User Acceptance and Adoption of a Clinical Reminder System: A Developmental Trajectory Approach

Kai Zheng [A], Rema Padman [B], Michael P. Johnson [C]

[A] School of Public Health and School of Information, The University of Michigan, Ann Arbor, MI 48109
[B] The H. John Heinz III School of Public Policy and Management, Carnegie Mellon University, Pittsburgh, PA 15213
[C] John W. McCormack Graduate School of Policy Studies, The University of Massachusetts Boston, MA 02125

Abstract In this study, we assess user acceptance and adoption of a clinical reminder system for chronic disease and preventive care management in an ambulatory care environment. We use a novel developmental trajectory approach for data clustering. This group-based, semi-parametric statistical modeling method identifies distinct groups, following distinct usage trajectories, among those who recorded use of the reminder system within an evaluation period of 10 months. We trace system use within these groups over time using computer-generated logs and user satisfaction surveys. Our analysis delineates three categories of users based on their demographics and computer literacy backgrounds. These user categories are also correlated with reminder compliance, and these reminders have triggered follow-up actions that could be otherwise missed. We conclude that this developmental trajectory approach combined with qualitative assessments has considerable promise to provide new insights into system usability and technology adoption issues that may benefit clinical decision support systems (CDSS) as well as information systems more generally.

Keywords: Diffusion of innovation; Usage analysis; Decision support; Clinical decision support systems; CDSS; Evaluation; Organizational change and impact; Evidence-based medicine

I. Background

Clinical cueing systems (CCS) are a class of computer based clinical decision support systems (CDSS) that send just-in-time alerts to clinicians when potential errors or deficiencies in patient management are detected. Beneficial outcomes of CDSS have been documented in many studies along a number of dimensions, including compliance with treatment standards, reduced costs, and improved health outcomes [1-5]. Several systematic reviews have also shown evidence that CDSS can be an effective means of implementing medical guidelines to enhance clinical performance in a wide range of aspects of medical care [6-10]. However, it has also been noted that most of these studies either focus on accuracy and relevance of the computer-aided recommendations, or use experimental or Randomized Controlled Clinical Trials (RCTs) designs to assess system or clinical performance [3-10]. Few involve a naturalistic design in routine settings with real patients [7, 11]. It is not clear whether a CDSS that has been shown to be effective in a laboratory setting will be fully utilized by end-users in clinical environments, and whether these users will adapt their practice style to efficiently accommodate computer-generated reminders.

Even if the design of a CDSS were optimized as much as technically possible, practitioners may not acknowledge that use of the system would add value to their medical practice and thus may be reluctant to incorporate it into their daily routine [12]. A variety of factors contribute to this resistance, such as insufficient computer ability, diminished professional autonomy, lack of insights into long-term benefits, and change of conventional behaviors. As the systematic reviews have indicated, few evaluation studies have considered these factors [7, 11]. Methodological characteristics shared by these studies may have limited the scope. RCTs and other experimental designs, although widely accepted as "gold standard" in evaluation studies, provide little insight into acceptance and adoption issues: under controlled conditions it can be difficult to investigate a variety of human, contextual, and cultural factors that affect system use, and a focus on pre-specified outcome measures precludes examining processes of the actual use during daily activities [11]. Researchers therefore have called for methodological pluralism for evaluating informatics applications [11-14].

Actual use of technology has also been recognized as being a critically important evaluation factor. In a recent information systems (IS) evaluation study, actual use is found to be an important variable in explaining the impact of information technology (IT) on organizational performance [15]. The study posits that "the driver of IT impact is not the investment in the technology, but the actual usage of the technology". Although their finding is derived from studying a hospital decision support system for financial management, it provides helpful insights into evaluation studies of CDSS. Measuring actual usage of a

Correspondence: Kai Zheng, PhD, School of Public Health & School of Information, The University of Michigan, M3531, 109 S Observatory Street, Ann Arbor, MI 48109; email: <kzheng@umich.edu>.

CDSS requires full integration of the application into routine settings as well as a longitudinal data collection strategy to allow users to achieve stability in usage. Unfortunately, there is limited use of CDSS despite demonstrated or potential benefits [7, 11]. In the absence of systems that are deployed and regularly used in the field, researchers tend to use user perception as a means of studying technology acceptance and adoption. For instance the widely-accepted method in IS research, the Technology Acceptance Model (TAM), relates perceived usefulness to use [17]. However, as other studies suggest, perceived usefulness is poorly correlated with actual use [18]. In addition to this, self-reported usage, or intent to use, may not be an appropriate surrogate for actual use because users are poor estimators of aspects of their own behaviors [19].

Users of an IT application differ in many ways. It has been well recognized that individual users play a crucial role in that their experiences, and their opinions or reactions to a technology make a difference in whether or not the technology will be adopted [18]. Nevertheless, few studies have used individual-level data to measure the magnitude of user differentiation, and the impact of such differentiation on technology adoption. Furthermore, voluntary use has not been adequately addressed in the existing informatics evaluation studies due to the prevalent experimental designs, while voluntariness has received close attention in evaluation studies in other disciplines for its value in assessing IT impacts [15, 16, 20].

In the present study, we aim to address these limitations. First, the system under evaluation, Clinical Reminder System (CRS), has been integrated into the daily operations of an ambulatory clinic. Its reminder component has been regularly used by clinicians in their patient encounters. Second, use of the system is strongly recommended, but not mandatory, providing a means of assessing voluntary use coherent to the decision support nature of CDSS. Third, longitudinal usage data for an evaluation period of 10 months were collected from computer logs, providing an objective and non-intrusive measure of the actual use over time. Fourth, we apply to the usage data a novel developmental trajectory method to identify groups that demonstrate distinct adoption behaviors, and to relate estimated group configurations to a variety of user characteristics. Although our analysis of user profiling is preliminary, due to sample size, the results can be interpreted as suggestive of trends whose presence may be verified in a larger user population. Finally, we conduct qualitative analyses to identify factors that may account for observed adoption behaviors.

In this paper, we report our assessment of user acceptance and adoption of CRS based on actual use. Parallel studies are currently being conducted to further examine and understand the observed behaviors and trends, and to identify solutions to minimize unexpected adoption behavior. The major goal of this stream of research is to determine the critical factors for CDSS success and the impact of CDSS on patient and organizational outcomes.

The next section describes the Clinical Reminder System and its basis in evidence-based medicine concepts. This is followed by a presentation of the methods for evaluating acceptance and adoption of CRS. The results of trajectory analysis are presented next, followed by a discussion of our findings in the larger context of evaluation studies of health informatics applications and information systems more generally. The final section presents some concluding remarks.

II. Clinical Reminder System and the Study Site

In this section we review evidence-based medicine principles used in the design of Clinical Reminder System, characteristics of CRS itself, details of the study site, and data collection methods.

A. Evidence-Based Medicine and CRS

Evidence-based medicine is the distillation of a large volume of medical research and standards into treatment protocols for diseases and preventive care procedures that represent the most accurate knowledge available [21]. There is a widely acknowledged gap between knowledge by physicians of appropriate treatments and physicians' consistent application of treatment standards in practice. Evidence-based medicine principles have been incorporated into some health informatics applications that assist in patient management decisions in contexts such as chronic disease management [21-24]. Evidence-based medicine requires four steps to be effective: current knowledge of patient health status, periodic reminders to patients to undergo appropriate treatments, decision support to enable physicians to provide appropriate treatments, and patient follow-up and monitoring, based on complex medical treatment histories amassed over a long period of time [21].

The clinical decision support application discussed in this paper, Clinical Reminder System, integrates the hospital's administrative, laboratory, and clinical records systems into one single database, and uses patients' current medical status to provide reminders to clinicians at the point of care that reflect evidence-based medicine guidelines. Reminders generated by CRS take the form of recommendations to have tests scheduled and performed, receive vaccinations, alert clinicians to review abnormal test results, or closely monitor patients with medical conditions that require unscheduled intervention. Patient data essential to specific protocols, but not stored in the database, are collected during the physician-patient encounter via pop-up windows containing "on the fly questions". Table 1 shows a sample of reminders and "on the fly" questions generated by CRS. At the time this evaluation study was conducted, CRS intended to improve the quality of care for two chronic diseases: diabetes and hyperlipidemia, and five preventive care categories: steroid-induced osteoporosis, influenza, pneumonia, breast cancer, and cervical cancer.

Table 1. Sample Questions and Reminders from CRS

Sample patient data:

Patient characteristics: Male, Age 50, Height 5'10", Weight 230 lbs, BP 140/90, Smoker.

Diagnosis: Diabetes.

Lab Tests: Urinalysis in last 12 months (value: -5), HBA1C in last 6 months (value: 7.5), Lipid test in last 3 months (LDL value =180).

Sample "on the fly" questions:

Has the patient received steroid treatments for more than three months? (Reason: to determine if patient meets criterion for prevention of Steroid-Induced Osteoporosis protocol.)

Sample reminders (reminder #, protocol, statement, rationale):

[309 - Pneumococcal Vaccine] Based on the medical history, the patient meets standard criteria to receive pneumococcal vaccine. (Reason: patient has chronic diseases and no history of the vaccine).

[109 - Diabetes] The patient's body mass index is above the desired level. Do you wish to discuss weight maintenance with the patient? (Reason: diabetic patient, BMI over 30.)

[105 - Diabetes] The patient's HBA1C level is above the recommended target level. Do you wish to consider adjustments in treatment and/or additional patient education? (Reason: diabetic patient, latest HBA1C test within 6 months and result value over target level.)

CRS is the result of a collaborative effort between University-based researchers and hospital physicians and staff. The system was developed over three years, and has been rigorously tested in the laboratory by software developers and external IT experts. The CRS deployed at the time of this study is a distributed windows application based on client-server architecture. The clients are written in *Visual Basic* and communicate with an *Oracle* database server via the hospital's internal computer network. Evidence-based medicine guidelines are programmed in *Oracle PL/SQL* procedures. Interfaces have been established to automate data flows from other existing hospital information systems to CRS. Patient records such as registration information, lab test results, and previous diagnoses are downloaded online, some in real time, and stored in CRS' local database. To accommodate these data and functional requirements, CRS has emerged as a light-weight Electronic Medical Record (EMR) system as well as a clinic management system that supports the clinic's routine administrative operations. Details of CRS architecture and the algorithms implementing evidence-based medicine guidelines are discussed elsewhere [24].

B. Study Site and Data Collection

The study was conducted in the primary care clinic of an urban teaching hospital offering comprehensive healthcare services. This clinic serves as a rotation site for the hospital's residents. The clinical staff used the system for managing appointments, patient check-in and check-out, and recording of vital signs. Given the availability in every exam room of desktop computers installed with CRS, residents used the system during the patient encounters. The system had been implemented and routinely used by all nursing and clerical staff three months prior to the study, allowing it to be integrated into the clinic's work flow as well as to collect historical patient information from paper charts. Training for residents was provided in December 2001 and January 2002, and once again in July 2002 for new residents rotating into the clinic. The Department of Medicine, which manages the clinic, strongly encouraged use of the system: individual usage reports along with directives to increase usage were presented to all residents twice, in April and June, 2002, respectively. Nevertheless, rigorous daily use of the system by residents was not mandatory during this study period.

Table 2. Information about the Study Site

Patients	4646
Appointments scheduled	7533
Active resident users	44
Active nursing and clerical staff users	10
Visits	3923
Visit w/ resident encounters	2949
Encounters w/ system use	1199
"On the fly" questions generated	1965
Reminders generated	6235
Diagnosis code downloaded *	22006
Lab test results downloaded **	2795
Lab or other tests entered in clinic	835 ***

* Downloaded online from the hospital billing system
** During this study period CRS only downloaded the lab tests that were relevant to the seven protocols
*** About 35% were entered by residents

Data collection started on February 1, 2002. In this study, we report 10 months of usage data to identify distinct user groups, and maturation of their interaction with the reminder system. There were 44 active residents registered in the application. Activities generated by 3 of them were removed from the analysis since there were no visits recorded for these residents in 6 or more continuous months. Table 2 shows general information of the study clinic and its activities, as recorded by CRS during the evaluation period.

Rotation of residents occurred in July 2002: some fourth-year residents left the clinic while some first-year residents joined. Among the 41 residents who recorded usage in the system, 21 had 10 months of continuous use, and the remaining 20 had 5 months of use. We

sorted the data of the first-year residents accordingly to associate their use with the first month onwards. For them, "Month 1" is set to July 2002, so our 10-month timeline does not always correspond to sequential calendar months starting in February 2002. We evaluate the effects of this truncation in separate analysis.

We have included in the study several key attributes of users that may relate to their human, contextual and cultural characteristics that in turn affect system use. These attributes are: general demographics including gender and citizenship (U.S. versus non-U.S.); computer literacy identified by use, knowledge, and optimism of computer systems; and frequency of encounters (work load). These attributes are found to influence adoption behavior in some other contexts [17, 19, 20, 25]. We exclude a few other potentially influential attributes because they are 1) invalid in the context, such as medical specialty (all users are internal medicine residents), or 2) lack of variation in the sample, such as age (mean 29.6, standard deviation 2.1), or 3) sample size too small for defined subgroups, such as year of residency, or 4) unable to access, such as evaluation of residents' clinical performance (due to confidentiality concerns).

Computer literacy of residents is assessed using Cork's instrument for measuring physicians' use of, knowledge about, and attitudes towards computers [25]. This instrument has been widely used in health informatics evaluation studies. We also administrated two user satisfaction surveys to assess users' perceived system usability and usefulness. The first survey was conducted in March 2002, one month after the system implementation, and the second one was given in February 2003, after the 10-month evaluation period was over. 32 residents participated in the first survey, and 37 residents participated in the second. These two surveys are based on IBM satisfaction questionnaire, developed by J. R. Lewis [26], which is comprised of 19 items (15 items to assess general system usability, 3 items to assess system interface factors, and 1 item to provide an overall satisfaction rating), with several open-ended questions. All quantitative items are rated using a seven-point scale, with both ends of the rating scale anchored.

III. Methods

A. Measurement of System Usage

We measure use of the system in a number of ways, including 1) percentage of patient encounters in which the system was used to generate physician directed reminders (PPE), 2) percentage of reminders for which users recorded favorable responses (PFR), and 3) percentage of reminders for which text comments were entered (PTC).

PPE measures the frequency of using the system for advisories. We only treated "generating reminders" as a signal of valid use; other uses of this application, such as browsing patient charts, are not considered in this study. Some proportion of generated reminders were skipped by users, however, since skipping reminders does not provide explicit insight into the motive for doing so, and the occurrence of skipping all reminders was rare (less than 8%), we deem "generating reminder" as the primary indicator of system use.

PFR indicates whether the reminders generated may have assisted in providing follow-up treatment actions. CRS allows users to record a variety of actions to specific reminders. The choice of actions is context-specific, i.e. not all choices appear for all reminders. Among possible actions, favorable responses include "Ordered", "Ordered, results pending", "Performed", and "Not indicated", and unfavorable responses include "No", "Refused", and "Noncompliant". If a user decides not to respond to a reminder, "No action" can be chosen or the reminder may be skipped without explicit response.

PTC is of particular interest in assessing users' enthusiasm towards the system, since entering textual comments is time-consuming but potentially valuable during the patient's follow-up visits.

B. Developmental Trajectory Analysis

We use a novel method, developmental trajectory analysis (DTA), to study adoption behavior of the reminder system. DTA is a semi-parametric, group-based approach for identifying distinctive groups of individual trajectories within the population and for profiling the characteristics of group members [27]. It has provided valuable insights into studying physical aggression among youth [28, 29], criminal careers [28], and web utilization and saturation patterns [30].

A "developmental trajectory" describes the course of a developmental behavior over age or time. Conventional methods of studying developmental behavior, such as hierarchical modeling and latent curve analysis, model variation in the parameters of developmental trajectories using continuous multivariate density function. They are designed to identify average developmental tendencies, to calibrate variability about the average, and to explain that variability in terms of covariates of interest. By contrast, this group-based approach uses a multinomial modeling strategy and is designed to identify relatively homogeneous clusters of developmental trajectories, to calibrate the probability of population members following each such cluster, and to relate those probabilities to covariates of interest [27].

DTA assumes that the population is composed of a mixture of distinct groups defined by their developmental trajectories. Whereas conventional methods are suitable for modeling typical patterns of growth that vary regularly through the population, this method is useful for modeling unobserved heterogeneity in a population, where trajectories vary greatly across population subgroups both in terms of the level of behavior at the outset of the measurement period and in the rate of growth and decline over time [27].

DTA enables users to choose the probability distribution that is best suited to the data at hand. The censored normal model is useful for modeling the conditional distribution of psychometric scale data that tend to cluster at the minimum and the maximum of the scale. DTA is also appropriate for continuous data that are approximately normally distributed, with or without censoring. We model usage of CRS with the censored normal model by specifying a minimum and a maximum that lie outside the range of the observed data values, because we do not observe clustering at extremes.

A brief overview of the statistical theory underlying the method is given below. Let the vector $Y_i = \{y_{i1}, y_{i2}, \ldots, y_{iT}\}$ denote the longitudinal sequence of individual i's behavioral measurement during the T periods of measurement. Let $P^j(Y_i)$ denote the probability of observing Y_i given membership in group j, and π_j denote the proportion of the population comprising group j. The unconditional probability of observing Y_i equals the sum across the j groups of the probability of Y_i given membership in group j, weighted by the proportion of the population in group j:

$$P(Y_i) = \sum_j \pi_j P^j(Y_i) \quad (1)$$

Let $p^j(y_{it})$ denote the probability distribution function of y_{it} given membership in group j. For given j, the conditional independence is assumed for y_{it} over the T periods of measurement, thus:

$$P^j(Y_i) = \prod^T p^j(y_{it}) \quad (2)$$

The likelihood for the entire population of N individuals is:

$$L = \prod^N P(Y_i) \quad (3)$$

DTA models the linkage between time and behavior by assuming polynomial relationships. For the censored normal model, a quadratic relationship is given as:

$$y_{it}{}^j = \beta_0{}^j + \beta_1{}^j Month_{it} + \beta_2{}^j Month_{it}{}^2 + \varepsilon_{it} \quad (4)$$

where ε_{it} is a disturbance assumed to be normally distributed with zero mean and constant variance σ^2.

For the censored normal distribution, we can write the probability distribution function of y_{it}, given membership in group j as:

$$p^j(y_{it}) = \frac{1}{\sigma} \phi\left(\frac{y_{it} - \beta^j x_{it}}{\sigma}\right) \quad (5)$$

ϕ is the density function of a normal random variable with mean $\beta^j x_{it} = \beta_0{}^j + \beta_1{}^j Month_{it} + \beta_2{}^j Month_{it}{}^2$ and standard deviation σ.

The parameters of interest, $\beta_0{}^j, \beta_1{}^j, \beta_2{}^j, \pi_j$ and σ can thus be estimated by the method of maximum likelihood. The maximization is performed using a general quasi-Newton procedure.

Note that the model parameters, $\beta_0{}^j, \beta_1{}^j, \beta_2{}^j$, may differ from cluster to cluster, which is the key feature of this method since it allows for easy identification of population heterogeneity not only in the level of behavior at a given stage but also in its development over time.

The posterior group membership $\widehat{P}(j \mid Y_i)$ can be then calculated as:

$$\widehat{P}(j \mid Y_i) = \frac{\widehat{P}(Y_i \mid j)\widehat{P}(j)}{\widehat{P}(Y_i)} = \frac{\widehat{P}(Y_i \mid j)\widehat{\pi}_j}{\sum_j \widehat{P}(Y_i \mid j)\widehat{\pi}_j} \quad (6)$$

where $\widehat{P}(Y_i \mid j)$ is the estimated probability of observing Y_i given membership in j, and $\widehat{\pi}_j$ is the estimated proportion of the population in group j.

The method also provides a multivariate procedure to evaluate impacts of individuals' characteristics on group membership probability. Assume that π_j varies based on individual, familial, or environmental factors. Let x_i denote a vector of factors that are potentially linked to group membership assignment. Then a multinomial logit model to estimate π_j is:

$$\pi_j(x_i) = \frac{e^{x_i\theta_j}}{\sum_j e^{x_i\theta_j}} \quad (7)$$

where parameters θ_j captures the impact of the covariates of interest, x, on the probability of membership in group j. θ_j is typically set to zero for one contrast group, and the coefficient estimates for other groups are interpreted as measuring the impact of covariates on group membership relative to the contrast group.

DTA has a distinctive advantage over classical clustering methods through the use of a Bayes factor to compare models, that is, to determine the optimal number of clusters as well as appropriate order of the polynomial used to model each group's trajectory. The Bayes factor measures the odds of each of the competing models being the correct model, which is difficult to calculate but can be reasonably approximated by exp (BIC_i - BIC_j). The Bayesian Information Criterion (BIC) for a given model is calculated as follows:

$$BIC = \log(L) - 0.5 * \log(n) * k \quad (8)$$

When prior information on the correct model is limited, selection of the model which has the maximum BIC is recommended. Schwarz also provided a related metric

for comparing more than two models [35]:

$$p_j = \frac{e^{BIC_i - BIC_{max}}}{\sum_j e^{BIC_i - BIC_{max}}} \quad (9)$$

where p_j denotes the probability that model j is the correct model, and BIC_{max} is the maximum BIC score of the competing models.

IV. Results

A. Aggregated Usage and Simple Grouping Scenario

We first treat the clinic as the unit of analysis to evaluate the level of aggregated usage. Figure 1 depicts this use rate along the 10-month timeline recorded as PPE (percentage of encounters in which the system was used to generate reminders). On average this rate stays at approximately 40% level, with some fluctuations.

Without further exploration, Figure 1 may lead to an impression that the developmental trend of user adoption of this reminder system can be interpreted as a typical three-phase learning curve: Month 1 to Month 3 represents a "burn-in" period, in which users started to learn the system and the usage constantly increased as more skills were acquired; this rate adjusted itself when "burn-in" effects diminished, and stayed stable afterwards. With the clinic as the unit of analysis, it may be reasonable to conduct further analysis along this direction; however, we shall see that the developmental trajectory method that we applied provides different and more valuable insights after breaking down the aggregated usage into heterogeneous clusters.

Similarly, Figure 2 shows the trends of aggregated usage after splitting the user population by a simple grouping scenario: 21 users who had mean use rates above the average are labeled as "High" users, and the remaining 20 are labeled as "Low" users. Although this grouping scenario conveys more information regarding the heterogeneity of system usage within the population, it fails to explain some of the phenomena that we have observed in the field. For instance, several residents were enthusiastic about the system after the training and the initial usage reports showed that this enthusiasm resulted in higher level of use and reminder compliance. However, their recorded usage and compliance rate at the end of the study period dropped to a level comparable with those who resisted the system from the very beginning. We therefore need a more sophisticated tool to distinguish such subtle patterns in order to better understand the distinctive adoption behaviors.

B. Developmental Trajectories

Based on Bayes factor model selection rule and the study setting, we have chosen to cluster users into three
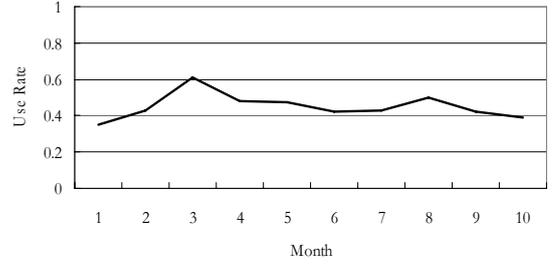

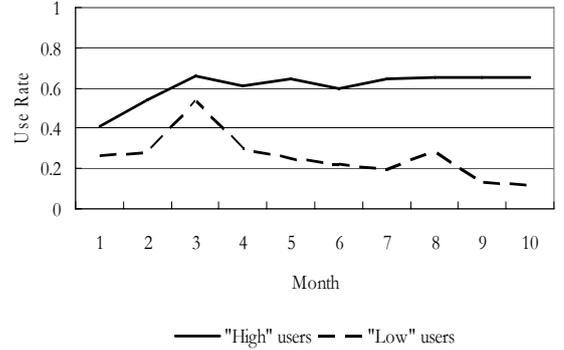**Figure 1.** Aggregated System Usage

Figure 2. Simple Grouping Scenario

groups and set the order of polynomial models to 1-2-1 (BIC: -37.54, N=41, probability of being the correct model among competing models: 93%). The developmental trajectories obtained thereby are depicted in Figure 3. Bold and light lines denote observed and predicted trends, respectively. Observed data values are computed as the mean use rate of users assigned to each of these groups identified by estimation, and expected values are computed using DTA model coefficient estimates.

We label these three groups as "Light" users, "Moderate" users and "Heavy" users, each comprising 17, 15 and 9 individuals (41.46%, 36.59%, and 21.95% of all users), respectively.

Alternate sets of trajectories, produced by two- and four-group models, are plotted in Figures 4 and 5. The two-group model is analogous to the simple grouping scenario (Figure 2): it conveys little information about the subtle developmental patterns as revealed by the three-group model. The four-group model breaks down the low-usage users further into two subgroups. As shown in Figure 5, users in group 1 and group 2 differ only slightly in their usage levels. We conclude that the three-group model we have selected not only meets the model selection criterion, but also results in trajectories that best describe the usage data. Henceforth, our analyses are based on the three-group model.

Table 3 summarizes the mean posterior group membership probability for each of the three groups. The mean probability for the 17 residents assigned to the light user group is high: .98, and the counterpart averages for the moderate user group and the heavy user group are relatively lower: .90 and .82, respectively.
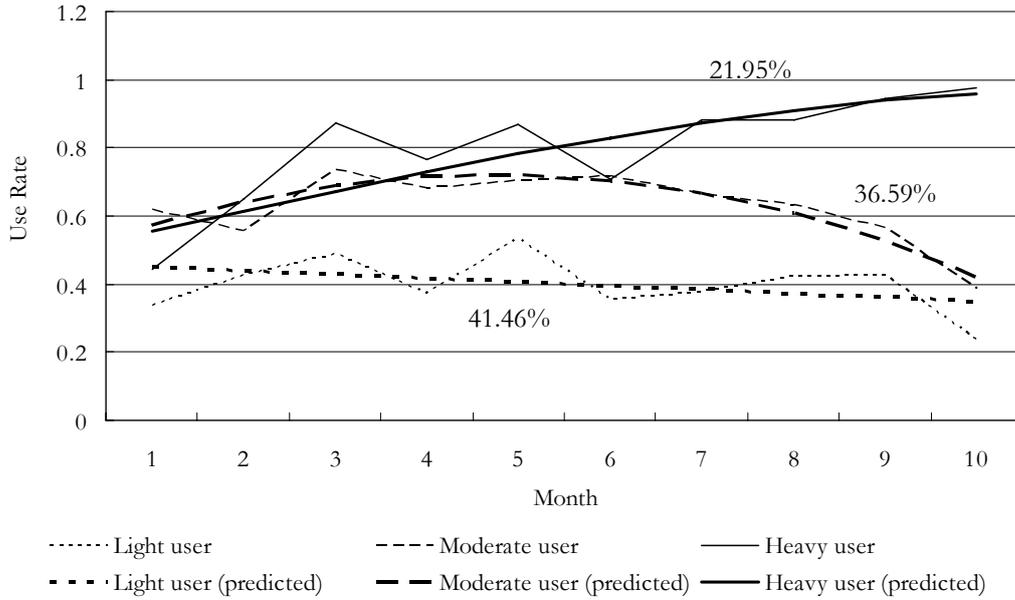
Figure 3. Developmental Trajectories of Three Distinct Groups

Off-diagonal cells show the probability of committing Type I errors. For example, .15 can be interpreted as "the probability of a 'Heavy' user being mislabeled as a 'Moderate' user is 0.15". Table 3 confirms that our model fits the data well. It also indicates that those "Light" users are relatively easy to identify, whereas ambiguity exists at the boundary between the heavy user group and the moderate user group.

Figure 3 also reveals developmental trends of the adoption behavior. Users classified as "Light" initially utilized the system in about 35% of their encounters, and this rate remained steady over the 10-month study period. "Moderate" users had the highest initial use rate, about 70%, but this rate consistently decreased over the study period to a level comparable with that of the "Light" users. "Heavy" users had an initial use rate of approximately 50%, and this rate increased consistently to about 100% at the end of the study period. Changes in the use rate for members of the moderate group are of particular interest since it indicates "Moderate" users demonstrated strong enthusiasm initially, while followed by a gradual decline in their use of the system. Table 4 summarizes these use rates.

Table 3. Group Composition

Group	Mean posterior probability of group membership		
	Light user (%)	Moderate user (%)	Heavy user (%)
Light user	.98	.0084	.008
Moderate user	.043	.90	.052
Heavy user	.024	.15	.82

Table 4. Summary of System Use Rate

Group	Mean rate (%)	Initial rate (%)	Ending rate (%)
Light user	32.42	31.24	23.77
Moderate user	61.73	70.23	34.66
Heavy user	67.15	38.46	94.44

C. Group Profiling

We further relate the adoption behavior to a set of user characteristics to identify the potential causal relations between the two. Due to the small sample size of this study, results can be interpreted as suggestive of trends that may obtain in a larger user population. Group profiles are shown in Table 5. For example, female users account for two-thirds of the heavy user group, whereas male residents account for two-thirds of the moderate user group.

The lower portion of Table 5 shows real-scale measures. On average, each resident saw 8.44 patients each month during the study period (standard deviation 1.89); the maximum count of daily patient encounters is 5, but this only occurred a few times. Since the frequency of patient encounters, monthly or daily, was low and relatively evenly distributed among residents, we do not deem this variable useful and will exclude it from further analysis.

The last three rows of Table 5 present the residents' computer literacy assessments obtained from Cork's instrument, measuring physicians' use of, knowledge about, and attitudes toward computers. These three items can be roughly interpreted as how

often physicians currently use computers, how much they know about computers, and their relevant beliefs and attitudes [25]. As shown in Table 5, "Light" users have the highest computer knowledge score (significantly higher than that of the moderate user group at .05 level), and "Heavy" users have the highest computer use and computer optimism scores (the optimism score of heavy user group is significantly higher than that of the light user group at .001 level). One interpretation of this result is that, based on their facility with IT, "Light" users may be less forgiving of inadequacies in trial applications such as CRS and may believe that these applications will not add value to their work. On the other hand, "Heavy" users, while modest regarding their IT skills, may believe that these applications could help them provide better service to patients. These results presage similar results regarding reminder response rate, reminders compliance rate, and qualitative assessments of attitude by usage group.
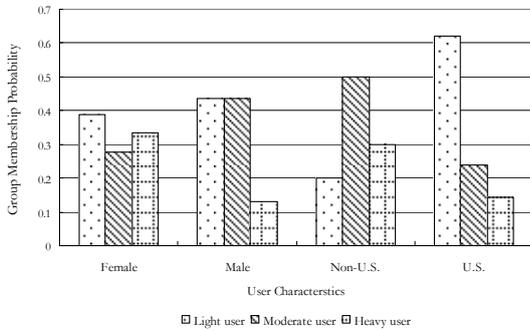

Figure 6. Mean Group Membership Probabilities for Categorical Variables

Figure 6 shows the group membership probabilities for each of the categorical variables. For example, male users tend to cluster in the light and the moderate user groups, and non-U.S. citizens are more likely to be present in the moderate user group, whereas U.S. citizens are more likely to be present in the light user group.

Group profiles (see Table 5) are a collection of univariate contrasts. A multivariate procedure is provided in the DTA model to construct a more parsimonious list of predictors to sort out redundant variables as well as to control for potential confounds. Table 6 shows the impact of the covariates of interest on group membership probabilities. The upper panel shows coefficient estimates and t-statistics. We use the light user group to serve as the contrast group, that is, for the light user group, the impacts of the covariates are set to zero, and the coefficient estimates for other groups are interpreted as measuring the impact of the covariates on group membership probabilities relative to that of the contrast group.

For example, if a user is a female, the probability of her being in the heavy user group is increased while the probability of her being in the moderate user group is decreased, compared to the probability of

membership in the light user group. The small number of statistically significant impact of covariates of interest on group membership probabilities is reflective of the small sample size in this study; the relationships listed here should be interpreted as only suggestive of the likelihood of the actual impacts.

The lower panel of Table 6 shows the predicted membership probabilities based on multinomial logit model coefficient estimates. The probabilities are calculated by substituting the estimated coefficients into Equation (7), and then computing group membership for assumed values of x_i. For each row, the impact of a single factor is evaluated. For instance, being a female alone increases the probability of membership in the heavy user group. Similarly, having non-U.S. citizenship alone decreases the probability of membership in the light user group and increases the probability of being in the moderate user group. The DTA model is also capable of estimating membership probabilities from a valid combination of multiple factors, but such analysis is inapplicable in this study given the small sample size.
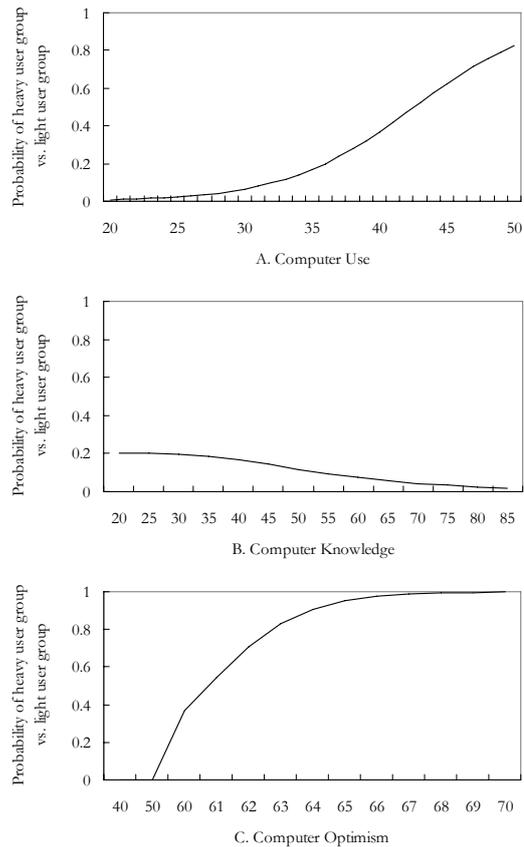

Figure 7. Membership Probability of Heavy User Group versus Light User Group as a Function of Computer Literacy Scores

Impact of computer literacy also emerge: relative to the light user group, higher score on previous computer use or computer optimism increases the probability of being in the heavy user group, and higher score on computer

optimism increases the probability of being in the moderate user group. In contrast, computer knowledge does not seem to affect the usage significantly. Figure 7 illustrates the marginal relationships between computer literacy scores and the likelihood of being in the heavy user group versus being in the light user group. While increase in computer use or optimism score dramatically boosts the probability of being in the heavy user group, increase in computer knowledge score has a negative effect; nevertheless, the magnitude of this effect is very small.

D. Analysis of Truncated-Usage Scenarios

A significant portion of potential system usage did not occur due to the rotation of residents and various site management issues. Residents who left or joined the clinic in the middle of the study only had 5 months of continuous usage data recorded.

To evaluate the effects of truncated usage due to scheduling, we use the DTA model to analyze two sub-datasets derived from the complete dataset. They are labeled as 1) 10-month dataset (DS10), comprising the usage data of a subset of 21 residents who stayed active for the whole 10-month study period, and 2) 5-month dataset (DS5), composed of 5 month continuous usage data of those residents who rotated in or out of the program. We have chosen to separate the data at Month 5 because it was at that point (July, 2002) the rotation of residents occurred.

Trajectories obtained from these two sub-datasets are depicted in Figure 8 and Figure 9, respectively. Similarity measures between the resulting clusters of these two sub-datasets comparing to those obtained from the complete dataset are shown in Table 7.
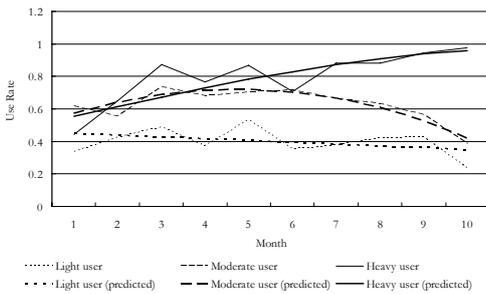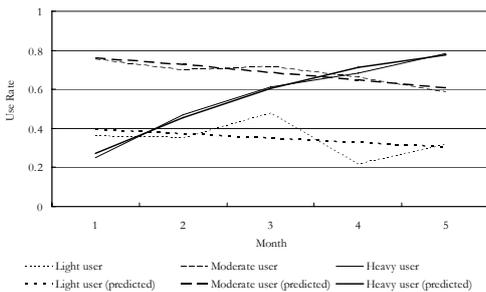

Figure 8. Trajectories based on DS10

Figure 9. Trajectories based on DS5

Analyses of DS10 and the complete dataset result in very similar clustering: chi-square tests are significant and Kappa statistics are good for all groups. The shapes of the trajectories are also consistent. Truncated usage, although not a random occurrence, does not have a significant impact on results. This indicates that the amount of information contained in the complete dataset is sufficient for the DTA model to track patterns, and correctly and consistently calculate posterior group membership probabilities.

Two noticeable spikes, in Month 3 and Month 5 respectively, are visible in the trajectories derived from the DS10 dataset (Figure 8). These spikes correspond to the points at which individual usage reports, along with directives to increase usage, were issued by the clinic management. Although effects of the directives were significant, they faded away very quickly: both spikes are followed by a decrease in usage that occurred in Month 4 and Month 6. Also note that these two spikes cannot be observed in all the trajectories depicted in Figure 3 (based on the complete dataset) or in Figure 9 (based on the DS5 dataset), because we sorted the usage data based on the first month of use for the residents who joined the clinic in July 2002. This rearrangement does not affect the DS10 dataset, however.

Results of the DS5 dataset and the complete dataset differ significantly, in particular, on group memberships for "Moderate" and "Heavy" users, whereas it is very similar for those "Light" users. This indicates that although the initial 5-month usage data are insufficient to reliably distinguish users with dynamic behaviors (those in the moderate and the heavy user groups), the information contained in DS5 is reasonably sufficient to detect the relatively stable "Light" users. This is further verified by the posterior probabilities of group membership: residents assigned to the light user group have the highest mean probability, .97, of correct membership, versus .86 for the moderate group and .85 for the heavy user group.

In addition, shapes of the trajectories obtained from DS5 (Figure 9) and from the complete dataset (Figure 3) are not consistent. The trajectories based on DS5 indicate that "Heavy" users had the lowest initial use rate, yet this rate steadily increased to a level of about 80%. In contrast, "Moderate" users had a constantly decreasing trajectory, starting at about 80% level and ending at about 60% level. The disparate group membership assignments and the discrepancy between resulting trajectories suggest that users' adoption behavior had not achieved steady-state by the end of the first 5-month period. As we see from the trajectories in Figure 8, which are derived from the complete 10-month observations, use rate of "Heavy" users keeps rising after the first 5-month period and reaches about 100% in Month 10, while use rate of the moderate user group keeps decreasing and eventually drops to a level comparable with that of "Light" users.

E. Impact on Patient Encounters

We can infer an early impact of the system on patient encounters by evaluating the proportion of all responses associated with a specific action. These are shown in Table 8 for the two chronic diseases and five preventive care categories addressed by the system. CRS also allows residents to enter notes/comments on each reminder. This comment rate appears in the last column of Table 8. Across all disease types and preventive care categories, the percentage of "Favorable" responses is significantly higher than that of any other response type (p<.001). It is also significantly higher than that of "Unhelpful" responses (consisting of responses of "Unfavorable" and "Reviewed but no action", p<.05). The high proportion of "Favorable" responses suggests that if the system is in use, residents' compliance with the CRS recommendations, or the implemented evidence-based medicine guidelines, is satisfactory. Accuracy and relevance of the reminders generated are thus verified in this field setting.

Reminders concerning pneumococcal vaccine received the lowest "Favorable" rate and the highest "Unfavorable" rate. We hypothesize that it is because pneumococcal vaccines administered elsewhere were not captured in any of the hospital systems, causing these reminders to be considered irrelevant. Comments entered with responses to these reminders also show that "vaccine was already done", and was followed by a date entry. Reminders concerning steroid-induced osteoporosis were skipped most often, which we anticipated in advance: for educational purpose the clinic required this category of "on the fly" questions and reminders to appear in all patient encounters, regardless of whether they were relevant or not.

On average, residents entered notes for about 20% of all reminders that they evaluated. Examples of such notes include: "she doesn't want take vaccine", "Mammo already performed 6 months ago - will enter in system". About 1,300 such notes were recorded in the application during the 10-month study period. These notes were analyzed using semantic *SQL* queries in *Oracle*. 37.95% of them indicate actions recommended by CRS were taken during the encounter, 30.91% indicate suggested actions had been taken elsewhere, and 19.05% indicate suggested actions were scheduled or deferred to a future date.

Despite the fact that the incidence of reminders by disease or preventive care type was relatively constant across the three usage groups (Table 9), users of distinct groups responded to reminders differently (Table 10). For instance "Heavy" users recorded favorably responses and entered notes more often, whereas "Light" users had the highest skipping rate.

F. Qualitative Assessments of Barriers

In this study we administrated two user satisfaction surveys. In the first survey, usable and completed forms were obtained from 25 out of 32 residents who participated. The second survey generated 29 valid responses out of 37 residents who participated. Table 11 summarizes quantitative results from these surveys. Ratings on various aspects of the system do not vary dramatically between the two surveys, and residents' attitudes towards the system are relatively neutral. In both surveys, about two-thirds of the residents provided feedback in response to the optional open-ended questions. We use the constant comparative method [31] to analyze these qualitative data.

Residents' feedback is first grouped corresponding to the questions asking for positive or negative feedback. Two major themes emerge from the positive comments, labeled as "practice implications" and "ease of use", and three major themes emerge from the negative comments, labeled as "time and efficiency issues", "heavy data entry duty", and "physician-patient communication".

Most of the positive feedback is grouped under the theme of "practice implications" that include specific comments suggesting that the system had positive implications for medical practice. Examples are: "It does move me to think about some preventive measurements" and "Hard to miss things we usually tend to". This theme shows a general consensus among residents that applications such as CRS can potentially enhance clinical performance and thus lead to better quality of care. Another major theme of positive feedback is "ease of use", including comments that suggest use of the system required very minimal skills. Difficulty of use or lack of computer proficiency does not appear to be a significant barrier for incorporating the system into users' routine practice.

Most of the negative feedback is grouped under the theme of "time and efficiency issues". This feedback focuses on perceptions that use of the system was time consuming and inefficient. Examples of such feedback are: "time consuming and detrimental to efficiency", and "takes too much time to review reminders". The next theme, "heavy data entry duty", is directly related to these efficiency issues: residents complained that use of the system required considerable amount of time for data entry, for example, the system requires residents to enter results of the lab tests performed offsite. Typical comments of the "heavy data entry duty" theme are "It's very tedious to put in all of the work" and "Takes too much time to enter patient". A final category of negative feedback is grouped under the theme of "physician-patient communication", suggesting that use of the system during patient encounters was distracting and diminished the quality of physician-patient interaction, for example, "Using the system is disruptive during patient encounters". This concern has been confirmed by videotaped physician-patient encounters in the clinic, which will be the basis for a more detailed assessment.

These themes resulted from qualitative analysis have provided valuable insights into reengineering the application and adapting it better to the clinical

workflow and functionality requirements. Next version of CRS is now under development with specifically designed features such as: a fully web-enabled interface for better user experience; full integration into CRS of physician notes already used in the clinic for less data entry; elimination of "on-the-fly" questions that users have found intrusive, and finally, certain reminders are carefully rephrased as "recommended orders", implying immediate action, for which responses are highly expected. Other process modifications to improve acceptance of CRS include specialized psychologist directed training in physician-patient communication, and increased emphasis on CRS use by physician supervisors and administrative staff. Details of this system reengineering are documented elsewhere [33].

V. Discussion

Resistance exists to use of the Clinical Reminder System. Although CRS has demonstrated the potential for beneficial patient health outcomes and reduction in medical errors, and has been accepted and routinely used by some of the residents, a significant proportion of users only utilized the system in slightly over 30% of their patient encounters. Although "Light" users skipped nearly one-third of the reminders, "Moderate" and "Heavy" users evaluated the recommendations more carefully, and over half of such advisories led to follow-up actions: lab tests were ordered, medications were prescribed, and future appointments were scheduled. These actions could be otherwise missed due to the high volume of patient data needs to be processed in a limited period of time. Meanwhile, "Heavy" users' ever-increasing usage reaching a level of 100% compliance is a strong indication that CRS has the potential to be fully utilized to improve clinical practice.

To determine the differences between "Heavy", "Moderate", and "Light" users, we relate system use and developmental trends to various user characteristics. We find that gender, citizenship, computer use, and computer optimism significantly influence the group membership probabilities. Computer use and optimism, rather than computer knowledge, are found to be positively correlated with the actual use. As compared with senior clinicians, the residents who participated in this study are younger and have had more extensive exposure to computer technologies. The primary barrier to routine use of computers and information systems such as CRS is not, according to our analyses, lack of computer-specific skills but rather a user's lack of involvement with various types of computer applications as well as skeptical attitudes towards such applications. Although our sample is relatively small and specialized (41 internal medicine residents), our identification of effects of user characteristics on system utilization can still provide value for evaluation studies of CDSS and information systems more generally.

Developmental trajectories based on actual use also show that "Moderate" users started with the highest level of initial use while this use rate consistently decreased to a level comparable with that of "Light" users. In contrast, the heavy user group started moderately but ended up with nearly 100% compliance. This observation suggests that when studying acceptance and adoption issues of information systems, a snapshot of usage immediately after the system implementation may not be a reliable indicator of future usage levels. A considerable length of period needs to be anticipated by system evaluators to allow users to achieve any levels of saturation. We also show that developmental trajectory analysis, which identifies distinctive groups of individual trajectories within a population and describes the course of a developmental behavior over time, provides valuable insights into studying technology acceptance and adoption behavior. Researchers and managerial personnel may utilize this approach to identify distinct trends as to precisely differentiate users at different levels, so that tailored training or other just-in-time incentive strategies can be initiated to encourage system use. As the two usage spikes observed in this study revealed, however, intervention strategies, such as administrative directives, do influence usage, but the effects of such short-term strategies may vanish very quickly.

Qualitative analysis based on user satisfaction surveys reveals some causes of the observed resistance. Decreased efficiency, increased data entry effort, and diminished quality of physician-patient communication are major themes that emerge from the negative feedback. These contextual factors appear to account for the majority of resistance to this reminder system. This finding is aligned with the results of some other hospital information system implementation studies [34].

These contextual factors affecting application usage appear to have differing impacts on different type of users, resulting in distinct adoption behaviors. For example, some users were able to overcome all barriers and achieve frequent use and satisfactory compliance to reminders, while others simply refused to change. Further qualitative research is planned to investigate these different levels of effects. The three groups identified based on actual system use provide a basis for initiating such qualitative investigation, for instance, to selectively interview representatives of each of the groups and assess their distinct views on the system.

A number of limitations, both technology- and workflow-based, of the version of CRS evaluated are the subject of current development efforts. Software and process modifications are underway to improve utilization of the system and reminder compliance.

VI. Conclusions

In this study, we assess medical residents' acceptance and adoption of a clinical reminder system for chronic disease and preventive care management in an ambulatory care environment. We use a novel

developmental trajectory approach to identify distinct groups, following distinct usage trajectories, among those who recorded use of the reminder system within an evaluation period of 10 months. We find that users in this study can be clustered into three groups: "Light" users (41.46% of all users) who used the system steadily over time for about 35% of their patient encounters, "Moderate" users (36.59%) whose initial use rate was the highest (70%) among all groups but declined steadily to a level comparable with that of "Light" users, and "Heavy" users (21.95%) whose use rate, initially moderate (50%), increased to nearly 100% at the end of the evaluation period. Compliance with reminders also varies across usage groups: "Heavy" users tend to respond more favorably to the reminders while "Light" users skipped a large proportion of them. This clustering is related to user characteristics: gender, citizenship, computer use and optimism are found to be correlated with level of actual use. Customized training and novel incentive strategies can thus be developed for groups of users with distinct adoption behaviors. We conclude that this developmental trajectory approach combined with qualitative assessments has considerable promise to provide new insights into system usability and technology adoption issues that may benefit CDSS as well as information systems more generally

Actual use is a key variable in assessing success of an information system after it moves from laboratory to field. Despite acknowledged benefits, various contextual factors, such as decreased efficiency, increased data entry effort, and diminished quality of physician-patient communication, affect the actual use of this reminder system. Extensions of this study will determine the extent to which the enduring use of this system and its reengineered versions results in increased compliance rates for medical guidelines and thereby improved quality of care.

This study has several limitations. First, evaluation of the reminder system is conducted by its developers, which may introduce bias in interpreting the system success. Second, the study is based on a single clinical reminder system. The findings may not be generalizable to other types of CDSS, or information systems more generally. Finally, the user population is relatively small, and composed of only internal medicine residents. The findings may not be generalizable to other types of clinicians who may have different views on such systems.

Acknowledgements

References

[1] Curtin K, Hayes BD, Holland CL, Katz LA. Computer-generated intervention for asthma population care management. Effective Clinical Practice. 1998;1:43-46.

[2] Nilasena DS, Lincoln MJ. A computer-generated reminder system improves physician compliance with diabetes preventive care guidelines. Proceedings of the annual symposium of computing applications to medical care. 1995:640-45.

[3] McDonald CJ, Hui SL, Smith DM, Tierney WM, Cohen SJ, Weinberger L, McCabe GP. Reminders to physicians from an introspective computer medical record. Annals of Internal Medicine. 1995;100:130-38.

[4] Tierney WM, Hui SL, McDonald CJ. Delayed feedback of physician performance versus immediate reminders to perform preventive care: effects on physician compliance and medical care. 1986;24:659-66.

[5] Litzelman DK, Dittus RS, Miller ME, Tierney WM, Requiring physicians to respond to computerized reminders improves their compliance with preventive care protocols. J. Gen. Intern. Med. 1993;8:311-317.

[6] Hunt DL, Haynes RB, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes. JAMA. 1998;280:1339-46.

[7] Kaplan B. Evaluating informatics applications - clinical decision support systems literature review. Int J Med Info. 2001;64:15-37.

[8] Shea S, DuMouchel W, Bahamonde L. A metaanalysis of 16 randomized controlled trials to evaluate computer-based reminder systems for preventive care in the ambulatory setting. J Am Med Inform Assoc. 1996;3:399-409.

[9] Balas EA, Austin SM, Mitchell JA, Ewigman BG, Bopp KD, Brown GD. The clinical value of computerized information services. Arch. Fam. Med. 1996;5:271-78.

[10] Van der Lei J, Musen MA, van der Does E, Man in't Veld AJ, van Bemmel JH. Comparison of computer-aided and human review of general practitioners' management of hypertension. Lancet. 1991; 338:1504-1508.

[11] Kaplan B. Evaluating informatics applications - some alternative approaches: theory, social interactionism, and call for methodological pluralism. Int J Med Info. 2001;64:39-56.

[12] Rousseau N, McColl E, Newton J, Grimshaw J, Eccles M. Practice based, longitudinal, qualitative interview study of computerised evidence based guidelines in primary care. Br. Med. J. 2003; 326(7384):314.

[13] Heathfield HA, Wyatt J, Philosophies for the design and development of clinical decision-support systems. Meth. Inform. Med. 1993; 32:1-8.

[14] Heathfield H, Pitty D, Hanka R. Evaluating information technology in health care: barriers and challenges. Br. Med. J. 1998;316:1959-1961.

[15] Devaraj S, Kohli R. Performance impacts of information technology: is actual usage the missing link? Management Sci. 2003; 49:273-289.

[16] Devaraj S, Kohli R. Information technology payoff in

the health-care industry: a longitudinal study. J. Management Info. Sys. 2000; 16:41-67.

[17] Berner ES, Maisiak RS. Influence of case and physician characteristics on perceptions of decision support systems. J Am Med Inform Assoc 1999; 6:428-434.

[18] Straub D, Limayem M, Karahanna-Evaristo E. Measuring system usage: Implications for IS theory testing. Management Sci. 1995; 41:1328-1342.

[19] Szajna B. Empirical evaluation of the revised technology acceptance model. Management Sci. 1996; 42:85-92.

[20] Agarwal R, Prasad J. The role of innovation characteristics and perceived voluntariness in the acceptation of information technologies. Decision Sci. 1997; 28:557-582.

[21] Sackett DL, Rosenberg MC, Muir Gray JA, Haynes RB, Richardson WS. Evidence based medicine: What it is and what it isn't. BMJ. 1996; 312:71-72.

[22] Sim I, Gorman P, Greenes RA, Haynes RB, Kaplan B, Lehmann H, Tang PC. Clinical decision support systems for the practice of evidence-based medicine. J Am Med Inform Assoc. 2001;8:527-34

[23] Hunt DL, Jaeschke R, McKibbon KA. Using electronic health information resources in evidence-based practice. JAMA. 2000; 283:1875-1879.

[24] Diamond H, Johnson MP, Padman R, Physician reminder system: a relational database application for evidence-based medicine practice, Working paper series 2000-19, The H. John Heinz III School of Public Policy and Management, CMU, Pittsburgh, PA 15213.

[25] Cork RD, Detmer WM, Freidman CP. Development and initial evaluation of an instrument to measure physicians' use of, knowledge about, and attitudes toward computers. J Am Med Inform Assoc. 1998; 5:164-76.

[26] Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. Int J of Human-Computer Interaction. 1995; 7:57-78.

[27] Nagin, DS. Analyzing developmental trajectories: A semi-parametric, group-based approach. Psychological Methods. 1999; 4:139-157.

[28] Nagin, DS, Tremblay RE. Trajectories of boys' physical aggression, opposition and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. Child Development. 1999; 70: 1181-1196.

[29] Nagin DS, Tremblay RE. Analyzing developmental trajectories of distinct but related behaviors: A group-based method. Psychological Methods. 2001; 6:18-34.

[30] Christ M, Krishnan R, Nagin DS, Kraut R, Guenther O. Trajectories of individual web usage: Implications for electronic usage. Proc. 34th Hawaii International Conference on System Science 2001.

[31] Friedman CP, Wyatt JG. Evaluation methods in medical informatics. New York: Springer 1996.

[32] Massaro TA. Introducing physician order entry at a major academic medical center: I. Impact on organizational culture and behavior. Academic Medicine. 1993; 68:20-25.

[33] Zheng K, Padman R, Johnson MP, Diamond HS. Understanding Technology Adoption in Clinical Care: Clinician Adoption Behavior of a Point-of-Care Reminder System. Heinz School Working Paper.

[34] Schwarz, G.. Estimating dimensions of a model. Annals of Statistics. 1978; 6:461-464.

Table 5. Group Profiles

User Characteristics	All	User Groups		
		Light	Moderate user	Heavy
Female (%)	43.90	41.2	33.3	66.7
Male (%)	56.1	58.8	66.7	33.3
Non-U.S. citizen (%)	48.78	23.5	66.7	66.7
U.S. citizen (%)	51.22	76.5	33.3	33.3
Number of appointments *	14.04	14.13	14.12	13.75
Number of visits *	8.44	8.50	8.67	7.95
Number of visits w/ system use *	4.53	3.02	5.45	5.87
Computer use score	33.54	33.0	33.56	34.13
Computer knowledge score	32.96	36 **	31.33	31.38
Computer optimism score	53.19	48.56	52.78	58.88 ***

* Monthly average
** Significantly higher than that of the moderate user group at .05 level
*** Significantly higher than that of the light user group at .001 level

Table 6. The Impact of User Characteristics on the Probability of Assignment to Usage Groups

Variable condition	User Group		
	Light	Moderate	Heavy
Multinomial logit coefficients			
Gender - female	-	-.56 (-0.62)	.27 (0.24)
Non-U.S. citizen	-	1.67 (1.87)	.68 (0.6)
Computer use	-	.045 (.39)	.23 (1.46)
Computer knowledge	-	-.09 (-.985)	-.06 (-.546)
Computer optimism	-	.31 (2.18) *	.98 (1.59)
Predicted membership probabilities based on multinomial logit model coefficient estimates			
Without user characteristics	.49	.36	.15
Gender - female only	.54	.27	.19
Gender - male only	.46	.41	.12
Non-U.S. citizen only	.29	.57	.14
U.S. citizen only	.63	.22	.15

* $p < .05$

Table 7. Group Membership Agreement between FD, F5 and Complete Dataset

Dataset	Measure	Group		
		Light user	Moderate user	Heavy user
F10	Chi-square	16.8 *	13.75 *	15.81 *
	Kappa	.92	.81	.93
F5	Chi-square	23.14 *	.32	1.45
	Kappa	.78	.033	.30

* Significant at .001 level.

Table 8. Response Rate by Protocols

Reminder type	Total	Response Rate of different types (%)				
		Favorable response	Unfavorable response	Reviewed but no action	Skipped	Commented
Diabetes	3299	54.41	10.22	19.22	16.16	16.58
Hyperlipidemia	73	61.64	5.48	26.03	6.85	23.29
Influenza vaccine *	264	53.79	14.02	15.15	17.05	12.88
Pneumococcal vaccine	581	39.41	16.18	27.02	17.38	23.92
Breast cancer	1228	46.01	12.21	24.76	17.02	25.08
Cervical cancer	615	52.36	10.89	18.86	17.89	25.69
Steroid-induced osteoporosis	175	52.57	8.00	14.29	25.14	8.57

* 1 month out of the 10-month study period was within the flu season

Table 9. Proportion of Different Disease Types * Treated by Each Group

Disease type	User Group		
	Light (%)	Moderate (%)	Heavy (%)
Diabetes	53.24	54.00	52.83
Pneumococcal vaccine	10.05	10.72	9.58
Breast cancer	18.89	19.95	19.16
Cervical cancer	10.26	9.48	10.18

* Protocols receiving less than 200 reminders are excluded.

Table 10. Response Rates of Each of the Groups

User Group	Response rate of different types (%)				
	Favorable response	Unfavorable response	Reviewed but no action	Skipped	Commented
Light	42.87	9.53	14.74	32.87 *	11.85
Moderate user	49.86	14.20	22.46	13.48	18.92
Heavy	61.98 *	10.35	19.54	8.13	26.38 **
Average	46.37	11.36	18.91	10.81	15.39

* Significantly higher than that of the other two groups at .001 level
** Significantly higher than that of the light user group at .01 level

Table 11. User Satisfaction Survey Results

Type of evaluator	Mean of system items		Mean of interface items		Overall rating	
	Survey 1	Survey 2	Survey 1	Survey 2	Survey 1	Survey 2
N	25	29	24	29	25	29
Mean	3.38	3.36	4.20	3.99	3.28	3.23
Std. Deviation	1.16	.97	1.15	1.05	1.43	1.42
Median	3.53	3.47	4.00	4.00	4.00	3.00