Research Paper ■

A Randomized Controlled Trial of a Computer-based Physician Workstation in an Outpatient Setting: Implementation Barriers to Outcome Evaluation

BARRY L. ROTMAN, MD, ANDREA N. SULLIVAN, AB, THOMAS W. MCDONALD, MD, BYRON W. BROWN, PHD, PHILIPPE DESMEDT, MS, DON GOODNATURE, MS, MICHAEL C. HIGGINS, PHD, HENRI J. SUERMONDT, PHD, CHARLES YOUNG, PHD, DOUGLAS K. OWENS, MD, MSC

Abstract **Objective:** A research prototype Physician Workstation (PWS) incorporating a graphical user interface and a drug ordering module was compared with the existing hospital information system in an academic Veterans Administration General Medical Clinic. Physicians in the intervention group received recommendations for drug substitutions to reduce costs and were alerted to potential drug interactions. The objective was to evaluate the effect of the PWS on user satisfaction, on health-related outcomes, and on costs.

Design: A one-year, two-period, randomized controlled trial with 37 subjects.

Measurements: Differences in the reliance on noncomputer sources of information, in user satisfaction, in the cost of prescribed medications, and in the rate of clinically relevant drug interactions were assessed.

Results: The study subjects logged onto the workstation an average of 6.53 times per provider and used it to generate 2.8% of prescriptions during the intervention period. On a five-point scale (5 = very satisfied, 1 = very dissatisfied), user satisfaction declined in the PWS group (3.44 to 2.98 $p = 0.008$), and increased in the control group (3.23 to 3.72, $p < 0.0001$).

Conclusion: The intervention physicians did not use the PWS frequently enough to influence information-seeking behavior, health outcomes, or cost. The study design did not determine whether the poor usage resulted from satisfaction with the control system, problems using the PWS intervention, or the functions provided by the PWS intervention. Evaluative studies should include provisions to improve the chance of successful implementation as well as to yield maximum information if a negative study occurs.

■ **JAMIA.** 1996;3:340–348.

Affiliations of the authors: Veterans Affairs Palo Alto Health Care System, Palo Alto, CA (BLR, ANS, TWM, DKO); Division of General Internal Medicine, Stanford University, Stanford, CA (BLR, TWM, DKO); Department of Health Research and Policy, Stanford University, Stanford, CA (BWB, DKO) Hewlett-Packard Laboratories, Palo Alto, CA (PDS, DG, MCH, HJS, CY).

Dr. Owens is supported by a Career Development Award from the VA Health Services Research and Development Service. Drs. Rotman and McDonald were supported by the VA Ambulatory Care Fellowship Program. Supported by a grant from the Ambulatory Care Education Initiative, Department of Veterans Affairs.

Correspondence and reprints: Douglas K. Owens, MD, MSc, Section of General Internal Medicine (111A), Veterans Affairs Medical Center, 3801 Miranda Avenue, Palo Alto, CA 94304. e-mail: owens@camis.stanford.edu

Received for publication: 12/6/95; accepted for publication: 5/14/96.

Collecting, processing, and recording information during an outpatient visit is already difficult. It is growing more complicated because more health care services are delivered in the outpatient arena, an aging population has more chronic diseases, and expanding medical technologies have created more diagnostic and therapeutic options to know and use. In addition, with the advent of restrictions on governmental funding for health care and the growing prominence of managed care, health care providers are facing increasing pressure to maximize their efficiency. New clinical information systems can assist providers with this challenge by enhancing data management and providing decision support.

Prior evaluations of clinical information systems and expert systems that provide patient-specific advice have found both successes and failures.[1-6] A systematic review found that three of four studies of computer-assisted drug dosing, one of five studies of computer-aided diagnosis, seven of nine studies of computer-aided active medical care, and four of six studies of computer-assisted preventive care showed improvement in clinicians' performance.[5] Only three of ten studies that examined patient outcomes found improvements, however.[5] Information systems also reduce costs in certain settings. For example, Tierney and colleagues conducted a randomized controlled trial of computer-based feedback during order entry in a hospital that showed a reduction in drug charges.[3]

Rigorous evaluation of clinical information systems is important for two reasons.[7] Simply providing more information, or providing it in a more usable manner, may not be sufficient to improve clinical outcomes or to reduce costs. Provider behavior is notoriously difficult to influence.[8] Also, innovative systems may have unexpected limitations, even if these systems have certain clear advantages over current ones.[7] Laboratory evaluations of clinical information systems are insufficient for determining clinical utility.[9] Thus, it is essential to demonstrate that an information technology actually change health and economic outcomes.

As with studies of other medical interventions, evaluations of information systems may be subject to biases that undermine the validity of their results.[9] Two threats to validity warrant particular attention. In a study comparing two groups of users, the users in each group should be comparable in terms of skill, training, interest, and clinical needs. Otherwise, selection bias may occur because the users of the new system are those with particular a interest or aptitude. This type of selection bias is particularly relevant to evaluations of clinical information systems, and it may lead to speciously optimistic results. Randomization of users to experimental and control groups, the traditional approach for clinical evaluations, is the most suitable method for ensuring comparability. However, experimental designs must allow user access to the existing system unless the new system contains all the clinically relevant information available through the existing one. A second threat to the validity of an evaluation of clinical information systems is the secular trend: changes in the clinical environment that occur during the study period and that may affect the study outcomes. For example, the results of a study of the influence of a computer-based clinical information system on pharmacy costs may be confounded if changes in the hospital formulary during the study period reduce (or increase) drug costs. Appropriately designed control groups enable investigators to control and adjust for such confounding trends in the analysis of results.

In the Physician Workstation (PWS) project, we evaluated an innovative, computer-based PWS designed to facilitate essential information-processing activities in ambulatory care. An internally developed research prototype,[10-16] the PWS was designed based on ethnographic studies during a process of formative evaluations.[14,17] We evaluated the PWS in a two-period, parallel design that assigned subjects randomly, controlled for secular trends, and maximized statistical power. The objective of our study was to evaluate the effects of PWS on user satisfaction, on health-related outcomes, and on costs. Because the intervention system was used infrequently, we were not able to evaluate whether such an intervention, if used, would affect health or economic outcomes. This paper reports the study design, describes the implementation problems that contributed to low usage, and recommends ways to avoid these problems in evaluative studies.

Methods

Clinical Environment

We performed our study in the General Medical Clinic (GMC) at the Veterans Affairs (VA) Palo Alto Health Care System from July 1994 to June 1995. The GMC provides continuity care for a patient population of predominantly elderly men who have multiple chronic illnesses. During the study, 2,071 patients were seen in the GMC. Our study subjects were 37 Stanford University Internal Medicine residents. Each resident cared for GMC patients one half-day per week. Each examination room contained a computer terminal connected to the hospital's medical information system. For the intervention period, we added PWS terminals and printers to the examination rooms

of the intervention physicians. The internal review boards of the Palo Alto VA and the Stanford University Medical Center approved the study.

Computer Systems

The Palo Alto VA uses a comprehensive hospital information system, the Decentralized Hospital Computer Program (DHCP), which has a menu-driven, command-line interface that allows the user to choose between simple data retrieval (including laboratory tests, radiology reports, appointments, and discharge summaries) and data entry (including test and consult ordering). Also, DHCP provides comprehensive health summaries, which include all of the most recent study results, as well as medication lists, which residents use to review and renew patients' medications. The residents staffing the GMC receive a printout of these reports at each patient visit.

DHCP does not contain a mechanism for monitoring drug interactions. Physicians write prescriptions on prescription pads or renew medications by using a preprinted medication summary. These prescriptions are read by pharmacists, who can recognize potential interactions. However, the system contains no comprehensive screening to detect drug interactions.

The PWS system uses seven Hewlett-Packard 700 X-Terminals connected to two Hewlett-Packard 735 UNIX workstations, employing a client-server architecture; the PWS system uploads data from the hospital information system via a local-area Ethernet network using 10 base-T and 10 base-2 cabling.

The PWS system offers features that are not available in the current hospital information system. It presents data uploaded from the DHCP using a graphical user interface that allows data stored in separate files to be presented simultaneously in an integrated context. The system allows users to track medications, problems, and laboratory values in a graphical format that displays changes over the course of time. Using the PWS's drug-ordering module, physicians can renew, cancel, or order new medications, and they can print out prescriptions on a printer installed in their examination rooms. The module allows physicians to point and click with a mouse to select medications and instructions from a template, or to use a keyboard in order to enter free text. However, if physicians renew prescriptions or prescribe information by using the paper-based system available in the clinic, the PWS cannot provide alerts; the information from the paper-based system is not immediately available to the PWS.

Intervention

The PWS system contains features designed to reduce prescription-drug costs and to reduce the number of adverse drug interactions. To reduce costs, the PWS system contains a knowledge base that provides advice and information, such as drug costs, institutional prescription policies, and recommendations for cost-effective drug substitutions. Thus, when a provider prescribes a drug using the PWS prescription module, the PWS alerts the physician if another effective but less expensive drug is available in the formulary. For example, if a physician prescribes terfenadine for allergic rhinitis, PWS will automatically alert the provider that chlorpheniramine is available and will provide the following advisory: "Terfenadine is nonformulary. Please first try more cost-effective antihistamines (e.g., chlorpheniramine). Among nonsedating agents, astemizole is least expensive." The message appears in a yellow outlined box across the screen; the user can either acknowledge the message or erase it or receive an expanded text pertaining to the alert.

The PWS system provides alerts about potential adverse drug interactions. For example, while using the drug-ordering module to prescribe warfarin for a patient who is already taking aspirin, a physician will be warned by a yellow outlined message appearing in a box across the screen that the combination might increase the patient's risk of bleeding. Users can simply acknowledge the alert or receive additional text. Physicians receive alerts about potential drug interactions during the patient visit only if they renew medications using the PWS system; if physicians write prescriptions with the paper-based system available in the clinic, the PWS receives the updated medication list after pharmacy staff has entered the new medications into the DHCP (usually the next day). A more comprehensive description of the PWS has appeared in earlier publications.[10-17]

Study Design

We chose a two-period parallel design (Fig. 1), with the study subjects randomly divided into two groups: the PWS group and the DHCP group.[18] We used a stratified block randomization method, stratifying by year of training, and balancing the total number and percent of first-year physicians in each group. During time period 1 (July 1, 1994, to March 30, 1995), both groups used the standard hospital system, the DHCP. During time period 2 (April 1, 1995, to June 30, 1995), the PWS group used the PWS, while the DHCP group continued to use the DHCP. The study was terminated on June 30, 1995, at the end of the academic year, when one-third of the subjects completed residency

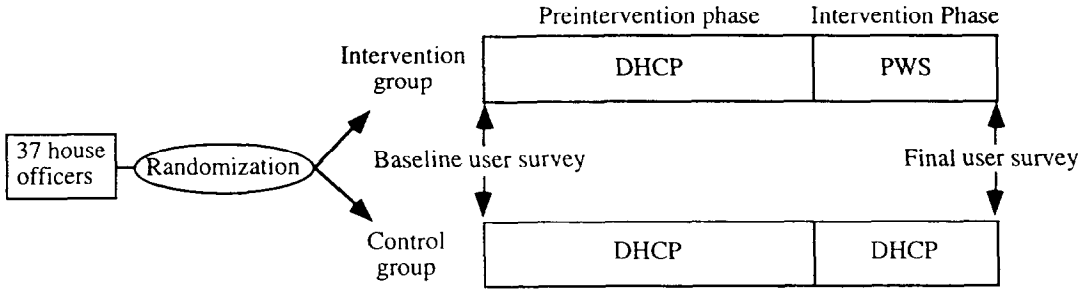

Figure 1 Design of the study. A two-period parallel design with the study subjects randomly divided into two groups, the PWS group and the DHCP group. DHCP = Decentralized Hospital Computer Program, PWS = Physician Workstation. During the nine-month preintervention phase, both groups used the standard hospital system, the DHCP. During the three-month intervention phase, the PWS group used the PWS, while the DHCP group continued to use the DHCP.

training. For each physician, we calculated the difference in the outcome variables between the two time periods.

We chose to assess the effect of the PWS system on drug costs, drug–drug interactions, and user satisfaction. We analyzed GMC prescription data from the 1992–93 academic year to select outcome variables that would be feasible for analysis and sensitive to our intervention.[18,19] Because our intervention targeted physicians' prescribing behavior, we expressed our drug cost and drug interaction variables in terms of quantities prescribed by each physician.

To derive drug costs, we performed a two-step process, generating a mean daily cost per patient and then aggregating each physicians' panel of patients to generate a mean daily cost per physician per patient. This procedure allowed us to correct for the fact that different physicians cared for different numbers of patients and for patients who were not enrolled in the clinic for the entire study. We used drug purchase prices at our institution. We calculated each patient's mean daily drug cost by dividing the patient's total prescription cost by the number of days that the patient was followed in GMC, as defined by an algorithm based on appointment and prescription data. We obtained patient-visit information, prescription data, and drug costs from the DHCP.

We determined the influence of recommendations for cost-effective drug substitutions by measuring compliance with the recommendations. We defined compliance as the proportion of prescriptions in which a provider accepted the recommended drug substitution. We calculated this proportion as the number of recommended drug prescriptions accepted by the provider divided by the total number of prescriptions for both recommended and substitutable drugs.

We developed a definition of drug–drug interactions that emphasized clinically relevant interactions. Using the Drug Therapy Monitoring System (DTMS), a commercial database developed by Medispan, we selected only level-1 interactions, the most dangerous and best-documented category of the five possible levels. Additionally, we defined a clinically relevant interaction as one with no laboratory evidence that the provider monitored the patient for adverse events related to the interaction.[19] Many interactions, even those at level 1, are inherent in the practice of medicine and may cause no morbidity if the provider carefully monitors for potential deleterious effects by performing laboratory studies. For example, if a physician generated a level-1 interaction by concurrently prescribing warfarin and allopurinol, the increased risk of bleeding would require laboratory monitoring. For this interaction, we examined the patient's laboratory data for evidence of a prothrombin-time study within 14 days of issuing the interacting prescriptions. If these data were absent, we assumed that the provider was unaware of the potential for a serious drug–drug interaction. We calculated the percentage of drug interactions as the number of clinically relevant interactions generated by each physician divided by the total number of prescriptions written by each physician.

To determine users' reactions to the PWS and DHCP computer systems, we designed a questionnaire that evaluated user satisfaction with each computer system, usage patterns, and computer-based and other sources of information retrieval. At the beginning of the study, both groups answered the questionnaire concerning the DHCP computer system. At the end of the study, we administered the same questionnaire, but the DHCP group replied about the DHCP, and the PWS group replied about the PWS. We used a series

Table 1 ■

Demographic Data for Study Subjects

	PWS Group	DHCP Group	p Values
Average age (years)	30.2	29.6	p = 0.408*
Gender			
Males	11	11	
Females	8	7	p > 0.9999†
Year of training			
Intern	6	6	
Resident	13	12	p > 0.9999†
Undergraduate or graduate degree			
Computer Science or Engineering	0	0	
Other	16	17	p > 0.9999†
Home computer use (hours/week)			
0 to 5	10	14	
6 or more	5	4	p = 0.697
Work computer use (hours/week)			
0 to 5	9	10	
6 or more	8	7	p > 0.9999†
Computer training			
Formal training	13	14	
No formal training	4	4	p > 0.9999†

*Two-sample t test (two-tailed)
†Fisher's exact test

of 25 Likert-scale items to assess the physicians' satisfaction with the relevant computer system. We included items relevant to the process of providing care (e.g., "I know how to find the lab results I need.") as well as to the performance of the computer system (e.g., "PWS displays data clearly on the screen."). Additionally, we transcribed subjective remarks users made during the trial.

To measure how well the users thought the computer met their clinical information needs, we asked them how many other sources of information they required for various queries. Also, we asked users to check which information source (DHCP or PWS, medical chart, health summary, telephone call, or interview with the patient) they used to retrieve various types of clinical information (laboratory values, medication lists, etc.)

To assess PWS performance, we randomly sampled 50 clinic patients and recorded the time that lapsed from when the request was entered until (1) all laboratory data were available for display and (2) all clinical data, including medications, appointments, and radiographic results, were available for display. We could not perform a comparable assessment of DHCP performance because of the difference in data display for the two systems. The PWS employs a graphical format that displays current laboratory data, appointments, and medications simultaneously without additional user requests, whereas the DHCP displays only requested data and requires users to navigate through a hierarchy of menus.

Subject Education

We provided individual instructional demonstrations to each physician in the PWS group to familiarize them with the functions and limitations of the PWS during a two-month run-in period. A research assistant explained how to use the PWS, provided each user with a password, and helped the physician use the PWS for the first time. In addition to signing a consent form, users were also asked to read and sign a form detailing the PWS's differences from the DHCP, such as specific laboratory tests or procedures that could not be found in the PWS. The physicians were told to use alternative sources of information, such as the DHCP, if they could not locate relevant patient information with the PWS. During the trial period, the users were encouraged but not required to write prescriptions using the PWS.

The DHCP group was treated in a similar but more abbreviated fashion. These physicians were told that they would not have access to the PWS and would be expected to continue using the DHCP. They were also offered a brief training session in its use.

Data Analysis

For our analyses of baseline data, we used Fisher's exact test and Student's t test. To analyze the outcome variables, we calculated the difference in the variable between each time period for each physician. We compared the means of the differences for the PWS and DHCP groups using Student's test.

We performed power calculations for a two-sample t test and assumed n = 16 per group (total sample size = 32), a single-tailed alpha = 0.05, and power (1-beta) = 0.80. We estimated that our two-period parallel design would have an 80% chance of detecting a change in our primary outcome variable, the mean drug cost per physician per day of 10.9¢ (the effective size), or 11.0% of the total drug cost per physician per patient per day.[18]

For our survey instrument, we assessed the internal consistency of the Likert scale items on the baseline, as well as postintervention DHCP and PWS responses with Cronbach's alpha on Testat (Systat, Inc.). We performed data processing and statistical analysis using Excel (Microsoft) and Statview (Abacus Concepts) software.

Results

Our randomization procedure distributed subjects equally according to gender, level of residency training, and prior computer experience (Table 1). Fifty-nine percent of our study subjects were men; none held undergraduate or graduate degrees in computer science or engineering. Fifty-six percent of subjects used computers for fewer than five hours per week at work, and 73% used computers fewer than five hours per week at home. However, 77% of the subjects had received formal instruction in computer use.

Our assessment of system performance indicated that the PWS required a median time of 20 seconds until all laboratory data were available and 50 seconds until all clinical data were available. The time required for full retrieval differed substantially among patients, depending on the amount of data available on a patient. Interruptions of PWS functionality were caused by network instability, interruption of DHCP functionality, and PWS system failures. We estimate, subjectively, that the PWS system was unavailable for the entire user session in at least 10% to 15% of sessions. Users' comments reflected the importance of such interruptions (e.g., "I couldn't gain access to the computer twice."). Also, PWS users expressed frustration with the speed of data retrieval (e.g., "The slow speed makes [me] 'uneasy': kind of high stress just sitting there waiting for stuff to come up.").

During the intervention period, the physicians in the PWS group wrote 2,645 prescriptions; of these, the physicians wrote 75 (2.8%) prescriptions using the PWS system and 2,570 (97.2%) prescriptions using the paper-based system. There were no significant differences in health or economic outcomes between the control and intervention groups.

User Satisfaction

All of our subjects completed the initial questionnaire; 95% completed the final questionnaire. Our 25-item Likert satisfaction scale demonstrated internal consistency, with Cronbach's alpha equal to 0.868 for all subjects at the beginning of the study. For the postintervention questionnaires, the PWS group had a Cronbach's alpha equal to 0.906, and the DHCP group had a Cronbach's alpha equal to 0.724. At baseline, there was no significant difference ($p = 0.11$) between the mean satisfaction score for the control and intervention groups (based on a five point scale, with 1 = very dissatisfied and 5 = very satisfied). The physicians in the control group reported increased satisfaction with their clinical information system during the intervention period, while the PWS physicians demonstrated reduced satisfaction (Table 2). After physicians used the PWS, their user-satisfaction score decreased by 0.34 Likert-scale units (approximately one half of one standard deviation of the mean score, $p = 0.008$). In contrast, the mean satisfaction in the control group (DHCP) increased by 0.49 Likert-scale units ($p < 0.0001$). Overall, the two groups diverged with a difference of 0.83 Likert-scale units between the two groups ($p < 0.0001$). In responses to individual Likert-scale items about the usability of the two systems, PWS subjects responded favorably to detecting clinical trends and locating study results.

Discussion

We sought to evaluate whether an outpatient computer workstation containing a drug-ordering module would modify provider prescribing behavior and satisfaction. Our study is unique in comparing two computer systems in a randomized controlled trial in an outpatient setting. Despite the innovative features of the PWS—such as a graphical user interface, the ability to graph laboratory data chronologically, and point-of-care decision support—we found that user satisfaction decreased and that changes in health outcomes and costs were not statistically significant. Pro-

Table 2 ■

User-Satisfaction Rating*

	Preintervention Phase (mean ± SD)	Intervention Phase (mean ± SD)	Difference (mean ± SEM)	p Value
PWS group (N = 16)	3.44 ± 0.533	2.98 ± 0.547	−0.343 ± 0.112	p = 0.008†
DHCP group (N = 18)	3.23 ± 0.400	3.72 ± 0.333	0.488 ± 0.084	p < 0.0001†
Group difference comparison (PWS vs. DHCP)			0.831 ± 0.138	p < 0.0001‡

*Values are based on responses to a five-point Likert scale, with 1 = strongly disagree and 5 = strongly agree.
†Paired t test (two-tailed) for difference between pre- and post-satisfaction ratings for PWS and DHCP.
‡Two-sample t test (two-tailed) for difference between PWS group and DHCP group (PWS group difference minus DHCP group difference).

viders used the PWS system for writing only a small fraction (2.8%) of their prescriptions, so it is not surprising that the number of drug interactions and the prescription costs were not affected. The implementation of the intervention at our institution precluded requiring the intervention subjects to use only the PWS. Thus, the low level of usage may reflect the difficulty users encountered in using the PWS, or it may reflect our clinicians' satisfaction with the comprehensive computer-based and paper-based systems that were already available in our outpatient setting. The DHCP gained many new features during the one-year study; these changes probably explain the increased level of satisfaction in our control group. The low PWS usage limits our ability to draw conclusions about the usefulness of its novel features, such as a graphical interface and decision support.

How can researchers design evaluative studies that maximize the chance of observing changes in the relevant outcomes yet also provide useful information if the intervention fails? The first task requires identifying and reducing potential impediments in evaluations and selecting an appropriate design. The second task requires incorporating a "just-in-case" mentality in experimental design to ensure that information capable of explaining a negative study is collected. Our experience illustrates crucial areas that require consideration.

First, investigators and developers should specify the performance characteristics and support for both the experimental and control systems required for the evaluation. These specifications should include parameters that assess completeness of data, speed of operation, reliability, and accessibility of the system. This step will help reduce problems related to the new information system itself. The PWS project was an independent study that was not part of the Palo Alto VA's overall information-management strategy. The DHCP was upgraded frequently during the study period, and these upgrades often changed the DHCP's underlying schema, which led to further development requirements and to significant downtime for the PWS.

Second, investigators and developers should assess the reliability and speed of the computing infrastructure. The PWS system used terminals connected to a server over a local-area network (LAN). The server was connected to the DHCP. Thus, an interruption in DHCP function, a problem with the LAN, or a problem with the PWS itself could potentially lead to the PWS being unavailable for the user. In a client-server architecture, network stability is of paramount importance. The network infrastructure in our clinic was better suited to the menu-based data transfers of the DHCP (in which small amounts of patient data were transmitted at a time) than to the PWS model (in which a patient's entire record was transmitted initially). Users' perceptions of the PWS's lack of speed may have been related to this initial delay, when the entire record was transmitted. In addition, the synchronization between the DHCP's Massachusetts General Hospital Utility Multi-Programming System (MUMPS)-based database and the PWS's object-oriented database was difficult to maintain. The nightly uploads of new patient data were disrupted by machine or network downtime, and frequent mapping changes between databases required a manual restoration of the data bridge, which also led to interruptions in PWS availability.

Third, evaluators should investigate what works and does not work with the current clinical workflow. A major component in our intervention involved modifying the prescription-writing workflow in our clinic. We assumed that the potential reduction in drug interactions and drug costs would be sufficient motivation for users to adapt to a new system. However, we underestimated the value of our current paper-based system. We did not assess users' perceptions of problems in their work environment and, thus, may not have been providing the most beneficial technology for the users.

Fourth, investigators should provide sufficient training time for users. Our users were in the clinic one afternoon per week and had about a fourth of the clinics canceled because of vacations and rotations to other hospitals. We provided training over a two-month period before our intervention. This may not have been enough exposure to the PWS, especially because users were already familiar with the DHCP and continued to use it.

Finally, investigators should choose an experimental design that maximizes statistical power to detect differences in the relevant outcomes. We selected a two-period parallel design that improved our experimental power compared with a traditional randomized controlled trial.[18] Depending on the variances of the outcome variable, the two-period parallel design may reduce the number of study subjects required.[18] Randomization should account for all providers, including attendings, if possible. Other experimental designs, such as the crossover design,[20] may further enhance statistical power but may introduce sources of bias in interpreting results.

In addition to these steps, we recommend that investigators design evaluative studies to be informative,

even if a study does not find positive effects on health or economic outcomes. A negative study is useful if the investigators can determine why their system failed to affect the outcomes of interest and how to correct the problem. We suggest the following guidelines, with examples from our study. First, the design of the study should enable the investigators to distinguish between an intervention that was not used and one that was used but was ineffective. For example, we documented that the number of prescriptions written with the PWS was too low to have a substantial influence on costs. Had the system been used frequently and costs were still unaffected, an optimal design would have enabled the investigator to determine which aspects of the intervention required improvement. For example, were recommended drug substitutions considered inappropriate? Did the providers notice the recommendations?

The study should link user feedback to objective measurements of system performance. We suggest incorporating objective measurements into the study that help corroborate and explain user feedback. For example, our subjects commented on the system's "slow speed" and "how it needs some reports it didn't have." Interpretation of our results would have been aided by a systematic quantification of performance parameters, which could have complemented user perceptions. These parameters include comprehensive measurements of system availability; of the time from initiation of a command until completion; and of user interaction with the system, including a record of requests that could not be satisfied. In addition, our study would have benefited from a more systematic approach for assessing subjective responses, including open-ended questions in our survey instrument. Finally, the study should enable investigators to prioritize system improvements.

As several observers have suggested,[5,7] clinical information systems should be evaluated with the same rigor applied to other medical interventions. The randomized controlled trial is the gold standard of evaluation. We agree that controlled trials will provide the most rigorous evaluation of clinical information systems, and that more are needed.[7] However, evaluation of a clinical information system in an outpatient setting is a time-consuming, labor-intensive project that requires careful attention to workflow and personnel as well as to problems of system performance, security, and functionality.

The authors thank Jeff Shyshka, Dean Robinson, Minerva Sims, and the other staff at the VA Information Resources Management Service for help with computer support. We thank Robert Coleman and the other staff at the Palo Alto VA pharmacy service for assistance with prescription analysis; Merlynn Bergen for assistance with questionnaire design and analysis; Andria Cardinalli for help with preparation of the manuscript and data analysis; Lyn Dupré for comments on the manuscript; and the reviewers for their helpful suggestions.

References ■

1. Tape TG, Campbell JR. Computerized medical records and preventive health care: success depends on many factors. Am J Med. 1993;94:619–25.
2. McDonald CJ, Hui SL, Tierney WM. Effects of computer reminders for influenza vaccination on morbidity during influenza epidemics. MD Comput. 1992;9:304–12.
3. Tierney WM, Miller ME, Overhage JM, McDonald CJ. Physician inpatient order writing on microcomputer workstations. JAMA. 1993;269:379–83.
4. Tierney WM, Miller ME, McDonald CJ. The effect on test ordering of informing physicians of the charges for outpatient diagnostic tests. N Engl J Med. 1990;322:1499–1504.
5. Johnston ME, Langton KB, Haynes RB, Mathieu A. Effects of computer-based clinical decision support systems on clinical performance and patient outcome: a critical appraisal of research. Ann Intern Med. 1994;120:135–42.
6. Safran C, Rind DM, Davis RM, et al. An electronic medical record that helps care for patients with HIV infection. In: Safran C (ed). Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care, Washington, DC. New York, NY: McGraw-Hill, 1994;224–8.
7. Tierney WM, Overhage JM, McDonald CJ. A plea for controlled trials in medical informatics. J Am Med Inform Assoc. 1994;1:353–5.
8. Massaro TA. Introducing physician order entry at a major academic medial center: I. Impact on organizational culture and behavior. Acad Med. 1993;68:20–5.
9. Wyatt J, Spiegelhalter D. Field trials of medical decision-aids: potential problems and solutions. In: Clayton PD (ed). Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care, Washington, DC. New York, NY: McGraw-Hill, 1992;3–7.
10. Annevelink J, Young CY, Tang PC. Heterogeneous database integration in a physician workstation. In: Clayton PD (ed). Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care, Washington, DC. New York, NY: McGraw-Hill, 1992;368–72.
11. Pham TQ, Young CY, Tang PC, Suermondt HJ, Annevelink J. Implementing a Physician's Workstation using client/server technology and the distributed computing environment. In: Ozbolt JG (ed). Proceedings of the Eighteenth Annual Symposium on Computer Applications in Medical Care, Washington, DC. Philadelphia, PA: Hanley & Belfus, 1995;626–30.
12. Suermondt HJ, Tang PC, Strong PC, Young CY, Annevelink J. Automated identification of relevant patient information in a physician's workstation. In: Safran C (ed). Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care, Washington, DC. New York, NY: McGraw-Hill, 1994;229–32.
13. Tang PC, Annevelink J, Fafchamps D, Stanton WM, Young CY. Physician's Workstations: integrated information management for clinicians: In: Clayton PD (ed). Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care, Washington, DC. New York, NY: McGraw-Hill, 1992;569–73.

14. Tang PC, Annevelink J, Suermondt HJ, Young CY. Semantic integration of information in a physician's workstation. Int J Biomed Comput. 1994;35:47–60.
15. Young CY, Tang PC, Annevelink J. An open systems architecture for development of a physician's workstation. In: Clayton P (ed). Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care, Washington, DC. New York, NY: McGraw-Hill, 1992;491–5.
16. Stanton WM, Tang PC. Knowledge-based support for a Physician's Workstation. In: Clayton P (ed). Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care, Washington, DC. New York, NY: McGraw-Hill, 1992;649–53.
17. Fafchamps D. Ethnographic workflow analysis: specifications for design. In: Bullinger H. Human Aspects in Computing: Design and Use of Interactive Systems and Work with Terminals. New York: Elsevier Science Publishers, B.V.; 1991;709–15.
18. Rotman BL, Sullivan AN, McDonald T, et al. A randomized evaluation of a computer-based physician's workstation: design considerations and baseline results. In: Gardner RM (ed). Proceedings of the Nineteenth Annual Symposium on Computer Applications in Medical Care, New Orleans, LA. New York, NY: McGraw-Hill, 1996;693–7.
19. Rotman BL, McDonald TW, Suermondt HJ, et al. Computer-based detection of drug interactions in an ambulatory care practice. Med Decis Making. 1994;14:440.
20. Brown B. The crossover experiment for clinical trials. Biometrics. 1980;36:69–79.