International Journal of Technology Assessment in Health Care, 16:1 (2000), 35–49.

PATIENT, PHYSICIAN, AND PAYER PERCEPTIONS AND MISPERCEPTIONS OF WILLINGNESS TO PAY FOR DIAGNOSTIC CERTAINTY

Richard A. Hirth
University of Michigan

Bernard S. Bloom
University of Pennsylvania

Michael E. Chernew
University of Michigan

A. Mark Fendrick
University of Michigan

Abstract

Little is known about the value patients, physicians, and payers place on intangible attributes of care. Differences in valuations among these groups and misperceptions of value of intangible attributes to other groups can contribute to conflicts about treatment recommendations or coverage decisions. We surveyed patients, physicians, and managed care executives to assess their willingness to pay (WTP) for diagnostic certainty for peptic ulcer disease (PUD) and gastroesophageal reflux disease (GERD). To determine if patients, physicians, and payers accurately perceive each other's valuations of diagnostic certainty, participants were also asked to estimate the WTP of each of the other types of respondents. Patients were most likely, and executives least likely, to value diagnostic certainty. For PUD, 84% of patients, 61% of physicians, and 43% of executives expressed a positive WTP. Median WTP was low for all three groups ($1–9 for patients and physicians; $0 for payers). Physicians and executives both correctly predicted patient WTP. For GERD, 87% of patients, 52% of physicians, and 29% of executives expressed a positive WTP. Executives underestimated patient WTP. For both diseases, physicians' WTP was overestimated by patients and underestimated by executives. The inconsistency in the value that patients, physicians, and managed care executives place on diagnostic certainty indicates the potential for conflict over practice guidelines or access to services. WTP surveys can provide information to aid in anticipating and addressing areas of disagreement.

Keywords: Cost-benefit analysis, Managed care programs, Peptic ulcer, Reflux, Gastroesophageal, Diagnostic test, routine

Supported in part by an unrestricted grant from Astra-Merck, Inc. to the University of Pennsylvania. The authors thank John Marquez for research assistance and Seema Sonnad for helpful suggestions. Earlier versions of this work were presented at the annual meetings of the International Society of Technology Assessment in Health Care and the American Gastroenterology Association.

Under traditional fee for service (FFS) health insurance in the United States, insurers were passive payers of claims that resulted from decisions made by physicians and patients. In this environment, it was natural to conceptualize care decisions as arising from an agency relationship between the doctor and patient. An agency relationship is created when one party (the principal) enlists another party (the agent) to act on his or her behalf (13;32;33;38). In this case, the patient (principal) seeks advice from the physician (agent) and expects the physician to act in the patient's best interests.

Incentives faced by patients and physicians under FFS insurance were closely aligned at the time of the clinical encounter: with nearly complete insurance coverage, patients have an incentive to consent to tests and procedures even when expected benefits (relative to risks) are small, and physicians have an incentive to provide these services. However, there are two main reasons why physicians might not act in patients' best interests in a FFS system. First, physicians have a financial incentive to recommend more aggressive intervention than an insured, fully informed patient would demand (the supplier-induced demand controversy) (10;12;19). Second, physicians may not understand patients' preferences, and these misperceptions can lead to suboptimal care plans even if the physician is attempting to act in the patients' best interests.

This close (but not perfect) alignment of physician and patient incentives in an environment with passive third-party payers generated considerable inflationary pressure on healthcare costs (41). As a result, the market share of managed care insurers who take an active role in managing the delivery of care has risen dramatically (25). Unlike traditional FFS insurers, managed care organizations (MCOs) attempt to control costs by employing techniques such as selective contracting, financial incentives, monitoring, and clinical practice guidelines (15;20;23).

Adding the payer as a third type of active participant in care decisions complicates the agency relationships. The physician's new position as a "double agent," simultaneously acting on behalf of the patient and the payer, has increased the potential for conflicts about clinical decisions (7). Such disagreements can arise both from the conflicting incentives and objectives of the three stakeholders (patients, physicians, and payers) and from misperceptions that each may have about the preferences of the other stakeholders.

One area where conflicts can arise between patients, physicians, and payers is with respect to diagnostic testing. A diagnostic test can provide both tangible and intangible value. Tangible value is created when the diagnostic test results are used to alter the treatment path in a way that improves expected clinical outcomes. Intangible value is created when, for example, patients or their caregivers place some value on establishing a definitive diagnosis independent of any effect of the information on treatments and subsequent clinical outcomes ("knowing for the sake of knowing") (4). Several previous studies have considered patients' desires or willingness to pay (WTP) for diagnostic information in situations where the information may not change clinical outcomes (5;6;31;35).

It is difficult to assess how much value others assign to intangibles, increasing the likelihood of misperceptions (e.g., physicians and insurers may underestimate the importance of diagnostic certainty to patients). Further, practice guidelines often fail to account for this value. In part, this omission arises because the cost-effectiveness analyses (CEA) that underlie many guidelines ignore intangibles. While an ideal CEA would account for all attributes of the alternatives being compared, the difficulty of quantifying and valuing intangibles has led to their

routine exclusion from analysis (14). When the intangibles differ across the alternatives being compared in a CEA (e.g., empiric therapy systematically generates less diagnostic information than a strategy that requires testing before treatment), both the absolute cost-effectiveness ratios and the rank ordering of alternatives can be incorrect.

Although the patient's perspective is perceived as the primary point of view from which the value of diagnostic certainty should be assessed, physicians and, increasingly, payers also influence treatment decisions. Physicians and payers may value diagnostic information from their own perspectives, and their desires may or may not be consistent with patients' desires. In the case of an inconsistency that leads to a restriction of a diagnostic test desired by patients, a tension may arise that ultimately leads to patient dissatisfaction—even if clinical outcomes are not compromised. Previous studies of the value of diagnostic information have not explored the perspectives of physicians and payers (5;6;31;35). Patients, physicians, and payers are likely to have unique perspectives and reasons for valuing information provided by a diagnostic test even if it does not affect the management of the patient's condition or change the clinical outcome. For patients, this value may arise from psychological benefits of reduced uncertainty about the underlying cause of their symptoms or a more accurate prognosis. For physicians, this value may represent comfort from having objectively documented their patient's underlying diagnosis rather than inferring it from symptoms, or from a perceived reduction of litigation risk. For payers, the benefits may include improved patient or physician perception of quality of care or greater acceptance of practice guidelines by patients or physicians.

We examine WTP for diagnostic certainty in two upper gastrointestinal diseases that are frequently encountered both by primary care physicians and specialists: peptic ulcer disease (PUD) and gastroesophageal reflux disease (GERD). These diseases have also been a target of the movement toward guideline-based medicine, with a number of guidelines and cost-effectiveness analyses recommending that physicians reserve definitive diagnostic testing for those patients who have failed empiric therapy (1;16;17;36;39). However, these studies reached their conclusions without quantifying the intangible value of diagnostic testing.

Even though 20% or fewer of individuals with symptoms suggestive of PUD have an active ulcer, pharmacological treatment without a diagnostic test to confirm the presence of an ulcer has been widely accepted (1). Previously, we demonstrated that a strategy reserving diagnostic testing for patients who failed empiric therapy was cost-effective relative to requiring a definitive ulcer test before treatment because it reduced costs without compromising clinical outcomes (16). Although we recognized that less diagnostic information was generated by strategies that reserved the "gold standard" test (endoscopy) for patients who failed empiric treatment, we did not quantify the value of diagnostic certainty. By recommending treatment without a confirmed ulcer diagnosis, we presumed that the value of this information to the average patient (net of any inconvenience or discomfort associated with the test) was less than the demonstrated savings of at least $423 per patient generated by empiric therapy.

Gastroesophageal reflux disease (GERD) is a common condition responsible for significant morbidity and healthcare expenditures (29). GERD symptoms, frequently heartburn, are usually caused by the backflow of acidic stomach contents into the esophagus. The diagnosis of GERD can be based on symptoms or made definitively by performing endoscopy to visualize injury to the esophagus. While

endoscopic diagnosis may influence the choice of drug therapy, the role of objective testing in GERD is largely to rule out rare but serious conditions. As with PUD, most existing GERD clinical guidelines recommend initial empiric therapy for patients with an uncomplicated clinical presentation, and limit diagnostic testing to those whose symptoms persist after therapy (17).

We explore two reasons why disagreements might arise between MCOs, physicians, and patients with respect to diagnostic tests that have an intangible value. First, the three groups may place different values on the diagnostic test from their own perspectives. If such differences exist and if decision makers are able to act primarily on the basis of their own valuations (e.g., if patients are unlikely to switch insurers or physicians in response to not receiving a desired test), it becomes more likely that the expectations of some groups will not be met. Second, MCOs and physicians may misperceive the value of information to patients. Misunderstandings and unmet expectations can arise if patients misperceive physicians' or payers' desire for diagnostic certainty. To our knowledge, no previous study has compared the value of diagnostic certainty among different stakeholders.

METHODS

Setting

To assess WTP for diagnostic certainty from the patient, physician, and payer perspectives, we conducted a WTP (contingent valuation) survey of patients, physicians, and managed care executives. In addition, each group was asked to estimate the WTP of the other groups. One participating MCO had 175,000 enrollees in a primarily rural area of the northeastern United States. It was a nonprofit, mixed group and staff model health maintenance organization that owned its facilities and paid its physicians by salary or capitation. In the second participating MCO, only physicians and executives were surveyed. This MCO was a nonprofit, academically affiliated health plan with 100,000 enrollees in an urban setting in the northeastern United States.

Sample Selection and Survey Administration

At the first MCO, patients being treated for a suspected or confirmed diagnosis of PUD or GERD were identified in the MCO's information system. Patients with a presumed or confirmed diagnosis of PUD or GERD were chosen in order to ensure that they had sufficient knowledge of the disease to place a reasonable estimate on the value of diagnostic certainty. A random sample of these patients was drawn for the study. Their next scheduled clinic visit (not necessarily related to PUD or GERD) was identified (nearly all patients had a scheduled visit within 3–4 months of when the sample was drawn), and a survey instrument was placed in their file at the clinic. When the patients arrived for their appointment, they were asked to participate in the study and fill out the survey instrument for the disease for which they had a confirmed or suspected diagnosis (PUD or GERD) while in the waiting room. At the same time, the physician who the patient was scheduled to see was also given the physician survey. Physicians were only eligible to fill out the survey once for each disease (that is, if more than one sample patient per disease had their next scheduled appointment with the same physician, the physician only received the instrument at the time of the first sample patient's visit). All executives at the level of chief executive officer, chief operating officer, chief financial officer, vice

president, medical director, and associate medical director received survey instruments for both diseases. At the second MCO, which provided data only for physicians and executives, the executive sample was defined in the same manner (all individuals with the listed job titles). The physician sample consisted of all physicians at one of the MCO's largest multispecialty clinics. The survey instruments were administered at clinical encounters (patients and physicians in the first MCO) or at several staff meetings (physicians in the second MCO and executives in both MCOs) to ensure high response rates. The responsible personnel at both sites reported that there were no refusals to participate. However, because the sample selection and survey administration processes were performed by MCO personnel (following our explicit instructions), we cannot independently confirm that every targeted patient, physician, and executive was approached (e.g., an executive may not have been present at any of the meetings at which the survey was administered or clinic staff may have neglected to present the questionnaire to a patient or physician). Because both MCOs provided the anticipated number of respondents, we have no reason to believe that approach rates and response rates were not very high.

Survey Instrument

We designed questionnaires that presented a hypothetical vignette to try as best as possible to isolate the value of diagnostic certainty from the test's clinical utility or perceived discomfort. The vignette stated that symptoms indicated a 20% probability of disease (PUD or GERD, depending on the disease for which the patient had a confirmed or suspected diagnosis), that the hypothetical test (which was not named or described) entailed no pain or risk, and that the duration and severity of symptoms depended on the presence of PUD or GERD, but establishing the diagnosis would not alter the course of symptoms. Patients were told that their responses would not affect their actual therapy and would not be communicated to their physician or insurer. The questionnaires made clear to respondents that the utility of the test results was solely to establish a diagnosis, *not* to guide treatment decisions or influence the ultimate resolution of symptoms. The patient vignette for PUD is in Appendix 1.

The questionnaires elicited participants' maximum WTP for the diagnostic information from their own perspectives by choosing from eight categories: $0 (would pay nothing), $1–9, $10–49, $50–99, $100–249, $250–499, $500–999, or greater than $1,000. All three types of respondents received the same clinical vignette, but the WTP question was tailored to the type of respondent. Patients were told that their response represented an out-of-pocket payment for the test. Physicians were told that their response represented an expense for which they were at risk from a capitation payment received for caring for the patient (that is, if physicians ordered the test for their patients, they would effectively have to pay for it). MCO executives were told that their response represented the maximum amount they would recommend their health plan to pay to include the test in a clinical practice guideline. In order to isolate the value of diagnostic certainty from the perspective of physicians and payers, physicians and executives were told to assume the patient was indifferent between receiving the test or not.

Each group was also asked to estimate WTP by the other two groups. Managed care executives and physicians were asked to estimate how much the average patient would be willing to pay out of pocket for the diagnostic information. Managed care executives and patients were asked to estimate the maximum amount they thought

physicians would pay for the test if the test cost had to be covered out of a capitation payment (to patients, a capitation payment was described as "the most your doctor would pay for this test if your doctor had to pay for it from the money received from your insurance company for your medical care"). However, we followed a different approach with respect to patients' and physicians' assessments of payers' WTP for the information. Rather than asking a positive question about how much the payer *would* be willing to pay for the information, we asked a normative question about how much an insurer *should* be willing to pay to cover the test. This approach was chosen because coverage expectations, and hence patient or physician dissatisfaction, appear more likely to be related to beliefs about what an insurer should pay rather than beliefs about what an insurer would actually pay. For example, suppose insurers would be willing to pay $10 for the test. Further suppose that patients correctly predict that insurers would pay $10 but nonetheless believe that insurers should pay $100 for the test. Patients' beliefs that insurers should pay significantly more than they actually would pay for the test create a potential conflict that could not have been identified had patients been asked about how much the insurer would pay.

Analysis

Pairwise comparisons were made between the proportion of respondents of different types (e.g., patient vs. physician) expressing a positive WTP. The statistical significance of differences between groups was assessed using a two-tailed t test for proportions.

RESULTS

Descriptive statistics for the combined PUD and GERD samples are reported in Table 1. The patients varied widely across most dimensions (e.g., gender, income, education, and marital status), but were racially homogeneous since they were drawn from a racially homogeneous area. Several physician specialties, particularly general internal medicine and gastroenterology, were well represented. The most common titles of the surveyed executives were at the vice president level and the next most frequent titles were medical director or associate medical director. Slightly more than half of the executives were nonphysicians.

Patients' WTP for the diagnostic information for PUD and GERD, as assessed by patients, physicians, and managed care executives, is shown in Table 2. Likewise, physicians' WTP is reported in Table 3, and managed care executives' WTP is reported in Table 4.

We first compare WTP from each group's own perspective. For both diseases, patients were most likely and executives were least likely to value knowing for knowing's sake. For PUD, 84% of patients, 61% of physicians, and 43% of executives expressed a positive WTP. Patients and executives ($p = .003$), and patients and physicians ($p = .060$) differed at the 10% significance level or better in the probability of expressing a positive WTP. The difference between physicians and executives was not significant ($p = .220$). Median WTP was low for all three groups ($1–9 for patients and physicians; $0 for payers). For GERD, 87% of patients expressed a positive WTP and assigned the largest value to test information (median, $10–49) while only 29% of executives valued the test information (median, $0). Physicians were between patients and executives in their probability of expressing a positive WTP (52%) and their median WTP ($1–9). Patients and executives ($p <$

Table 1. Descriptive Statistics

Patients (n = 53)	
Age, yrs	53.8
Female, %	47.2
White, %	100.0
Education	
High school or less, %	49.1
Some college, %	20.8
College graduate, %	30.1
Marital status	
Married, %	71.7
Single, %	13.2
Widowed or divorced, %	15.1
Family income	
<$20,000, %	35.4
$20,000–$40,000, %	33.3
>$40,000, %	30.9
Physicians (n = 34)	
Age, yrs	43.0
Female, %	32.6
Specialty	
Gastroenterology, %	32.3
General internal medicine, %	45.2
Other, %	22.6
Executives (n = 26)	
Age, yrs	49.6
Female, %	4.0
Nonphysician, %	52.2

.001), patients and physicians ($p = .005$), and physicians and executives ($p = .094$) differed in the probability of expressing a positive WTP at the 10% level of significance or better.

Physicians did not significantly underestimate patients' WTP for either disease. Conversely, the value of GERD information to patients was underestimated by MCO executives. MCO executives believed that patient WTP was lower ($1–9) than the median value stated by patients ($10–49) and patients were significantly more likely to express a positive WTP than executives had anticipated ($p = .029$) (Table 2). For PUD, the pattern was similar. Thirty-five percent of executives believed that patients did not value diagnostic certainty, whereas only 16% of patients did not assign a positive value to the information, but this difference was not statistically significant.

Table 2. Patients' Willingness to Pay for PUD and GERD Test Information

	$0	$1–9	$10–49	$50–99	$100–249	$250–499
PUD test information as assessed by:						
Patients, n (%)	4 (16)	10 (40)	8 (32)	1 (4)	2 (8)	0 (0)
Physicians, n (%)	5 (18)	9 (32)	9 (32)	1 (4)	3 (11)	1 (4)
Executives, n (%)	8 (35)	1 (4)	12 (52)	1 (4)	1 (4)	0 (0)
GERD test information as assessed by:						
Patients, n (%)	3 (13)	7 (29)	9 (38)	5 (21)	0 (0)	0 (0)
Physicians, n (%)	7 (23)	7 (23)	7 (23)	8 (26)	2 (6)	0 (0)
Executives, n (%)	10 (40)	3 (12)	8 (32)	2 (8)	2 (8)	0 (0)

Table 3. Physicians' Willingness to Pay for PUD and GERD Test Information

	$0	$1–9	$10–49	$50–99	$100–249	$250–499	$500–999
PUD test information as assessed by:							
Patients, n (%)	7 (29)	3 (13)	7 (29)	3 (13)	2 (8)	1 (4)	1 (4)
Physicians, n (%)	11 (39)	9 (32)	1 (4)	4 (14)	3 (11)	0 (0)	0 (0)
Executives, n (%)	14 (61)	3 (13)	4 (17)	0 (0)	2 (9)	0 (0)	0 (0)
GERD test information as assessed by:							
Patients, n (%)	3 (15)	4 (20)	4 (20)	5 (25)	3 (15)	0 (0)	1 (5)
Physicians, n (%)	15 (48)	3 (10)	6 (19)	4 (13)	3 (10)	0 (0)	0 (0)
Executives, n (%)	17 (71)	2 (8)	5 (21)	0 (0)	0 (0)	0 (0)	0 (0)

Table 4. Payers' Willingness to Pay for PUD and GERD Test Information

	$0	$1–9	$10–49	$50–99	$100–249	$250–499	$500–999	⩾$1,000
PUD test information as assessed by:								
Patients, n (%)	6 (25)	1 (4)	7 (29)	2 (8)	4 (17)	2 (8)	1 (4)	1 (4)
Physicians, n (%)	9 (32)	7 (25)	3 (11)	3 (11)	5 (18)	1 (4)	0 (0)	0 (0)
Executives, n (%)	13 (57)	1 (4)	6 (26)	1 (4)	2 (9)	0 (0)	0 (0)	0 (0)
GERD test information as assessed by:								
Patients, n (%)	1 (5)	2 (10)	6 (29)	7 (33)	3 (14)	0 (0)	1 (5)	1 (5)
Physicians, n (%)	12 (39)	4 (13)	2 (6)	4 (13)	9 (29)	0 (0)	0 (0)	0 (0)
Executives, n (%)	17 (71)	3 (13)	3 (13)	1 (3)	0 (0)	0 (0)	0 (0)	0 (0)

For both diseases, physicians' median WTP ($1–9) was overestimated by patients (median patients thought their physician would pay $10–49) and underestimated by executives (median executive thought the average physician would not pay any positive amount out of their capitation) (Table 3). In terms of expressing a positive WTP, these differences were statistically significant at the 10% level or better only for patients' assessment of physicians' WTP for the GERD test (p = .015) and executives' assessment of physicians' WTP for the GERD test (p = .094).

In Table 4, executives' recommendations as to the maximum amount their organization should pay to include the test in a practice guideline are contrasted with patients' and physicians' opinions on how much an insurer should be willing to pay to cover the test. Substantial differences emerged among the groups. Most patients (75% for PUD and 95% for GERD) and physicians (68% for PUD and 61% for GERD) believed an insurer should pay something to cover the test. Further, several patients and one physician stated that an insurer should pay more than $250 for such coverage, an amount in excess of the highest WTP of any individual patient or physician. Conversely, the majority of executives (57% for PUD and 71% for GERD) stated that, from their own perspectives, they placed no value on including these tests in a guideline for treating patients with suspected PUD or GERD. For PUD, the proportion of executives recommending their organization pay some positive amount to include the test in a guideline was significantly lower than the proportions of patients (p = .028) and physicians (p = .080) who thought such a test should be covered at some positive price. For GERD, the proportion of executives recommending their organization pay to include the test was also significantly lower than the proportions of patients ($p < .001$) and physicians (p = .018) who thought such a test should be covered. Physicians and patients also differed with respect to WTP for the GERD test (p = .006).

DISCUSSION

Much of the research on the intangible value of diagnostic certainty has focused on issues such as testing for a genetic predisposition to a serious disease for which effective preventive measures or treatments are unavailable (e.g., Huntington's disease or multiple sclerosis) (31;35). However, the value of diagnostic information, aside from its direct impact on clinical management, has not been well assessed for routine diagnostic tests for conditions that are often treated empirically, such as PUD, GERD, or urinary tract infections (normal pregnancies are an exception [6], and experimental evidence indicates that diagnostic certainty is valued even when there is no effect on clinical decisions and outcomes of care [5]). Thus, the value of knowing for the sake of knowing in many common conditions is poorly understood.

Our findings suggest that for the two diseases studied, most patients and physicians place some value, albeit low, on diagnostic certainty. Conversely, most managed care administrators did not value diagnostic certainty. This indicates the potential for conflict over the content of practice guidelines if each stakeholder does not recognize and account for the perspectives of the other actors when making decisions. Further, because empirical treatment strategies generate less diagnostic information than strategies that include definitive diagnostic studies prior to initiating therapy, CEAs will be biased in favor of empirical therapy if this intangible value of information is ignored. However, in the case of PUD, the value of diagnostic certainty found in this survey was too low to change the cost-effectiveness rankings of the strategies considered by Fendrick et al. (16;24).

The role of payers and physicians as agents for patients depends on the timing of their interactions. At the time of insurance purchase, the insurer interacts with a potential enrollee who faces some probability of illness but does not yet know if he/she will actually become ill (*ex ante*). In exchange for a lower insurance premium, an expected utility maximizing enrollee would forgo coverage for tests whose value is less than their full cost. However, after the illness has occurred (*ex post*), the insured patient would face little or no out-of-pocket cost for the test. Thus, at the time of the actual clinical encounter, the patient would want the test even if its perceived value were low. A physician acting as an agent for the *ex post* patient would provide the test as long as its value exceeded the patient's out-of-pocket cost. In this scenario, even if managed care executives and physicians attempt to act in the best interests of enrollees/patients, an inherent conflict arises because MCOs respond to the preferences of the *ex ante* enrollee while physicians respond to the preferences of the *ex post,* insured patient.

MCOs or physicians would be most likely to violate their agency relationships with patients when the WTP for diagnostic certainty varies across groups. For example, if capitated physicians place less value than patients on diagnostic certainty, the potential for conflict exists because physicians may not order a test in some situations in which the patient would like to have the test performed. Likewise, if managed care executives place less value on the information than do physicians or patients, guidelines recommending empirical treatment may not win acceptance due to physician dissatisfaction with the MCO's "interference" in clinical decision making or patient dissatisfaction with access to care.

For both diseases, managed care executives were less likely than either patients or physicians to value diagnostic information. Although this suggests that there may be difficulty in gaining patient and physician acceptance of guidelines that exclude the diagnostic test, the absolute values of the differences were small for these diseases. Because patients and physicians expressed similar willingness to pay for the information, conflicts are not likely to arise in the doctor/patient relationship because of asymmetric valuation of this intangible.

In order to act on the patient's behalf, MCOs and physicians must be well informed about patients' desires. For PUD, physicians and executives did not significantly misestimate patients' WTP. However, executives underestimated patients' WTP for diagnostic information in the case of GERD. Thus, an MCO developing a GERD guideline could be biased against including a diagnostic test even if its value to patients was sufficient to warrant inclusion.

Misperceptions also arose with respect to physicians' WTP, which was overestimated by patients and underestimated by executives for both diseases. If patients are aware of the existence of the test, these misperceptions could lead patients to expect to receive it and executives to underestimate physician resistance to adopting empiric treatment strategies.

Finally, patients' and physicians' responses to the question about how much an insurer should be willing to pay to include the test in a practice guideline indicated the potential for significant unmet expectations. Most patients and physicians stated that the insurer should pay some positive price to cover the test, while most executives indicated that, from their own perspective, the test was not valuable. Interestingly, the amount patients thought the insurer should pay often exceeded the value of the information to patients. This implies that even if physicians and MCOs are attempting to act in patients' best interests and are well informed about how much value patients assign to diagnostic certainty, it may still be difficult to meet patients'

expectations. This divergence between the value patients place on the test and what they believe insurers should pay could reflect patients' general attitudes about managed care or a "deep pockets" effect under which they believe the insurer can afford to spend a substantial amount to provide services even though they provide little value.

This study provides the first data on the value of diagnostic information from different perspectives. This study is limited by its small samples and the fact that all patients were drawn from only one MCO and physicians and executives were drawn from only two MCOs. It would be useful to have data on patients, physicians, and executives from a wider variety of MCOs to determine whether responses differ across settings. Larger and more geographically and racially diverse populations would facilitate multivariate modeling of the predictors of patient, physician, and executive WTP for diagnostic certainty, allowing additional research questions to be addressed.

Additional research is also required to generalize these findings to other diseases. The value of diagnostic certainty for PUD and GERD was quite low, but willingness to pay for diagnostic information for more life threatening diseases (e.g., coronary diseases) is likely to be greater. Thus, the potential for conflicts in the event of WTP disparities or misperceptions among stakeholders is also likely to be greater for more life-threatening illnesses.

Exploring other methodologies of eliciting WTP could also be useful. Some studies have concluded that open-ended WTP questionnaires suffer from biases or poor response rates (3;11;27). These difficulties might be avoided by asking participants whether they would be willing to pay $X for the good or service in question, varying $X randomly across participants, and then applying logistic regression techniques to estimate WTP from the observed probabilities of accepting the offer at various prices (3;9;21). The primary drawback of this method is that it requires large samples. Therefore, this method was infeasible in the context of our study. Instead, we avoided using an open-ended WTP question by asking for categorical responses. Pre-testing indicated that some respondents had difficulty stating a precise dollar WTP but could easily commit to a range within which their WTP fell. No mention was made of test cost, since other studies have found that subjects tend to anchor their WTP responses to the cost when such information is given (3;11), but a study using the varying price offers would be necessary to definitively rule out any anchoring of responses to our choice of categories.

Another limitation of this study is that we cannot be certain that the survey instruments successfully isolated the value of diagnostic certainty from perceptions about the clinical utility of the test results and discomfort associated with the test. If the hypothetical vignettes were taken literally, then subjects would have assumed that the test had neither clinical utility nor unpleasant side effects. If respondents nonetheless assumed that the test would have some clinical utility, their reported valuations would be biased upward as an estimate of the value of diagnostic certainty. Conversely, a downward bias on the responses could arise if respondents disregarded our instructions to assume that there was no pain or risk associated with the unnamed test. This possibility cannot be ruled out, because many respondents undoubtedly knew that the definitive test is endoscopy, which does involve some discomfort. Thus, perceptions about endoscopy may have entered patients' thinking about the value of the unnamed test.

A strength of our study sample is that all patients with a PUD or GERD diagnosis code were eligible, regardless of whether the diagnosis was presumed or

confirmed. Had the sample been restricted to *confirmed* diagnoses, there would likely have been an upward bias on the estimated WTP for diagnostic information. Because all patients with a confirmed diagnosis necessarily consented to a definitive test, they would be expected to have a greater than average desire for diagnostic information. However, to avoid biases due to recall or hindsight, it would have been ideal to survey patients with undiagnosed symptoms suggestive of either PUD or GERD (that is, survey patients at the time of the diagnostic testing decision). However, it is impossible to identify such patients prospectively using an MCO's information system. The alternative of identifying a sufficient sample of such patients at the point of the clinical encounter was also infeasible as it would have required substantial staff time to screen all patients arriving at the clinic to identify those eligible for the study.

IMPLICATIONS FOR POLICY AND PRACTICE

Despite these limitations, we believe that WTP surveys can provide useful information to policy makers, health plans, and healthcare providers. The potential sources of conflict identified in our survey can have significant managerial implications in areas such as adoption and implementation of practice guidelines (26;40). Likewise, knowledge about these sources of disagreement can contribute to the understanding and resolution of policy issues such as public dissatisfaction with managed care (30;37), legislative mandates to cover specific services (18;22), congressional proposals for broad regulation of managed care (2;34), and litigation over coverage denials (8;28).

This information might aid in decisions about whether clinical practice guidelines should recommend empiric therapy and might help decision makers anticipate where and why disagreements about diagnostic testing are likely to arise. For example, potential strains on the doctor/patient relationship arising from diagnostic testing decisions could be identified. Assessing the desire for diagnostic certainty from multiple perspectives could prevent these issues from taking decision makers by surprise and could facilitate addressing them explicitly. Likewise, conflicts that arise simply because of misperceptions about the value of diagnostic certainty to other stakeholders could be avoided altogether if such information were available.

REFERENCES

1. American College of Physicians, Health and Public Policy Committee. Endoscopy in the evaluation of dyspepsia. *Annals of Internal Medicine* 1985, 102, 266–69.
2. Annas, G. J. A national bill of patients' rights. *New England Journal of Medicine*, 1998, 338, 695–99.
3. Arrow, K., Solow, R. , Portney, P. R. , et al. Report of the NOAA Panel on Contingent Valuation. *Federal Register*, 1993, 58, 4601–14.
4. Asch, A. D., Patton, J. P., & Hershey, J. C. Knowing for the sake of knowing: The value of prognostic information. *Journal of Medical Decision Making*, 1990, 10, 47–57.
5. Baron, J., Beattie, J., & Hershey, J. C. Heuristics and biases in diagnostic reasoning, II: Congruence, information, and certainty. *Organizational Behavior and Human Decision Processes*, 1988, 42, 88–110.
6. Berwick, D. M., & Weinstein, M. C. What do patients value? Willingness to pay for ultrasound in normal pregnancy. *Medical Care*, 1985, 23, 881–93.
7. Blomqvist, A. The doctor as double agent: Information asymmetry, health insurance, and medical care. *Journal of Health Economics*, 1991, 10, 411–32.
8. Brushwood, D. B. Challenging denial of coverage for innovative therapy. *American Journal of Health-System Pharmacy*, 1997, 54, 572–74.

9. Cameron, T. A. A new paradigm for valuing non-market goods using referendum data: Maximum likelihood estimation by censored logistic regression. *Journal of Environmental Economics and Management*, 1988, 13, 255–68.
10. Cromwell, J., & Mitchell, J. B. Physician-induced demand for surgery. *Journal of Health Economics*, 1986, 5, 293–313.
11. Donaldson, C., Shackley, P., & Abdalla, M. Using willingness to pay to value close substitutes: Carrier screening for cystic fibrosis revisited. *Health Economics*, 1997, 6, 145–59.
12. Dranove, D., & Wehner, P. Physician-induced demand for childbirths. *Journal of Health Economics*, 1994, 13, 61–73.
13. Dranove, D., & White, W. D. Agency and the organization of health care delivery. *Inquiry*, 1987, 24, 405–15.
14. Drummond, M. F., O'Brien, B., Stoddart, G. L., & Torrance G. W. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press, 1997.
15. Ellis, R. P., & McGuire, T. G. Supply-side and demand-side cost sharing in health care. *Journal of Economic Perspectives*, 1993, 7, 135–51.
16. Fendrick, A. M., Chernew, M. E., Hirth, R. A., & Bloom, B. S. Alternative management strategies for patients with suspected peptic ulcer disease. *Annals of Internal Medicine*, 1995, 123, 260–68.
17. Fennerty, M. B., Castell, D. O., Fendrick, A. M., et al. The diagnosis and treatment of GERD in a managed care environment. *Archives of Internal Medicine*, 1996, 156, 477–84.
18. Frank, R. G., Koyanagi, C., & McGuire, T. G. The politics and economics of mental health 'parity' laws. *Health Affairs*, 1997, 16, 108–19.
19. Fuchs, V. R. The supply of surgeons and the demand for surgical operations. *Journal of Human Resources,* 1978, 13, S6–56.
20. Gold, M., Hurley, R., Lake, T., Ensor, T., & Berenson, R. Arrangements between managed care plans and physicians: Results from a 1994 survey of managed care plans. *Physician Payment Review Commission*, 1995.
21. Hanemann, M. W. Welfare evaluations in contingent valuation experiments with discrete responses. *American Journal of Agricultural Economics*, 1984, 66, 332–41.
22. Hellinger, F. The expanding scope of state legislation. *JAMA*, 1996, 276, 1065–70.
23. Hillman, A. L., Pauly, M. V., & Kerstein, J. J. How do financial incentives affect physicians' clinical decisions and the financial performance of health maintenance organizations. *New England Journal of Medicine*, 1989, 321, 86–92.
24. Hirth, R. A., Bloom, B. S., Chernew, M. E., & Fendrick, A. M. Willingness to pay for diagnostic certainty: A comparison among patients, physicians, and managed care executives. *Journal of General Internal Medicine*, 1999, 14, 193–95.
25. *The InterStudy competitive edge 5.2.* Minneapolis: InterStudy, 1995.
26. James, P. A., Cowan, T. M., Graham, R. P., & Majeroni, B. A. Family physicians' attitudes about and use of clinical practice guidelines. *Journal of Family Practice*, 1997, 45, 341–47.
27. Johannesson, M., Jonsson, B., & Borgquist, L. Willingness to pay for antihypertensive therapy: Results of a Swedish pilot study. *Journal of Health Economics*, 1991, 10, 461–74.
28. Johnson, A. L. Temporomandibular joint litigation: Resolving issues of medical necessity and contract ambiguity. *Seminars in Orthodontics*, 1997, 128–32.
29. Kahrilas, P. J. Gastroesophageal reflux disease. *JAMA*, 1996, 276, 983–88.
30. Kerr, E. A., Hays, R. D., Lee, M. L., & Siu, A. L. Does dissatisfaction with access to specialists affect the desire to leave a managed care plan? *Medical Care Research and Review*, 1998, 55, 59–77.
31. Kessler, S., Field, T., Worth, L., & Mosbarger, H. Attitudes of persons at risk for Huntington disease toward predictive screening. *American Journal of Medical Genetics*, 1987, 26, 259–70.
32. McLean, R. A. Agency costs and complex contracts in health care organizations. *Health Care Manage Review*, 1989, 14, 65–71.

33. Mooney, G., & Ryan, M. Agency in health care: Getting beyond first principles. *Journal of Health Economics*, 1993, 12, 125–35.
34. Moran, D. W. Federal regulation of managed care: An impulse in search of a theory? *Health Affairs*, 1997, 16, 7–21.
35. Mushlin, A. I., Mooney, C., Holloway, R. G., et al. The cost-effectiveness of magnetic resonance imaging for patients with equivocal neurological symptoms. *International Journal of Technology Assessment in Health Care*, 1997, 13, 21–34.
36. Ofman, J. J., Etchason, J., Fullerton, S., Kahn, K. L., & Soll, A. H. Management strategies for *Helicobacter pylori*-seropositive patients with dyspepsia. *Annals of Internal Medicine*, 1997, 126, 280–91.
37. Patrick, D. L., Martin, D. P., Madden, C. W., et al. Dissatisfaction and disenrollment in a subsidized managed care program. *Medical Care Research and Review*, 1997, 54, 61–79.
38. Sappington, D. E. Incentives in principal-agent relationships. *Journal of Economic Perspectives*, 1991, 5, 45–66.
39. Silverstein, M. D., Petterson, T., & Talley, N. J. Initial endoscopy or empirical therapy with or without testing for *Helicobacter pylori* for dyspepsia: A decision analysis. *Gastroenterology*, 1996, 110, 72–83.
40. Tunis, S. R., Hayward, R. S., Wilson, M. C., et al. Internists' attitudes about clinical practice guidelines. *Annals of Internal Medicine*, 1994, 120, 956–63.
41. Weisbrod, B. A. The health care quadrilemma: An essay on technological change, insurance, quality of care, and cost containment. *Journal of Economic Literature*, 1991, 29, 523–52.

APPENDIX 1

Clinical Vignette for PUD

Hypothetical Clinical Scenario. Based on your symptoms, there is a 20% chance that you have an ulcer (peptic ulcer disease) and an 80% chance that your symptoms are due to some other cause. How long the symptoms last and how severe they are depend on whether or not you have an ulcer, but knowing whether or not you have an ulcer will have no effect on how long your symptoms last or how severe they are. A test is available to determine whether or not you have an ulcer. Therefore, *the test is ONLY used to tell you if you actually have an ulcer*. There is no pain or risk to you with getting the test.

We are trying to discover how much patients like you value the information from this test (whether you have an ulcer or not). *REMEMBER:* The test results will not change your symptoms.

Your responses to this survey are *completely confidential* and will not be shown to anyone outside the research team. Your physician and insurer will NOT be informed of your answers, and your answers will NOT influence your actual treatment.