RATIONAL DELAY IN APPLYING FOR POTENTIALLY LIFE-SAVING DIAGNOSIS

by

Gideon Yaniv
RATIONAL DELAY IN APPLYING FOR POTENTIALLY
LIFE-SAVING DIAGNOSIS

by

Gideon Yaniv

ABSTRACT

The self-discovery of a suspicious symptom is often subject to an emotional turbulence: while recognizing the importance of having the symptom diagnosed promptly, individuals frequently delay diagnosis, fearing to hear that they are developing a serious illness (and hoping that the symptom will disappear by itself). On the one hand, delaying diagnosis yields the benefit of not knowing for sure that one is seriously ill, as well as the opportunity to avoid painful treatment. On the other hand, delaying diagnosis entails not only the loss of relief brought about by finding out that one is actually healthy, but also the risk of incurring increased health damage or dying before getting life-saving treatment. The present paper proposes a rational, economic-oriented approach to explaining individuals' delay behavior, inspired by health psychologists' contention that people hold organized, cognitive representations of particular diseases that influence their reactions to symptoms. A major psychological premise underlying the analysis is that one is presently better off not knowing whether he is actually ill or not than knowing for sure that he is ill and must undergo painful treatment. Delay behavior may then be perceived as reflecting a rational process of weighing the cost and benefit involved in delayed diagnosis, or of balancing the fear of being told the suspected truth against the fear of consequences of further procrastination.

Assuming that the real costs of diagnosis and medical treatment are borne by health insurance, the paper first constructs a multi-period expected-utility maximization model to inquire into the individual's decision of whether or not to delay diagnosis, as well as to determine the optimal time of applying for diagnosis. The results rationalize a variety of observed behavior concerning individuals' reaction to suspicious symptoms that vary with the likelihood of illness and the severity of the damage to health incurred by delayed diagnosis. The paper proceeds to examine, through employing a multi-period expected-cost minimization model, the desirability of delayed diagnosis to the health insurer, comparing the implications of the two models. The individual's delay decision is finally compared with the socially desired solution, and cost-sharing measures to induce the individual to approach the social optimum are suggested.

Key Words: Symptom, Diagnosis, Severe Illness, Health Damage, Delay Behavior.

JEL Classification: I12, I18, D81.

PsycINFO Classification: 3360, 3361.
I. Introduction

The self-discovery of a suspicious physical or mental symptom is often subject to an emotional turbulence: while recognizing the importance of having the symptom diagnosed promptly, individuals frequently delay diagnosis, fearing to hear that they are developing a serious illness (and hoping that the symptom will disappear by itself). On the one hand, delaying diagnosis yields the benefit of not knowing for sure that one is seriously ill, as well as the opportunity to avoid painful treatment. On the other hand, delaying diagnosis entails not only the loss of relief brought about by finding out that one is actually healthy, but also the risk of incurring increased health damage or dying before getting life-saving treatment.

Delaying diagnosis for recognized symptoms has been extensively researched by health psychologists, unveiling "rational" (e.g., perceived expenses and fear of treatment) as well as "irrational" (e.g., sense of invulnerability or fatalism) reasons. Economists, however, have not yet addressed this issue, despite its involvement of typical economic features such as intertemporal tradeoffs, uncertainty (of illness) and risk (of dying), self-induced health damage, and costs of diagnosis and treatment. The present paper proposes a rational, economic-oriented approach to explaining individuals' delay behavior, inspired by health psychologists' contention that people hold organized, cognitive representations of particular diseases (i.e., beliefs about their causes, symptoms, treatment, and likelihood of being fatal) that influence their reactions to symptoms.

---

1 Antonovsky and Hartman (1974) concluded that at least three-fourth of cancer patients delayed visiting a physician for at least one month after first noticing a suspicious symptom, and that somewhere between 35 to 50 percent of patients delayed seeking treatment for over three months.

2 For a review of the health psychology literature on delay behavior, see Antonovsky and Hartman (1974) and Taylor (1991, Ch. 9).

3 A closely-related exception is Francis (1997), who has recently extended Brito's et al (1991) vaccination decision model to take account of dynamic considerations. Delaying vaccination, while entailing the risk of catching the lethal disease and dying, has not been associated with the discovery and assessment of symptoms, nor has it been assumed to incur health damages (besides dying) or to involve treatment costs.

underlying the analysis is that one is presently better off not knowing whether he is actually ill or not than knowing for sure that he is ill and must undergo painful treatment. Delay behavior may then be perceived as reflecting a rational process of weighing the cost and benefit involved in delayed diagnosis, or of balancing the fear of being told the suspected truth against the fear of consequences of further procrastination.5

The paper begins with addressing the individual’s problem, assuming that the real costs of diagnosis and medical treatment are borne by health insurance. A multi-period expected-utility maximization model is constructed to inquire into the individual’s decision of whether or not to delay diagnosis, as well as to determine the optimal time of applying for diagnosis (Section II). The results rationalize a variety of observed behavior concerning individuals’ reaction to suspicious symptoms that vary with the likelihood of illness and the severity of the damage to health incurred by delayed diagnosis. The paper proceeds to examine, through employing a multi-period expected-cost minimization model, the desirability of delayed diagnosis to the health insurer, comparing the implications of the two models (Section III). The individual’s delay decision is finally compared with the socially desired solution, and cost-sharing measures to induce the individual to approach the social optimum are suggested (Section IV). A summary of the main results concludes the paper (Section V).

I. The Individual’s Problem

Consider an individual who at a certain point of time, denoted by 0, becomes aware of the presence of a suspicious physical or mental symptom, which, to the best of his knowledge, has the probability \( \lambda \) of indicating a serious illness. Suppose that \( 0 < \lambda < 1 \) (i.e., \( \lambda \) is

---

5 Antonovsky and Hartman (1974) conclude that the general suggestion emerging from studies dealing with styles of coping with fear and anxiety is that some people are immobilized by fear, adopting the defense mechanisms of denial and repression, whereas others respond to fear by mobilizing energy for positive coping action. A clear-cut dichotomy as such implies, however, that people either promptly apply for diagnosis or do not apply at all. The observed variability in delay behavior suggests rather a conflict between the two coping modes (or between two opposing fears), which the present paper rationalizes.
strictly positive and less than unity), so that the individual does not know for sure whether he is ill or healthy and must undergo a diagnosis to find this out.\footnote{On the social and psychological factors affecting individuals' recognition of symptoms see, for instance, Mechanic (1972), Scheier et al (1979), Pennebaker (1983), and Fillingin and Fine (1986). On factors affecting individuals' interpretation of symptoms, see, for instance, Safer et al (1979), Leventhat et al (1982), Stoller (1984), and Jemmot et al (1988).} Having the symptom diagnosed thus bears the probability $\lambda$ of yielding a ‘positive’ result (‘ill’), and the probability $1-\lambda$ of yielding a ‘negative’ one (‘healthy’). Suppose further that the individual’s well-being is dependent upon knowing whether he is ill or healthy and denote his utility levels at the alternative states of knowledge by $v^p$ and $v^N$, respectively, where $v^N > v^p$.\footnote{The diagnosis of severe illness has been reported in the health psychology literature to result in lowered body image and self-esteem [e.g., Schwab and Hameling (1968)], which is often accompanied by anxiety and depression [e.g., Rodin and Voshart (1986)].} Suppose also that not knowing for sure whether he is ill or healthy is inferior to knowing for sure that he is healthy, but superior to knowing for sure that he is ill. Denoting the utility level attained at the initial state of uncertainty by $v^0$, assume therefore that $v^N > v^0 > v^p$.

Suppose now that the suspicious symptom, while life-threatening, is not painful or incapacitating. The individual may thus consider the possibility of delaying diagnosis, fearing to find out that he is actually ill, and hoping that the symptom will disappear by itself. Given that the symptom does not indicate severe illness, suppose that there is a differentiable cumulative probability distribution, $F(t)$, of the symptom disappearing by itself at or before time $t$. However, given that the symptom does indicate severe illness and that the individual avoids treatment, suppose that there is a differentiable cumulative probability distribution, $P(t)$, of dying at or before time $t$, and that after-death utility is zero.\footnote{For simplicity, suppose that the individual is a lonely self-sufficient person, whose premature death inflicts no economic or emotional losses on his family, which he might otherwise take into account in his delay decision.} Suppose further that the functions $F(t)$ and $P(t)$, as well as their time derivatives, $\dot{F}(t)$ and $\dot{P}(t)$, are known to the individual. If a diagnosis is made at some time $\theta$ in the future, and the result turns out to be ‘positive’, the individual is assumed to follow doctors’
orders concerning immediate and future treatment. Following doctors' orders ensures, by assumption, that the individual sustains his life. However, the longer the delay in diagnosis, the greater the irreversible damage, \( m(\theta) \), incurred to the individual's health, thus the greater the intensity of treatment required constantly, at each time \( t \) that follows, to sustain life.

Suppose now that the damage to health incurred by severe illness consists of a fixed component, \( g \geq 0 \), reflecting damages which cannot be avoided by prompt diagnosis of the symptom, and a self-induced, variable-with-delay component, \( \mu(\theta) \), where \( \mu(0) = 0 \), \( \mu'(\theta) > 0 \), and \( \mu''(\theta) > 0 \). Hence, by assumption, \( m(\theta) = g + \mu(\theta) \). Suppose also that the greater the damage to health (and thus the required intensity of treatment), the greater the pain and discomfort involved in obtaining treatment. Suppose further that the pain and discomfort of treatment are proportionate to the accumulated health damage, thus expressible as \( sm(\theta) \), where \( s > 0 \) is a coefficient which transforms the damage into the disutility derived from the pain and discomfort involved in its treatment. Diagnosing the symptom is assumed to inflict no pain or discomfort. The money costs of diagnosis and treatment are assumed to be covered by health insurance.

Suppose now that the individual chooses the time of applying for diagnosis, \( \theta^* \), so as to maximize the expected present value of the (infinite lifetime) utility stream, \( V \), resulting from delayed diagnosis. That is, suppose that the individual seeks to maximize

\[
V = (1 - \lambda) \left[ \int_0^\theta F(t) \left[ \int_0^t v^0 e^{-\delta t} d\tau + \int_0^\theta v^N e^{-\delta t} d\tau \right] dt \right]
+ \left[ 1 - F(\theta) \right] \left[ \int_0^\theta v^0 e^{-\delta t} dt + \int_0^\theta v^N e^{-\delta t} dt \right]
+ \lambda \left[ \int_0^\theta \dot{P}(t) \left[ \int_0^t v^0 e^{-\delta t} d\tau \right] dt + \left[ 1 - P(\theta) \right] \left[ \int_0^\theta v^0 e^{-\delta t} dt + \int_0^\theta (v^R - sm(\theta)) e^{-\delta t} dt \right] \right],
\]

(1)

where \( \delta(\leq 1) \) denotes a subjective discount rate.
Notice that the expected value of the utility stream comprises two major terms, one multiplied by \(1 - \lambda\) and the other by \(\lambda\). The former term relates to the possibility that the symptom does not indicate severe illness. In this case, the symptom either disappears, with probability \(\tilde{F}(t)\), at any time \(t\) preceding time \(\theta\), or, with probability \(1 - \tilde{F}(\theta)\), remains intact until time \(\theta\) when the individual applies for a diagnosis. The latter term relates to the possibility that the symptom does indicate severe illness. In this case the individual either dies, with probability \(\tilde{P}(t)\), at any time \(t\) preceding time \(\theta\), or survives, with probability \(1 - \tilde{P}(\theta)\), to apply for diagnosis at time \(\theta\). Expression (1) attaches the alternative utility levels, \(v^0, v^N, v^P - sm(\theta)\), or zero to the appropriate cases in accordance with the time of revelation.

Simplifying (1), the individual’s problem can be reformulated as that of maximizing\(^9\)

\[
V = \frac{1 - \lambda}{\delta} \left\{ \int_0^\theta \tilde{F}(t)[v^0(1 - e^{-\delta t}) + v^Ne^{-\delta t}] dt + [1 - \tilde{F}(\theta)][v^0(1 - e^{-\delta \theta}) + v^Ne^{-\delta \theta}] \right\} \\
+ \frac{\lambda}{\delta} \left\{ \int_0^\theta \tilde{P}(t)[v^0(1 - e^{-\delta t})] dt + [1 - \tilde{P}(\theta)][v^0(1 - e^{-\delta \theta}) + (v^P - sm(\theta))e^{-\delta \theta}] \right\}.
\]

(2)

Differentiating (2) with respect to \(\theta\), the first-order condition for the maximization of the expected utility stream is given by

\[
\frac{dV}{d\theta} = -e^{-\delta \theta}(1 - \lambda)[1 - \tilde{F}(\theta)](v^N - v^0) + \\
e^{-\delta \theta} \lambda \left\{ [1 - \tilde{P}(\theta)][\delta(v^0 - v^P + sm(\theta)) - s\mu'(\theta)] - \tilde{P}(\theta)(v^P - sm(\theta)) \right\} = 0,
\]

(3)

which, rearranging, yields

\(^9\) Notice that \(\int_0^\theta e^{-\delta t} dt = 1/\delta\), \(\int_0^\theta e^{-\delta t} dt = (1/\delta)(1 - e^{-\delta \theta})\), and \(\int_0^\theta e^{-\delta t} dt = (1/\delta)e^{-\delta \theta}\).
\[ MB_\theta^i = e^{-\theta_0} \lambda (1 - P(\theta))[v^0 - (v^P - sm(\theta))] = e^{-\theta_0} (1 - \lambda) [1 - F(\theta)](v^N - v^0) \]
\[ + \frac{e^{-\theta_0}}{\delta} \lambda \{ \hat{P}(\theta)[v^P - sm(\theta)] + [1 - P(\theta)]s\mu'(\theta)\} = MC_\theta^i. \quad (4) \]

That is, the optimal time of applying for diagnosis is that level of \( \theta \) for which the marginal benefit to the individual (\( i \)) from delaying diagnosis, \( MB_\theta^i \), equals the marginal cost, \( MC_\theta^i \).

The marginal benefit of delaying diagnosis at time \( \theta \) arises in view of the probability \( \lambda \) that the symptom indicates severe illness, given the probability \( 1 - P(\theta) \) of surviving time \( \theta \). Delaying diagnosis for an additional unit of time yields the present value of the utility difference at time \( \theta \) between not knowing the exact state of one’s health and knowing that he is ill (and bearing the pain and discomfort of treatment), \( e^{-\theta_0}[v^0 - (v^P - sm(\theta))] \). The marginal cost of delaying diagnosis at time \( \theta \) arises, on the one hand, in view of the probability \( 1 - \lambda \) that the symptom does not indicate severe illness, given the probability \( 1 - F(\theta) \) that it does not disappear until time \( \theta \). Delaying diagnosis for an additional unit of time incurs the loss of the present value of the utility difference at time \( \theta \) between knowing that one is actually healthy and not knowing the exact state of his health, \( e^{-\theta_0}(v^N - v^0) \). On the other hand, given the probability \( \lambda \) that the symptom indicates severe illness, delaying diagnosis bears the risk \( \hat{P}(\theta) \) of dying at time \( \theta \), thus losing the present value of the utility stream that could have been attained from time \( \theta \) to eternity if not delaying, \( (e^{-\theta_0}/\delta)(v^P - sm(\theta)) \). Most importantly, given the probability \( 1 - P(\theta) \) of surviving time \( \theta \), delaying diagnosis for an additional unit of time increases the present value of future pain and discomfort by \( (e^{-\theta_0}/\delta) s\mu'(\theta) \), avoidable if applying for diagnosis at time \( \theta \).

Alternatively, condition (4) may be interpreted as balancing, at the margin, a struggle between two opposing fears: on the one hand, the fear of finding out that one is actually ill (which, at the optimum, must be greater than the hope that one is actually healthy).

---

\[ ^{10} \text{Notice that the marginal benefit from delaying diagnosis is positive even if } g = 0 \text{ [and thus } m(\theta) = \mu(\theta)], \text{ which implies that all damages (and pain of treatment) can be avoided if the symptom is promptly diagnosed. Still, delay may be desirable if } v^0 - v^P \text{ is sufficiently large (that is, if one's self-esteem decreases sufficiently due to finding out that he is actually ill).} \]
encourages further delay in diagnosis, at the risk of dying or of incurring increased health damage; on the other hand, fearing the consequences of further procrastination discourages further delay, at the risk of finding out that one is actually ill. Equilibrium is reached when these opposing forces balance.

Proposition 1: (a) If the damage to health incurred by a slight delay in diagnosis is sufficiently high, delay is never desirable to the individual, irrespective of the probability that the symptom indicates severe illness. (b) If the damage to health incurred by a slight delay in diagnosis is sufficiently low, delay is desirable to the individual only if the probability that the symptom indicates severe illness is sufficiently high. (c) Given that delay of diagnosis is desirable to the individual, optimal delay will be longer the higher the probability that the symptom indicates severe illness and the greater the unavoidable health damage inflicted by the illness.

Proposition 1 is proved in Appendix 1. Evaluating equation (4) at $\theta=0$, it is shown that there exist cutoff levels, $\mu^*$ and $\lambda^*$, such that if $\mu'(0) \geq \mu^*$, delay is never desirable, and if $\mu'(0) < \mu^*$, delay is desirable only if $\lambda > \lambda^*$. In the latter case, $\lambda^*$ is lower, and thus the incentive to delay is greater, the smaller is $\mu'(0)$ but the greater is the unavoidable damage, $g$. If delaying, $d\theta^*/d\lambda > 0$ and $d\theta^*/dg > 0$, thus $\theta^*$ increases with $\lambda$ and $g$. The former implications are summarized in Table 1, where D and N denote delay and no-delay decisions, respectively.

<table>
<thead>
<tr>
<th>Probability that the symptom indicates severe illness</th>
<th>LOW</th>
<th>HIGH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Damage to health incurred by a slight delay in diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOW</td>
<td>N</td>
<td>D</td>
</tr>
<tr>
<td>HIGH</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

Table 1: Delay (D) / No-Delay (N) decisions implied by utility maximization
Proposition 1 provides a rational explanation for a variety of observed behavior concerning individuals’ response to the self-discovery of a potentially life-threatening symptom. At the one extreme lies the worried well, who frequently rushes to emergency rooms upon the discovery of minor symptoms which “rational” individuals tend to ignore. Proposition 1(a) suggests that if the damage to health resulting from the slightest delay in obtaining treatment is significant, delay is undesirable even if the probability that the symptom indicates severe illness is very low. Hence, it is perfectly rational to apply for an immediate diagnosis even when experiencing a minor chest or left-hand pain, since any delay in case that the pain does happen to indicate a developing heart attack is known to be crucial in physicians’ ability to save life or to prevent an irreversible heart damage. At the other extreme lies the seemingly health negligent who delays diagnosis even if the probability that the symptom indicates severe illness is very high. Proposition 1(b) suggests that such behavior is rational given that the damage to health resulting from delaying diagnosis is relatively small. A retired college professor who experiences frequent memory difficulties and gradual loss of analytical ability is likely to deny any problem, thus refuse seeing a neurologist who might diagnose the development of the Alzheimer’s disease. He may find comfort and justification in recognizing that delay in diagnosis has little or no effect on the progress of this illness. Similarly, individuals of any age do not rush to take blood tests upon the revelation that they have gained weight, although there is a high probability that gaining weight is accompanied by increased levels of triglycerides and cholesterol in their blood. They enjoy not knowing for sure that they are developing arterial sclerosis, reasoning that delay in diagnosis incurs relatively little damage, as this illness is known to progress relatively slow. Proposition 1(b) also suggests that if both the probability of severe illness and the damage to health incurred by a slight delay in diagnosis are low, prompt diagnosis, which is likely to yield an immediate sense of relief, is advantageous. Indeed, singles who do not belong to any of the AIDS high-risk groups would normally not

11 In a study of the worried well, Wagner and Curran (1984) found this group to be more concerned about their physical and mental health and to perceive minor symptoms as more serious than appropriate users of medical services.

12 See Devanand and Mayeux (1992).
hesitate to take the HIV antibody test upon the request of a new sex partner, even if suffering from frequent colds. Homosexuals, on the other hand, are known to delay the HIV antibody test for years (and even for months after beginning to experience suspicious symptoms such as intermittent fevers or loss of weight),\(^\text{13}\) reasoning that being diagnosed as a carrier of the virus would adversely affect their well-being while doing little to affect the progress of the illness. As suggested by Proposition 1(c), diagnosis is delayed significantly, since both the probability of illness and the unavoidable damage inflicted by this illness are very high. Finally, notice that Proposition 1(a) implies that if both the probability of illness and the damage incurred by a slight delay in diagnosis are high, avoiding a prompt diagnosis is irrational. Hence, a rational approach to diagnosis delay behavior is unable to explain individuals' delay in seeking treatment for suspicious symptoms of malignant melanoma, known to be curable if treated promptly,\(^\text{14}\) or for an irradiating chest pain. Delaying diagnosis in these cases may thus be attributed to 'irrationalities', such as feelings of invulnerability (i.e., denying the possibility that the symptom might indicate severe illness) or fatalism (i.e., disbelief in the effectiveness of treatment).

II. The Health Insurer’s Problem

Suppose now that the individual is insured by a health insurance firm which bears the money costs of diagnosis and medical treatment following a ‘positive’ result. Suppose that the money costs of treatment are proportionate to the accumulated health damage, thus expressible as \(cm(\theta)\), where \(c > 0\) is a coefficient which transforms the damage into the money costs involved in its treatment. Suppose further that the health insurance firm is interested in minimizing the expected present value of its total money costs, \(C\), emanating from the discovery of a suspicious symptom by the individual. What is the optimal time of

---

\(^{13}\) See Rabkin and Kertzner (1992).

\(^{14}\) Taylor (1991, Ch. 9) reports a study which found that patients waited, on average, 1 year from the time they first noticed a skin lesion to receiving a diagnosis of malignant melanoma.
diagnosis, \( \hat{\theta} \), for the firm? Is delay in diagnosis desirable to the firm? To answer these questions in a comparable manner to the individual's problem, the analysis will be restricted to non-contagious diseases, allowing the health insurer to abstract from the possibility that delay in diagnosis might infect other insurees. It will also be assumed that the health insurer's assessments regarding the probability that the symptom indicates severe illness and the damage to health incurred by delaying diagnosis coincide with those of the individual's.

Denoting by \( z \) the money cost of diagnosis, the expected present value of the firm's money costs stream, given diagnosis at time \( \theta \), will be

\[
C = (1 - \lambda)[1 - F(\theta)]e^{-\theta r} + \lambda[1 - P(\theta)][ze^{-\theta r} + \int_{0}^{\theta} cm(\theta)e^{-\theta t} dt], \tag{5}
\]

where \( r \) represents the market rate of interest. The first term above relates to the possibility that the symptom, with probability \( 1 - \lambda \), does not indicate severe illness. In this case, given the probability \( 1 - F(\theta) \) that the symptom does not disappear until time \( \theta \), the individual will apply for diagnosis. The second term above relates to the possibility that the symptom, with probability \( \lambda \), does indicate severe illness. In this case, given the probability \( 1 - P(\theta) \), that the individual survives time \( \theta \), he will apply for diagnosis and undergo life-saving treatment. Equation (5) attaches the appropriate costs, \( z \) and \( cm(\theta) \), to the alternative cases.

Differentiating (5) with respect to \( \theta \), the first-order condition for the minimization of the expected money costs stream is given by

\[
\frac{dC}{d\theta} = -e^{-\theta r}(1 - \lambda)[1 - F(\theta)]r + (1 - \lambda)\dot{F}(\theta))z -
\]

\[
e^{-\theta r} \lambda \left\{ [1 - P(\theta)][rz + cm(\theta) - \frac{c\mu'(\theta)}{r}] + \dot{P}(\theta)[z + \frac{cm(\theta)}{r}] \right\} = 0, \tag{6}
\]
which, rearranging, yields

\[ MB^f = e^{-r\theta} \{(1-\lambda)[1-F(\theta)] + \lambda[1-P(\theta)]\}rz + e^{-r\theta}[(1-\lambda)\hat{F}(\theta) + \lambda \hat{P}(\theta)]z + e^{-r\theta} \lambda \{[1-P(\theta)]cm(\theta) + \hat{P}(\theta) \frac{cm(\theta)}{r} \} = e^{-r\theta} \lambda \{1-P(\theta)\} \frac{c\mu'(\theta)}{r} = MC^f. \]  

That is, the optimal time of applying for diagnosis is that level of \( \theta \) for which the marginal benefit to the firm \( (f) \) from delaying diagnosis, \( MB^f \), equals the marginal cost, \( MC^f \).

The marginal benefit to the firm of delaying diagnosis at time \( \theta \) stems first from the postponement of diagnosis costs for an additional unit of time, if, with the probability of \( (1-\lambda)[1-F(\theta)] + \lambda[1-P(\theta)] \), the symptom remains intact and the individual survives time \( \theta \), as well as from the saving in diagnosis costs if, with the probability of \( (1-\lambda)\hat{F}(\theta) + \lambda \hat{P}(\theta) \), the symptom disappears or the individual dies at exactly time \( \theta \). The present values of the costs saved in these alternative cases are \( e^{-r\theta}rz \) and \( e^{-r\theta}z \), respectively.

Secondly, given that the symptom indicates severe illness, delaying diagnosis for an additional unit of time saves the firm the costs of treatment required if the individual survives time \( \theta \) and were to apply for diagnosis at that time, as well as the cost of treatment from time \( \theta \) to eternity if the individual dies at exactly time \( \theta \). The present values of the costs saved in these alternative cases are \( e^{-r\theta}cm(\theta) \) and \( e^{-r\theta}cm(\theta)/r \), respectively. The marginal cost of delaying diagnosis at time \( \theta \) arises solely in view of the probability \( \lambda \) that the symptom indicates severe illness. In this case, given the probability \( 1-F(\theta) \) that the individual survives time \( \theta \), the firm will bear an increment to its future treatment costs, the present value of which is \( e^{-r\theta}c\mu'(\theta)/r \).

**Proposition 2:** (a) If the damage to health incurred by a slight delay in diagnosis is sufficiently low, delay is always desirable to the health insurance firm, irrespective of the probability that the symptom indicates severe illness. (b) If the damage to health incurred by a slight delay in diagnosis is sufficiently high, delay is desirable to the health insurance firm only if the probability that the symptom indicates severe illness is sufficiently low.
(c) Given that delay of diagnosis is desirable to the health insurance firm, optimal delay will be longer the lower the probability that the symptom indicates severe illness and the greater the unavoidable health damage inflicted by the illness.

Proposition 2 is proved in Appendix 2. Evaluating equation (7) at $\theta=0$, it is shown that there exist cutoff levels, $\hat{\mu}$ and $\hat{\lambda}$, such that if $\mu'(0) \leq \hat{\mu}$, delay is always desirable, and if $\mu'(0) > \hat{\mu}$, delay is desirable only if $\lambda < \hat{\lambda}$. In the latter case, $\hat{\lambda}$ is higher, and thus the desirability of delay is greater, the smaller is $\mu'(0)$ but the greater is the unavoidable damage, $g$. Given that delay is desirable, $d\hat{\theta}/d\lambda < 0$ and $d\hat{\theta}/dg > 0$, thus $\hat{\theta}$ decreases with an increase in $\lambda$ but increases with $g$. The former implications are summarized in Table 2, where D and N denote delay and no-delay decisions, respectively.

<table>
<thead>
<tr>
<th>Probability that the symptom indicates severe illness</th>
<th>LOW</th>
<th>HIGH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Damage to health incurred by a slight delay in diagnosis</td>
<td>LOW</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>HIGH</td>
<td>D</td>
</tr>
</tbody>
</table>

Table 2: Delay (D) / No-Delay (N) decisions implied by cost minimization

Notice the qualitative differences between the cost-oriented Proposition 2 and the utility-oriented Proposition 1. While the latter suggests that if the damage to health resulting from a slight delay in diagnosis is sufficiently high delay is never desirable, the former maintains that delay is desirable as long as the probability that the symptom indicates severe illness is sufficiently low (the lower the probability, the longer the optimal delay). While the former suggests that if the marginal damage to health is sufficiently low delay is always desirable, the latter maintains that delay is desirable only if the probability of severe illness is sufficiently high (the higher the probability, the longer the optimal delay).
Health insurance is claimed to generate two major incentive effects on individuals' behavior (broadly viewed as a "moral hazard"): it encourages excessive consumption of medical services whether actually ill or not, and it discourages sufficient care to prevent illness or to detect its presence at sufficiently early stages.\footnote{See, for instance, Stiglitz (1988, Ch. 11).} Hence one would intuitively expect individuals to excessively complain about their symptoms if the risk (that they are actually ill) is low, and to deficiently complain if the risk is high. The health insurer, in contrast, is intuitively expected to favor the inverse behavioral pattern. Comparing Tables 1 and 2 reveals that when the probability of illness is low, rational decision-making requires indeed that the individual avoids delay of diagnosis (irrespective of the self-induced damage) and that the health insurer favors delay. Hence, contrary to Proposition 1, delaying diagnosis if suffering a minor left-hand pain is desirable by Proposition 2, so as to avoid the cost of diagnosing a symptom which probably has no connection to a heart problem and is likely to disappear by itself. However, when the probability of illness is high, Tables 1 and 2 reveal that both parties should, counter-intuitively, advocate a similar behavioral pattern: delaying diagnosis if the self-induced damage is low and diagnosing promptly if the self-induced damage is high. Thus, delaying an Alzheimer's diagnosis is also desirable by Proposition 2: aside from avoiding the present cost of diagnosis, delay would save the health insurer the immediate costs of treatment, while hardly incurring additional future costs that could have been avoided by undergoing an earlier diagnosis. The greater the unavoidable health damage inflicted by this illness, the longer the delay in diagnosis advocated by both propositions. Notice, however, that the cutoff levels determining whether the probability of illness or the self-induced damage are sufficiently high or low are not likely to be the same for the individual and the insurer (see Appendix 1 and 2, respectively), since each party's problem involves different components. Similarly, the lengths of delay implied by both propositions in the case of high probability of illness and low self-induced damage cannot be compared.
IV. The Social Planner's Problem

Consider now the socially desired time of applying for diagnosis, which incorporates both utility and cost considerations. Suppose that the money costs stream expecting the health insurer upon the diagnosis of a suspicious symptom reflects real resource costs and define the expected present value of the net social gain, \( G \), of applying for diagnosis at time \( \theta \), as

\[
G = V - \gamma [C - \int_{0}^{\theta} \hat{P}(t)be^{-\delta t} dt],
\]  

(8)

where \( V \) and \( C \) are given by (1) and (5), respectively, \( b \) represents burial costs, assumed to be borne by society (in case that the individual dies before applying for diagnosis), and \( \gamma \) is a coefficient transforming resource costs into utility terms.\(^{16}\) The socially optimal time of applying for diagnosis, \( \bar{\theta} \), is that which maximizes the net social gain (8), the first-order condition for which is

\[
\frac{dG}{d\theta} = \frac{dV}{d\theta} - \gamma \left[ \frac{dC}{d\theta} + \lambda \hat{P}(\theta)be^{-\delta \theta} \right] = MB_o' - MC_e' - \gamma (MC_o' - MB_o') = 0,
\]  

(9)

where \( dV/d\theta \) and \( dC/d\theta \) are given by (3) and (6), respectively, \( MB_o' \) and \( MC_e' \) are defined in (4), \( MC_o' = MC_e' + \lambda \hat{P}(\theta)be^{-\delta \theta} \), and \( MB_o' \) and \( MC_o' \) are defined in (7). For simplicity, assume that the individual's subjective discount rate, \( \delta \), is identical to the health insurer's objective rate, \( r \).

Proposition 3: (a) When the cost of diagnosis is sufficiently low, delay in diagnosis is never desirable to society if the damage to health incurred by a slight delay is sufficiently high; if the damage incurred by a slight delay is sufficiently low, delay is desirable to society given that the probability that the symptom indicates severe illness is sufficiently high. (b) When

\(^{16}\) In consistence with the health insurer's problem, the analysis in this section is restricted to non-contagious diseases.
the cost of diagnosis is sufficiently high, delay in diagnosis is always desirable to society if the damage to health incurred by a slight delay is sufficiently low; if the damage incurred by a slight delay is sufficiently high, delay is desirable to society only if the probability that the symptom indicates severe illness is sufficiently low (c) Given that delay is desirable to society, optimal delay will be longer the greater the unavoidable health damage inflicted by the illness, whereas the effect on optimal delay of a higher probability of illness is ambiguous.

Proposition 3 is proved in Appendix 3. Evaluating equation (9) at \( \theta = 0 \), it is shown that there exist cutoff levels, \( \bar{\mu} = \beta \mu^* + (1-\beta) \bar{\mu} \) and \( \bar{\lambda} = \alpha \lambda^* + (1-\alpha) \bar{\lambda} \), such that when \( z \leq (\nu^* - \nu^0) \gamma \delta \), delay is never desirable if \( \mu'(0) \geq \bar{\mu} \), but desirable if \( \mu'(0) < \bar{\mu} \) and \( \lambda > \bar{\lambda} \). In contrast, when \( z > (\nu^* - \nu^0) \gamma \delta \), delay is always desirable if \( \mu'(0) \leq \bar{\mu} \), but otherwise desirable only if \( \lambda < \bar{\lambda} \). Whenever delay is desirable, \( d \bar{\theta} / d \theta > 0 \) whereas \( d \bar{\theta} / d \lambda \) is ambiguous.

The interesting question now is how the social optimum compares with the individual’s optimum. Evidently, the social planner’s problem involves resource costs which the individual ignores. However, from a real-cost perspective, the same as from a utility-based one, delaying diagnosis is both beneficial and costly to society. Thus, taking account of resource costs would not produce a clear-cut effect on optimal delay. Still, Proposition 3 suggests that if the cost of diagnosis is sufficiently low, the socially desired solution obtains a format similar to that of the individual (Table 1), whereas if the cost of diagnosis is sufficiently high, the socially desired solution obtains a format similar to that of the health insurer (Table 2). The crucial role that the cost of diagnosis plays in this matter is due to the fact that this cost must be borne irrespective of the diagnosis’ result, whereas treatment costs are borne only if the diagnosis yields a ‘positive’ result, and burial costs are borne only if the individual dies before applying for diagnosis. However, even in the former case, where the social and individual solutions seem to overlap, the cutoff levels for the probability of illness and the self-induced damage determining whether delay is desirable or
not are not likely to be the same for the individual and society.\textsuperscript{17} Hence, it is still possible that delay is desirable to society while being undesirable to the individual, and vice versa, and, if mutually desirable, that the optimal time of diagnosis is not identical for them both.

Given a discrepancy between the individual’s and the social optimum, the social planner may attempt to deal with it by requiring that the individual share the real costs of diagnosis and treatment with the health insurer. As a first approximation to the socially desired solution, the individual could be required to share the cost of diagnosis in case that the social optimum obtains a format similar to that of the health insurer’s (Table 2). Since this occurs when $z > (v^* - v^0)/\gamma \delta$, the individual’s share in the cost of diagnosis should be set a little above $(v^* - v^0)/\gamma \delta$ (which suffices to turn Table 1 into Table 2). More generally, evaluating the individual’s marginal cost and benefit of delaying diagnosis, $MC_0 \ i$ and $MB_0 \ i$, at the socially desired time of diagnosis, $\bar{\theta}$, the social planner is bound to find out whether or not $\bar{\theta}$ coincides with the individual’s desired time of diagnosis, $\theta^*$. If not, the social planner would discover either that $MB_0 \ i < MC_0 \ i$ at $\bar{\theta}$ (implying that $\theta^* < \bar{\theta}$), or that $MB_0 \ i > MC_0 \ i$ at that point of time (implying that $\theta^* > \bar{\theta}$). In the former case, the individual may be required to share the cost of diagnosis, so as to increase the benefit associated with delaying. In the latter case, the individual may be required to bear the costs of treatment which can be identified as emanating from delay in applying for diagnosis.

V. Conclusions

The paper has applied the economic tools of rational choice under uncertainty to the individual’s decision of whether or not to delay diagnosis of a suspicious symptom, confronting the fear of being told the suspected truth with fearing the consequences of procrastination. The analysis reveals that delay is desirable to the (fully insured) individual

\textsuperscript{17} It is easily seen that $\mu^*$ and $\lambda^*$ are respectively greater (less) than $\mu \hat{\mu}$ and $\lambda \hat{\lambda}$, if they are respectively greater (less) than $\mu$ and $\lambda$. 
only if the probability that the symptom indicates severe illness is sufficiently high and the damage to health incurred by a slight delay in diagnosis is sufficiently low; the higher the probability of illness and the greater the unavoidable (i.e., delay unrelated) damage inflicted by the illness - the longer the optimal delay in diagnosis. The results have been shown to explain a variety of observed behavior concerning individuals’ response to symptoms that vary with the likelihood of indicating severe illness and the severity of damage to health incurred by delayed diagnosis, although failing to explain observed delay when both the probability of illness and the self-induced health damage are high. Since a rational approach to analyzing the latter case implies, as intuitively expected, that diagnosis should be carried out promptly, practicing delay when suspecting cancer or myocardial infarction may better be explained, in line with health psychologists’ findings, by either ignorance of the meaning of symptoms, or by feelings of invulnerability and fatalism.

Examining the desirability of delaying diagnosis to the health insurer reveals that prompt diagnosis is desirable only if both the probability that the symptom indicates severe illness and the damage to health incurred by a slight delay in diagnosis are sufficiently high. Otherwise, delay is always desirable to the insurer; the lower the probability of illness and the greater the unavoidable damage inflicted by the illness - the longer the optimal delay in diagnosis. The cost of diagnosis has been found to play a crucial role in determining the social desirability of delaying diagnosis: when the cost of diagnosis is sufficiently low, the socially desired behavior reflects, in general, the individual’s preferences, but when the cost of diagnosis is sufficiently high, the socially desired behavior reflects those of the insurer. Given a discrepancy between the individual’s and the social optimum, cost-sharing measures may be imposed to induce the individual to approach the socially desired solution: when the individual’s desired delay is shorter than the socially desired one, he may be required to share the cost of diagnosis; when the socially desired delay is shorter, the individual may be required to share the increased costs of treatment emanating from delaying diagnosis.
APPENDIX

1. Proof of Proposition 1

Delaying diagnosis is desirable to the individual if at time 0 the marginal benefit of delaying exceeds the marginal cost. Substituting \( F(0) = P(0) = \hat{P}(0) = 0 \) into (4), this occurs if

\[
MB_0^i = \lambda \left[ v^0 - (v^p - sg) \right] > (1 - \lambda)(v^N - v^0) + \lambda \frac{s\mu'(0)}{\delta} = MC_0^i,
\]

or, rearranging, if

\[
\lambda \left[ \delta(v^N - (v^p - sg)) - s\mu'(0) \right] > \delta(v^N - v^0).
\]

(a')

Given the probability \( \lambda \) that the symptom indicates severe illness, a slight delay in diagnosis increases health damage by \( \mu'(0) \), which can be avoided by applying immediately upon the discovery of the symptom. Focusing on the magnitude of \( \mu'(0) \), it is evident that if \( \mu'(0) \geq \delta(v^N - (v^p - sg))/s \), condition (a') cannot be satisfied, since its L.H.S. will be zero/negative. Furthermore, if \( \mu'(0) \) is lower, condition (a') will be satisfied if

\[
\lambda > \frac{\delta(v^N - v^0)}{\delta(v^N - (v^p - sg)) - s\mu'(0)} = \lambda^*.
\]

(b)

However, as long as

\[
\mu'(0) \geq \frac{\delta[v^0 - (v^p - sg)]}{s} = \mu^*.
\]

(c)

\( \lambda^* \) equals/exceeds unity, thus condition (b) cannot be satisfied, irrespective of the value of \( \lambda \). Hence, delaying diagnosis will not be desirable. This proves Proposition 1(a). For \( \mu'(0) < \mu^* \), condition (b) will be satisfied only if \( \lambda > \lambda^* \), which proves Proposition 1(b). The optimal time of diagnosis, \( \theta^* \), will be determined by equation (3), the second term of which (that multiplied by \( \lambda \)) must be positive at the optimum (since the first term is negative). Hence,
\[
\frac{d\theta^*}{d\lambda} = -\frac{1}{\Delta} \frac{d^2V}{d\theta d\lambda} > 0,
\]

and, since \(dV/d\theta\) is positively related to \(m(\theta)\), which is a positive function of \(g\),

\[
\frac{d\theta^*}{dg} = -\frac{1}{\Delta} \frac{d^2V}{d\theta dg} > 0,
\]

where \(\Delta = d^2V/d\theta^2 < 0\) denotes the second-order condition for the maximization of expected utility, assumed to be satisfied at \(\theta^*\). It thus follows that the higher the probability that the symptom indicates severe illness and the greater the unavoidable health damage inflicted by the illness, the longer the optimal delay in applying for diagnosis, which proves Proposition 1(c).

2. Proof of Proposition 2

Delaying diagnosis is desirable to the health insurance firm if

\[
MB_0 = rz + \lambda cg > \lambda \frac{c\mu'(0)}{r} = MC_0,
\]

or, rearranging, if

\[
\lambda \left( c\mu'(0) - rg \right) < r^2z.
\]

Given that \(\mu'(0) \leq rg\), condition (f) is clearly satisfied, since its L.H.S. will be zero/negative. Furthermore, given that \(\mu'(0)\) is higher, condition (f') will be satisfied if

\[
\lambda < \frac{r^2z}{c\mu'(0) - rg} = \hat{\lambda}.
\]

However, as long as
\[ \mu'(0) \leq \frac{r(cg + rz)}{c} = \hat{\mu}, \]  

(h) \[ \lambda \text{ equals/exceeds unity, thus (g) is satisfied, irrespective of the value of } \lambda. \text{ Hence, delaying diagnosis will be desirable. This proves Proposition 2(a). For } \mu'(0) > \hat{\mu}, \text{ condition (g) will be satisfied only if } \lambda < \hat{\lambda}, \text{ which proves Proposition 2(b). The optimal time of diagnosis, } \hat{\theta}, \text{ is determined by equation (6), the second term of which (that multiplied by } \lambda) \text{ must be positive at the optimum (since the first term is negative). Hence,} \]

\[ \frac{d\hat{\theta}}{d\lambda} = -\frac{1}{\lambda} \frac{d^2C}{d\theta d\lambda} < 0, \]  

(k) \[ \text{and, since } \frac{dC}{d\theta} \text{ is negatively related to } m(\theta), \text{ which is a positive function of } g, \]

\[ \frac{d\hat{\theta}}{dg} = -\frac{1}{\lambda} \frac{d^2C}{d\theta dg} > 0, \]  

(l) \[ \text{where } \Lambda = dC/d\theta^2 > 0 \text{ denotes the second-order condition for the minimization of expected costs, assumed to be satisfied at } \hat{\theta}. \text{ It thus follows that the lower the probability that the symptom indicates severe illness and the greater the unavoidable health damage inflicted by the illness, the longer the optimal delay in applying for diagnosis, which proves Proposition 2(c).} \]

3. Proof of Proposition 3

Delaying diagnosis is desirable to society if

\[ MB_o^i + \gamma MB_o^f = \lambda[v^o - (v^p - sg)] + \gamma(\delta z + \lambda cg) > \]

\[ (1 - \lambda)(v^N - v^0) + \lambda \frac{\delta \mu'(0)}{\delta} + \gamma \lambda \frac{c\mu'(0)}{\delta} = MC_o^i + \gamma MC_o^i. \]  

(m) \[ \text{or, rearranging, if} \]


\[ \lambda \{ \delta[v^N - v^0 + (s + \gamma c)g] - (s + \gamma c)\mu'(0) \} > \delta(v^N - v^0 - \gamma \delta z) \quad \text{(m')} \]

Suppose first that \( z \leq (v^N - v^0)/\gamma \delta \), so that the R.H.S. of (m') is zero/positive. Given that \( \mu'(0) \geq \delta[v^N - v^0 + (s + \gamma c)g] / (s + \gamma c) \), condition (m') cannot be satisfied, since its L.H.S. will be zero/negative. Furthermore, given that \( \mu'(0) \) is lower, condition (m') will be satisfied if

\[ \lambda > \frac{\delta(v^N - v^0 - \gamma \delta z)}{\delta[v^N - v^0 + (s + \gamma c)g] - (s + \gamma c)\mu'(0)} = \alpha \lambda^* + (1 - \alpha) \hat{\lambda} = \bar{\lambda}, \quad \text{(n)} \]

where \( \lambda^* \) and \( \hat{\lambda} \) are defined in (b) and (g), respectively, and

\[ \alpha = \frac{\delta[v^N - (v^p - sg)] - s\mu'(0)}{\delta[v^N - v^0 + (s + \gamma c)g] - (s + \gamma c)\mu'(0)}, \quad \text{(p)} \]

\[ 1 - \alpha = \frac{-\gamma c[\mu'(0) - \delta g]}{\delta[v^N - v^0 + (s + \gamma c)g] - (s + \gamma c)\mu'(0)}. \quad \text{(q)} \]

However, as long as

\[ \mu'(0) \geq \frac{\delta[v^0 - v^p + (s + \gamma c)g] + \gamma \delta^2 z}{s + \gamma c} = \beta \mu^* + (1 - \beta) \hat{\mu} = \bar{\mu}, \quad \text{(r)} \]

where \( \mu^* \) and \( \hat{\mu} \) are defined in (c) and (h), respectively, and

\[ \beta = \frac{s}{s + \gamma c}, \quad \text{(s)} \]

\[ 1 - \beta = \frac{\gamma c}{s + \gamma c}, \quad \text{(t)} \]

\( \bar{\lambda} \) equals/exceeds unity, thus condition (n) cannot be satisfied, irrespective of the value of \( \lambda \).

Hence, delaying diagnosis will not be desirable. This proves the first part of Proposition 3(a). For
\( \mu'(0) < \bar{\mu}, \bar{\lambda} \) will be lower than unity. Delaying diagnosis will be desirable if \( \lambda > \bar{\lambda} \), which proves the second part of Proposition 3(a).

Suppose now that \( \epsilon > (\nu' - \nu')/\gamma \delta \), so that the R.H.S. of (m') is negative. Given that \( \mu'(0) \leq \delta[\nu' - \nu' + (s + \gamma c)g] / (s + \gamma c) \), condition (m') is clearly satisfied, since its L.H.S. will be zero/positive. Furthermore, given that \( \mu'(0) \) is higher, condition (m') will be satisfied if \( \lambda < \bar{\lambda} \).

However, as long as \( \mu'(0) \leq \bar{\mu}, \bar{\lambda} \) equals/exceeds unity, thus delaying diagnosis will still be desirable, irrespective of the value of \( \lambda \). This proves the first part of Proposition 3(b). For \( \mu'(0) > \bar{\mu}, \bar{\lambda} \) will be lower than unity. Delaying diagnosis will be desirable only if \( \lambda < \bar{\lambda} \), which proves the second part of Proposition 3(b). Whenever delay is desirable, the optimal time of diagnosis, \( \bar{\theta} \), will be determined by equation (9), which differentiating with respect to \( \lambda \) and \( g \), yields

\[
\frac{d\bar{\theta}}{d\lambda} = -\frac{1}{\Omega} \left[ \frac{d^2V}{d\theta d\lambda} - \frac{d^2C}{d\theta d\lambda} \right], \tag{w}
\]

\[
\frac{d\bar{\theta}}{dg} = -\frac{1}{\Omega} \left[ \frac{d^2V}{d\theta dg} - \frac{d^2C}{d\theta dg} \right], \tag{x}
\]

where \( \Omega = d^2 G / d\theta^2 < 0 \) denotes the second-order condition for the maximization of net social gain, assumed to be satisfied at \( \bar{\theta} \). Since both \( d^2V / d\theta d\lambda \) and \( d^2C / d\theta d\lambda \) are positive, the sign of (w) is unclear. However, since \( d^2V / d\theta dg \) is positive and \( d^2C / d\theta dg \) is negative, the sign of (x) is positive. It thus follows that the greater the unavoidable health damage inflicted by the illness, the longer the optimal delay in applying for diagnosis, whereas the effect on optimal delay of a greater probability of illness is ambiguous, which proves Proposition 2(c).
REFERENCES


Safer, M. A., Jackson. Q. J. and H. Leventhal, 1979, Determinants of three stages of delay in seeking care at a medical care clinic, Medical Care 17, 11-29.


ניתן להזמין פרוшומם במועד ליווי להזמנה, מנהל התחבר וה hakkı
שד, יremium 13, ירושלים 9109, טל. 02(670)9579