The physician–patient relationship: The impact of patient-obtained medical information

Bin Xie, David M. Dilts and Mikhail Shor

Introduction

Recent studies indicate that patients obtain medical information from the Internet and other sources, and that this information influences their clinical medical decisions [1–3]. Physicians also respond to such patient-obtained medical information (POMI) by ordering treatments, tests, or procedures that they would not otherwise order [4,5]. POMI impacts the physician–patient relationship, as some health economists have predicted [6–10] and as studies in medical practice and health policy have suggested or confirmed [11–15]. These theories also suggest a link between the level of patient information and subsequent utilization of medical services.

A physician is unlikely to observe perfectly an individual patient’s POMI due to time constraints [16–18] or the patient’s collecting of information after receiving a physician’s treatment recommendation but before accepting the treatment [19–21]. Instead, the physician may rely heavily on the patient’s demographic, socioeconomic, and health characteristics as imperfect signals of the patient’s level of information [22–24]. In this study, we investigate the physician–patient interaction when the physician employs the POMI of a patient population as a surrogate for the POMI of individual patients. For example, although a physician may know that most college-educated White females in their early forties are likely to use online sources for medical information, there is no certainty that a specific patient with these characteristics does so [22]. It is this residual uncertainty that differentiates the present study from previous work where the physician is

Summary

We investigate the impact of patient-obtained medical information (POMI) on the physician–patient relationship when patients, as a group, are heterogeneously informed and a physician’s interests do not coincide with those of her patients. Introducing additional well-informed patients to the population discontinuously affects the physician’s strategy, having no effect unless a sufficient quantity is added. When few patients are well informed, increasing the precision of their information level has no effect on the physician’s strategy. Alternately, when a sufficient number of well-informed patients exists, increasing the precision of their information allows all patients to free-ride by receiving more appropriate treatment recommendations.

Counterintuitively, we also identify circumstances under which increasing the general level of information may potentially harm patients. Copyright © 2006 John Wiley & Sons, Ltd.

Keywords supplier-induced demand; patient-obtained medical information; physician–patient relationship; patient population
assumed to know the level of information of each individual patient.

The use of patient population characteristics as surrogates for the characteristics of a particular patient has been widely reported in the medical literature, as patients’ demographic and socioeconomic characteristics are often used in medical registries and practice guidance to steer physician practices [25–28]. In such instances, the physician is actually treating the patient as a representative member of a patient population with similar demographic, socioeconomic, or health characteristics. Using a game-theoretic approach, we find that a patient’s level of care depends on the overall medical literacy of his peer group. In populations with coexisting high and low levels of POMI, one information level will be “dominant” in the sense that a physician will treat all members of the population in the same way, determined only by the dominant level.

Which information level becomes dominant is determined by several factors, including the quality of information at each level, the proportion of patients who are well informed, and the physician’s and patients’ payoffs. In populations where few patients are well informed, the low information level is dominant. In these cases, introducing additional well-informed patients discontinuously affects the physician’s strategy, having no effect unless a sufficient quantity of patients is added. Additionally, any increase in the quality of information of the few well-informed patients has no effect on the physician’s strategy. Alternately, when the high information level is dominant (e.g. when a sufficient number of patients is well informed), increasing the quality of their information allows all patients to “free-ride” by receiving more appropriate treatment recommendations. Counterintuitively, we also identify circumstances under which increasing the general level of information may potentially harm patients.

**Literature**

A corpus of papers suggests that physicians may exploit their agency relationship with the patient when subject to incentives that diverge from the sole concern of maximizing patient well-being. For example, Evans [6] demonstrates that the physician induces the patient to buy more medical care than the patient would demand in the absence of information asymmetries. Farley [7] finds that even when the physician’s incentives directly incorporate the patient’s utility function, demand inducement beyond optimal levels still results.

When the patient is conferred the ability to act strategically, a patient–physician encounter becomes an interaction between asymmetrically informed decision makers, each maximizing his or her respective utility [8, 9]. Dranove [8] endows the patient with some (albeit imperfect) information, which is used to evaluate and potentially reject the demand set by the physician. Equilibrium demand induced by the physician is still higher than what the patient deems optimal, but the more accurate the patient’s initial information, the closer equilibrium demand is to the patient’s optimal level. Following on the ideas of strategic information transmission from Crawford and Sobel [29] and Pitchik and Schotter [30], De Jaegher and Jegers [9] model the physician–patient interaction as a game of strategic information transmission. While the patient is uninformed about his optimal treatment, his ability to make the final decision as to which treatment option to choose compels the physician to reveal some information in equilibrium.

Many of the above models represent the supplier-induced demand (SID) literature. In its earliest form, the SID hypothesis suggested that negative income shocks to physicians – be they the result of increased competition, a new pay structure, or a change in patient tastes – will result in the provision of excessive care. In its most extreme form, the SID hypothesis postulates that physicians will always induce sufficient demand to match utilization to available capacity. Techniques for validating the hypothesis have been critiqued on methodological grounds [31–33], and the underlying assumption that physicians induce demand contrary to patients’ interests has been widely debated [34]. However, we take from the SID literature the observation that physicians may have influences apart from a patient’s well-being, such as economic or ethical considerations or the furtherance of medical research.

We adopt the approach of De Jaegher and Jegers [9] in assuming that the physician knows with certainty a patient’s true medical state but recommends a treatment option strategically. In modeling patients’ medical information, we use an approach similar to Dranove [8], representing information quality by the accuracy with which a patient understands his medical state.
increases the likelihood that a patient is correct in his beliefs about the optimal treatment. De Jaeger [10] also incorporates patient information, but as a binary event—either information is present or it is not—into the single patient—single physician model of [9], but does not allow for the possibility of differing levels of information. In our model, we focus on how information distribution amongst the population impacts the physician—patient interaction. We extend both De Jaeger and Jegers [9] and Dranove [8] by introducing a patient population within which there are multiple levels of information. By taking a similar approach to existing models, we isolate the role of heterogeneously informed patient populations on the interaction.

**Conceptual overview of the model**

While most patients prefer to be well informed prior to a physician visit [3], there are certainly situations in which POMI is unlikely to be relevant. A patient may lack the incentive (perhaps when suffering from a minor cold) or time (perhaps if afflicted with a time-critical condition such as internal bleeding) to first obtain information. Alternately, POMI may have minimal impact on the interaction when the patient’s condition is so critical (e.g. heart attack) or when some treatment options are likely to lead to malpractice lawsuits (e.g. prescription of a large dose of narcotics for a minor ailment) that concerns over the patient’s welfare or potential malpractice lawsuits are likely to outweigh any personal preference of the physician. On the other hand, in some situations, POMI has the potential to significantly impact the physician—patient interaction. If a patient’s condition is serious but not urgent and no treatment option presents itself as obviously superior, the patient has incentive to gather POMI to preclude the physician from capitalizing on her superior information and pursuing her own preferences.

We envision a population of patients who share a common medical condition and a physician who treats these patients. Two treatment options ($A$ and $B$) are available, with the physician preferring one ($B$) to the other ($A$). While we remain agnostic as to the source of this preference, it may reflect scarcity of resources; concern for the global population of patients, perhaps reflected in an unwillingness to prescribe antibiotics for fear of raising population resistance levels; pursuit of scientific progress, if, for example, a new option is in clinical trials; the desire to avoid proposing treatments not covered by a patient’s insurance; and pursuit of increased monetary remuneration. Patients do not always share this preference, desiring the option that brings them the most individual benefit. Prior to the physician visit, patients gather information and make an estimation of the two available options. Patients may differ in their preferred treatment options, perhaps due to lifestyle, tolerance to side effects, or risk tolerance. During the encounter, the physician and patient exchange information and make a decision on which, if any, treatment option to implement. Agreement on a treatment option may result, or they may fail to reach agreement and the patient may pursue treatment with another physician [35].

The results of such a shared-decision-making process can depend on the structure of the bargaining process. We model the bargaining as a ‘take-it-or-leave-it’ process in which the physician recommends a treatment and the patient either accepts and receives the offered treatment or rejects and goes to another physician. This bargaining process reflects several institutional features. The physician as the offerer reflects her role as the expert and the gatekeeper of the healthcare system. Moreover, given the fact that a typical physician–patient interaction lasts about 15 min [16–18], it is reasonable to assume that the time cost of delaying agreement results in termination of bargaining after the first round if no agreement has been reached. Finally, restricting the patient’s strategy to a binary accept/reject decision reflects institutional barriers to surplus sharing; side payments in return for favorable treatment recommendations, for example, are ruled out.

Since our main focus is on the role of information in the physician–patient relationship, our simple model allows us to concentrate on the variables of interest without unduly complicating the theoretical analysis in the text. However, we demonstrate in Appendix D that our results correspond to the unique stationary equilibrium of a more general multi-round bargaining model. In the appendix, we also note that a class of bargaining processes may introduce additional screening equilibria in which the first treatment recommendation made by the physician serves to discriminate between differently informed patients. Such a screen allows the physician to determine precisely a patient’s information level, making the subsequent interaction one of full information on the part of the physician. Since the results under
full information have been widely studied [8, 9] and physicians’ time constraints put significant limitations on physicians’ ability and willingness to screen, this paper focuses on situations where screening does not occur.

Because patients may prefer treatment A or B depending on various factors, the physician actually faces two subgroups of patients: one initially preferring treatment A, the other B. Within each subgroup, patients may have different levels of confidence in their assessments. When these two subgroups of patients visit the physician, the interactions are quite different. For patients who initially prefer treatment B, in accord with the physician’s preferences, the interaction would be straightforward: the patient requests treatment B and the physician happily obliges. While it is possible that some patients in this subgroup would be better served by treatment A, there is no reason for the physician to recommend A since patients will certainly accept B. On the other hand, the interaction between the physician and patients whose preferences are opposite to hers exhibits tension: the physician has to persuade the patients to change their minds if she wants them to take her preferred option. Thus, the amount of information conveyed in the physician’s recommendation becomes of central importance. For this reason, we focus our attention on the interaction between the physician and the subgroup of patients whose initial preference is for treatment A, contrary to the physician’s interest. In the rest of this paper, ‘patient’ or ‘patient population’ will refer to this subgroup.

The likelihood that a patient will accept the physician’s recommendation depends on how certain the patient is about his state, which, in turn, depends on the quality of information the patient has gathered. While physicians are endowed with superior medical information, the possibility that the physician’s interests may differ from the patient’s leads the patient to be cautious when considering the physician’s recommendation. Such suspicion, in turn, makes the physician think twice before recommending a non-preferred option because doing so, regardless of an individual patient’s situation, may lead the patient to reject the physician’s services outright.

If all patients are uniformly informed, existing models describe the resulting physician–patient interaction [8, 9]. We envision different patients within the same socioeconomic group potentially having varying levels of information. Sacchetti et al. [36], for example, note the differences in quality of medical information patients obtained on the Internet. Alternatively, we can think of dissimilar qualities of information resulting from different predictive powers of various screening tests patients may undertake. Some patients may be quite informed about treatment options due to their access to information or their ability to adequately research and assimilate such information. These patients can have a fairly accurate assessment as to which option is best for them, and they may be unwilling to follow the physician’s recommendation if that recommendation differs from what they believe is best for them. Other patients may be particularly uninformed and will tend to follow the physician’s recommendation even if it differs from their initial beliefs. Certainly, the physician’s incentive to recommend the ‘right’ treatment from the patient’s perspective is greater when facing a more informed patient who is more likely to reject what he considers to be a less desirable suggestion.

Model

We consider the interaction between a patient population and a physician who cannot observe the information level of individual patients. We model the interaction as a Bayesian game of incomplete information [37–39]. The physician knows only the distribution of information levels in the patient population.

Each patient is in one of two states, either $x$ or $\beta$, corresponding to the patient’s best treatment option, $A$ or $B$. We denote the set of patient states by $S \equiv \{x, \beta\}$ with arbitrary element $s$ and the set of treatment options by $T \equiv \{A, B, 0\}$ with arbitrary element $t$ where $t = 0$ reflects no treatment. Patients are unaware of their state but the physician, who is assumed to have perfect information about patients’ medical conditions, assesses the state with complete accuracy. The patient’s utility from receiving treatment $t$ when he is in state $s$ is denoted by $u(t|s)$.

**Assumption 1**

\[
u(A|x) > u(0|x) = 0 > u(B|x)\quad\text{and}\quad u(B|\beta) > u(0|\beta) = 0 > u(A|\beta)
\]

We interpret the treatment $t = 0$ as rejecting the physician’s recommendation and normalize its utility to 0. The utility $u(0|x)$ may be thought of as the cost of rejecting the recommendation and going to another physician. The assumption that
both \(u(0 \mid z)\) and \(u(0 \mid \beta)\) are equal to zero is made to simplify the analysis but is not important to the nature of the equilibrium under study. Further, we assume that obtaining the less desirable treatment is strictly worse than rejecting the current recommendation and getting another chance for the desired treatment. Otherwise, the patient will accept any recommendation by the physician, which, in turn, implies that the physician will always recommend her preferred treatment.

Prior to their accepting any physician recommendations, patients obtain medical information that provides some imperfect information about their state. We consider two types of patients, high types with more accurate information and low types with poor information, denoted by \(i \in \{h, l\}\). A patient is a high type with probability \(q\) and a low type with probability \(1 - q\). A patient of type \(i\) has a likelihood of being in state \(z\) given by \(p_i \equiv \Pr\{s = z \mid i\}\). A patient with accurate information is more likely to be correct about his true state \((1 \geq p_h > p_l > 0)\) than a patient with poor information.

We are interested in situations in which the interests of the patients and the physician diverge. Thus, we assume that a physician always prefers treatment \(B\) while the patient population of interest initially prefers treatment \(A\).

**Assumption 2**

\[
p_h u(A \mid z) + (1 - p_h) u(A \mid \beta) > p_l u(A \mid z) + (1 - p_l) u(A \mid \beta) > 0
\]

\[
> p_l u(B \mid z) + (1 - p_l) u(B \mid \beta) > p_h u(B \mid z) + (1 - p_h) u(B \mid \beta)
\]

Assumption 2 implies that prior to the encounter with a physician, patients prefer treatment \(A\) to 0 to \(B\). Also, better-informed, high-type patients have stronger preferences than lesser informed, low types. Better information implies a greater certainty that one is in state \(z\), and thus a greater expected utility from accepting treatment \(A\). Denote by \(v(t)\) the physician’s utility from performing treatment \(t\). Since we wish to focus on disjointed patient and physician interests, we assume that \(v(B) > v(A) > v(0)\) and normalize \(v(0) = 0\). To summarize the timing of the game:

1. Nature determines the patient’s type, or information level, \(i \in \{h, l\}\) with \(\Pr\{i = h\} = q\).

The type is observed by the patient but not the physician. Nature also determines the patient’s state, \(s \in \{z, \beta\}\) with \(\Pr\{s = z \mid i\} = p_i\). The state is observed only by the physician;

2. The physician recommends a treatment option \(t\) from two possible options \(A\) and \(B\);

3. The patient either accepts or rejects this recommendation; and

4. The players (patient and physician) receive their payoffs \(u(t \mid s), v(t)\).

The extensive form representation (game tree) of this game is shown in Figure 1. In the next section we proceed to locate a perfect Bayesian equilibrium.

**Equilibrium analysis**

**Physician’s strategy space**

The physician has four possible pure strategies in the form of recommending one of two treatments.
likelihood that a physician prescribes treatment $A(B)$ when the patient is in state $x(\beta)$. We define the level of information transmission from the physician to the patient as

$$I = \pi(A|x) \times \pi(B|\beta)$$

Obviously, $0 \leq I \leq 1$. The case $I = 0$ implies that there is no information transmission in the interaction; $I = 1$ implies that the physician’s recommendation perfectly and accurately reveals the patient’s state. Note that when $\pi(A|x) = \pi(B|\beta) = 0$, we obtain $I = 0$ but one could argue that there is perfect information transmission since the patient can perfectly infer his true state. However, the physician is always recommending exactly the wrong treatment, and the patient would always reject this recommendation. Since this case can never arise in equilibrium and is unlikely to exist in practice, we can safely ignore such situations.

We use $I$ as a key indicator of physician behavior and of the physician–patient relationship. We demonstrate later (in Proposition 6) that $I$ is proportional to patient welfare, and thus captures both the level of information transmission and the level of patient welfare.

Because the physician has strictly ordered preferences over the treatments, $\nu(B) > \nu(A) > 0$, the following proposition demonstrates that perfect information transmission is never attainable.

**Proposition 1.** In equilibrium, $I < 1$.

The proof of this and all other propositions is in Appendix C.

### Patient’s strategy space

Each patient type (low and high information) has four possible pure strategies, in the form of accepting or rejecting each of the two possible recommendations. Denote by $\pi_i(t)$ the probability that a patient of type $i \in \{l, h\}$ accepts recommendation $t \in \{A, B\}$. Thus, a strategy profile for the patient consists of a four-tuple, $(\pi_h(A), \pi_h(B), \pi_l(A), \pi_l(B))$. The values of $\pi_i(t)$ reflect the level of trust patients place in the physician: larger values indicate a greater willingness to accept the physician’s recommendations. Because a high-type patient is more certain about his state than a low type ($p_h > p_l$), we have the following relationship when both options are recommended with positive probabilities:

**Proposition 2.** In any equilibrium in which both treatment options are recommended with positive probability, we have

- (i) $\pi_h(A) > 0 \Rightarrow \pi_h(A) = 1$
- (ii) $\pi_h(A) < 1 \Rightarrow \pi_h(A) = 0$
- (iii) $\pi_h(B) > 0 \Rightarrow \pi_h(B) = 1$
- (iv) $\pi_l(B) < 1 \Rightarrow \pi_l(B) = 0$

**Corollary 2.1.** In equilibrium, $\pi_h(A) \geq \pi_l(A)$, $\pi_h(B) \leq \pi_l(B)$.

Part (i) of Proposition 2 arises because, if $\pi_h(A) > 0$, a low type must view accepting treatment $A$ at least as favorably as rejecting it. Since the high type is even more likely to be in state $x$ (his prior level of information is higher than the low type, by Assumption 2), he can expect strictly positive expected utility from accepting treatment $A$. Similar arguments apply to parts (ii)–(iv).

**Corollary 2.2.** In equilibrium, $\pi_h(B) < 1$.

The better-informed patients must occasionally reject the physician’s preferred treatment option—a mechanism for keeping the physician honest. Intuitively, if a highly informed patient always accepted recommendation $B$, then low types would also accept recommendation $B$. Faced with $B$ being universally accepted, the physician will find it profitable to always recommend $B$. But then neither patient type will ever accept recommendation $B$, leading to a contradiction.

### Equilibrium

The results from the previous two subsections allow us to identify equilibria of the physician–patient interaction game. Signaling games often permit a multiplicity of equilibria because the players can entertain a whole range of out-of-equilibrium beliefs. However, as we demonstrate below, all but one class of equilibria are quite unappealing. For example, a physician may believe that patients will always reject any of her recommendations making the physician indifferent between all of her strategies. An equilibrium is
found in which the physician always proposes treatment $B$. All patients, of course, reject this recommendation, confirming the physician’s beliefs.

While such equilibria may seem absurd, most nevertheless survive multiple refinements, including perfection [40] and sequential rationality [41]. In what follows, we concentrate on one equilibrium in particular. In Appendix B, we carefully derive all alternate equilibria and demonstrate that these other equilibria fall into one of two classes. First are equilibria in which patients reject all physician recommendations. Thus, the physician–patient interaction game:

First are equilibria in which patients reject all physician recommendations. Thus, the physician–patient interaction game:

We concentrate on one equilibrium

First, we note several properties of this equilibrium

The equilibrium strategies depend on the value of $q$, the proportion of the population with high information levels. Specifically, the frequency with which the physician recommends option $A$ to patients in state $x$ and the probability with which patients accept treatment $B$ depends on whether the proportion of high types is above or below a critical threshold $q^*$. The implications of this equilibrium are presented in the next section. First, we note several properties of this equilibrium that are not shared by the other equilibria discussed above and presented in the appendix.

The equilibrium of Proposition 3 is the only equilibrium in which the physician’s strategy conveys any useful information, and it is the only equilibrium in which both treatment options are accepted by patients with positive probability. These properties suggest that the physician–patient relationship, while less than perfect (as evidenced by the information transmission level being less than one), is more than a buyer–seller relationship. The properties also suggest that when a patient goes to a physician, he is willing to listen to the physician, albeit suspiciously; and the physician is willing to take the patient’s demand into account, albeit not fully as there exists a conflict of interest. It is our belief that real-world physician–patient interactions, while not perfect because of
possible conflicts of interest and information asymmetry, do achieve a positive level of trust and information transmission as predicted by this equilibrium.

**Proposition 4.** The candidate equilibrium is the only equilibrium in which the information transmission level $I$ is strictly greater than 0.

The parameter $I$ is an indicator of physician behavior and the quality of the physician–patient relationship. When $I = 0$, the patient gains no information from the relationship, and the physician does not take the patient’s interest into account.

As this is the only equilibrium with positive information transmission, we may appeal to focal point theory to rule out other equilibria. If one equilibrium has some property, such as Pareto optimality, that conspicuously distinguishes it from other equilibria, and if this property is common knowledge among the players, then this equilibrium is likely to be the unique outcome [42]. The assumption that players can coordinate on a Pareto optimal (or payoff dominant) outcome has been proven useful in many applications. For example, see Katz and Shapiro [43] in a network context and De Jaegher and Jegers [9] in a model underlying this work.

The candidate equilibrium has two focal properties. First, the physician receives higher expected payoffs in this equilibrium than in any other. As the first mover, it may be reasonable to conceive of the physician as selecting the equilibrium. Second, under specific conditions, this equilibrium Pareto dominates all others.

**Proposition 5.** The physician receives higher expected utility under the candidate equilibrium than any other equilibrium. Further, the candidate equilibrium Pareto dominates other equilibria when $p_b(1 - p_t)u(A|x)u(B|\beta) < p_t(1 - p_b)u(A|\beta)u(B|x)$.

More important to policy makers than the level of information transmission is patient welfare. If we define patient welfare, $W$, as the expected utility of patients in equilibrium, then in the candidate equilibrium, $W$ is proportional to $I$, as indicated in Proposition 6 below.

**Proposition 6.** In the candidate equilibrium, $W = u(A|x)I[p_b + (1 - q)p_t]$.

Thus, increasing information transmission, $I$, is analogous to improving patient welfare.

---

**Information, patient welfare and the physician–patient relationship**

In this section, we examine the properties and implications of the equilibrium outcome. Foremost, we are concerned with patient welfare (or with $I$, which is proportional to $W$). Also, an efficient physician–patient encounter requires that the physician uncovers a patient’s true type and the patient accepts the associated treatment option. Hence, the information transmission level, $I$, reflects the quality of the physician–patient relationship as it denotes the amount of candor in the physician’s recommendation. Further, the value of $I$ may reflect the interests of public policy institutions for which $q$, $p_b$, and $p_t$, may serve as policy instruments that can be influenced by such institutions to achieve desired outcomes. The following four theorems establish the influence of POMI on the physician–patient relationship.

**Theorem 1.** In equilibrium, $\frac{\partial I}{\partial q} = 0$ when $q \neq q^*$.

Holding information levels $p_b$ and $p_t$ constant, a local change in the proportion of well-informed patients will not impact the information transmission level except when $q$ crosses the threshold value $q^*$ (Figure 2). Since the physician cannot observe a specific patient’s information level, she treats every patient in the population identically. How exactly she treats each patient depends only on whether the likelihood of encountering a well-informed patient is above or below $q^*$. If most patients are poorly informed (i.e. $q_1$ in Figure 2), there is little harm in inducing demand. If most are
well informed (i.e. $q_2$ in Figure 2), then the losses accrued from well-informed patients refusing such treatments loom large.

In a sense, the physician elects one of two pre-defined strategies, represented by $I_h$ and $I_l$, based on the perceived POMI characteristics of the patient population. When $q > q^*$, the physician’s equilibrium strategy makes high types indifferent between accepting and rejecting recommendation $B$ resulting in a level of information transmission, $I_h$. The high information level is ‘dominant’ in the sense that $I$ depends on $p_h$ but not on $p_l$. When $q < q^*$, the opposite case results. The level of information transmission is equal to $I_l$ and is independent of $p_h$. Thus, the physician–patient relationship exists as one of two regimes, which we term the ‘well-informed regime’ and the ‘ill-informed regime.’ A change in $q$ has no effect on information transmission ($I$) within a regime but may potentially shift us from one regime to the other.

Efforts to elevate the quality of care (increase $I$) by increasing the number of well-informed patients (direct-to-consumer advertising, public education programs, etc.) are effective only if the number of informed patients reaches a certain critical mass. Otherwise, the physician will just ignore the well-informed patients. The following result notes the impact of information precision on this critical value $q^*$.

**Theorem 2.** The threshold value $q^*$ increases with $p_l$ and decreases with $p_h$.

(i) $\frac{\partial q^*}{\partial p_l} > 0$; (ii) $\frac{\partial q^*}{\partial p_h} < 0$

Theorem 2 suggests that the threshold value is endogenously determined and may be lowered either by increasing $p_h$ or by decreasing $p_l$. Changes in information levels potentially have two effects. The direct effect, of course, is that better information leads to greater honesty on the part of the physician, and thus a higher level of information transmission. However, a secondary effect is that changes in information levels may also influence the critical value $q^*$ and bring about a regime change. These two effects are summarized in Theorems 3 and 4.

**Theorem 3.** In equilibrium,

(i) $\frac{\partial I}{\partial p_l} > 0$ when $q > q^*$

(ii) $\frac{\partial I}{\partial p_h} = 0$ when $q < q^*$

(iii) for small $\varepsilon > 0$, there exists a $q < q^*$ such that $I|_{p_h} < I|_{p_h+\varepsilon}$

A change in $p_h$ both increases the level of information transmission to the right of $q^*$ and lowers the threshold value of $q$ for the well-informed regime. Thus, part (iii) of the theorem indicates that a small change in the precision of information may change the regime, leading to a marked increase in the level of information transmission. Figure 3 illustrates this result.

Increasing $p_h$ makes well-informed patients even more informed. For example, if the Internet is used to distribute information about clinical trial openings or new treatment options, then patients with Internet access become more informed while the information level of those without Internet access is unchanged. Theorem 3 indicates that such increases in information precision will influence the physician–patient relationship only when we are in or quite close to the well-informed regime. The effect is most profound when the proportion of well-informed patients is just below the critical value. By raising the accuracy of information, physicians now find it optimal to concern themselves with the well-informed patients. This regime change can substantially increase the overall level of information transmission.

**Theorem 4.** In equilibrium,

(i) $\frac{\partial I}{\partial p_l} > 0$ when $q < q^*$

(ii) $\frac{\partial I}{\partial p_l} = 0$ when $q > q^*$

(iii) for small $\varepsilon > 0$, there exists a $q < q^*$ such that $I|_{p_l} > I|_{p_l+\varepsilon}$

Increasing $p_l$ corresponds with raising the level of information amongst the least informed patients,

![Figure 3. Impact of an increase in the precision of high-type information ($p_h$) on information transmission ($I$)](image-url)
possibly through health literacy education. Theorem 4 suggests that increasing information precision among lesser-informed patients will influence the physician only in the ill-informed regime, or when the proportion of high types is sufficiently small. However, the effect need not be positive. By increasing the general level of information, the physician may begin to show more concern for the ill-informed patient, when before, she concerned herself only with the well informed. This regime change, bringing us from the well-informed regime to the ill-informed regime, implies that the physician will switch her focus from well-informed patients, instead treating all patients as if they possess the lower level of information (Figure 4).

While this result seems counterintuitive, it is actually similar to the reason for regime change due to a change in $q$ discussed earlier. Increasing the information level of the ill-informed patients is equivalent to increasing the relative number of ill-informed patients among those requesting the less desirable treatment from the physician’s standpoint. This makes the physician more attuned to this population, and as a result, makes the well-informed patients less willing to accept the physician’s recommendation of $B$. The reason well-informed patients are more cautious is that they believe that physicians may be favoring ill-informed patients (and treatment $B$), to their detriment. When $q$ is just above $q^*$, the well-informed patients are just marginally willing to accept recommendation $B$ (i.e. $\pi_{ill}(B)$ is just above $0$), and an increased $p_i$, just as a decreased $q$, will make the well-informed patients unwilling to accept a recommendation of $B$, causing a regime change.

For concreteness, consider a medical clinical trial of a novel treatment for a particular condition. Clinical trials, while valuable for the advancement of medicine, require significant additional resources and physician time, hence, they are typically not the preferred choice in a busy physician practice (the clinical trial is treatment $A$ in our context). The normal course of treatment would be more time efficient and rewarding for the physician (treatment $B$ in our context). However, some well-informed patients will request to be placed on a clinical trial, while most ill-informed patients will be unaware of such a study. Increasing the information level of ill-informed patients makes more ill-informed patients aware of the trial, and thus increases the number who begin to ask for the study as their preferred treatment option. Before this increase in information, well-informed patients are willing to accept the traditional treatment if they believe that their physician is reasonably looking out for their best interest, and that she is suggesting not going on study because she honestly believes that the study will not assist them. These well-informed patients are aware, however, that an increase in the number of ill-informed patients asking for the trial will make the physician more likely to recommend the traditional treatment even when the study will be helpful to them, and as a result, these patients become less willing to accept the traditional treatment. They may reason that almost all of the patients that ask for the clinical trial are not well informed and easily dissuaded. Thus, the physician suggests the traditional treatment because most patients will accept it whether or not it is the right option. This trend will continue as the ill-informed patients become more informed up to the point where the well-informed patients are no longer willing to accept the traditional treatment at all. In other words, a regime change has happened due to increased information level of ill-informed patients.

Our results share certain features with models that consider a single patient type. In De Jaegher and Jegers [9], the nature of the interaction induces the physician to reveal some information about the patient’s state, but this revelation is necessarily imperfect. Similarly, Dranove [8] examines a single treatment option that can be offered in different quantities, and finds that the level of demand induction is a function of the patient’s information level. Thus, the quality of the physician–patient relationship depends directly on the level of the POMI. In our setting with heterogeneous levels of POMI, this conclusion still holds but in a qualified way. When the physician cannot observe
an individual patient’s information level, it is the quality of information of the dominant patient type that matters to the physician. The physician responds to changes in the accuracy of information of the dominant patient type but does not respond to a change in the level of information of other patients, unless such a change switches the dominant patient type.

Policy implications

Many in the medical community envision a patient-centered medical system and a shared-decision-making model as the future of medical practice [44]. Yet, physicians sometimes fail to recommend the treatment option most desirable for a patient with a specific condition. As many as one-third of physicians ‘sometimes,’ ‘often,’ or ‘very often’ do not offer a useful service to a patient because of health plan rules [45]. Not surprisingly, physicians whose own rewards were closely tied to controlling costs were more likely to withhold information. Beyond the financial incentives, other concerns may further distort physicians’ interests away from those of the patient. For example, if doing so would require an undue amount of a physician’s time, the physician may not supply all the information a patient demands. Partly due to these reasons, patients are increasingly using computer-mediated tools to gather information about their conditions. Our study verifies that better-informed patients help improve the physician–patient interaction, but in a qualified way. Together with efforts to align physicians’ incentives with patients’ objectives, the targeted provision of high-quality medical information to the public positively shapes the patient–physician relationship. While informed patients are better suited to share in decision making [46, 47], we know that access to outside medical information is linked to a patient’s socioeconomic status [22–24]. Those with poorer socioeconomic status usually have less access to information, and have less ability to understand and use such information in the medical encounter [19–21]. Further, we find that the educational makeup of a given demographic group can impact the healthcare of all in that group. How, then, should policy makers target educational efforts, potentially to decrease or eliminate the persistent healthcare disparities among different socioeconomic groups [48–52]? Our study suggests that we should target education efforts towards different subgroups within a population depending on the existing information structure of the population.

For example, when the low information level is dominant (e.g. for patient population with low socioeconomic status), the most effective way to enhance information transmission and patient welfare is to increase the ratio of highly informed patients so that a ‘regime change’ occurs by focusing on the provision of uncomplicated, understandable medical information [53]. Efforts to increase the information level of those already well informed would have no impact. On the other hand, when the high information level is dominant (e.g. for patient population with high socioeconomic status), the most effective way to enhance information transmission and patient welfare is to increase the information level of those already informed. Efforts to increase the information level of those not informed within such a population would have no impact.

Our study also suggests that when the physician cannot distinguish the information level of individual patients within the same population, every patient may be treated the same, regardless of his information level. The physician may respond only to changes in the overall information structure of the population, not to a change in the information level of a particular patient. Patients in a certain socioeconomic group may find themselves in a situation where they cannot alter the behavior of their physicians even if they become better informed. Our model provides further explanation for empirical studies that show widespread disparities in access and utilization of health care among patients with different socioeconomic characteristics [48–52].

Acknowledgements

We thank two anonymous referees for their comments. We also thank participants of the 4th iHEA conference in San Francisco for helpful comments. We are also grateful to Mythreyi Bhargavan, Kris De Jaegher, and Jeff Rubin for their comments.

Appendix A: Notation

As in the text, we denote by \( p_i(t) \) the probability that a patient of type \( i \in \{h, l\} \) accepts recommendation \( t \in \{A, B\} \). Denote by \( \mu_i(s|t) \) the probability
that a patient of type $i$ assigns to being in state $s$ upon receiving a treatment recommendation $t$. If treatment $t$ is prescribed with positive probability, we have

$$
\mu_i(x|t) = \frac{p_i \pi(t|x)}{p_i \pi(t|x) + (1 - p_i) \pi(t|\beta)} \\
\mu_i(\beta|t) = \frac{(1 - p_i) \pi(t|\beta)}{p_i \pi(t|x) + (1 - p_i) \pi(t|\beta)}
$$

If a treatment option is not on the equilibrium path, then a Bayesian equilibrium permits the patient to hold any belief. Further, denote by $U_i(t)$ the expected utility of a patient of type $i$ from accepting treatment $t$ and by $V(t|s)$ the expected utility of the physician from recommending treatment $t$ to a patient in state $s$. These are given by

$$
U_i(t) = \mu_i(x|t)u(t|x) + \mu_i(\beta|t)u(t|\beta)
$$

$$
V(t|x) = v(t) \frac{q_p h \pi_h(t) + (1 - q)p \pi_a(t)}{q_p h + (1 - q)p}
$$

$$
V(t|\beta) = v(t) \frac{q(1 - p_h) \pi_h(t) + (1 - q)(1 - p) \pi_a(t)}{q(1 - p_h) + (1 - q)(1 - p)}
$$

### Appendix B: Equilibrium analysis

As discussed in the text, the model permits multiple equilibria. In this section, we enumerate all equilibria and demonstrate that the list is exhaustive. In Table B1, we classify equilibria into three categories. Equilibrium C1 (which is a continuum of equilibria when $q = q^*$) is the candidate equilibrium analyzed in the text. Equilibria in C2 share a common feature that no patient will ever accept any recommendation. Because in this case the physician always receives an expected payoff of 0 regardless of which state the patient is in or which option the physician recommends, she is always indifferent between recommending the two options, so she can choose $\pi(A|x)$ and $\pi(B|\beta)$ freely as long as her choice does not make any patient better off accepting any recommendation than rejecting it. Equilibria in C3 share a common feature that only treatment $A$ is recommended. Because patients' prior beliefs are that option $A$ yields positive utility, both types of patients will always accept this recommendation. To support these equilibria, out-of-equilibrium beliefs (when option $B$ is recommended) must lead to rejection of $B$ with sufficiently high probability. Prior to demonstrating that the list of equilibria is exhaustive, we establish a lemma specifying an equilibrium condition.

**Lemma B.1.** In any equilibrium in which both treatment options are recommended with positive probability,

(i) If $\exists i \in \{h, l\}$ such that $U_i(A) \leq 0$, then $\pi_i(B) = 0, j \in \{h, l\}$

(ii) If $U_i(A) > 0, i \in \{h, l\}$, then $\exists j \in \{h, l\}$ such that $\pi_j(B) > 0$

**Proof.** (i) We proceed by contradiction. Assume that, for some $i$ and $j$, $U_i(A) \leq 0$ and $\pi_j(B) > 0$.

$$
U_i(A) \leq 0 \Rightarrow \mu_i(x|A) < p_i \text{ by Assumption 2} \\
\Rightarrow \pi(A|x) + \pi(B|\beta) < 1
$$

$$
\pi_j(B) > 0 \Rightarrow U_j(B) \geq 0 \\
\Rightarrow \mu_j(\beta|B) \geq 1 - p_i \text{ by Assumption 2} \\
\Rightarrow \pi(A|x) + \pi(B|\beta) \geq 1
$$

leading to a contradiction. Therefore, $\pi_i(B) = 0$.

(ii) $U_i(A) > 0, i \in \{h, l\} \Rightarrow \pi_i(A) = 1 \Rightarrow V(A|x) = V(A|\beta) = v(A)$

But for $B$ to be recommended with positive probability, $\exists s \in \{x, \beta\}$ such that $V(B|s) \geq V(A|s) = v(A)$, which implies $\exists j \in \{h, l\}$ such that $\pi_j(B) > 0$.

We now proceed to prove the main result of this subsection.

**Proposition B.1.** Table B1 contains an exhaustive list of equilibria of the game.

**Proof.** We classify possible equilibria into four exhaustive categories:

(i) Equilibria in which only treatment $B$ is recommended in equilibrium;

(ii) Equilibria in which only treatment $A$ is recommended in equilibrium;

(iii) Equilibria in which both treatments are recommended with positive probability and in which either $U_i(A) \leq 0$ or $U_j(A) \leq 0$ (or both);

(iv) Equilibria in which both treatments are recommended with positive probability and $U_h(A), U_h(A) > 0$. 

Copyright © 2006 John Wiley & Sons, Ltd.
<table>
<thead>
<tr>
<th>Equilibrium</th>
<th>Physician strategy</th>
<th>Patient strategy</th>
<th>Patient strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\pi(A</td>
<td>x)$</td>
<td>$\pi(B</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\pi_1(A)$</td>
</tr>
<tr>
<td></td>
<td>$\pi_2(A)$</td>
<td>$\pi_3(A)$</td>
<td>$\pi(B)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\pi_4(B)$</td>
<td>$\pi(B)$</td>
</tr>
</tbody>
</table>

**C1**

$q > q^*$

$1 + \frac{(1 - p_h) u(B|\beta)}{p_h u(B|x)}$

$1 + \frac{(1 - p_l) u(B|\beta)}{p_l u(B|x)}$

$v(A) - \frac{p_l (1 - q)}{p_h} q$ \left[ 1 - \frac{v(A)}{v(B)} \right]$

$1$

$q < q^*$

$1 + \frac{(1 - p_l) u(B|\beta)}{p_l u(B|x)}$

$v(A) - \frac{p_l (1 - q)}{p_h} q$

$1 + \frac{p_h}{p_l} \frac{q}{p_l (1 - q)}$

$1$

$q = q^*$

$1 + \frac{(1 - p_l) u(B|\beta)}{p_l u(B|x)}, \ldots, 1 + \frac{(1 - p_h) u(B|\beta)}{p_h u(B|x)}$

$v(A) - \frac{p_l (1 - q)}{p_h} q$

$1 + \frac{p_h}{p_l} \frac{q}{p_l (1 - q)}$

$1$

$q^* = \frac{v(B) - v(A)}{pr v(A) + [v(B) - v(A)]}$

**C2**

$\pi_1(A|x) u(A|x) \leq (1 - p_l) (1 - \pi_2(B|\beta)) [-u(A|\beta)]$

and

$(1 - p_l) \pi_2(B|\beta) u(B|\beta) \leq p_l (1 - \pi_1(A|x)) [-u(B|x)]$

**C2.1**

$0$

$0$

$0$

$0$

**C2.2**

$0$

$0$

$0$

$0$

**C3**

$0 \leq \pi_0(B) \leq \min \left\{ \frac{v(A)}{v(B)} + \frac{(1 - q) p_l}{p_h} \left[ \frac{v(A)}{v(B)} - \pi_1(B) \right], \ldots \right\}$

$1$

$1$

$1$
We construct equilibria in each category.

(i) Equilibria in which only treatment $B$ is recommended in equilibrium.

\[ \pi(A|x) = 0, \pi(B|\beta) = 1 \implies \mu_i(\beta|B) = 1 - p_i \]
\[ \Rightarrow U_i(B) < 0, \pi_i(B) = 0, \]
\[ i \in \{h, l\}, \text{ by Assumption 2} \]
\[ \Rightarrow V(B|x) = 0 \]

Since $\pi(A|x) = 0$, $V(A|x) \leq V(B|x) = 0$. But for $V(A|x)$ to equal 0, the probability of either type accepting $A$ must be zero; thus, $\pi_i(A) = 0$, corresponding to equilibrium C2.2.

(ii) Equilibria in which only treatment $A$ is recommended in equilibrium.

\[ \pi(A|x) = 1, \pi(B|\beta) = 0 \implies \mu_i(x|A) = p_i \]
\[ \Rightarrow U_i(A) > 0, \pi_i(A) = 1, i \in \{h, l\}, \]
\[ \text{by Assumption 2} \]
\[ \Rightarrow V(A|x) = V(A|\beta) = v(A) \]

\[ \pi(A|x) = 1, \ \pi(B|\beta) = 0 \text{ also imply that} \]
\[ V(B|x) \leq V(A|x) \text{ and } V(B|\beta) \leq V(A|\beta), \]

These conditions are equivalent to:

\[
0 \leq \pi_b(B) \leq \min\left\{ \frac{\pi(A)}{v(B)} + \frac{\left(1 - q\right) p_i}{q p_i} \left[ \frac{\pi(A)}{v(B)} - \pi_i(B) \right], \frac{v(A)}{v(B)} + \frac{\left(1 - q\right) (1 - p_i)}{q (1 - p_i)} \left[ \frac{v(A)}{v(B)} - \pi_i(B) \right], 1 \right\}
\]

which correspond to equilibria in C3. Such equilibrium always exist. For example, \[ \pi_b(B) \leq \pi_i(B) = \frac{v(A)}{v(B)} \] satisfies the above condition.

(iii) Equilibria in which both treatments are recommended with positive probability and in which either $U_i(A) \leq 0$ or $U_b(A) \leq 0$ (or both).

By Lemma B.1, $\pi_b(B) = \pi_i(B) = 0$. Since neither type accepts treatment $B$, and both options are recommended with positive probability, we have

\[ V(B|x) = V(B|\beta) = 0 \]
\[ \Rightarrow V(A|x) = V(A|\beta) = 0, \ \pi_b(A) = \pi_i(A) = 0 \]

But, for $i \in \{h, l\}$,

\[ U_i(A) \leq 0 \Rightarrow p_i \pi(A|x) u(A|x) \]
\[ \quad + (1 - p_i)(1 - \pi(B|\beta)) u(A|\beta) \leq 0 \]
\[ U_i(B) \leq 0 \Rightarrow p_i(1 - \pi(A|x)) u(B|x) \]
\[ \quad + (1 - p_i)\pi(B|\beta) u(B|\beta) \leq 0 \]

which correspond to equilibria in C2.1. Note that these equilibria always exist since the above conditions are satisfied in a neighborhood around $\pi(A|x) = \pi(B|\beta) = 0$.

(iv) Equilibria in which both treatments are recommended with positive probability and $U_h(A), U_i(A) > 0$.

By Lemma B.1, $U_h(A), U_i(A) > 0$ imply $\pi_j(B) > 0$ for some $j \in \{h, l\}$. From Proposition 2 and Corollary 2.2, either:

(iv-a) $0 \leq \pi_b(B) < 1, \pi_i(B) = 1$; or

(iv-b) $\pi_b(B) = 0, 1 \geq \pi_i(B) > 0$

We consider both of these cases:

(iv-a) $U_h(A) > 0, U_i(A) > 0$, $0 \leq \pi_b(B) < 1$, and

\[ \pi_i(B) = 1 \implies \pi_b(A) = \pi_i(A) = 1, \]
\[ \pi_i(B) = 1 \implies U_i(B) \geq 0 \]
\[ \Rightarrow p_i[1 - \pi(A|x)] u(B|x) + (1 - p_i)\pi(B|\beta) u(B|\beta) \geq 0 \] (B1)
\[ U_b(A) > 0 \implies \pi(A|x) > 0 \]
\[ 0 \leq \pi_b(B) < 1 \implies U_b(B) \leq 0 \]
\[ \Rightarrow p_b[1 - \pi(A|x)] u(B|x) \]
\[ \quad + (1 - p_i)\pi(B|\beta) u(B|\beta) \leq 0 \]
\[ \implies \pi(A|x) < 1 \] (B2)

Thus, $0 < \pi(A|x) < 1$, which implies indifference on the part of the physician:

\[ V(B|x) = V(A|x) = v(A)[qph + (1 - q)p_i] \]
\[ = v(B)[\pi_b(B) qph + (1 - q)p_i] \]
\[ \Rightarrow \pi_b(B) = \frac{v(A)}{v(B)} - \frac{p_i(1 - q)}{p_i q} \left[ 1 - \frac{v(A)}{v(B)} \right] \]

for which no solution exists when $q < q^*$. But when $q \geq q^*$, $\pi(B|\beta) = 1$ since

\[ V(B|\beta) - V(A|\beta) = \frac{(1 - q)(p_b - p_i)[v(B) - v(A)]}{p_b q (1 - p_b) + (1 - q)(1 - p_i)} > 0 \]

Substituting into (B1) and (B2), we obtain

\[ p_f[1 - \pi(A|x)] u(B|x) + (1 - p_f)u(B|\beta) \geq 0 \] (B1')
\[ p_b[1 - \pi(A|x)] u(B|x) + (1 - p_b)u(B|\beta) \leq 0 \] (B2')

When \( q > q^* \), \( \pi_h(B) > 0 \) implies that \( U_h(B) \geq 0 \). From above, \( U_h(B) \leq 0 \). Thus, \( U_h(B) = 0 \) and (B2') holds with equality. Since \( p_h > p_l \), (B1') holds with strict inequality. Thus,  
\[
\pi(A|x) = 1 + \frac{(1 - p_l) u(B|\beta)}{p_l} u(B|z)
\]

When \( q = q^* \), (B1') and (B2') require that  
\[
1 + \frac{(1 - p_l) u(B|\beta)}{p_l} u(B|z) \leq \pi(A|x) \leq 1 + \frac{(1 - p_h) u(B|\beta)}{p_h} u(B|z)
\]

The above conditions correspond to equilibrium C1 when \( q \geq q^* \).  

(iv-b) \( U_h(A) > 0 \), \( U_l(A) > 0 \), \( \pi_h(B) = 0 \) and \( 0 < \pi_l(B) \leq 1 \) imply that \( \pi_h(A) = \pi_l(A) = 1 \), and  
\[
0 < \pi_l(B) \leq 1 \implies U_l(B) \geq 0 
\]

\[
\Rightarrow p_l[1 - \pi_l(A|x)]u(B|x) + (1 - p_l)\pi_l(B)u(B|\beta) \geq 0 \quad (B3)
\]

\( U_h(A) > 0 \implies \pi(A|x) > 0 \), and  
\[
\pi_h(B) = 0 \implies U_h(B) \leq 0 
\]

\[
\Rightarrow p_h[1 - \pi(A|x)]u(B|x) + (1 - p_h)\pi_l(B)u(B|\beta) \leq 0 \quad (B4)
\]

\[
\Rightarrow \pi(A|x) < 1
\]

Thus, \( 0 < \pi_l(A|x) < 1 \), which implies indifference on the part of the physician:  
\[
V(B|x) = V(A|x) 
\]

\[
\Rightarrow v(A)(q p_h + (1 - q) p_l) = v(B)(1 - q)p_l \pi_l(B) 
\]

\[
\Rightarrow \pi_l(B) = \frac{v(A)}{v(B)} \left[ 1 + \frac{p_l q}{p_l(1 - q)} \right]
\]

for which no solution exists when \( q > q^* \). But when \( q \leq q^* \), \( \pi_l(B|\beta) = 1 \) since  
\[
V(B|\beta) - V(A|\beta) = \frac{v(B)(1 - q)(1 - p_l)\pi_l(B) - v(A)(q(1 - p_h) + (1 - q)(1 - p_l))}{q(1 - p_h) + (1 - q)(1 - p_l) - q v(A) > 0}
\]

Substituting into (B3) and (B4),  
\[
p_l[1 - \pi_l(A|x)]u(B|x) + (1 - p_l)\pi_l(B|\beta) \geq 0 \quad (B3')
\]

\[
p_h[1 - \pi_l(A|x)]u(B|x) + (1 - p_h)\pi_l(B|\beta) \leq 0 \quad (B4')
\]

When \( q < q^* \), \( \pi_l(B) < 1 \) implies that \( U_l(B) \leq 0 \). From above, \( U_l(B) \geq 0 \). Thus, \( U_l(B) = 0 \) and (B3') holds with equality. Since \( p_h > p_l \), (B4') holds with strict inequality. Thus,  
\[
\pi(A|x) = 1 + \frac{(1 - p_l) u(B|\beta)}{p_l} u(B|z)
\]

When \( q = q^* \), \( \pi_l(B) = 0 \), (B3') and (B4') require that  
\[
1 + \frac{(1 - p_l) u(B|\beta)}{p_l} u(B|z) \leq \pi(A|x) \leq 1 + \frac{(1 - p_h) u(B|\beta)}{p_h} u(B|z)
\]

The above conditions correspond to equilibrium C1 when \( q \leq q^* \).  

\[\square\]

### Appendix C: Proofs

**Proposition 1.** In equilibrium, \( I < 1 \).  

**Proof.** We proceed by contradiction. Assume that \( I = \pi(A|x) \times \pi(B|\beta) = 1 \). This means that \( \mu_i(x|A) = \mu_i(B|B) = 1 \), \( i \in \{h,l\} \), implying (by Assumption 1) that both types of patients will accept either recommendation. But \( \pi_l(A) = \pi_l(B) = 1 \), \( i \in \{h,l\} \), implies that the physician will always recommend option \( B \). Then, \( \pi(A|x) = 0 \), leading to a contradiction.  

**Proposition 2.** In any equilibrium in which both treatment options are recommended with positive probability, we have  

(i) \( \pi_l(A) > 0 \implies \pi_l(A) = 1 \)  

(ii) \( \pi_h(A) < 1 \implies \pi_h(A) = 0 \)  

(iii) \( \pi_h(B) > 0 \implies \pi_h(B) = 1 \)  

(iv) \( \pi_l(B) < 1 \implies \pi_l(B) = 0 \)  

**Proof.** We demonstrate parts (i) and (ii) as the others are obtained analogously. If \( \pi_l(A) > 0 \), the low-type patient must weakly prefer accepting recommendation \( A \) to rejecting it. Since \( p_h > p_l \), we have  
\[
\pi_l(A) > 0 \implies U_l(A) \geq 0 
\]

\[
\Rightarrow p_l\pi_l(A|x)u(A|x) + (1 - p_l) \times (1 - \pi_l(B|\beta))u(A|\beta) \geq 0
\]

\[
\Rightarrow p_l\pi_l(A|x)u(A|x) + (1 - p_l)
\]

\[
\times (1 - \pi_l(B|\beta))u(A|\beta) > 0
\]

\[
\Rightarrow U_l(A) > 0
\]

\[
\Rightarrow \pi_l(A) = 1
\]
Note that (ii) is logically equivalent to (i). Analogous arguments demonstrate (iii) and (iv).

**Corollary 2.1.** In equilibrium, \( \pi_h(A) \geq \pi_i(A) \), \( \pi_h(B) \leq \pi_i(B) \).

**Proof.** Follows immediately from Proposition 2. □

**Corollary 2.2.** In equilibrium, \( \pi_h(B)<1 \).

**Proof.** If \( \pi_h(B) = 1 \), then by Proposition 2, \( \pi_i(B) = 1 \). Thus, both patient types accept treatment \( B \). Since \( v(B) > v(A) \), the physician would always recommend treatment \( B \), which would, by Assumption 2, be rejected by both patient types, leading to a contradiction. □

**Proposition 3.** C1 in Table B1 is an equilibrium of the patient–physician interaction game.

**Proof.** Follows directly from Proposition B.1. □

**Proposition 4.** The candidate equilibrium (equilibrium C1 from Table B1) is the only equilibrium in which the information transmission level \( I \) is strictly greater than 0.

**Proof.** For equilibria in C2 and C3, \( I = 0 \). For equilibrium C1,

\[
I \geq 1 + \frac{(1-p_i)u(B|\beta)}{p_i} > 0
\]

where the first inequality follows from \( p_h > p_i \) and \( u(B|\beta) > 0 > u(B|x) \) and the second inequality follows from Assumption 2. □

**Proposition 5.** The physician receives higher expected utility under the candidate equilibrium than any other equilibrium. Further, the candidate equilibrium Pareto dominates other equilibria when \( p_h(1-p_i)u(A|x)u(B|\beta) < p_i(1-p_h)u(A|\beta)u(B|x) \).

**Proof.** First, note that in C2, the physician and both types of patients receive an expected payoff of 0, while in C1, all players receive positive expected payoffs. Next, denote by \( E[v] \) the physician’s expected utility in equilibrium, and by \( E[U_h] \) and \( E[U_i] \) the high- and low-type patients’ expected utilities in equilibrium. We use \( E[v]|_{C1} \) and \( E[v]|_{C3} \) to denote the physician’s expected utility in equilibrium C1 and C3; similarly, \( E[U_i]|_{C1} \) and \( E[U_i]|_{C3} \) denote the expected utility of a patient of type \( i \) in equilibrium C1 and C3. By direct computation, we obtain

\[
E[v]|_{C1} = v(A)[q_p + (1-q)p_i]\pi(A|x)
\]

\[
+ v(B)[\pi_h(B)q(1-p_h) + \pi_i(B)(1-q)(1-p_i)]
\]

\[
+ v(B)[\pi_h(B)q p_i + \pi_i(B)(1-q)p_i][1 - \pi(A|x)]
\]

\[
= v(A)[q_p + (1-q)p_i] + v(B)[\pi_h(B)q(1-p_h)
\]

\[
+ \pi_i(B)(1-q)(1-p_i)]
\]

\[
= \begin{cases} 
  v(A)[q_p + (1-q)p_i] + v(B) \left[ \frac{(1-q)p_h - p_i}{p_h} \right] & q > q^* \\
  v(A)[q_p + (1-q)p_i] + v(B) \left[ \frac{q_p + (1-q)p_i}{p_i} \right] & q \leq q^* 
\end{cases}
\]

\[
> v(A)
\]

\[
= E[v]|_{C3}
\]

Hence, the physician receives higher expected utility under C1 than C3. In equilibrium C3, patients’ payoffs are given by

\[
E[U_i]|_{C3} = p_i u(A|x) + (1-p_i)u(A|\beta) \quad \text{for } i \in \{h,l\}
\]

In equilibrium C1,

\[
E[U_i]|_{C1} = \begin{cases} 
  p_i u(A|x) + \frac{(1-p_h)p_i u(A|\beta)u(A|x)}{p_i u(B|x)} & q > q^* \\
  p_i u(A|x) + \frac{(1-p_i)p_i u(A|\beta)u(A|x)}{p_i u(B|x)} & q \leq q^*
\end{cases}
\]

and

\[
\frac{p_i u(A|x)}{p_i u(B|x)} + \frac{(1-p_h)p_i u(A|\beta)u(A|x)}{p_i u(B|x)} \leq E[U_i]|_{C1} \leq \frac{p_i u(A|x)}{p_i u(B|x)} + \frac{(1-p_h)p_i u(A|\beta)u(A|x)}{p_i u(B|x)}
\]

Because \( p_h > p_i \), we have

\[
E[U_i]|_{C1,q=q^*} \geq E[U_i]|_{C1,q=q^*} \geq E[U_i]|_{C1,q<q^*}
\]

So we need only to show that \( E[U_i]|_{C1,q<q^*} > E[U_i]|_{C3} \).

\[
E[U_i]|_{C1,q<q^*} - E[U_i]|_{C3} = p_i u(A|x) + \frac{1-p_i}{p_i} \frac{u(B|\beta)}{u(B|x)} u(A|x)
\]

\[
- p_i u(A|x) - (1-p_i)u(A|\beta)
\]

which is positive when \( p_h(1-p_i)u(A|x) u(B|\beta) < p_i(1-p_h)u(A|\beta)u(B|x) \). □
Proposition 6. In the candidate equilibrium, $W = u(A|z)I[q_{ph} + (1 - q)p_i]$. 

Proof. As we just showed, in the candidate equilibrium,
\[
E[U_i]|_{C_1} = \begin{cases} 
 p_iu(A|z) + \frac{(1 - p_h)p_iu(B|\beta)u(A|z)}{p_hu(B|\beta)} & q > q^* \\
 p_iu(A|z) + \frac{(1 - p_h)p_iu(B|\beta)u(A|z)}{p_hu(B|\beta)} & q < q^*
\end{cases}
\]

and
\[
I = \begin{cases} 
 1 + \frac{(1 - p_h)u(B|\beta)}{p_iu(B|\beta)} & q > q^* \\
 1 + \frac{(1 - p_h)(1 - p_i)u(A|z)}{p_iu(B|\beta)} & q < q^*
\end{cases}
\]

So
\[
W = qE[U_i] + (1 - q)E[U_j] = q[p_iu(A|z) + (1 - 1)p_iu(A|z)] + (1 - q)[p_iu(A|z) + (1 - 1)p_iu(A|z)] = u(A|z)I[q_{ph} + (1 - q)p_i]
\]

Theorem 1. In equilibrium (C1), $\frac{\partial I}{\partial q} = 0$ when $q \neq q^*$. 

Proof. Follows directly from the form of C1. 

Theorem 2. The threshold value $q^*$ increases with $p_i$ and decreases with $p_h$.

(i) $\frac{\partial q^*}{\partial p_i} > 0$; (ii) $\frac{\partial q^*}{\partial p_h} < 0$

Proof. This follows directly from the form of
\[
q^* = \frac{p_iV(A) - V(B)}{p_i}
\]

Theorem 3. In equilibrium (C1),

(i) $\frac{\partial I}{\partial q_{ph}} > 0$ when $q > q^*$;

(ii) $\frac{\partial I}{\partial p_h} = 0$ when $q < q^*$

(iii) for small $\epsilon > 0$, there exists a $q < q^*$ such that $I|_{p_h} << I|_{p_h + \epsilon}$

Proof. (i) and (ii) follow directly from the form of C1.

(iii) Let $\delta \equiv q^*(p_h) - q^*(p_h + \epsilon) > 0$. Then, there exists a $q$ such that $q^* - \delta < q < q^*$. 

Theorem 4. In equilibrium (C1),

(i) $\frac{\partial I}{\partial p_i} > 0$ when $q < q^*$; (ii) $\frac{\partial I}{\partial p_i} = 0$ when $q > q^*$.

(iii) for $\epsilon > 0$, there exists a $q > q^*$ such that $I|_{p_i} >> I|_{p_i + \epsilon}$

Proof. (i) and (ii) follow directly from the form of C1.

(iii) Let $\delta \equiv q^*(p_i) - q^*(p_i + \epsilon) > 0$. Then, there exists a $q$ such that $q^* - \delta < q < q^* + \delta$. 

Appendix D: Generalized bargaining model

In the body of this manuscript, we model the bargaining procedure over treatment options as a take-it-or-leave-it offer proposed by the physician. In this section, we discuss more robust bargaining models. Traditional models of two-player bargaining assume both continuity of the set of possible outcomes and a cost of delay, which encourages earlier agreements. For example, Rubinstein’s [54] classic model of sequential bargaining considers a division of a pie of unit size, with the pie ‘shrinking’ at a predetermined rate after each round of unsuccessful negotiation. In our scenario, since bargaining is likely to occur in a span of minutes rather than years, discounting is unlikely to be relevant. Instead, we adopt the ‘uncertain termination’ approach of Binmore et al. [55] in which a rejection of the current proposal implies some positive probability that bargaining will terminate rather than continue into another period. This probabilistic termination model is more intuitive in the present context and, in many settings, is isomorphic to a model in which the surplus vanishes with time [55].

The discrete set of bargaining outcomes is either treatment $A$ or treatment $B$. Institutionally, there is no ability for the patient and physician to split the surplus through side payments, such as through a monetary payment by the physician to the patient in consideration of the patient accepting a less desirable treatment option. Further, any such side payments would limit the generality of our model as they would require interpersonal comparisons of utility. Instead, we ‘convexify’ the set of treatment options by considering that a proposal by the physician is of the form $\pi(A|x)$ and
\(\pi(B|\beta)\); the physician proposes a probability of treatment \(A\) to a patient in state \(z\) and a probability of treatment \(B\) in state \(\beta\).

For concreteness (though quite divorced from reality), we imagine a computer program or some other intermediary carrying out the proposal according to the following rules: if the patient accepts a proposal \(\{\pi(A|z), \pi(B|\beta)\}\), the physician verifies the patient’s state and the intermediary recommends a treatment drawn from the state-appropriate distribution. By accepting a proposal, the patient also commits to accepting any recommendation that is a result of the proposal. More realistically, envision the physician sharing her ‘treatment philosophy’ prior to examining the patient: ‘For patients who exhibit conditions similar to yours, if the test results are positive, I recommend a treatment drawn from the state-period, \(b\) bargaining model with incomplete information. In each period, \(t\), the physician proposes a probability pair \(\{\pi(A|z), \pi(B|\beta)\}\). A patient of type \(i \in \{h, l\}\) accepts the proposal with probability \(\text{Acc}_i\). If it is accepted, the implemented treatment option \((A \text{ or } B)\) is randomly drawn according to the probability \(\pi(A|z)\) or \(\pi(B|\beta)\) corresponding to the patient’s state. If the offer is rejected, bargaining terminates with probability \(1 - \lambda\). With complementary probability \(1 - \lambda\), the physician makes another offer in the following period. The termination probability \(\lambda\) represents the chance of breakdown in negotiations; either the physician or the patient may find such seesaw negotiations awkward or costly and decide to end further bargaining. This model is a (simple) instance of the sequential bargaining with one-sided incomplete information model of Ausubel et al. [56], to which we refer the reader for a more general discussion. In what follows, we solve for the equilibria of this model.

Denote the history of the game through period \(t - 1\) by

\[h^t = \{\pi^t(A|z), \pi^t(B|\beta), \text{Acc}^t\}^t_{t=1}\]

We concentrate on stationary equilibria of this model in which a player’s strategy depends only on whether or not an acceptance has occurred, and not on the period, \(t\). For the patient, since an acceptance implies the end of the game, his strategy is independent of history. The physician’s strategy, on the other hand, is independent of \(t\) but may take on a different value if the patient has previously rejected an offer. Thus, an equilibrium strategy for the physician may be denoted by the four-tuple: \(\{\pi^1(A|z), \pi^1(B|\beta), \pi^2(A|z), \pi^2(B|\beta)\}\), in which \(\pi^1(A|z), \pi^1(B|\beta)\) is the offer in period 1, and \(\pi^2(A|z), \pi^2(B|\beta)\) is the offer in all subsequent periods \((t > 1)\) if previous offers have been rejected.

Denote by \(U_i[\pi(A|z), \pi(B|\beta)]\) the expected utility received by a type \(i\) patient when he accepts proposal \(\{\pi(A|z), \pi(B|\beta)\}\); by \(V_i[\pi(A|z), \pi(B|\beta)]\) the expected utility the physician receives when only a type \(i\) patient accepts the proposal; and by \(V_i[\pi(A|z), \pi(B|\beta)]\) the expected utility the physician receives when both types of patients accept. The following proposition characterizes the unique stationary equilibrium.

Proposition D.1. Define

\[
\pi^h(A|z) = 1 + \frac{(1 - p_h)u(B|\beta)}{p_h} u(B|z)
\]

\[
\pi^l(A|z) = (1 - \lambda)\pi^h(A|z) - \frac{u(B|\beta) - p_i[u(B|\beta) - u(B|z)]}{p_i[u(A|z) - u(B|z)]}
\]

\[q^* = \frac{\lambda V_i[\pi^l(A|z), 1] + V_i[\pi^h(A|z), 1]}{\lambda V_i[\pi^h(A|z), 1] + V_i[\pi^l(A|z), 1] - V_i[\pi^h(A|z), 1]}
\]

The following is a unique stationary equilibrium of the model:

(i) if \(q < q^*\):

\[
\pi^1(A|z) = \pi^2(A|z) = \pi^h(A|z)
\]

\[
\pi^1(B|\beta) = \pi^2(B|\beta) = 1
\]

\[\text{Acc}^i = 1 \forall t, \ i \in \{l, h\}\]

(ii) if \(q > q^*\):

\[
\pi^1(A|z) = (1 - \lambda)\pi^h(A|z) - \frac{u(B|\beta) - p_i[u(B|\beta) - u(B|z)]}{p_i[u(A|z) - u(B|z)]}
\]

\[
\pi^2(A|z) = \pi^h(A|z)
\]

\[
\pi^1(B|\beta) = \pi^2(B|\beta) = 1
\]

\[\text{Acc}^l = 1 \forall t, \ \text{Acc}^h = 0, \ \text{Acc}^h = 1 \ t > 1\]

We note several features of the equilibrium. First, similar to the equilibrium discussed in the text, its
nature depends on a threshold value of $q$ (i.e. $q^{**}$). When $\lambda = 1$, and thus negotiations do not last past one period, $q^{**}$ corresponds to the $q^*$ defined in the text. Second, the physician’s strategy is similar in this infinite horizon model as in the take-it-or-leave-it model discussed in the text in that she basically chooses between two possible strategies to maximize her expected utility. Thus, the equilibrium of this model resembles that of the simple model in the text. However, this model also opens the possibility to screening. When the ratio of well-informed patients is above a critical value, the physician plans on making two offers. The first is less advantageous but is accepted by the less informed. If rejected, the physician is aware that the patient is more informed and offers a better (from the patient’s perspective) treatment option. Whether such screening actually occurs is an empirical question deserving of future research efforts.

The rest of this section develops several lemmas to prove the above proposition.

**Lemma D.1.** In a stationary equilibrium, a patient accepts within the first two periods.

**Proof.** Suppose $\{\pi^i(A|x), \pi^i(B|\beta), \text{Acc}^i\}$ is an equilibrium and consider a period $t > 1$ such that a patient of type $i$ has rejected previous offers. If a patient rejects the physician’s offers $\{\pi^i(A|x), \pi^i(B|\beta)\}$ in period $t$, the physician earns 0, since the same offer will exist in all future periods (by definition of a stationary equilibrium). However, the physician can offer $\pi^i(A|x) = \pi^* = 1$ which the patient will accept (since it yields maximal possible payoffs), and which results in positive payoffs for the physician. Thus, a $\{\pi^i(A|x), \pi^i(B|\beta)\}$, $t > 1$, which is rejected cannot be part of an equilibrium. □

**Corollary D.1.** In a stationary equilibrium, a type $l$ patient will accept in period 1.

**Proof.** By contradiction. If both types were to reject in period 1 and (by Lemma D.1) accept in period 2, then letting $\pi^1(A|x) = \pi^2(A|x)$ and $\pi^1(B|\beta) = \pi^2(B|\beta)$, that is, changing the first period offer to be the same as the second period offer, will lead both types to accept in period 1, yielding higher payoffs for the physician. Lastly, by construction, if a low type rejects an offer then so will a high type, implying that the low type must accept in period 1. □

**Lemma D.2.** In a stationary equilibrium, $\pi^i(B|\beta) = 1$ for all $i$.

**Proof.** By construction, the patient’s utility function is increasing in $\pi^i(B|\beta)$. Also, as long as one type of patient accepts, the physician’s utility function is increasing in $\pi^i(B|\beta)$. Suppose in equilibrium that for some period $t$, $\pi^i(B|\beta) < 1$. Consider an alternate physician strategy where $\pi^i(B|\beta) = \pi^i(B|\beta)$ for all $t \neq t$, and $\pi^i(B|\beta) = 1$. The patient is not less likely to accept and the physician earns a higher payoff. Thus, $\pi^i(B|\beta) < 1$ cannot be part of an equilibrium. □

**Lemma D.3.** In a stationary equilibrium, if both types do not accept in period 1, then $U_i[\pi^i(A|x)] = (1 - \lambda)U_i[\pi^i(A|x)]$.

**Proof.** (i) Suppose $U_i[\pi^i(A|x)] > (1 - \lambda)U_i[\pi^i(A|x)]$. Then there exists an $\varepsilon$ such that $U_i[\pi^i(A|x) - \varepsilon] > (1 - \lambda)U_i[\pi^i(A|x)]$; the physician can lower the offer a little bit and still have type $l$ patients accept, but resulting in a higher payoff for the physician. (ii) Suppose

$$U_i[\pi^i(A|x)] < (1 - \lambda)U_i[\pi^i(A|x)]$$

Then, a type $l$ patient rejects $\pi^i(A|x)$ in round 1 and accepts $\pi^2(A|x)$ in round 2, contradicting Corollary D.1. □

Next, denote by

$$\pi^h(A|x) = 1 + \frac{(1 - p_h)u(B|\beta)}{p_hu(B|\beta)}$$

and

$$\pi^l(A|x) = 1 + \frac{(1 - p_l)u(B|\beta)}{p_lu(B|\beta)}$$

the offers that result in type $h$ and type $l$ patients earning a payoff of zero if accepted.

**Lemma D.4.** In a stationary equilibrium

(i) $\pi^1(A|x) \geq \pi^i(A|x)$

(ii) $\pi^2(A|x) = \pi^i(A|x)$

**Proof.** (i). Follows from Corollary D.1 since a low type must find the offer acceptable. (ii). An offer of $\pi^2(A|x)$ implies that a rejection occurred in period 1. By Corollary D.1, only the high-type patient rejected the first offer. Thus, $\pi^h(A|x)$ places a minimum bound on offers acceptable in period 2. Any higher offer would also be accepted but yield a lower payoff for the physician. □

Thus, two possibilities exist. Either both types of patients accept in period 1, or only low types
accept in period 1 and high types accept in period 2. The first period offer is \( \pi^*(A|x) = \pi^h(A|x) \). In the latter case, Lemma D.3 implies a first period offer of

\[
\pi^*(A|x) = (1 - 2\pi^h(A|x)) - \frac{u(B|\beta) - p_u[u(B|\beta) - u(B|x)]}{p_u[u(A|x) - u(B|x)]}
\]  

(D1)

The physician effectively selects between these two strategies depending on which maximizes her payoffs.

**Lemma D.5.** In a stationary equilibrium, both types accept in period 1 if and only if

\[
V[\pi^h(A|x), 1] \geq (1 - q)V[\pi^*(A|x), 1] + q(1 - \lambda)V[\pi^h(A|x), 1]
\]  

(D2)

**Proof.** Suppose \( V[\pi^h(A|x), 1] < (1 - q)V[\pi^*(A|x), 1] + q(1 - \lambda)V[\pi^h(A|x), 1] \) and both types accept in period 1. But then an offer of \( \pi^*(A|x) \) in period 1 and \( \pi^h(A|x) \) in period 2 would result in a higher payoff. Equation (D2) may be rewritten in terms of \( q \). The game will end in the first round if:

\[
q > q^{**} = \frac{V[\pi^*(A|x), 1] - V[\pi^h(A|x), 1]}{\lambda V[\pi^h(A|x), 1] + V[\pi^*(A|x), 1] - V[\pi^h(A|x), 1]}
\]

**Proof of Proposition D.1.** follows from Lemmas D.1–D.5. \( \square \)

**References**


