34

An Analysis of Physicians' Attitudes

Randy L. Teach and Edward H. Shortliffe

Despite the promise of medical computing innovations, many health care professionals have expressed skepticism about the computer's role as an aid to clinicians. A number of barriers have been noted. For example, Friedman and Gustafson (1977) have suggested that system designers tend to develop systems that are neither convenient for physicians nor responsive to their needs. Glantz (1978) has questioned the trade-off in costs and benefits for most medical computing applications, including computer-assisted consultations. Schwartz (1970) has noted that physicians are wary of formal decision aids because they perceive such tools to be a threat to their jobs and to their professional stature. He has also suggested that physicians are concerned about their ability to learn how to use computer systems (Schwartz, 1979), but that they simultaneously fear the prospect of being "left behind" if they fail to keep current. Other observers (Eisenberg, 1974; Weizenbaum, 1976) have questioned the role of computers as clinical consultation systems, suggesting that computer-based consultants may be an inappropriate use of computing technology that will inevitably degrade and debase the human function.

Observations such as these are generally based on personal experience without benefit of formal studies of physicians' attitudes. The few available studies have sought physicians' opinions regarding computing technology in general, but have tended not to specifically examine attitudes regarding the clinical introduction of computers. One early study (Mayne et al., 1968) found little physician interest or faith in the role of computing technology. However, Startsman and Robinson (1972) and others (Day, 1970; Resnikoff

---

This chapter is based on an article originally appearing in *Computers in Biomedical Research* 14: 542–558 (December 1981).

et al., 1967) have reported supportive physician attitudes. A follow-up to the Startsman and Robinson study by Melhorn and coworkers (1979) produced almost identical results, but also noted that physicians might be reluctant to accept the clinical use of computing technology.

Motivation for the Current Study

Our study was motivated by the belief that the future of research in medical computing, particularly the development of computer-based consultation systems, depends on improving our understanding of the needs, expectations and performance demands of clinicians. The previous studies had not specifically addressed these issues. Our study used a questionnaire, similar in format to the instrument developed by Startsman and Robinson (1972) but different in content. One modification was to limit the scope of our survey by focusing only on physicians' attitudes regarding clinical consultation systems. Previous studies had been more general in their focus and had surveyed a broader range of opinion. We chose this more limited focus because several research groups currently developing medical consultation systems are concentrating on physician users and have recognized the need for better information about the concerns and performance demands of clinicians. Another change was the inclusion of statements designed to ascertain the performance capabilities that physicians consider necessary for a consultation program to be clinically acceptable. Previous studies had not addressed this important aspect. We hoped that with these modifications the study would yield results from which guidelines could be formulated to help medical computing experts design more acceptable clinical consultation systems.

Relationship Between Physicians' Characteristics and Attitudes

A second objective of the study was to test the common assumption that prior experience with computers affects attitudes about the clinical use of computing technology. We therefore included measures of both computing experience and *knowledge* of computing concepts in the questionnaire. A number of other demographic variables were also included.

Impact of a Medical Computing Course on Attitudes

A third objective was to assess the impact of an intensive medical computing course on physicians' attitudes. The authors of both of the previous major studies (Startsman and Robinson, 1972; Melhorn et al., 1979), as well as others (Levy, 1977), had speculated that intensive educational efforts might result in increased acceptance of medical computing by physicians. Partly to test this assumption, we designed a medical computing tutorial and measured its impact on the attitudes of the physician attendees.[1] The tutorial faculty consisted of 15 physicians and computer scientists who are active researchers in the development of computer-based clinical consultation systems. Presentations encompassed the researcher's work, goals, and perspective on the role of computer-assisted decision making in clinical medicine. An introductory session was included to introduce physicians to general computing concepts and terminology.

34.1 Methods

34.1.1 Instrument

A survey instrument (questionnaire) was developed to measure physicians' attitudes regarding computer-based consultation systems. Attitudes were measured by the instrument along three dimensions: (1) the *acceptability* of different medical computing applications; (2) *expectations* about the effect of computer-based consultation systems on medicine; and (3) *demands* regarding the performance capabilities of consultation systems. Every effort was made to include items representative of the design issues that are currently being considered by medical computing experts. We performed extensive pilot testing of the questionnaire prior to its use in the study.

Acceptance was measured by asking physicians about eight real or imagined medical computing applications. The applications ranged from computer-based medical records to the use of computers as substitutes for physicians in underserved areas (Table 34-1). The Expectation- and Demand-scales included statements about medical computing, emphasizing the potential role of computer-based consultation systems. Each statement used a Likert-type scale in which respondents were instructed to mark one of five categories: (1) strongly disagree, (2) somewhat disagree, (3) not sure, (4) somewhat agree, (5) strongly agree.

The Expectation-scale (E-scale) included 17 statements and was designed to measure physicians' opinions about how computer-based consultations are likely to affect the practice of medicine (i.e., how computers *will* affect medical practice).[2] The Demand-scale (D-scale) of 15 statements

[1]The tutorial was offered by the Departments of Medicine and Computer Science at Stanford University in August of 1980. It was organized in conjunction with the Sixth Annual Workshop on Artificial Intelligence in Medicine, which was sponsored by the Division of Research Resources of the NIH.

[2]The statements are shown in Table 34-3. For identification purposes in this paper, each is identified by an E followed by a number. The letter E denotes that the statement belongs to the Expectation-scale.

sought physicians' opinions regarding the most desirable performance capabilities for computer-based consultation systems (i.e., what computers *should* be able to do).[3] The possible range of ratings for statements on both the E- and D-scales is -2 to $+2$. On the E-scale a positive rating means that respondents felt that the stated effect is not likely to occur, and a negative rating means that they felt that the effect is likely. On the D-scale a positive rating means that the item was judged to be an important capability for computer-based clinical systems, and a negative rating means that it is judged to be unimportant.

A set of background questions was also included on the questionnaire. These included items about medical specialty, type of practice (academic medicine or private practice), number of years since receiving the M.D. degree, percentage of time devoted to research, and extent of prior experience with computers. All questions in this group contained fixed response categories. A second set of 22 questions asked respondents to indicate their (self-reported) level of *knowledge* about computers and computer science concepts.

34.1.2 Participants

Two samples of physicians were included in the study. One included registrants for the tutorial mentioned above. The 85 physicians who filled out the questionnaire represented 90% of the physicians registered for the tutorial. Twenty-nine nonphysician attendees who were engaged in either basic medical research or medical computing also returned survey forms.

By announcing that the course was appropriate for physicians with little or no knowledge of medical computing, we hoped to attract a cross section of physicians. Although continuing medical education (CME) credit was also available, we were aware that the backgrounds and attitudes of these physicians might contrast with those who chose not to attend the tutorial. Therefore, a second sample of physicians was selected from Stanford Medical School clinical faculty and from Stanford-affiliated physicians practicing in the surrounding community.

34.1.3 Procedure

The questionnaire was included in the preregistration packet that was mailed to all tutorial registrants approximately one month before the course. A cover letter asked respondents to complete and return the questionnaire as soon as possible so that the results could be used to guide the

[3]The Demand-scale statements are shown in Table 34-5. Each statement is identified by a D followed by a number.

speakers' presentations. At the end of the tutorial, participants were asked to complete the same questionnaire for a second time. A respondent-selected code number facilitated matching of pretutorial and posttutorial results. To encourage open and unbiased responses, the respondents were assured of anonymity.

The second sample, stratified by medical specialty, was randomly selected from the roster of Stanford Medical School faculty and affiliated community physicians. These individuals, 57 faculty members and 92 affiliated physicians, received a questionnaire with a cover letter requesting their help with the research study and assuring them of anonymity. The letter also invited them to participate in the tutorial and instructed them to return the registration form instead of the questionnaire if they wished to do so. None chose to register.[4] A follow-up letter was sent to the entire 149-member sample three weeks after the original mailing to maximize questionnaire return. Sixty-one questionnaires of the original 149 were eventually returned (41%).

Nonparametric Chi-square analysis was used to compare the tutorial and nontutorial samples. Reliability of the attitude scales was determined on a subsample of ten subjects (Cronbach, 1970). Internal consistency of the scales was calculated by correlating odd and even items and correcting the resulting correlations using the Spearman-Brown formula (Cronbach, 1970). Means and standard deviations were computed for each of the individual statements included on the three attitude scales. The Expectation- and Demand-scales were subjected to factor analysis to identify meaningful subgroupings of statements. Principal factoring with iteration was employed (Nie et al., 1975). Simplification of the factor structure was obtained by oblique rotation with delta set equal to zero. Analysis of variance was used to compare the attitudes of physicians with different backgrounds and knowledge of medical computing. Analysis of variance was also used to compare pretutorial and posttutorial ratings.

34.2 Results

34.2.1 Characteristics of Physicians Studied

The final sample of 146 physicians included subsamples of 85 tutorial participants and 61 physicians who were associated with Stanford University Medical Center but who chose not to participate in the tutorial (control group). Of the combined sample, 43% were in medical fields (internal

[4]All recipients had also received an initial announcement for the course several weeks earlier, and none had registered in response to the initial mailing.

medicine, family practice, pediatrics, general practice), 27% were from surgical fields (general surgery, surgical subspecialties, obstetrics/gynecology, anesthesiology), and 30% were from other specialties (primarily radiology and pathology). There was no significant difference between the two subsamples (Chi-square = 5.16, $p > .05$).

Of the combined sample, 44% were academicians, 45% were in private practice, and 11% were Stanford house staff.[5] Differences between the subsamples (Chi-square = 6.28, $p < .01$) were due to the separation of the house staff group from the academic subgroup. A separate analysis of house staff responses to the questionnaire items revealed that they had response patterns almost identical to those of the academicians. Incorporation of the house staff into the academic category resulted in comparable frequencies for the attendees and controls (Chi-square = 4.93, $p > .05$).

Of the combined sample, 31% had fewer than 10 years of experience since graduating from medical school, 22% 10 to 20 years, and 47% more than 20 years. Differences between the attendees and controls were not significant (Chi-square = 3.24, $p > .20$). While 43% of subjects reported that they devoted no time to research, 27% devoted less than a third of their time, and only 30% devoted more than a third of their time to research. The difference between attendee and control groups was not significant (Chi-square = 5.73, $p > .05$). Finally, 46% reported no computing experience, 32% had had some experience (i.e., at least running "canned" computer programs), and 22% reported extensive experience including the design of computing systems. There was no significant difference between the tutorial attendees and the controls (Chi-square = 3.17, $p > .20$).

34.2.2 Acceptance Ratings

The options for the Acceptance question are shown in Table 34-1. Physicians had an average Acceptance rating of 5.5 applications out of the 8 included on the scale. The table shows that support for the 5 major applications exceeded 80% of respondents.

Medical speciality was the only characteristic that was significantly predictive of a respondent's Acceptance of computing applications. Table 34-2 shows that surgeons were less accepting of medical computing applications than either of the other two subgroups. There was no significant difference in the Acceptance rating between tutorial and nontutorial participants, private practice and academic physicians, those with several years in practice and those who had recently graduated, physicians engaged in research and those who were not, or physicians with and without computing experience.

[5]All house-staff subjects were tutorial attendees rather than members of the control group.

TABLE 34-1 **Physicians' Acceptance of Medical Computing Applications**

		97%						
	83%	XXXX			83%			
	XXXX	XXXX	80%	81%	XXXX			
%	XXXX	XXXX	XXXX	XXXX	XXXX			
Acceptance	XXXX	XXXX	XXXX	XXXX	XXXX			
	XXXX	XXXX	XXXX	XXXX	XXXX			52%
	XXXX	XXXX	XXXX	XXXX	XXXX	36%		XXXX
	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	32%	XXXX
	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
	Medical Records	HIS*	Patient Monitoring	Diagnostic Consults	Therapy Consults	Physical Exams	M.D. Substitute†	M.D. Licensure‡

*Hospital Information Systems
†Substituting for physicians in medically underserved areas
‡Testing physicians for relicensure

TABLE 34-2 Scheffe Comparison of Acceptance Ratings for Subgroups of Medical Specialists

Specialty	*Mean*	*Standard deviation*	*Significance*
1. Medical	6.03	1.55	1 vs. 2 → $p < .01$
2. Surgical	4.35	1.82	2 vs. 3 → $p < .01$
3. Other	5.67	1.84	
Total	5.45	1.84	

34.2.3 Expectation Ratings

Table 34-3 displays the ratings and standard deviations for each statement on the Expectation-scale. The statements are listed in order of their average ratings, from those outcomes that physicians thought were the most likely to occur to those that were expected to occur less frequently. The average Expectation rating for physicians was slightly positive ($X = .42$). This was comparable to that of the nonphysician sample, shown in the right-hand column. Only 3 of the 17 statements received negative ratings (i.e., were judged likely to occur), including fears about the possibility that consultation systems will increase government control of medicine, concerns that systems will increase the cost of care, and expectations that patients will blame the computer program for ineffective treatment decisions. On the other hand, physicians felt strongly that consultation systems would neither interfere with their efficiency nor force them to adapt their thinking to the reasoning process used by the computer program. They also felt that the use of consultation systems would *not* reduce the need for either specialists or paramedical personnel.

Subgroups of physicians displayed significant differences in their Expectations about how computer-assisted consultations will affect medical practice. The means and standard deviations for all the significant findings are summarized in Table 34-4. A significance level of .01 was used for each analysis in order to maintain an overall significance level of less than .06. The Expectations of tutorial registrants were on the average more positive than those of the nontutorial group, although neither group thought that consultation programs would adversely affect medical practice. Physicians in academic settings and those in training indicated overall positive Expectations, whereas private practice physicians tended to hold slightly negative Expectations. Young doctors expressed more positive Expectations than did physicians with 10 to 20 years of experience, although the recent graduates were no more positive than physicians with at least 20 years experience. Experience with computers was positively related to Expectations, as was Knowledge about computing concepts.

TABLE 34-3 Means and Standard Deviations (in Parentheses) for Ratings of Expectation Statements

		Physicians *n = 146*	*Nonphysicians* *n = 29*
E1.	*Will* increase government control of physicians' practices	−.26 (1.23)	.15 (.95)
E2.	*Will* be blamed by patients for errors in management	−.23 (1.15)	−.30 (1.10)
E3.	*Will* increase the cost of care	−.14 (1.07)	.44 (1.09)
E4.	*Will* threaten personal and professional privacy	.02 (1.41)	.50 (1.45)
E5.	*Will* result in serious legal and ethical problems (e.g., malpractice)	.32 (1.06)	−.04 (.98)
E6.	*Will* threaten the physician's self-image	.32 (1.23)	.15 (1.01)
E7.	*Will* be hard for physicians to learn	.34 (1.17)	.85 (.95)
E8.	*Will* result in reliance on cookbook medicine and diminish judgment	.43 (1.34)	.92 (1.14)
E9.	*Will* diminish the patient's image of the physician	.45 (1.16)	.74 (1.10)
E10.	*Will* be unreliable because of computer "malfunctions"	.51 (1.09)	1.07 (.83)
E11.	*Will* dehumanize medical practice	.53 (1.34)	1.04 (1.09)
E12.	*Will* depend on knowledge that cannot easily be kept up to date	.53 (1.20)	1.00 (1.00)
E13.	*Will* alienate physicians because of electronic gadgetry	.62 (1.03)	.41 (1.08)
E14.	*Will* force physician to think like computer	.73 (1.15)	1.19 (1.00)
E15.	*Will* reduce the need for paraprofessionals	.83 (.91)	.82 (1.08)
E16.	*Will* reduce the need for specialists	.99 (1.07)	1.11 (1.09)
E17.	*Will* result in less efficient use of physician's time	1.05 (.84)	1.56 (.58)
	Total scale =	.42	.68

TABLE 34-4 Scheffe Comparisons of Expectations for Physicians with Different Characteristics

Characteristic	*Groups*	*Mean*	*Standard deviation*	*Significance*
Totals		.41	.59	
Professional orientation	1. Academic	.55	.58	1 vs. 2 → $p < .01$
	2. Private	.22	.59	3 vs. 2 → $p < .01$
	3. Training	.64	.48	
Clinical experience	1. < 10 yrs.	.59	.52	
	2. 10 to 20 yrs.	.18	.54	1 vs. 2 → $p < .01$
	3. > 20 yrs.	.39	.63	
Computing experience	1. Little or none	.24	.62	1 vs. 3 → $p < .01$
	2. Moderate	.50	.58	
	3. Extensive	.63	.47	

34.2.4 Demand Ratings

Table 34-5 depicts statements on the Demand-scale, ordered from most to least important according to the average rating each received. Physicians' Demands were significantly less than those of the nonphysicians, although the ranked ordering of each Demand statement was almost the same for the two groups. A system's ability to explain its advice was thought to be its most important attribute. Second in importance was a system's ability to understand and update its own knowledge base. Improving the cost effectiveness of tests and therapies was also important. Physicians did not think that a system has to display either perfect diagnostic accuracy or perfect treatment planning to be acceptable. On the other hand, they would not accept the use of a consultation system as a standard for acceptable medical practice, nor would they recommend reducing the amount of technical knowledge that physicians have to know just because a consultation system is available. The differences found among physician subgroups on the Expectation-scale were not evident on the Demand-scale.

A test-retest reliability coefficient of $r = .94$ was obtained across two administrations of the three scales: Acceptance, Expectations, and Demands. The split-half reliability for the D-scale was only $r = .70$, and that of the E-scale was $r = .83$. These rather modest split-half reliabilities suggested to us that the scales were measuring more than one aspect of physicians' attitudes. In order to better understand the structure of physicians' attitudes measured, these scales were subjected to factor analysis. Five major groups of statements (factors) were extracted from the combined scales and are described below. Correlations among them were low, ranging from .01 to .19, except for Factors 1 and 5, which correlated at .31. The factors accounted for 45% of the total variance of the combined scales.

TABLE 34-5 Means Ratings and Standard Deviations (in Parentheses) for Demand Statements

	Physicians *n = 146*	*Nonphysicians* *n = 129*
D1. *Should* be able to explain their diagnostic and treatment decisions to physician users	1.42 (.80)	1.78 (.42)
D2. *Should* be portable and flexible so that physician can access them at any time and place	1.14 (.81)	1.52 (.51)
D3. *Should* display an understanding of their own medical knowledge	.99 (.94)	1.48 (.80)
D4. *Should* improve the cost efficiency of tests and therapies	.85 (.99)	1.11 (1.58)
D5. *Should* automatically learn new information when interacting with medical experts	.84 (1.02)	1.41 (.75)
D6. *Should* display common sense	.75 (1.20)	1.11 (.97)
D7. *Should* simulate physicians' thought processes	.64 (1.16)	.93 (1.07)
D8. *Should* not reduce the need for specialists	.46 (1.18)	.70 (1.07)
D9. *Should* demand little effort from physician to learn or use	.35 (1.20)	1.19 (.92)
D10. *Should* respond to voice command and not require typing	.26 (1.23)	.56 (1.05)
D11. *Should* not reduce the need for paraprofessionals	.26 (1.06)	.85 (1.03)
D12. *Should* significantly reduce amount of technical knowledge physician must learn and remember	−.08 (1.34)	.00 (1.49)
D13. *Should* never make an error in treatment planning	−.25 (1.33)	−.22 (1.34)
D14. *Should* never make an incorrect diagnosis	−.45 (1.31)	−.26 (1.46)
D15. *Should* become the standard for acceptable medical practice	−.80 (1.13)	.00 (1.07)
Total scale =	.44	.81

TABLE 34-6 Intercorrelation of Physicians' Computing Knowledge, Acceptance, Expectations, and Demands

	Demands	Expectations	Knowledge
Acceptance	.27*	.26*	.27*
Knowledge	.08	.26*	
Expectations	.05		

*$p < .001$

Factor 1 includes statements E7, E8, E11, E13, and E17 (Table 34-3). It relates to Expectations about how physicians might be personally affected by a consultation system. All of these statements received positive ratings (i.e., the outcomes were judged to be unlikely) ranging from .34 to 1.05. Factor loadings for the statements ranged from .43 to .59.[6]

Factor 2 includes statements D1, D2, D3, D5, and D6 from the D-scale (Table 34-5). The factor is composed of the performance Demands thought by physicians to be the most important. Ratings of the statements ranged from .75 to 1.42. Factor loadings for the statements ranged from .41 to .65.

Factor 3 relates to Demands about system accuracy. It includes statements D13 and D14, which were rated relatively unimportant by the respondents. Factor loadings were .84 and .89, respectively.

Factor 4 includes statements from both scales and relates to physicians' attitudes regarding the effect of computing systems on the need for health care personnel. It includes statements E15, E16, D8, and D11. The factor reflects the opinion that consultation systems will not and should not affect the need for either specialists or paraprofessionals.

Factor 5 includes statements E1, E4, E5, E6, E8, E9, and E11 from the E-scale. It is similar to Factor 1 because statements E8 and E11 relate to both factors; however, its focus appears to be slightly different. Whereas Factor 1 related to the individual practitioner, Factor 5 is concerned with the effect of consultation programs on medical practice in general. Factor loadings ranged from −.70 to −.41.

Nearly the same pattern of differences among physicians was found for the factors as was found for the full-scale ratings. Individual differences in Expectations on Factors 1 and 5 were related to differences in knowledge about computer concepts, experience with computers, time in medical practice, professional orientation, and tutorial participation. Individual differences were not found on ratings of the other three factors.

Table 34-6 shows the relationship between the scale ratings and Knowledge about computers and medical computing concepts. Acceptance was

[6]Factor loadings can range from −1.0 to +1.0 and indicate the degree of relationship between each statement and the factor.

moderately related to Knowledge, Expectations, and Demands. Knowledge was also related to Expectations but not to Demands, and Expectations were unrelated to Demands. These results are consistent with the differences reported above for the analyses of variance.

34.2.5 Tutorial Findings

Of the tutorial participants, 50% completed the posttutorial questionnaire. The posttutorial sample did not differ from the pretutorial group on any of the sample characteristics including medical specialty, professional orientation, years of medical experience, time devoted to research, or computing experience.

The tutorial affected physicians in two ways. First, it significantly increased their self-reported knowledge about computing concepts from a mean of 15.0 concepts to a mean of 25.5 concepts ($p < .001$). Second, it raised the level of their performance Demands from a mean of .44 to a mean of .72 ($p < .01$), although the relative importance of the individual statements did not change. Physicians' Expectations did not change overall; although Factor 1 did show a slight change in the positive direction (i.e., the outcomes were judged less likely than they had been before the course), the difference was not enough to be statistically significant. The mean posttutorial Acceptance rating of 6.0 was not significantly different from the tutorial registrants' pretutorial rating of 5.8. Also, participation in the tutorial did not alter the relatively low pretutorial Acceptance ratings of the surgical specialists.

34.3 Discussion

The study we have described had three principal goals: (1) to measure physicians' attitudes regarding consultation systems, (2) to compare the attitudes of subgroups of physicians, including those who chose to attend a medical computing tutorial and those who did not, and (3) to assess the impact of the continuing education course on the attitudes and knowledge of the physicians who enrolled. In this section, we discuss some of the results relevant to each of these goals.

34.3.1 Attitudes of Physicians

There was no significant difference in demographics or computing knowledge between the tutorial attendees and the control group. The overall analysis of physicians' attitudes was therefore based on responses from all physicians surveyed. The respondents were selective in their Acceptance of computing applications. Applications that were presented as aids to clinical practice were more readily accepted than those that involved the automation of clinical activities traditionally performed by physicians themselves. The distinction between a clinical *aid* and a *replacement* seems to be important to physicians and suggests design criteria and preferred modes for the introduction of computing innovations. This perspective is consistent with historical attitudes regarding the adoption of other kinds of technological innovation. For example, computerized axial tomography has been widely accepted largely because it functions as a remarkably useful clinical *tool*, providing physicians with faster and more reliable information, but it in no way infringes on the physician's patient-management role. In contrast, automated history-taking systems have not received widespread acceptance, despite their accuracy and reliability. We suspect that one reason physicians have resisted their use is because they are perceived as a threat to a traditional clinical function.

Some observers have speculated that many physicians oppose computer-based decision aids because they fear a loss of job security and prestige. The study results do not support this viewpoint. The physicians surveyed believe that consultation systems will not reduce the need for either specialists or paraprofessionals. Furthermore, they do not feel that either a physician's self-image or the respect he or she receives from patients will be reduced by the use of this kind of system. They are worried that consultation systems may increase the cost of care, although they believe that the programs should be designed to decrease costs. This Expectation may reflect past experience with new technologies that have generally increased cost, at least initially, but have eventually been accepted because of perceived improvement in patient care. In light of the generally positive Expectations of physicians, as demonstrated in this study, it is unlikely that the acceptance of a medical consultation system will depend solely on its ability to reduce the cost of care; the crucial factor, rather, is likely to be the system's ability to improve the quality of patient care or to simplify its delivery.

The results from the Demand-scale indicate, however, that for a system to improve patient care in an acceptable fashion, it must be perceived as a tool that will *assist* physicians with management decisions. It is clear that physicians will reject a system that dogmatically offers advice, even if it has impressive diagnostic accuracy and an ability to provide reliable treatment plans. Physicians seem to prefer the concept of a system that functions as much like a human consultant as possible.

34.3.2 Comparisons Among Subgroups

Physicians' Expectations about the effect of computer-assisted consultation systems on medical practice were generally positive, although considerable differences among physicians were noted. The finding that physicians with prior computing experience have more positive Expectations regarding the effects of consultation systems supports the belief of other investigators, although even the groups with little or no experience generally had positive attitudes. The slightly more positive Expectations of academic physicians may be a source of encouragement to medical computing researchers because this kind of system development typically depends on support from the academic community. However, the more negative Expectations of private practice physicians and of those who chose not to attend the tutorial are worrisome. These groups represent the majority of practitioners in the country and are, in particular, the physicians for whom many of the research systems are designed.[7] Furthermore, although many of their concerns, such as worries about increased government control of medical practice, defy direct attention by the medical computing researcher, an increased awareness of them may lead to more sensitive design decisions and more tactful introduction of new systems.

34.3.3 Effect of the Tutorial

The tutorial experience had a small but significant effect on physicians' Demands and also produced a substantial increase in their knowledge about computing concepts. The results from the Demand-scale were of particular interest. Physicians apparently gained new insights from the tutorial into the potential use and capabilities of medical computing and increased their performance Demands accordingly. These opinions regarding the attributes of acceptable computing systems were surprisingly uniform across physician subgroups both before and after the tutorial. Our interpretation of this result is that physicians are serious about these Demands and that consultation systems are not likely to be clinically effective, regardless of the accuracy of their advice, until these capabilities have been incorporated.

On the other hand, the tutorial had no significant effect on physicians' Acceptance of computer applications or on their Expectations regarding the effect of consultation systems on medical practice. The failure of the tutorial to change the Acceptance rating is not surprising because the pretutorial ratings were already very high. It is possible that an expanded set of applications on the Acceptance scale, particularly applications that involve the automation of traditional physician functions, would have produced a different result. Similarly, the Expectations of the tutorial registrants were markedly positive prior to the tutorial and were not significantly changed as a result of the course. Before the survey we were concerned that the Expectations of the course participants might decline

[7] Although our study included physicians with different backgrounds and interests (e.g., medical specialty, time devoted to research), we cannot generalize with certainty from our results to the national community of physicians. Our self-selected tutorial participants were almost all academic or academically affiliated, and our nontutorial (control) sample was selected from a similar population.

on the posttutorial questionnaire; it was possible that the physicians in the audience would begin to worry about the effects of certain applications after being exposed to the problems and uncertainties experienced by the medical computing researchers. Instead, the attendees apparently understood both the potential and the problems associated with designing consultation programs and took a more positive approach by increasing their Demands for more humanlike performance from the systems.

Although physicians with positive Expectations could be distinguished from those with negative ones on the basis of their knowledge about computing concepts prior to the tutorial, increasing their knowledge about these concepts did not change their Expectations. Since physicians with negative Expectations were also the least likely to participate voluntarily in our CME program, the effectiveness of CME in increasing the acceptance of clinical computing among the most resistant physicians is questionable. However, the study results indicate that computing applications have already obtained a strong core of support among some physicians. This support may even be deeper than we had expected because, for the physicians we surveyed, it extended to the belief that medical computing should be considered an area of basic medical research, comparable to biochemistry and immunology. In response to a question on this subject included at the end of the questionnaire, 75% of the pretutorial and control group physicians agreed that medical computing should be considered an area of basic medical research, and another 14% were undecided. We believe that this uniformly positive response may have been influenced by the administration of the questionnaire, and physicians asked the same question without the context provided by the survey instrument might respond less favorably. On the other hand, even physicians with minimal computing experience seem likely to accept the fundamental research component of medical computer science if it is pointed out to them. This suggests a strong educational message that must be conveyed to the medical community regarding the research role of the discipline.

34.4 Recommendations

The results of this survey counter the common impression that physicians tend to be resistant to the introduction of clinical consultation systems. Although we have polled physicians only from the immediate vicinity of our medical center, there is no reason to assume that a nationwide survey would achieve markedly different results. We have found that a significant segment of the medical community believes that assistance from computer-based consultation systems will ultimately benefit medical practice. However, a major concern at present is whether system developers can respond adequately to physician demands for performance capabilities that extend beyond currently available computer science techniques. In light of these results, the following recommendations may be helpful.

1. Strive to *minimize changes to current clinical practices.* The system should ideally replace some current clinical function, thereby avoiding the need for an *additive* time commitment by the physician. The system should ideally be available when and where physicians customarily make decisions.
2. Concentrate some of the research effort on *enhancing the interactive capabilities* of the expert system. The more natural these capabilities, the more likely it is that the system will be used. At least four features appear to be highly desirable:
   a. *Explanation.* The system should be able to justify its advice in terms that are understandable and persuasive. In addition, it is preferable that a system adapt its explanation to the needs and characteristics of the user (e.g., demonstrated or assumed level of background knowledge in the domain). A system that gives dogmatic advice is likely to be rejected.
   b. *Common sense.* The system should "seem reasonable" as it progresses through a problem-solving session. Some researchers argue that the program's operation should therefore parallel the physician's reasoning processes as much as possible. There is a growing body of knowledge about the psychological underpinnings of medical problem solving (Elstein et al., 1978), and systems that draw on these insights are likely to find an improved level of acceptance by the medical community.
   c. *Knowledge representation.* The knowledge in the system should be easy to bring up to date, and this often seriously constrains the format for storing information in the computer. A challenging side issue is the automatic "learning" of new knowledge of the domain, either through interaction with expert physicians or through "experience" once the system is in regular use.
   d. *Usability.* The system should be easy to learn and largely self-documenting. The mode of interaction may be the key to acceptability, and effective methods for understanding text or spoken language should dramatically increase the utility of clinical systems. For routine activities, it is preferable that use of the system be as easy as pressing a button.
3. Recognize that *100% accuracy is neither achievable nor expected.* Physicians will accept a system that functions at the same level as a human expert, as long as the interactive capabilities noted above are a component of the consultative process.
4. Consider carefully the *most appropriate criteria for assessing a clinical consultation system.* Not all medical computer programs should be judged

on the same basis, and cost-effectiveness may appropriately be a secondary concern when a system can be shown to significantly improve the quality of patient care or the efficiency of its delivery.

5. When designing systems, *consider the concerns and demands that physicians express* about consultation systems. These should be used to guide both the development and the implementation of the systems of the future. It is increasingly recognized that it takes only one shortcoming to render an otherwise well-designed system unacceptable.

The considerations outlined here place severe demands on current computing capabilities. Many of the issues that we have cited, and that were included on the Demand-scale in the survey, are capabilities that are beyond the current state of the art in computer science. They thus help delineate some of the important basic research issues for future work in medical computing.