Risk Adjustment and Prevention

Karen Eggleston, Randall P. Ellis, and Mingshan Lu *

January 4, 2011

Abstract

Widespread integration of market-based incentives into healthcare systems calls for – and has elicited – increasing adoption of risk adjustment. By deterring selection, risk adjustment helps to assure fair and efficient payments among health insurers or capitated provider groups. However, since conventional risk adjustment allocates funds among regions or insurers according to current population health status, it does not reward – indeed, it penalizes – preventive efforts that improve population health. This prevention penalty of risk adjustment represents a hidden cost of unclear magnitude, undermining provider incentives for health promotion. We develop a theoretical model of selection and prevention demonstrating this problem with conventional risk adjustment and suggesting a simple alternative: risk adjustment should be linked to pay-for-performance for prevention.

JEL classification: I1

Keywords: prevention; health promotion; risk adjustment; pay-for-performance; payment incentives

*Eggleston, Stanford University, karene@stanford.edu; Ellis, Boston University, ellisrp@bu.edu; Lu, University of Calgary, lu@ucalgary.ca. Correspondence: Mingshan Lu, Department of Economics, University of Calgary; Ph: (403) 220-5488; Fax: (403) 282-5262; E-Mail: lu@ucalgary.ca; 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4.
1 Introduction

Appropriate allocation of health care resources has been a priority issue for Canadian health decision-makers so as to ensure equitable, efficient resource allocation and eliminate disparities. The most fundamental question in health care resource allocation is how to measure population health needs. As Eyles and Birch (1993) note, the need for health care is not restricted to curing disease; it should be interpreted in terms of ability to benefit from health care, including “reducing the risks of deterioration in health status or improving the probability of improvements to health status.” The Romanow report also emphasized that “for too long, Canada’s health care system has been overly focused on treatment rather than prevention. A central focus of primary health care must be on preventing illness and injury and helping Canadians stay healthy” (Romanow report, November 2002).

Currently, Alberta Health and Wellness (AHW) uses a population-based funding formula to allocate annual operating budgets to health regions, as do many other Canadian provinces. The AHW formula uses the age and gender distribution and a socioeconomic factor that allocates higher per capita funding to regions that have lower socio-economic status (SES). A primary concern is that the current SES adjuster is only a weak proxy for health care utilization, not health needs. Other jurisdictions in Canada and the US have considered and applied various methodologies to adjust resource allocation based on case-mix (i.e., level of health need). AHW has therefore been evaluating the feasibility of using existing risk adjustment models to improve its funding formula. However, all existing risk adjustment models allocate funds according to current population health status. Such methods fail to account for and reward any improvements in population health and therefore fail to provide adequate incentives for efficient operations of health regions and providers.

We study how to pay for population health improvements when payments to healthcare providers are risk adjusted to compensate for underlying differences in population risk. Risk adjustment is critical to deter selection and to assure fair and efficient payments across differing population groups, including competing insurers or capitated provider groups (van de Ven and Ellis 2000). As healthcare systems around the world experiment with market-based incentives and competition (Cutler 2002), risk adjustment becomes more important. However, existing risk adjustment models are not designed to encourage health promotion and prevention.

Although no theoretical models have yet addressed this dynamic incentive problem associated
with risk adjustment, some researchers and policymakers have noted the dilemma. For example, in discussing challenges for risk adjustment in the Netherlands, van de Ven, van Vliet, and Lamers (2004) observe that “if (in the future) sickness funds pay providers for their performance – for example, measured by the change in health status of their patients over time – there might be an incentive problem. The better the providers perform in terms of improving health status, the more a sickness fund pays to providers but the lower the next year’s premium subsidies that the sickness fund receives” (p.53).

This paper presents a simple two-period, two-type model of provider risk selection and prevention effort to analyze the prevention disincentives of risk adjustment. Our stylized model is designed to illustrate the incentives of a given region, insurer, or provider, and does not attempt to model overall general equilibrium effects. We conceive of prevention broadly as any health services that slow the pace of health deterioration associated with aging and the natural course of chronic diseases. Thus our model applies to secondary and tertiary forms of prevention as well as primary prevention activities.1 We illustrate how conventional risk adjustment discourages both selection and prevention.

By focusing on current population health, conventional risk adjustment resembles static optimization. In contrast, combining conventional risk adjustment with pay-for-performance on prevention helps payers to distinguish between provider activities that make existing patients healthier (prevention) and provider activities that differentially attract healthier enrollees (risk selection). Refining payment systems along these lines could help to deter selection while rewarding prevention. This approach is akin to dynamic optimization, encouraging “discovery” of health promotion strategies that are not ex ante contractible. For example, encouraging flexible team response to disease progression – such as allowing a pharmacist to adjust medications – may improve metabolic control for patients with diabetes (Shojania et al. 2006). Alternatively, physician-led diabetes management in primary care also may improve health outcomes while lowering spending (Stock et al. 2010). To encourage adoption of effective disease management “technologies,” payers need to align payment incentives with prevention and quality improvement goals. For example, when Germany implemented nation-wide disease management programs in 2002, policymakers allowed for extra incentive payments from the Risk Compensation Scheme to sickness funds for setting up special care programs for the chronically ill. In 2009, Germany moved instead toward a “comprehensive

---

1 See Kenkel (2000).
‘morbidity-oriented’ Risk Compensation Scheme, plus an additional administrative fee of 180 euros, or approximately US$262, per year for each disease management participant” (Stock et al. 2010, p.2199). This combination of risk adjustment with extra incentives for prevention activities well illustrates the policy implications of our model, although it remains unclear if the most recent payment change strengthened or reduced incentives for sickness funds to promote noncontractible forms of health promotion in an optimal way.

In the next section we begin developing our model of provider investment in risk selection and prevention. We show that while risk adjustment is necessary to avoid rewarding cream skimming and dumping, it also introduces an incentive problem, discouraging prevention. In our central proposition we suggest an alternative payment paradigm that combines conventional risk adjustment with a bonus payment for better-than-“expected” prevention performance. The final section discusses related literature, empirical estimation, and theoretical extensions left for future research.

2 A Simple Model of Selection and Prevention

Consider a simple two-period, two-risk type model. In period 0, fraction $\lambda$, $0 < \lambda \leq 1$, of the population are L-type patients, and fraction $1 - \lambda$ are H-types. L-types have lower expected costs ($c_L$) than H-types do ($c_H$): $c_H > c_L$. The two risk categories can capture a broad range of patient heterogeneity – healthy vs. having a chronic illness, or chronically ill without complications vs. chronically ill with complications. The fraction of the total population who remain low-risk in period 1 depends on the incentives for providers to invest in prevention.

2.1 Selection and prevention

We focus on the selection and prevention choices of a single insurer-provider organization that serves a small fraction of the large population risk pool with average risk characterized by $\lambda$. The insurer’s total enrollment is assumed to be stable and normalized to 1. In period 0, a fraction $\lambda_0$ of the insurers’ patients are the relatively healthy L-types. For example, the risk mix in a given province may vary from the Canadian national average risk of $\lambda$ because of multiple factors, including differences in fertility rates, migration, and economic conditions. Provinces with higher-than-average fertility will tend to have younger populations and a higher-than-average proportion

\[2\text{ We refer to the single decision-maker on the supply side as the insurer or provider. Equally appropriate terms would be “regional health authority,” “sickness fund,” or “managed care organization.”}\]
of low risks, $\lambda_0 > \lambda$. Within and between provinces, patients may move for higher education, employment, marriage, closer proximity to aging parents, or other reasons, leading provider A to have a more favorable risk mix, $\lambda^A_0$, than that of provider B (i.e., $\lambda^A_0 > \lambda > \lambda^B_0$).

In addition to natural variations in risk because of demographics and patient locational decisions, the insurer can invest in measures to attract low-cost patients. Examples of selection effort include locating in a healthier community, selectively advertising, or (in a classic illustration given by Newhouse 1996), stinting on oncologists and employing numerous pediatricians, to avoid expensive cancer patients and attract young families who are better risks. We assume that such selection effort, $e$, improves the risk-mix of patients, with diminishing returns: $\lambda'_0 (e) > 0$, $\lambda''_0 (e) < 0$.

Imagine an insurer-provider that invests more than the average amount of effort in selection. It would then have a higher than average value for $\lambda_0$, such that $\lambda_0 > \lambda$.\(^3\) We are not saying that in equilibrium all insurer-providers could have a higher fraction of low risks than exist in the population, or that all variations in risk mix arise from nefarious provider selection efforts. Rather, we are focusing on the incentives that could arise if some providers for exogenous reasons (e.g., say those with fewer ethical constraints) undertake more risk selection than other providers in the region or country do. The model applies to the fraction of providers who would undertake selection efforts given sufficient financial incentive to do so.

For simplicity we conceptualize risk selection as a pure social welfare loss, since creaming and dumping efforts merely redistribute risk among insurer-providers. Thus the socially optimal level of selection is zero.\(^4\)

In contrast, prevention spending, $m$, has a true social value: it reduces morbidity loss and saves treatment cost. Define $x(m)$ to be the probability that low risks remain healthy (L-types) in the second period. This probability is increasing and strictly concave in prevention: $x'(m) > 0$; $x''(m) < 0$.

Prevention can be primary, secondary, or tertiary. For example, a provider can encourage obese

---

\(^3\)In analysis of how premium inertia and employee turnover discourage small firms from offering health insurance, Ellis and Ma (2010) make a similar assumption that new employees are healthier than average. Here, we are assuming that the provider may actively seek to attract younger, healthier patients and/or refer sicker and older patients to alternative providers.

\(^4\)We assume that demographic factors such as fertility rates and patient locational decisions are independent of provider prevention efforts. An interesting case to consider would be when some prevention efforts differentially attract low risks and thus constitute selection. Although not explicitly modeled here, this case would be consistent with our model, which (as we explain further below) allows providers to use prevention to “buy some selection.” Indeed, the likelihood of overlap between prevention and selection efforts reinforces our argument for linking conventional risk adjustment with payment for prevention.
patients to exercise and lose weight to decrease the probability of developing heart disease, diabetes and other related conditions. A provider can also decrease the likelihood of a diabetic person developing complications by regularly monitoring hemoglobin A1c and blood lipid levels, ordering renal screens, and so on. Such prevention activities cost the provider \( m \). That is, our measure of preventive actions is the provider’s spending on health promotion and screening activities, which exhibits diminishing returns.

The simplest way to reward prevention is to pay the provider FFS for prevention activities, \( m \). However, this solution does not work when \( m \) is noncontractible. It may be difficult to distinguish treatment from prevention (such as medications for a patient with diabetes both treat the current disease and prevent future complications). Noncontractibility also arises if preventive services are identifiable, but the conditions under which the proposed prevention activities will prove cost effective are not well defined. For example, a diagnostic full body MRI may be worthwhile for some patients, but once covered, will tend to be over used. Therefore we focus on observed outputs – improvements in patient health – rather than reimbursement for specific inputs. Our analysis applies to an important subset of prevention: prevention activities that are \textit{ex ante} noncontractible and thus cannot be rewarded through straightforward FFS payment.

Figure 1a illustrates the evolution of population risk mix over the two periods of the model, with prevention \( m \) increasing the fraction \( x(m) \) of L-types who remain L-types in period 1. Adding selection would make the fraction of L-types dependent on selection effort, \( e \).

For concreteness, think of selection in this model as an up-front investment in the characteristics of the provider-insurer. In addition to the aforementioned example of hiring many pediatricians but few oncologists, selection could mean investing in excellent acute care, but avoiding a reputation for good treatment of depression or HIV/AIDS. As a result of these investments, selection \( e \) attracts the same mix of L and H in both periods. The fraction of the provider’s patients who are low risk in period 1 is given by

\[
\lambda_1 (e, m) = \lambda_0 (e) x (m). \tag{1}
\]

Prevention and selection both improve period 1 risk-mix \( \frac{\partial \lambda_1}{\partial m} = \lambda_0 (e) x' (m) > 0; \frac{\partial \lambda_1}{\partial e} = \lambda' (e) x (m) > 0 \).

In sum, the overall timing of the model is as follows. Before the game starts, the payer sets the terms of payment. Then the provider chooses selection and prevention, risk-mix evolves accordingly,
Figure 1a. The population risk mix, with prevention (m>0)

λ
L
1-λ
H
x(m)
1-x(m)
L
H
H

Figure 1b. A provider’s risk mix, with prevention (m>0), selection (e>0), and turnover (τ>0), assuming other providers choose m=0

λ’<sub>0</sub>(e) >0
λ’<sub>1</sub>(m) >0

λ<sub>0</sub>(e) [ (1-τ) x(m) + τ x(0) ] = λ<sub>1</sub>(e, m, τ)
and the provider receives payments, as follows:

- **Period 0**
  - Provider choice of selection determines initial risk-mix $\lambda_0(e)$;
  - Provider chooses prevention $m$;
  - Provider incurs treatment costs $(c_L, c_H)$ and receives payment;

- **Period 1**
  - Provider prevention and selection results in risk mix $\lambda_1(e, m)$;
  - Provider incurs costs $(c_L, c_H)$ and receives payment, possibly with bonus $(\alpha)$.

### 2.2 Optimal prevention

The socially optimal level of prevention would maximize population benefits $B$ over the two periods, net of the resource cost of producing those health benefits. Low risks enjoy both higher quality of life ($B_L > B_H$) and lower treatment costs ($c_L < c_H$), so that net benefits are higher for low risks: $[B_L - c_L] > [B_H - c_H]$. Recall that the population fraction of low risks is $\lambda$. (The remaining patients, $1 - \lambda$, are already high risks in period 0 and therefore would not benefit from prevention.)

Social net benefits in period 1 resulting from period 0 prevention are given by

$$B(m) = \lambda x(m) \{[B_L - c_L] - [B_H - c_H]\}. \tag{2}$$

Our assumptions guarantee that social benefits are increasing and strictly concave in prevention expenditures: $B'(m) > 0$, $B''(m) < 0$.

Thus a social planner should choose $m$ and $e$ to maximize\(^5\)

$$W = B(m) - m - e. \tag{3}$$

Let $m^*$ and $e^*$ denote the unique optimal levels of selection and prevention, respectively.

Clearly risk selection is counter-productive: selection effort, $e$, does not yield any social benefit, because it merely redistributes risk among insurer-providers. One provider’s favorable selection is other providers’ adverse selection. At the social optimum, therefore, providers should not choose to invest in risk selection: $e^* = 0$.\(^6\)

---

\(^5\)If the planner wants to include weight on producer surplus, the uniform profit margin $\pi$ can be increased above the minimum level needed to assure provider participation.

\(^6\)We could also include choice of treatment. Treatment should be given up to the point where the marginal treatment benefit equals its marginal cost, regardless of the level of prevention.
The first-order condition defining optimal prevention is

$$
\lambda ([B_L - c_L] - [B_H - c_H]) x'(m^*) = 1.
$$

Equation (4)

The marginal benefit of prevention includes better health (i.e., more quality-adjusted life years per low risk, $B_L - B_H$) as well as saving resources on future treatment costs ($c_H - c_L$) for each low risk. Providers should invest in prevention up to the point where the marginal benefit of avoided future morbidity and treatment cost equals the marginal cost of prevention, 1. Re-writing (4), we see that at the social optimum each provider should serve a patient pool reflecting the population average risk, $\lambda$, and undertake prevention effort per patient according to

$$
x'(m^*) = \frac{1}{\lambda ([B_L - c_L] - [B_H - c_H])}.
$$

Equation (5)

The socially optimal level of prevention is increasing in the fraction of patients who can benefit and the difference in net benefits enjoyed by the healthy relative to the sick. Our model focuses on how to align provider payment incentives with this goal.

### 2.3 Patient turnover: A tax on prevention investment

In this section we allow for turnover of a fraction $\tau$ of patients in period 1. Figure 1b illustrates how the provider’s risk mix in period 1 depends on three factors: selection in period 0, the productivity of prevention, and the rate of turnover. Assume that a fraction $\tau$ of patients leave after period 0, with $0 \leq \tau < 1$ (i.e., turnover is less than 100%). Since we are modeling a ‘steady state,’ each person who leaves is replaced by a new patient who was treated by a different provider in period 0.

The level of prevention among other providers is assumed to be a baseline level, normalized to zero. The effect of turnover occurs through differences in prevention between the modeled provider (who chooses $m$) and other providers (who choose baseline prevention). Turnover means the modeled provider looses some L types who are quite likely to remain L types, and in their place enrolls L types who received a lower level of prevention in period 0 and therefore are more likely to become H-types in period 1. Ellis and Ma (2010) make a similar assumption that new employees are healthier than average in their analysis of how employee turnover and premium inertia motivates small firms to not offer health insurance. Here we model how turnover reduces the incentive to invest in prevention, a similar weakening in an activity that improves patient health.
With these assumptions, the fraction of the provider’s patients who are low risk in period 1 changes from (1) to include both the L who stay and remain L, \( x(m) (1 - \tau) \), and those new L who remain L, \( x(0) \tau \):

\[
\lambda_1(e, m, \tau) = \lambda_0(e) [x(m) (1 - \tau) + x(0) \tau]. \tag{6}
\]

Prevention and selection still both improve period 1 risk-mix \( \frac{\partial \lambda_1}{\partial m} = \lambda_0(e) x'(m) (1 - \tau) > 0 \); \( \frac{\partial \lambda_1}{\partial \epsilon} = \lambda' (e) [x(m) (1 - \tau) + x(0) \tau] > 0 \), but turnover lessens their effectiveness \( \frac{\partial \lambda_1}{\partial \mu \partial \tau} = -\lambda x' < 0 \); \( \frac{\partial \lambda_1}{\partial \mu \partial \tau} = \lambda' [x(0) - x(m)] \leq 0 \). If the provider invests in above-average prevention \( (m > 0) \), then turnover worsens the risk mix: \( \frac{\partial \lambda_1}{\partial \tau} = \lambda_0(e) [x(0) - x(m)] \leq 0 \), with equality if and only if \( m = 0 \). In other words, patient turnover acts as a tax on providers who invest in (above-average) prevention.

Thus, in the modeled ‘steady state,’ turnover represents “leakage” of prevention investments, the rewards of which accrue to other providers in period 1. This externality from turnover undermines incentives for prevention. An example would be US Medicare benefiting from chronic disease management for the non-elderly. Knowing that Medicare will inherit the risks of the elderly, private provider-insurers have less incentive to invest in disease management that will prevent complications after enrollees turn 65 than the providers would if they themselves continued to bear the risks of the elderly. Health care systems that integrate rather than separate the elderly and non-elderly risk pool(s) have an advantage in avoiding this turnover leakage and the disincentive for preventive efforts that it implies.

The combination of turnover and risk adjustment may doubly reduce providers’ financial returns from investments in better future health of their patients. Any payment system designed to mitigate the prevention-deterring effect of risk adjustment, therefore, also needs to take account of the prevention-deterring effect of patient turnover. We return to this issue in the concluding discussion regarding empirical implementation of pay-for-performance for prevention combined with risk adjustment.

### 2.4 Provider objectives

We follow the conventional assumption that the provider seeks to maximize expected net revenue, \( V(e, m) \).\(^7\) Assume net revenues are \( \pi_L \) and \( \pi_H \) for low- and high-risk patients, respectively. Let

\(^7\)Our working paper included a model of altruistic providers who seek to maximize patient benefits subject to breaking even. The qualitative conclusions remain the same: conventional risk adjustment embodies a financial
\( \pi_{LH} \) define the difference in net revenue between low and high risks: \( \pi_{LH} \equiv \pi_L - \pi_H \). With uniform capitation payment and higher expected costs of high risks, there would be a differential in net revenue (\( \pi_{LH} > 0 \)) that gives the provider the financial incentive to serve low risk patients.

The provider-insurer chooses to invest in selection and prevention according to the following program:

\[
\begin{align*}
Max_{<e,m>} V(e, m) &= \lambda_0(e) \pi_L + (1 - \lambda_0(e)) \pi_H - e - m + \\
&\quad \lambda_1(e, m, \tau) \pi_L + (1 - \lambda_1(e, m, \tau)) \pi_H.
\end{align*}
\] (7)

The first line represents net revenue per patient in period 0: \( \pi_L \) per low risk and \( \pi_H \) per high risk, less spending on selection and prevention \( (e + m) \). The second line represents the resulting net revenues in period 2. We abstract from discounting.

### 2.5 Perfect risk adjustment removes incentives for prevention

We model risk adjustment as a technology that enables the payer to differentiate prospective payments based on risk type. This technology can compensate providers for natural variations in patient mix as well be used to discourage providers from risk selection behavior. Accurate risk adjustment is consistent with any given level of net revenue per patient, \( \pi = \pi_L = \pi_H > 0 \), set to fulfill the provider’s participation constraint. Risk adjustment merely removes the differential profitability of low risks compared to high risks.

We start by assuming the simplest case, that of perfect risk adjustment: \( \pi_{LH} = 0 \). A later section will explore the implications of imperfect risk adjustment (\( \pi_{LH} > 0 \)), and how mixed forms of payment can compensate for inaccuracies of risk adjustment while preserving incentives for prevention.

When risk adjustment technology is so accurate as to match expected costs of each risk type precisely, implementing conventional risk adjustment entirely removes incentives for both selection and prevention. To see this, note that when \( \pi_L = \pi_H = \pi \), the provider’s maximization problem is

\[
Max_{<e,m>} V(e, m) = \pi - e - m + \pi.
\] (8)

Neither selection nor prevention increase net revenue. In this case, the provider would choose not penalty on prevention, and can induce even a fully benevolent provider to under-invest in prevention.
to invest in selection \((e = e^* = 0)\), as hoped. However, in the absence of a reward for prevention, the provider would also choose not to invest in prevention \((m = 0 < m^*)\). Thus, risk adjustment removes incentive for prevention.

### 2.6 Pay-for-prevention

As we have noted above, risk adjustment has the unintended side effect of financially penalizing providers who invest in health promotion. This prevention penalty of risk adjustment will only become more and more salient as three inter-related trends converge – aging societies, chronic disease epidemics, and wider adoption of conventional risk adjustment.

An approach to overcome this incentive problem is to link implementation of risk adjustment with introduction of pay-for-performance on successful prevention. Assume the provider-insurer receives a bonus payment for effective prevention activities that are \(ex~ante\) noncontractible. We model this approach with a per-patient bonus of \(\alpha [x(m) - x(0)]\) for above-average prevention \((x(m) > x(0))\). The payer may decide whether or not to levy penalties for below-average prevention. Note that the prevention bonus is based on the difference between risk progression of the provider’s period 0 enrollees, \(x(m)\), and that of new enrollees, \(x(0)\). This means that the payer should base the pay-for-performance on the health improvement or deterioration of continuous enrollees (across two periods), netting out the effect of turnover.\(^8\)

Adding pay-for-prevention to perfect risk adjustment, the provider’s maximization problem becomes

\[
\max_{\langle e,m \rangle} V(e,m) = \pi - e - m + \pi + \alpha [x(m) - x(0)].
\]

The provider will invest in prevention according to the magnitude of the prevention payment \(\alpha\): \(\alpha x'(m) = 1\). If the payer sets the bonus \(\alpha\) so that it equals the social marginal benefit of prevention, then optimal prevention results (i.e., setting \(\alpha = \lambda ([B_L - c_L] - [B_H - c_H])\) implements the first best).\(^9\)

---

\(^8\)As long as turnover is not too great, estimating the risk mix based on those continuously enrolled should be feasible. Note that empirical implementation would have to take account of non-random turnover, since low risks are more likely to switch providers than high risks are. Adding this wrinkle to the model would allow separate analysis of adverse selection from movers compared to what Altman, Cutler and Zeckhauser (1998) call adverse retention.

\(^9\)The optimal alpha is increasing in turnover, since reaching the first-best level of prevention (given in equation 5) requires compensating for the impact of turnover. Exploring the implications of nonlinear payment schemes for prevention would also be useful; we thank a referee for this point.
3 Imperfect Risk Adjustment and Mixed Provider Payment

The more accurate risk adjustment technology is, the more closely payments can match expected costs, reducing the differential net revenue \( \pi_{LH} \) toward zero. However, all risk adjustment technology remains imperfect. Let \( \beta \) reflect the accuracy of risk adjustment, \( 0 \leq \beta \leq 1 \), with \( \beta = 0 \) representing no risk adjustment (implying maximum \( \pi_{LH} = c_H - c_L \)), and \( \beta = 1 \) representing perfect risk adjustment as assumed above (implying \( \pi_{LH} = 0 \)). The payer would like to implement perfect risk adjustment, but is constrained by the accuracy of currently available risk adjustment technology, as represented by the parameter \( \beta \).

However, policymakers are not confined to using capitation payment. Indeed, when risk adjustment is imperfect, purchasers can increase social welfare by using mixed forms of provider payment, as we illustrate in this section.

3.1 Supply-side cost sharing

Incentives for selection and prevention depend on the basis of payment, i.e., how much is bundled in one prospective payment or unbundled as retrospective reimbursement for claims. Let total payment include a prospective payment, \( R_i \), and reimbursement of fraction \( (1 - s) \) of costs, so that the per-patient total payment \( P_i \) is given by

\[
P_i = R_i(s, \beta) + (1 - s) c_i, \quad i \in \{L, H\};
\]

where

\[
R_i(s, \beta) = \pi + \beta sc_i + (1 - \beta) sc_{avg},
\]

\[
c_{avg} = \lambda c_L + (1 - \lambda) c_H. \tag{12}
\]

The parameter \( s \) represents the degree of supply-side cost sharing, with \( 0 \leq s \leq 1 \). Full supply-side cost sharing, \( s = 1 \), represents capitation payment: \( P_i = R_i \) and no costs are reimbursed. Cost reimbursement \( (s = 0) \) is the opposite extreme: \( P_i = c_i \). Mixed payment such as \( s = 0.5 \) implies that the provider receives both a (risk-adjusted) prospective payment and reimbursement for a fraction (in this case, 50%) of actual costs: \( P_i = R_i + 0.5c_i \). We assume that payment is set to be expenditure-neutral with respect to the degree of supply-side cost sharing. Accordingly, in the above formula, the higher the fraction of costs that the provider bears, \( s \), the higher the prospective payment \( R_i \). Without risk adjustment \( (\beta = 0) \), the prospective payment reflects the average cost.
of the population, \( c_{\text{avg}} \). Risk adjustment moves the prospective payment closer to the expected cost of the actual patient according to risk type, \( c_i \).

With these assumptions, it is straightforward to show that net revenue of each patient is given by \( \pi_i = \pi + s (1 - \beta) [c_{\text{avg}} - c_i] \); the difference in net revenues between low and high risks depends on both the level of supply-side cost sharing and the accuracy of risk adjustment, according to the following simple formula:

\[
\pi_{\text{LH}}(s, \beta) = s (1 - \beta) (c_H - c_L).
\] (13)

Under these payment rules, the provider chooses to invest in selection and prevention according to the following program:

\[
\begin{align*}
\max_{e, m} V(e, m) &= \lambda_o(e) \pi_L + (1 - \lambda_o(e)) \pi_H - e - m + \\
&\quad \lambda_1(e, m, \tau) \pi_L + (1 - \lambda_1(e, m, \tau)) \pi_H + \\
&\quad \alpha [x(m) - x(0)].
\end{align*}
\] (14)

This expression can be simplified as follows: \( V(e, m) = 2\pi_H + \lambda_o(e) \pi_{\text{LH}} - e - m + \lambda_1(e, m, \tau) \pi_{\text{LH}} + \alpha [x(m) - x(0)] \), where \( \lambda_1(e, m, \tau) = \lambda_0(e) [x(m) (1 - \tau) + x(0) \tau] \).

### 3.2 Provider choice of selection and prevention

Seeking to maximize net revenue according to (14), the provider chooses \( e \) and \( m \) to balance the marginal benefits and costs of each investment, as defined by the following first-order conditions:

\[
\begin{align*}
\lambda'(e) \pi_{\text{LH}}(s, \beta) [1 + x(0)\tau + x(m) (1 - \tau)] &= 1, \quad \text{and} \\
x'(m) [\lambda_o(e) (1 - \tau) \pi_{\text{LH}}(s, \beta) + \alpha] &= 1.
\end{align*}
\] (15)

Since the provider objective function (14) is by assumption strictly concave, these equations define unique levels of selection and prevention for each set of payment parameters and patient turnover: \( e(\alpha, s; \beta, \tau) \) and \( m(\alpha, s; \beta, \tau) \). The marginal benefit to the provider of selection effort includes two terms: the first, \( \lambda'(e) \pi_{\text{LH}}(s, \beta) \), represents higher profits per enrolled low risk in period 0; the second, \( \lambda'(e) \pi_{\text{LH}}(s, \beta) [x(0)\tau + x(m) (1 - \tau)] \), represents the profits from the fraction \( (1 - \tau) \) of low risks who are retained in period 1 and who remain low risk, in part thanks to prevention efforts \( (x(m) > x(0)) \). When the provider is paid prospectively \( (s > 0) \) and risk
adjustment is not sufficiently accurate to narrow the difference in profits between low and high risks \( (\pi_{LH} (s, \beta) > 0) \), the provider has financial incentive to invest in selection: \( e > e^* = 0 \).

The provider invests in prevention up to the point where the marginal benefit of reduced period 1 treatment cost, net of losses from enrollee turnover, \( x'(m)(1 - \tau)\pi_{LH} \lambda(e) \), plus the pay-for-prevention bonus, \( \alpha x'(m) \), equals the period 0 marginal cost of 1. The provider will under-invest in prevention if the provider’s marginal benefit is less than the social marginal benefit; that is,

\[
m < m^* \text{ if } [\lambda(e)(1 - \tau)\pi_{LH} (s, \beta) + \alpha] < \lambda([B_L - c_L] - [B_H - c_H]).
\]

This may occur if the productivity of provider prevention efforts \( (x'(m)) \) is low, the enrollee turnover rate \( (\tau) \) is high, the pay-for-prevention bonus is too small, and/or the cost savings from maintaining a low risk constitute only a small fraction of the total social benefit from prevention \( (\pi_{LH} \ll [B_L - c_L] - [B_H - c_H]) \).

Comparative statics reveal the main points of the model (see table). To analyze how differences in the productivity of selection and prevention affect provider choice of investments, let \( \lambda' \) be replaced by \( \sigma \lambda' \), and \( x' \) be replaced by \( \theta x' \). Low levels of \( \sigma \) indicate that provider selection effort is not very successful in attracting a favorable risk mix among new enrollees. Similarly, as \( \theta \) becomes arbitrarily small, provider investment in prevention becomes less and less productive, so that for any given \( m \) fewer and fewer low risks enjoy complication-free health into period 1.

For notational convenience, define the following: \( \kappa(m, \tau) \equiv [1 + x(0)\tau + x(m)(1 - \tau)] > 0 \), \( \psi(e, \tau, \alpha, s, \beta) \equiv [\lambda(e)(1 - \tau)\pi_{LH} + \alpha] > 0 \), and \( c_H - c_L \equiv c_{HL} \). \( D \) denotes the Hessian and is assumed to be positive by the assumption that provider utility is strictly concave; \( D = \sigma^{*h} \pi_{LH} \kappa \theta x'' \psi - [\sigma^{*h} \pi_{LH} (1 - \tau) \theta x'']^2 > 0 \).

**Table of Comparative Static Results.**

<table>
<thead>
<tr>
<th>Selection, ( e )</th>
<th>Prevention, ( m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{de}{dx} = -\sigma^{*h} \pi_{LH} (x(0) - x(m))\theta x'' \psi + (1 - \tau)[\theta x']^2 \lambda \pi_{LH} ) ( &lt; 0 )</td>
<td>( \frac{dm}{dx} = \frac{\theta x'' \pi_{LH} \kappa \sigma \lambda^{*h} (1 - \tau) \theta x'']^2 (x(0) - x(m)) ) ( &lt; 0 )</td>
</tr>
<tr>
<td>( \frac{de}{dx} = -\sigma^{*h} \pi_{LH} \kappa \theta x'' \psi ) ( &gt; 0 )</td>
<td>( \frac{dm}{dx} = \frac{[\sigma^{*h} \pi_{LH}]^2 \kappa \theta x'(1 - \tau)}{D} ) ( &gt; 0 )</td>
</tr>
<tr>
<td>( \frac{de}{ds} = \sigma^{*h} \pi_{LH} (1 - \tau) \theta x'' \psi ) ( &gt; 0 )</td>
<td>( \frac{dm}{ds} = -\sigma^{*h} \pi_{LH} \kappa \theta x'' \psi ) ( &gt; 0 )</td>
</tr>
<tr>
<td>( \frac{de}{\beta} = \frac{\sigma^{*h} \pi_{LH} \kappa \theta x'' \psi - \pi_{LH} (1 - \tau) \theta x'' \psi \lambda}{D} ) ( &lt; 0 )</td>
<td>( \frac{dm}{\beta} = \frac{\pi_{LH} \kappa \theta x'(1 - \tau) \sech \pi_{LH} [\sigma \lambda^{*h} - [\sigma \lambda]^{*2}]}{D} ) ( &lt; 0 )</td>
</tr>
<tr>
<td>( \frac{de}{\alpha} = \frac{\sigma^{*h} (1 - \beta) \pi_{LH} \kappa \theta x'' \psi + \pi_{LH} (1 - \tau) \theta x'' \psi \lambda}{D} ) ( &gt; 0 )</td>
<td>( \frac{dm}{\alpha} = \frac{\pi_{LH} \kappa \theta x'(1 - \tau) \sech \pi_{LH} \pi_{LH} \kappa \sigma \lambda^{*h} + [\sigma \lambda]^{*2}}{D} ) ( &gt; 0 )</td>
</tr>
<tr>
<td>( \frac{de}{\tau} = \frac{\theta x'' \pi^{*h} \pi_{LH} (1 - \tau)}{D} ) ( &gt; 0 )</td>
<td>( \frac{dm}{\tau} = -\theta x'' \sigma^{*h} \pi_{LH} \kappa ) ( &gt; 0 )</td>
</tr>
</tbody>
</table>
Factors that encourage both prevention and selection are high supply-side cost sharing (e.g., capitation); high productivity of prevention and selection efforts; large difference in net revenue between high and low risks; and low enrollee turnover (high retention rate). Not surprisingly, then, factors that discourage both selection and prevention include low or no supply-side cost sharing (such as cost reimbursement or fee-for-service); low productivity of prevention and selection efforts; accurate risk adjustment, lowering the net revenue difference between low and high risks; and high enrollee turnover (low retention).

Proposition 1 Conventional risk adjustment discourages both selection and prevention.

Proof. In the model, an increase in $\beta$ reflects implementing, or increasing the accuracy of, conventional risk adjustment. Such an increase lowers the net revenue difference between low and high risks: as $\beta \to 1$, $\pi_{LH} \to 0$. Totally differentiating the first order condition (15) gives that

$$\frac{d_e}{d\beta} = \frac{\sigma \lambda' s (c_H - c_L) \left[ \kappa \theta x'' \psi - \pi_{LH} (1 - \tau)^2 \left[ \theta x' \right]^2 \lambda \right]}{D} < 0. \quad (17)$$

As designed and intended, conventional risk adjustment deters selection. However, it also discourages prevention, since

$$\frac{d_m}{d\beta} = \frac{\pi_{LH} \kappa \theta x' (1 - \tau) s (c_H - c_L) \left[ \sigma \lambda'' - [\sigma \lambda']^2 \right]}{D} < 0. \quad (18)$$

QED.

By paying more as the population becomes less healthy, conventional risk adjustment discourages prevention. The strength of this disincentive depends on the productivity of provider investments in prevention. Conventional risk adjustment will not be a deterrent to prevention if, as Newhouse (2002) suggests, prevention even in the absence of risk adjustment is at minimal levels because provider efforts are a poor substitute for (or an ineffective complement with) consumer behavioral change. Moreover, many primary prevention efforts, such as immunizations, are readily contractible, so the disincentive from risk adjustment can be readily offset by contracted standards for prevention (ibid).

However, some forms of prevention may be both highly productive and ex ante noncontractible. We invoke here a broad definition of prevention, encompassing all interventions that reduce the pace
of health deterioration over time. New processes that promote secondary and tertiary prevention (such as team definition and coordination for disease management) may be far less contractible than primary prevention efforts such as immunizations, yet quite productive for reducing the morbidity and treatment cost burden of chronic diseases. In this case, the negative effect of risk adjustment on prevention presents a serious policy dilemma.

4 Linking risk adjustment to pay-for-prevention

Proposition 2 Combining risk adjustment with pay-for-performance on prevention deters selection while rewarding prevention.

Proof. The previous proposition confirmed that risk adjustment discourages selection. Comparative statics also reveal that pay-for-performance bonuses for above-average prevention counter the effects of risk adjustment on $m$ and can restore incentives for prevention:

$$\frac{dm}{d\alpha} = \frac{-\theta x' \sigma \lambda' \pi_{LH} k}{D} > 0. \quad (19)$$

QED.

Generally, achieving the two policy goals of no selection and optimal prevention requires two policy instruments: risk adjustment, and pay-for-prevention. Risk adjustment technology has been designed to meet the first goal, eliminating incentives for selection. Payment for prevention addresses the second. We discuss the design and rationale for each instrument in turn.

The payer should choose payment parameters in light of the current level of risk adjustment technology. As previous contributors have noted (Newhouse 1996), the payer can compensate for inaccuracy of conventional risk adjustment by softening supply-side cost sharing. In terms of our model, the payer could set supply-side cost sharing $s$ to increase along with the accuracy of risk adjustment $\beta$, according to

$$s^* = \beta. \quad (20)$$

When risk adjustment is very inaccurate ($\beta \approx 0$), the payer relies on reimbursing costs ($s^* \approx 0$) to deter selection. When risk adjustment technology improves, the payer can increase supply-side cost sharing without inducing selection. With perfect risk adjustment ($\beta = 1$), the payer can use full supply-side cost sharing, such as capitation ($s^* = 1$). In this way, the difference in profits from
serving a low risk instead of a high risk, (13), will remain a small value regardless of risk adjustment technology, leaving selection a *de minimus* problem.\(^{10}\) However, prevention will also be minimal: conventional risk adjustment (or cost reimbursement, as suggested here to supplement inaccurate risk adjustment) removes incentive for selection and prevention.

Our contribution is to highlight the penalty on prevention that conventional risk adjustment introduces and to suggest a straightforward remedy. A payer can supplement conventional risk adjustment with an extra payment based on health improvement, \(\alpha |x(m) - x(0)|\). Improvement is defined relative to what would be expected under contractible standards of cost-effective preventive care for that population, \(x(0)\). If the actual fraction of consumers who develop a chronic condition is less than that which would be predicted based on contractible prevention standards, then the provider receives a bonus payment. The optimal payment reflects the patient risk mix (or the severity of disease) in the initial period, as well as changes in disease states from the first period to the second period. A positive bonus rewards better-than-“expected” prevention.

Under conventional risk adjustment, a provider that attracts lower-than-average risks receives a lower risk-adjusted payment, to discourage ‘creaming’. By contrast, under the proposed payment approach, such a provider may not have risk-adjusted payments lowered at all if the health status of continuously enrolled patients reveals that the provider is investing in prevention adequately. In effect, the provider can “buy some selection” by investing in prevention. The pay-for-performance bonus will offset, partially or even fully, the lower payments under conventional risk adjustment. (The tendency of ‘stayers’ to be of higher risk than ‘movers’ works in the payer’s favor, to the extent that the provider’s disease management is better targeted on those of higher baseline risk.)

5 Related Literature and Discussion

Theoretical and empirical work on risk adjustment covers many important issues, but so far has not addressed incentives for prevention, the topic we highlight. Numerous articles present empirical research on risk adjustment in the context of the US multi-payer system, noting the need for refinement along several dimensions (e.g., Burgess 2000; Frank, Glazer and McGuire 2000; Newhouse 2002; Shen and Ellis 2002b; Stafford, Li, Davis and Iezzoni 2004; Zhao et al. 2005; Sessler et al. 2006).

\(^{10}\)Setting \(s\) equal to \(\beta\) has the extra benefit of making the financial return to selection, proportional to \(s(1-\beta)\), nonlinear in \(\beta\). This mirrors the empirical finding of Eggleston and Bir (2006) that returns to selection appear to be nonlinear (see their Figure 3): reducing supply-side cost sharing from 1 to 0.5 more than halves the incentives to risk select. Setting \(s = \beta\) assures that provider incentives are high-powered only when risk adjustment is very accurate.
2010). A related literature explores why adoption of risk adjustment among private payers was slow, at least until recently. Growing experience with risk adjustment in Europe and elsewhere also illustrates its usefulness and challenges (e.g., van de Ven and Ellis 2000; Schut, Gress and Wasem 2003; Yuen, Louis, Di Loreto and Gonnella 2003; Antioch and Walsh 2004; van de Ven, van Vliet, and Lamers 2004; Nuscheler and Knaus 2005; Schut and Van de Ven 2005).

The theoretical literature on risk selection and risk adjusting payments to health plans and providers is more limited, but growing. Some contributions touch upon dynamic or multi-period measures (Luft and Dudley 2002 and 2004; Marchand, Motohiro and Schokkaert 2003), quality and pay-for-performance (Ma and McGuire 1997; Chalkley and Malcomson 1998; Rosenthal et al. 2004; Eggleston 2005; Glazer and McGuire 2006; Miller, Eggleston and Zeckhauser 2006). However, none highlight the focus of our paper: the need to link conventional risk adjustment to pay-for-performance for prevention. Indeed, the conceptual literature on prevention (e.g., Kenkel 2000 and sources cited therein; Byrne and Thompson 2001; Barros and Martinez-Giralt 2003; Dor 2004; Ellis and Manning 2007) has evolved independently from that on risk adjustment. We argue that this separation does a disservice to both literatures and obscures important interactions in practice.

Many factors can motivate providers and insurers to undertake preventive efforts to improve patients’ health: professionalism; an altruistic “warm glow” from helping to prevent suffering; ability to charge a higher price or premium for demonstrated prevention quality; or even differentially attracting healthier patients or enrollees. We focus on a separate and complementary motivation: net revenue from lower (future) treatment costs when paid prospectively, and a payment bonus for better-than-expected health improvements.

Since providers and regions vary considerably in how well they implement health promotion and prevention, aligning payment with the goal of better prevention should be a policy priority. The traditional FFS approach would be to reimburse providers directly for contractible prevention. Unfortunately, this approach fails to promote valuable efforts that are currently noncontractible, and may encourage excessive provision of some services. Conventional risk adjustment formulas and capitated payment are intended to eliminate the overprovision incentive of FFS payment, and

---

11 See articles in a 2001 edition of Inquiry (volume 38 number 3), as well as Newhouse 1998 and Blumenthal 2005. For discussion of specific state and federal program experiences with risk adjustment adoption, see for example Dunn 1998 and Weissman, Wachterman and Blumenthal 2005.

do not require provider actions to be contractible. Moreover, state-of-the-art diagnosis-based risk adjustment is quite effective in achieving its primary goal, deterring risk selection and compensating for adverse selection and adverse retention. However, conventional risk adjustment does not create the correct incentives for insurers and providers to invest adequately in preventive care: insurers that experience deterioration in the health status of their population will receive higher payments in the future.

This incentive problem can take many guises. Consider, for example, a provider choosing whether to allocate funds to a high-risk procedure or a low risk procedure. Suppose that both procedures have the same cost and expected health outcomes, but that they differ in their variances. When an outcome turns out badly, under conventional risk adjustment the provider receives higher payment. This effect may bias decisions toward riskier procedures.

Our simple model suggests a straightforward solution to this incentive problem: linking risk adjustment to pay-for-prevention. Providers should be paid the conventional risk adjusted amount as long as the disease progression observed is as good as expected. But when the progression of diseases is better than expected, given the health status of a population group, then the incremental payment received for those who are healthier than expected should be higher than the full cost differential expected for that group.

Since a certain amount of deterioration of low cost to high cost individuals is to be expected even with exemplary provider effort, we do not propose that a provider or region receive a reduced payment for all individuals whose health deteriorates. Only the difference between the actual and the expected rate of health transition should be used to reward or penalize a provider or region. Moreover, the payer may tailor the payment method to local conditions or phase it in gradually. For example, the payer could choose to use rewards only, and never levy penalties.

Note that the prevention bonus should depend on the probability of provider actions translating into improved health. This is consistent with the Institute of Medicine’s definition of quality health care as increasing the likelihood of desired health outcomes (IOM 2001). The magnitude of the

---

13 Adverse retention is “the tendency for people who stay put to magnify cost differentials between plans, as they will if they differ in age and costs are more than linear in age” (Altman, Cutler and Zeckhauser 1998). Although most of the risk adjustment literature focuses on adverse selection from “movers” (and provider risk selection activities), adverse retention from “stayers” is also an empirically important phenomenon. For example, Altman, Cutler and Zeckhauser (1998) found that adverse retention accounted for about 60 percent as large an effect on premium differences among insurers as adverse selection.

14 The IOM defines quality as “the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge” (http://www.iom.edu/CMS/8089.aspx).
reward for health improvement should depend on the quality of the signals used, and the degree to which transition probabilities between health states can be influenced by health care. For many disease transitions, there is little that a provider or region can do to slow down disease progressions. When progressions are random or reflect patient lifestyle choices that are not easily influenced by medical practice, conventional risk adjustment should suffice. For other disease transitions, however, provider actions can play an important role. Ideally these transitions should be identified empirically and the conventional risk adjustment formula modified to provide a greater incentive for providers to invest in prevention. Specific preventive efforts do not need to be observable; only the outcomes (as measured by disease state transitions) must be. This focus on outcomes rather than process is potentially easier to implement than a system that rewards observable efforts.¹⁵

Indeed the key challenge to implementing pay-for-prevention is how to measure health improvement. One possible solution lies in constructing measures of disease transition. Two approaches for measuring disease transitions are available. The first approach measures whether or not people in each category of a certain disease, say patients with diabetes, get better, remain the same, or get worse solely in terms of their diabetes severity. This approach, similar to disease staging, will only model progressions in the severity of the mutually exclusive diabetes conditions recorded. The second approach looks at the progression in the overall disease score of each individual, taking advantage of the concurrent relative risk score generated by risk adjustment models as a metric. Risk scores within a given body system provide a simple metric, at the individual level, of whether a person has become sicker, remains the same, or has become worse, according to the change in the risk score.

Our conceptual model suggests that payments to providers or regions should be based not only on levels of disease, but also on differences between observed disease transitions and expected rates. The magnitude of the incentive to reward providers or regions for doing better than expected in their disease transitions is in part an empirical decision and in part a policy decision. Future empirical work is urgently needed to develop a methodology that could potentially be used to refine and improve fund allocation and payment formulas by improving incentives to keep people healthy.

Of course, patients may also differ in their responsiveness to preventive efforts, leaving pay-for-prevention with a second-order selection problem: financially rewarding provider-insurers that

¹⁵ Of course, providers will still have incentives to game the system, particularly regarding initial budget allocations, just as UK primary care physicians increased hospital-based activity prior to becoming GP fundholders, thereby inflating their budgets (Crosson, Propper, and Perkins 2001). Also see Lu 1999.
attract and retain low risk patients who are more receptive to prevention efforts (such as better-educated patients with chronic disease who are more likely to adhere to treatment regimens). Therefore, any full model of optimal incentives for prevention needs to take account of both provider and patient incentives and the remaining selection issues that arise from heterogeneity in patient response to provider prevention efforts.\footnote{We thank an anonymous referee for emphasizing this point.}

In sum, we highlight the logic linking two increasingly prevalent approaches to payment for healthcare services: risk adjustment and pay-for-performance. Our simple model illustrates how their combination can both deter socially wasteful risk selection and reward prevention. Fruitful theoretical extensions might include incorporating strategic interactions among competing providers; modeling appropriate dynamic incentives for consumers to maintain their own health (e.g., decrease obesity, increase exercise, stop smoking); or incorporating a “business case for quality,” allowing insurers ability to charge consumers more for proven ability to keep them healthier.

Acknowledgment. The authors thank Albert Ma and anonymous referees for comments on an earlier draft. Lu thanks the Alberta Heritage Foundation for Medical Research and Institute of Health Economics, Alberta, for financial support. The usual disclaimer applies.

References


