larger odds ratios. No such trend is evident in the plot.

Computer versus Manual Reminders

As shown in Table 5, the 95% confidence intervals for computer reminder versus control and for manual reminder versus control overlapped for all six preventive practices, as well as for the six practices combined. A direct comparison of computer reminder versus manual reminder was done because of the statistical nonequivalence of these confidence intervals to the P values for direct comparisons. None of these direct comparisons of computer versus manual reminders showed a statistically significant difference (adjusted OR 1.13; 95% CI 0.85–1.49; P = 0.4 for all six preventive categories combined).

Computer Plus Manual Reminders versus Computer Reminders

As Table 5 also shows, the 95% confidence intervals for these treatment levels overlapped for five of the six preventive practices, the exception being other preventive care. The 95% confidence intervals also overlapped for the six practices combined. A direct comparison of these treatment conditions confirmed this pattern (adjusted OR 1.26; 95% CI 0.93–1.70; P = 0.13 for all six preventive categories combined). Thus, the data were most consistent with at most a small benefit from combining the two intervention strategies compared with computer reminders alone.

Table 5 ■

Odds Ratios and 95% Confidence Intervals for the Effect of Three Interventions versus Control Groups on Each of Six Classes of Preventive Services and From a Combined Analysis That Weights Each Outcome Equally

Intervention	Computer Reminder		Manual Reminder		Computer Plus Manual	
Preventive service measured	Odds ratio (95% CI)	P Value	Odds ratio (95% CI)	P Value	Odds ratio (95% CI)	P Value
Vaccinations	3.09 (2.39–4.00)	< .0001	2.46 (1.86–3.25)	< .0001	3.06 (2.25–4.16)	< .0001
Breast cancer screening	1.88 (1.44–2.45)	< .0001	1.63 (1.21–2.18)	.001	1.88 (1.44–2.45)	< .0001
Cervical cancer screening	1.15 (0.89–1.49)	.3	1.10 (0.82–1.46)	.5	1.12 (0.82–1.51)	.5
Colorectal cancer screening	2.25 (1.74–2.91)	< .0001	1.85 (1.39–2.47)	< .0001	2.71 (2.01–3.66)	< .0001
Cardiovascular risk reduction	2.01 (1.55–2.61)	< .0001	1.86 (1.41–2.47)	< .0001	2.57 (1.89–3.51)	< .0001
Other preventive care	1.02 (0.79–1.32)	.9	.99 (0.71–1.37)	.9	2.59 (1.73–3.86)	< .0001
Combined outcomes	1.77 (1.38–2.27)	< .0001	1.57 (1.20–2.06)	.001	2.23 (1.67–2.98)	< .0001

Estimates are computed using the mixed effects model described in the text (see Methods) and specified in Table 4. The odds ratio for *Outcome m, Intervention k* is $\exp(\beta_k + \gamma_{mk} - \gamma_{m1})$, where $k = 1$ denotes the control arm of a study. The odds ratio for the combined outcomes is $\exp[\beta_k + \Sigma_m(\gamma_{mk} - \gamma_{m1})/6]$.

Computer Plus Manual Reminders versus Manual Reminders

The adjusted OR for this comparison was 1.42 for all six preventive categories combined (95% CI 1.02, 1.97; P = 0.04), supporting an additional benefit from combining the two intervention strategies compared with manual reminders alone.

Sources of Heterogeneity

The degree of variability in results among the 16 studies is shown in Table 6. These odds ratios and confidence intervals were computed for each study from the overall model. While the magnitude of the intervention effect varied among studies, there was remarkable consistency across studies in the positive effects found. This is also shown in Figure 1. The differences in effect of reminder systems among different classes of preventive practices was noted earlier.

To test whether the intervention effect was significantly influenced by study design, we examined these effects in the overall model. There was a trend toward a larger effect size being described in studies using comparisons to historical control groups (d.f. = 3,240; F = 2.00; P = 0.1), possibly reflecting the additive effect of secular trends favoring more preventive services or contamination effects in the comparisons with concurrent control groups. Whether the study randomized patients or providers did not affect the size of the intervention effect (d.f. = 3,240; F = 0.24; P = 0.9).

Discussion

This overview of 16 randomized trials supports the effectiveness of computer-generated reminders as a method for increasing preventive services in the ambulatory setting. The overall increase in the odds ratio attributable to computer-generated reminders compared with the control condition across different preventive services was 77%. The interpretation of this effect in terms of an absolute increase in delivery of preventive services to patients will depend on the baseline prevalence of compliance with the recommended preventive service. For example, if the baseline is 50%, this increase in the odds ratio implies an increase to 64%; for a baseline of 75%, the implied increase is to 84%.

Computer-generated reminders were effective for increasing vaccinations, breast cancer screening, colorectal cancer screening, and cardiovascular risk reduction, but they were not effective for increasing cervical cancer screening or the other six specific forms of preventive care examined. This selectivity with respect to preventive practices is unlikely to be a characteristic of the computer-based intervention strategy, since the same pattern of results was found for manual reminders compared with the control condition. We speculate that patient resistance to pelvic examination and the time required for the primary care practitioner to perform a pelvic examination may be the limiting factors in relation to Pap smears.

In order to examine the overall effect of computer-generated reminders on preventive care, we combined studies that were heterogeneous with respect to several features, including the specific preventive maneuvers studied and some aspects of study design. The mixed effects model analytic method, in which terms representing both fixed (homogeneous) and random (heterogeneous) effects were included, permitted direct testing of the effects of these sources of heterogeneity. As noted, there were differences in effect across classes of preventive practice. Heterogeneity in study design did not appear to have a significant influence on the estimates of the intervention effect. We also examined the relationship between the size of the effect estimate and the sample size, a method of assessing for publication bias,[39] but we found no evidence to suggest that bias arising from exclusion of small negative studies was present.

Several limitations of our meta-analysis should be considered in interpreting the findings. We grouped preventive maneuvers by disease or organ system rather than by behavioral characteristics that might be associated with the likelihood of the maneuver being completed. We took this approach to categorization because it is widely used[13–15] and because, to our knowledge, there is no accepted alternative. The impact of manual reminder systems on physician completion of recommended preventive maneuvers was approximately equivalent to the impact of computer-generated reminders in the studies included in our analysis. This finding suggests that computer-generated reminders are as acceptable to physicians and other primary care providers as reminders generated by other sources; it also suggests that, in choosing a method for generating reminders, issues of cost, sustainability, and auditability will dominate. We interpret this finding with some caution, however, because our literature search was designed to be comprehensive for randomized controlled trials of computer-generated reminders for preventive care, not for such trials of manual reminders. We believe the complexity of the statistical model is a strength rather than a limitation of the analysis because a mixed-effects model permits an assessment of the overall effect of computer-based reminders for preventive care across a

Table 6 ■

Odds Ratios and 95% Confidence Intervals Comparing Interventions versus Control for Each Study, Adjusted for Specific Class of Preventive Service (Comparisons Not Directly Addressed in the Studies Are Shown as Open Spaces)

Intervention	Computer Reminder		Manual Reminder		Computer Plus Manual	
Study	Odds ratio (95% CI)	P Value	Odds ratio (95% CI)	P Value	Odds ratio (95% CI)	P Value
Barnett[17]	1.82 (1.03–3.20)	.04				
McDonald[18,19]	2.55 (1.79–3.63)	< .0001				
Tierney[20]	2.35 (1.76–3.13)	< .0001	1.75 (1.31–2.34)	< .0001	2.55 (2.16–3.00)	< .0001
Turner[21]	1.65 (1.09–2.52)	.02	1.66 (1.09–2.55)	.02	1.99 (1.27–3.14)	.003
McDowell[22–26]	2.62 (1.89–3.63)	< .0001	3.43 (2.59–4.53)	< .0001		
Chambers[27]	1.67 (1.08–2.57)	.02				
Becker[28]	1.99 (1.28–3.11)	.002			2.57 (1.64–4.03)	.0001
Ornstein[29]	0.93 (0.82–1.05)	.2	1.03 (0.91–1.18)	.6	1.37 (1.20–1.57)	.0001
Litzelman[30]						
McPhee[31,32]	1.47 (1.01–2.13)	.04	1.32 (0.92–1.92)	.1		
Szilagyi[33]			1.95 (1.05–3.60)	.03		
Landis[34]	1.82 (0.89–3.70)	.1	1.70 (0.86–3.36)	.1	2.64 (1.32–5.26)	.01
Moran[35]			0.83 (0.53–1.31)	.4		
McPhee[36]					1.93 (1.33–2.81)	< .001
Burack[37]					2.74 (1.84–4.06)	.0001
Frame[38]	1.75 (1.20–2.55)	.03	1.09 (0.75–1.59)	.6		

The study by Litzelman et al.[30] compared computer reminder vs. computer plus manual reminder. It therefore did not contribute information directly to the specific comparisons shown in this table. It was, however, included in the overall modeling procedure from which the estimates in both this table and Table 5 were derived. This study found that computer plus manual reminder compared with computer reminder alone improved preventive services with an adjusted odds ratio of 1.48 (95% CI 0.84–2.37; p = .19).

group of studies that are distinctly heterogeneous with respect to aspects of study design as well as the specific preventive maneuvers chosen as the study outcomes.

The computer-reminder systems examined in the 16 studies share a design philosophy in which the system uses clinical data, an expert system that monitors the clinical data base, and a knowledge base in which the logic that triggers the reminders is represented. The physician or other provider remains passive, and the reminders are data generated rather than sought by the user. This design addresses the seminal insight of McDonald[6] that even the best-trained clinicians have a measurable rate of omitting to do things that they know they should do. While the findings of this meta-analysis also support the use of manual reminders as a strategy for increasing preventive services, adding computer-generated reminders to manual reminders led to a significantly greater increase, while adding manual reminders to computer-generated reminders did not.

The findings imply that the physicians and other providers targeted by the computer-generated reminders accepted the recommendations implicit in the alerts. Preventive services is an area in which, despite many areas of ongoing controversy, a consensus exists regarding a substantial number of practices.[13,14] In addition, there is good evidence that many preventive services are underused, particularly by patients who are poor or members of racial minorities.[15] These groups constitute a disproportionate share of the patients cared for in the ambulatory facilities of academic health centers, where many of the 16 studies were conducted. System users must agree that the content of the computer-generated patient management advice is medically appropriate. Based on randomized[1,8–12,40,41] and nonrandomized[42] trials of computer-generated alerts in other medical contexts, it appears that, in extending this strategy to other aspects of managing medical costs and quality, it is more important that users reach a consensus on appropriate guidelines than that they accept computers as a way of delivering reminders.[43,44]

Formidable technical issues also must be addressed before computer-based reminder systems can have widespread use in health care systems. These include difficulties in capturing the necessary clinical data[45]; the need for standards for coded medical vocabulary, medical logic frames, and clinical and medical knowledge databases[46]; confidentiality and data security[47]; legal issues[48]; and the capital and operating costs of such systems.[49,50] Nonetheless, such systems have now progressed to the proof-of-concept stage, at least in the ambulatory setting with regard to preventive care.

References ■

1. Johnston ME, Langston KB, Haynes RB, Mathieu A. Effects of computer-based decision support systems on clinician performance and patient outcome. A critical appraisal of the literature. Ann Intern Med. 1994;120:135–42.
2. Sullivan F, Mitchell E. Has general practitioner computing made a difference to patient care? A systematic review of published reports. BMJ. 1995;311:848–52.
3. Leaper DJ, Horrocks JC, Staniland JR, deDombal FT. Computer-assisted diagnosis of abdominal pain using "estimates" provided by clinicians. BMJ. 1972;4:350–4.
4. Miller RA, Pople HE Jr, Myers JD. INTERNIST-1, an experimental computer-based diagnostic consultant for general internal medicine. N Engl J Med. 1982;307:468–76.
5. Baxt WG, Skora J. Prospective validation of artificial neural network trained to identify acute myocardial infarction. Lancet. 1996;347:12–15.
6. McDonald CJ. Protocol-based computer reminders, the quality of care and the non-perfectibility of man. New Engl J Med. 1976;295:1351–5.
7. Pryor TA, Gardner RM, Clayton PD, Warner HR. The HELP systems. J Med Syst. 1983;7:87–102.
8. Hripcsak G. The Columbia–Presbyterian Medical Center decision support system as a model for implementing the Arden Syntax. SCAMC Proc. 1991;248–51.
9. Tierney WM, Miller ME, Overhage JM, McDonald CJ. Physician inpatient order writing on microcomputer workstations: Effects on resource utilization. JAMA. 1993;269:379–83.
10. Tierney WM, McDonald CJ, Martin DK, Hui SL, Rogers MP. Computerized display of past test results: effects on outpatient testing. Ann Intern Med. 1987;107:569–74.
11. Tierney WM, McDonald CJ, Hui SL, Martin DK. Computer predictions of abnormal test results: effects on outpatient testing. JAMA. 1988;259:1194–8.
12. Tierney WM, Miller ME, McDonald CJ. The effect on test ordering or informing physicians of the charges for outpatient diagnostic tests. N Engl J Med. 1990;322:1499–1504.
13. Guide to clinical preventive services, 2nd ed. Report of the U.S. Preventive Services Task Force. Baltimore: Williams & Wilkins, 1996.
14. Sox HC, Jr. Preventive health services in adults. N Engl J Med. 1994;330:1589–95.
15. U.S. Department of Health and Human Services. Healthy people 2000: National health promotion and disease prevention objectives. DHHS Publication No. (PHS) 91-50213. Washington, DC: U.S. Government Printing Office.
16. Colditz GA, Burdick E, Mosteller F. Heterogeneity in meta-analysis of data from epidemiology studies: a commentary. Am J Epidemiol. 1995;142:371–82.
17. Barnett G, Winickoff R, Morgan M, Zielstorff R. A computer-based monitoring system for follow-up of elevated blood pressure. Med Care. 1983;21:400–9.
18. McDonald C, Hui S, Smith D, et al. Reminders to physicians from an introspective computer medical record. Ann Intern Med. 1984;100:130–8.
19. McDonald C, Hui S, Tierney W. Effects of computer reminders for influenza vaccination on morbidity during influenza epidemics. MD Comput. 1992;9:304–12.
20. Tierney W, Hui S, McDonald C. Delayed feedback of physician performance versus immediate reminders to perform prevention care. Med Care. 1986;24:659–66.
21. Turner B, Day S, Borenstein B. A controlled trial to improve delivery of preventive care: Physician or patient reminders. J Gen Intern Med. 1989;4:403–9.
22. McDowell I, Newell C, Rosser W. A randomized trial of computerized reminders for blood pressure screening in primary care. Med Care. 1989;27:297–305.
23. McDowell I, Newell C, Rosser W. Computerized reminders to encourage cervical screening in family practice. J Fam Pract. 1989;28:420–4.
24. Rosser W, Hutchison B, McDowell I, Newell C. Use of reminders to increase compliance with tetanus booster vaccination. Can Med Assoc J. 1992;146:911–17.
25. McDowell I, Newell C, Rosser W. Comparison of three methods of recalling patients for influenza vaccination. Can Med Assoc J. 1986;135:991–7.
26. Rosser W, McDowell I, Newell C. Use of reminders for prevention procedures in family medicine. Can Med Assoc J. 1991;145:807–13.
27. Chambers C, Balaban D, Lepidus Carlson B, Ungemack J, Grasberger D. Microcomputer-generated reminders: improving the compliance of primary care physicians with mammography screening guidelines. J Fam Pract. 1989;29:273–80.
28. Becker D, Gomez E, Kaiser D, Yoshihasi A, Hodge R. Improving prevention care at a medical clinic: How can the patient help? Am J Prev Med. 1989;5:353–9.
29. Ornstein S, Garr D, Jenkins R, Ruts P, Arnon A. Computer-generated physician and patient reminders: Tools to improve population adherence to selected prevention services. J Fam Pract. 1991;32:82–90.
30. Litzelman DK, Dittus R, Miller M, Tierney W. Requiring physicians to respond to computerized reminders improves their compliance with prevention care protocols. J Gen Intern Med. 1993;8:311–17.
31. McPhee S, Bird JA, Jenkins C, Fordham D. Promoting cancer screening: a randomized, controlled trial of 3 interventions. Arch Intern Med. 1989;149:1866–72.
32. Fordham D, McPhee S, Bird JA, Rodnick J, Detmer W. The cancer prevention reminder system. MD Comput. 1990;7:289–95.
33. Szilagyi P, Rodewald L, Savageau J, Yoos L, Doane C. Improving influenza vaccination rates in children with asthma: A test of a computerized reminder system and an analysis of factors predicting vaccination compliance. Pediatrics. 1992;90:871–5.
34. Landis S, Hulkower S, Pierson S. Enhancing adherence with mammography through patient letters and physician prompts. N C Med J. 1992;53:575–8.
35. Moran W, Nelson K, Wofford J, Velez R. Computer-generated mailed reminders for influenza immunization: A clinical trial. J Gen Intern Med. 1992;7:535–7.
36. McPhee S, Bird JA, Fordham D, Rodnick J, Osborn E. Promoting cancer prevention activities by primary care physicians. JAMA. 1991;266:538–44.
37. Burack R, Gimotty P, George J, Stengle W, Warbasse L, Moncrease A. Promoting screening mammography in inner-city settings: a randomized controlled trial of computerized reminders as a component of a program to facilitate mammography. Med Care. 1994;32:609–24.
38. Frame P, Zimmer J, Werth P, Hall J, Eberly S. Computerized-based vs. manual health maintenance tracking. Arch Fam Med. 1994;3:581–8.
39. Begg CB. Publication bias. In: Cooper H, Hedges LV (eds). The Handbook of Research Synthesis. New York: Russell Sage Foundation, 1994;399–409.
40. Shea S, Sideli RV, DuMouchel W, Pulver G, Arons RR, Clayton PD. Effect on length of hospital stay of computer-generated informational messages directed to physicians. J Am Med Inform Assoc. 1995;2:58–64.

41. Pestotnik SL, Classen DC, Evans RS, Burke JP. Implementing antibiotic practice guidelines through computer-assisted decision support: Clinical and financial outcomes. Ann Intern Med. 1996;124:884-90.
42. Safran C, Rind DM, Davis RB, et al. Guidelines for management of HIV infection with computer-based patient's record. Lancet. 1995;346:341-6.
43. Brook RH. Implementing medical guidelines. Lancet. 1995; 346:132.
44. Tunis SR, Hayward RSA, Wilson MC, et al. Internists' attitudes about clinical practice guidelines. Ann Intern Med. 1994;120:956-63.
45. Committee on Improving the Patient Record. Division of Health Care Services. Institute of Medicine. The computer-based health record: an essential technology for health care. Washington, DC: National Academy Press, 1991;81-2.
46. Jenders RA, Morgan M, Barnett GO. Use of open standards to implement health maintenance guidelines in a clinical workstation. Comput Biol Med. 1994;24:385-90.
47. Committee on Regional Health Data Networks, Division of Health Care Services, Institute of Medicine. Health data in the information age: Use, disclosure, and privacy. Washington, DC: National Academy Press, 1994.
48. Committee on Improving the Patient Record, Division of Health Care Services, Institute of Medicine. The computer-based health record: an essential technology for health care. Washington, DC: National Academy Press, 1991;156-79.
49. Clayton PD, Anderson RK, Hill C, McCormack M. An initial assessment of the cost and utilization of the Integrated Academic Information System (IAIMS) at Columbia-Presbyterian Medical Center. SCAMC Proc. 1991;109-13.
50. Heathfield HA, Wyatt J. Philosophies for the design and development of clinical decision-support systems. Methods Inf Med. 1993;32:1-8.

Statistical Appendix

Two of the studies[31,36] did not follow the paradigm assumed by the meta-analytic model in that they did not report compliance rates for cohorts of patients subjected to different interventions. Instead, data were collected at the provider level before and after intervention, and an analysis of covariance was performed on the provider-specific compliance rates. This design and analysis methodology is superior to that of the other studies for several reasons. The fact that providers are both the unit of randomization and the unit of analysis should reduce contamination of results across intervention groups. Also, the summarization of results by provider means that the correlation of outcomes of patients who have the same provider will not distort the statistical analysis, where outcome refers to whether or not the patient received the specific class of preventive service targeted by the intervention. Finally, the analysis of covariance incorporates baseline information into the analysis with a resulting reduction in uncertainty in intervention effects. In order to include these two studies in our analysis using the model shown in Table 4, we imputed values of n and p that would lead to unchanged values of intervention effects and standard errors of intervention effects. For example, consider the outcome "Rectal Exam" for the study by McPhee et al.[31] Table 3 in this report gives regression coefficients $b_{constant}$ = 16.2%, $b_{intervention}$ = 10.5%, P = 0.004 for significance of the intervention effect. The units of analysis for the regression were the n = 39 providers that were randomized. To make this result comparable with the other studies, we created values $p_{control}$, $p_{intervention}$ such that $p_{intervention} - p_{control} = .105$, with sample sizes chosen so that the difference between proportions would be significant at the (two-sided) 0.004 significance level. The formulas chosen to do this are:

$$P_{control} = b_{constant}/100$$

$$P_{intervention} = (b_{constant} + b_{intervention})/100$$

$$n_{constant} = 9999$$

$$n_{intervention} = Z^2 P_{intervention}(1 - P_{intervention})/(P_{intervention} - P_{control})^2$$

where z is the two-sided standard normal percentile corresponding to the observed significance level.

In the above example, the resulting value of $n_{intervention}$ = 148. In other words, the reported regression results are equivalent in statistical significance and effect to a study in which a known base rate = .162 in a very large population of patients is increased by .105 in a sample of 148 patients whose providers are exposed to the intervention.

In a similar fashion, the results reported by McPhee et al.[31] were converted to statistically equivalent patient counts and proportions of compliance. The actual formulas used are slightly more complicated than those used above because that study has three groups (control and two types of intervention) and because the response variables in the reported regressions are not raw percentages but have been scaled by American Cancer Society's recommended annual rates. Details of the conversion method are available from the authors.