# Risk adjustment in health care markets:

# concepts and applications

Randall P. Ellis

Boston University

February 2007

*Abstract*

This chapter summarizes the recent literature on risk adjustment, synthesizing the theoretical and empirical literatures, emphasizing areas of current research and policy interest. The paper relates optimal risk adjustment to the previous theory and empirical work. The empirical estimation section covers choice of an objective function, different types of information that can be used, conceptual issues, estimation, and validation. The paper concludes with detailed discussion of risk adjustment applications in Canada, US, Netherlands, and Germany.

# Introduction

This chapter provides an overview of the theory and practice of risk adjustment models for health care systems. After defining risk adjustment and discussing its importance, the chapter describes the theory, empirical estimation, and international applications of risk adjustment. The theoretical section attempts to motivate why risk adjustment is important, as well as to clarify the linkages between the theoretical and empirical risk adjustment literatures. The empirical section first summarizes the different information sets that can be used for risk adjustment and then describes how risk adjustment formulas are be estimated and used for predicting health care spending. The international applications section illustrates a few examples of how risk adjustment models are used around the world in practice, and helps identify both the opportunities and limitations of risk adjustment.

## *What is risk adjustment?*

For this chapter *risk adjustment* is broadly defined to mean "the use of patient-level information to explain variation in health care spending, resource utilization, and health outcomes over a fixed interval of time, such as a year." This definition focuses on what are often called "population-based models", which is to say the utilization of services by an individual rather than a group. Episode-based predictive models and models using grouped data are not considered here.

There is considerable diversity in the literature about what is meant by risk adjustment. In their review van de Ven and Ellis [1] focus on using risk adjustment to pay competitive health plans and define the term *risk adjustment* narrowly to mean "the use of information to calculate the expected health expenditures of individual consumers over a fixed interval of time (e.g., a month, quarter, or year)." Glazer and McGuire [2, 3] define risk adjustment as paying a provider a price "conditioned on observable characteristics of the enrollee or patient." According to the classic book

on risk adjustment, written by a US physician Lisa Iezzoni, the terms "risk adjustment" and "severity adjustment" can be used interchangeably, although risk adjustment is more commonly used. In the health services research literature, the terms "case mix adjustment" is often used instead. [4] Among economists, "risk adjustment" is often but not always limited to uses for payment: the creation of formulas to be used for capitation payment. Risk adjustment models used for payment purposes, have also been referred to as "health-based payment" [5, 6].  In Europe and Canada, the terminology "needs-based payment" is often used [7, 8]. Health-based payment is but one application of how explicit risk adjustment formulas can be used.  In recent years the terminology "predictive modeling" has come to be used in the US for models designed for predicting health care utilization without regard to whether the predictive model will be used for payment.

{Figure 1 goes about here}

Figure 1 organizes the different terminology that has been used to talk about risk adjustment, primarily from a US perspective.  The terminology "predictive modeling" is broader, and includes case identification models used to identify high cost cases without any regard to incentives.  For this purpose almost any information available can be used, including other endogenous variables such as lagged spending or use of procedures information.

In the risk adjustment literature, two different questions are often asked.  One question asked is: What is the predicted resource use for individual XYZ?  Having a good prediction model to answer this question can be useful for many reasons, and an abundant literature has developed classification systems and estimation methods that attempt to answer this question well.  A different question of interest is: What is the best payment formula to use for capitated payments for individual XYZ? Answering this question is a recent literature on optimal risk adjustment.[2] This paper discusses models useful for answering both of these types of questions: risk adjustment prediction models and risk adjustment payment models.

The different terminology is also related to the information used for prediction and payment. As elaborated upon in van de Ven and Ellis [1] there are many different types of information that can be used to predict health care service utilization. Variation in health care use can be decomposed conceptually and empirically into variation due to patient characteristics, provider characteristics (e.g. specialists, general practitioners, hospitals, nurses), and the nature of the services actually provided (such as their pricing, intensity and duration). Depending on the purpose, all of this information may be useful for prediction. Patient characteristics can be further decomposed into variations due to the underlying health status of the patient, socioeconomic variables such as income and education; enabling information such as benefit design and geographic location that affect access and utilization, and patient tastes. The best set of information to use depends upon the intended use. Health-based payment models and severity adjustment models restrict the information set to only use health status. Needs-adjusted payment, widely used in Europe, broadens the information to reflect further demographic variables such as income, race, geography, and access (e.g., distance). The choice of information may also affect the desired empirical specification to be used.

## How are health risk adjustment models used?

Economists tend to focus on the uses of risk adjustment models for purposes of payment, whereas the majority of use in practice is for other purposes. Risk adjustment models are used in the US commercial sector for measuring quality, case management, disease management, high cost case identification, underwriting, plan selection by employers, and provider profiling. The answer to the question: how many resources is person XYZ expected to use next year is valuable for all of these uses. One way of documenting how risk adjustment is used in practice is to quantify its use on the web. A web search using Google on September 6, 2006, found the 259,000 hits on "risk adjustment"+"health". Of these, the word "payment" appeared in 45% of the web sites, while the word "quality" appeared in 73%. More than half of the health risk adjustment web sites mention "quality" without mention of the word "payment". This highlights that risk adjustment

4

predictive models are increasingly being used extensively for many diverse purposes. Thomas et al. [9] provides a useful review of different models

{Figure 2 goes about here.}

There are many ways in which risk adjustment models can be used for payment. Figure 2 uses an organizational framework from van de Ven and Ellis [1] for thinking about the relationships between the various agents in health care markets. There are four primary contracting agents in each health care system: consumers, providers, sponsors and health plans. The solid lines reflect five possible contractual relationships among these agents. Consumers and providers are well-known participants in health care markets. Consumers receive health care services, choose providers and service quantities, and collectively contribute toward the cost of health care. Consumers generally contract with at least one sponsor, and may or may not contract directly with health plans. Providers provide the actual health services and accept payments from health plans. They contract with consumers and health plans.

A sponsor plays the critical role of accepting contributions from each consumer that may not be equal to their expected cost of health care. The sponsor makes payments to a health plan that need not be equal to the contribution by a given individual. As such, the sponsor plays the extremely important role of redistributing the burdens of premium contributions by consumers, generally subsidizing high risk and low income consumers by low risk and high income consumers.

The health plan also plays an important role. The health plan is an intermediary that takes payments from the sponsor or from consumers and pays providers. Any agent contracting with providers for services can be considered a health plan, and if the sponsor pays providers directly, then it has merged the role of sponsor and health plan. Health plans need not bear any financial risk. In the US, they may be simple financial intermediaries who transfer risk back to the sponsor.

Risk adjustment payment models can potentially be used for each of the five contracting relationships shown in Figure 1. The primary use in many countries is on payments from sponsor to health plans. This is commonly called capitation payment, since payments by the sponsor are calculated on a per capita (per person) basis. For example, in Germany, Netherlands and the US Medicare program, risk adjustment payment models are used to calculate payments by public or quasi-public organizations to health plans. In relatively few settings in the US, a single employer makes capitated payments to multiple health plans. [10, 11]

Risk adjustment can also be used for other contractual relationships. In the US, risk adjusted payments are often made by a health plan to networks of providers, such as hospital or primary care networks. In the UK and the Netherlands, primary care doctors are receiving capitated payments for broad sets of outpatient services. Risk adjustment also plays a key role in most "pay for performance" programs, which are discussed briefly below in the review of US experience.

As ven de Ven and Ellis [1] highlight, fees, risk-sharing and other non-capitated payments are often used in combination with risk adjusted capitation. For example, capitation payments may be calculated for only the facility-based spending by a government, or primary care providers may share the risk that costs are higher or lower than the capitated amounts. Although we discuss such partially capitated models briefly below in our empirical section, following the literature, the theoretical discussion focuses on pure capitation models.

## Theory

In recent years there has been a burgeoning of theoretical analyses of risk adjustment payment model incentives. This section does not attempt to provide an exhaustive review of that literature, but rather provides a conceptual overview of some of the key issues that have been studied and the insights from that work. Attention is focused here on the particular issues of perfect versus imperfect information,

6

imperfect signals available to the regulator, the incentive problems that health based-payments are intended to correct, and the strategic responses to risk adjusted payments.

This section uses a simplified framework using primarily a graphical approach.  I focus on the case where there are only two states of the world, two types of health care goods, two types of consumers, and two possible signals about consumer types. Following the recent literature, I assume that the only information that is *contractible*, and hence usable for paying capitated health plans, are the signals about consumer types.  Realized states of the world, actual levels of spending on each type, and true patient types are not observable to the payer.  Even if they are observable to all agents, including the sponsor, they may nonetheless not be contractible.

Assume that consumers can find themselves in one of only two possible states of the world. For concreteness I call them healthy and diabetic.  Healthy consumers use only general practitioner services (GP) and cost $\alpha$ per year.  Diabetic consumers use both GP and specialists services (SP) and cost $\alpha+\beta$. Hence $\alpha$ is the annual cost of healthy consumers under a given set of incentives, while $\beta$ is the incremental cost of diabetes. Using the terminology of Glazer and McGuire [12], GPs would be the acute care service, while specialists would be the chronic care service used more intensively by high risk types. In the terminology of Jack [13] GPs would be "regular care" while specialists would be "acute care." They key assumption is that one service (here called specialist care) is used more intensively by high risk types than the other service (GPs).

Assume there are only two types of consumers, called low and high risk types.  Initially I focus on the simple case where low risk types are always healthy and high risk types are always diabetic. While it is true that there is no uncertainty (risk) in this polar case, it is still convenient to think of each type as potentially having a distribution across different states of the world.  A richer model, as in Glazer and McGuire ([12] henceforth GM), has both high and low risk types having some nonzero probabilities of being diabetic and healthy, but the simple model developed here captures the essential features needed for

talking about risk adjustment. Most of the useful insights arise due to imperfect signals about patient types rather than due to uncertainty about how consumer types are related to which state of the world consumers find themselves in. For pedantic simplicity, it is also convenient to assume that both types are equally common in the population, thereby saving another parameter.

{Figure 3 goes about here.}

This situation is shown in Figure 3, where healthy (low risk) consumers consume at some point A using only GP services and diabetic consumers (high risk) consume at some point B. Later it will be significant that I assume that diabetics use more of both GP and SP, not just SP. Notice that so far only variation in consumer health status drives the variation in demand for different types of care. Differences in tastes and income may also affect the utility of different states of the world and hence the use of health care services. Various papers consider models where patients differ not only in health status but also in taste parameters. [12, 13, 14].

The optimal risk adjustment formula to use depends on the objective function of the sponsor, as well as the cost structure, competition, health plan and consumer objective functions, the information (signals) available to sponsor, consumers, and health plans, and the strategic behavior allowed by the health plans. In the next section I start by considering the polar cases where signals are either totally uninformative or perfectly informative about consumer type. Initially, health plans are assumed not to have any strategic behavior possible. Later I introduce imperfect signals and allow different forms of strategic behavior.

Throughout this section I need an assumption about the health plan's objective function. Although a more general framework might have the health plan caring about both patient benefits and profits [15], in a static model the conceptually straightforward assumption is that health plans maximize profits, or what is also known as net revenue.

The objective function of the sponsor is initially assumed to be very simple.  The sponsor wishes to simply pay each health plan the expected value of each signal for each consumer, which is what is GM have termed *conventional risk adjustment*. [12]  GM and others may note that this is a naïve and inferior objective function for the sponsor, but it is nonetheless the most common one.  Since this objective and the implied behavior assumption about the sponsor are central to many empirical models of risk adjustment, it is also an important one to start with.

## *No risk adjustment signals*

The simplest case is one in which the GP and SP services provided to each type of consumer are exogenous to (independent of) payments and information, and there are no signals about consumer type available. If quantities of care offered are unaffected by capitation payments, then it does not matter whether consumer types are observable or unobservable by the health plan or consumers at the time that plan enrollment decisions are made.   Under pure capitation, health plan budgets are set prospectively and will equal the expected amount of $R = \alpha + \beta/2$ per person, where this formula reflects the equal prevalence in the population of high and low risk types.

This no risk adjustment situation is shown in Figure 3, with X being the average quantity of services consumed.  As long as the proportion of high risk types enrolling in the plan is no greater than the population average, then the plan can at least break even by subsidizing the losses on high cost types with the savings on the low cost types.   Economists would point out that X is not incentive compatible for all possible plans: plans will not want to participate if they have a higher proportion of diabetics than the population average used for payment.  Accommodating such plans creates the desire for risk adjustment by the sponsor.

Of course a profit maximizing health plan would prefer not to have to enroll high risks. Under the stated assumptions plans do not have any tools available for avoiding diabetics, since no information is available

9

ex ante. With no risk adjustment, health plans are forced to carry out the key role of redistributing funds from low to high risk consumers. Shen and Ellis [16] explore how a profit-oriented health plan would prefer to exclude high risk types based on observable signals, but they assume that plans can perfectly select which consumers to avoid or "dump." Almost all countries and markets prohibit explicit dumping of high risk consumers. Settings where the exclusion (dumping) of high risk types is allowed include the individual and small employer markets in the US, commercial health plans in Chile, Columbia, and India, and supplementary health insurance policies in many countries.

## *Perfect signals*

Now assume that health status signals S are costlessly available to all agents (consumers, health plans and sponsors) at the beginning of each period, and that these signals are informative about each patient's risk type. In the simplest case there are only two possible signals, so that S takes on values of either 0 or 1. Assume initially that the signal S is perfectly informative, so that a value of 1 (0) perfectly classifies a consumer as a high (low) risk type. This is often called perfect information.

With perfect information by all agents, the sponsor wishing to pay health plans the expected cost will pay a risk adjusted payment of $R_i = \alpha + \beta S_i$ for each consumer i. Since signals are perfect, then the implied payment parameters are $R_1 = \alpha$ and $R_2 = \alpha + \beta$. This is the conventional risk adjustment solution, calculating $\alpha$ and $\beta$ so as to just pay the average cost of each signal. Even if health plans (or regions) vary in the proportions of high and low risk types, payments will equal costs. Since profits on each type of consumer are zero, a health plan should be just indifferent to enrolling consumers who have low or high risk signals. Conventional risk adjustment with perfect information and solves the objective of eliminating the incentive problem facing profit maximizing health plans to selectively avoid or "dump" unprofitable consumers when the signals used are the only information known to health plans, and costs of treating patients are exogenous.

The large literature on the calculation of conventional risk adjustment formulas, including the US

Medicare and the Dutch risk adjustment payment formulas are calculated while implicitly using

assumptions similar to those used here.  It is straightforward to calculate the α and β parameters

empirically using a regression approach. Recent theoretical research has greatly expanded the

understanding of risk adjustment incentives and possible corrections to it.  The following sections

highlight various extensions involving imperfect signals, provider distortions of services available, and

heterogeneous tastes.

## *Exogenous imperfect signals, exogenous service quantities*

The above model assumes that the signals are perfectly informative: once the signal is observed then the

true type of each consumer is known and perfect risk adjustment is feasible.  Unfortunately the norm is

that signals are highly imperfect. Empirical studies repeatedly find that even for serious chronic

conditions such as AIDS, diabetes, multiple sclerosis and quadriplegia, less than 75 percent of consumers

with these diagnoses coded in one year have the same information coded the following year. [17, 18]

Even signals of serious illness are highly imperfect.

There are two broad possibilities for why signals are imperfect.  Signals may exogenously misrepresent

true patient risk types, or signals may be endogenously (intentionally) chosen by the health plan or

providers so as to influence payments.  GM [12] examined the concept of exogenous imperfect risk

adjustment signals.  Expanding the notation of this paper, suppose that proportion $\gamma_i$ of type i consumers

have a signal S=1.  In order for the signal to be informative then we need proportion $0 \leq \gamma_L < \gamma_H \leq 1$.

The polar case $\gamma_L=0$ and $\gamma_H=1$ corresponds to the perfect information case just considered.  Empirically, it

is widely found that some low risk types have a false positive signal $(0 < \gamma_L)$ and many high risks have

false negative signals $(\gamma_H < 1)$.  This means that those with a signal S=1 will contain both H and L types,

and the average cost of those with S=1 will be less than α+β.  As shown in Figure 4, starting from the

point X where no information is available, improving risk adjustment signals will better differentiate

between low and high risk types, reducing the calculated payment for healthy signals (S=0) toward α and increasing the calculated payment for S=1 toward β. Under these assumptions the proportion of a plan with high risk signal S=1 would be $(\gamma_H+\gamma_L)/2$, the average cost of those signal S=1 would be $R_1= \alpha +\beta \gamma_H/ (\gamma_L+\gamma_H)$, while the average cost of signal S=0 would be $R_0= \alpha +\beta (1-\gamma_H)/(2- \gamma_L-\gamma_H)$ which yields the correct weighted average of low and high cost types.

{Figure 4 goes about here.}

Imperfect risk adjustment signals create problems for conventional risk adjustment. In the literature three types of problems have been emphasized. One common complaint of managed care plans in the use is that their decentralized payment systems mean that they are have less frequent high risk signals (i.e., more false negatives) than traditional health plans. Newhouse [15] discusses the profitability of selection by health plans when they have only a little bit of private information and finds that profits are highly nonlinear, with a little bit of private information being very profitable.[15] GM [12] and Frank, Glazer and McGuire [19] highlight how imperfect signals, in combination with strategic behavior, create significant distortions in health services, which we turn to after considering endogenous risk adjustment signals.

## *Endogenous risk adjuster signals, exogenous service quantities*

Among actuaries and policy makers, one of the great concerns with implementing risk adjustment is over the problem of endogenous signals, which is to say that the information used to make payments is influenced by the payment formulas themselves (e.g., Ingber, 1998). For example, following the introduction of DRGs in the US, there is convincing evidence of a meaningful increase in the apparent coded severity of hospitalized patients due to the new incentives which raised Medicare payments by a few percentage points. Similar concerns arise with the policy implementation of risk adjusted payments. The pattern of information used for calibrating the models may differ from the pattern that will arise after

risk adjusted payments are introduced. Concern over this endogeneity has been a major force in the US in its choice of risk adjusters for its Medicare program. (See for example Newhouse, 2002, p 56 for a discussion of Medicare's adoption of the PIPDCG model.) This concern was a major factor in the Medicare programs choice of payment formulas that intentionally ignored diagnoses predictive of lower cost conditions.

In terms of the simplified model used here, endogenous signals would mean that health plans would wish to increase the proportion of high risk types $\gamma'_L > \gamma_L$ and $\gamma'_H > \gamma_H$ reported beyond the levels used to calibrate the models, $\gamma_L$ and $\gamma_H$. The reported proportion of high risks enrollees has increased, and there will be an overpayment to the health plans until payments are recalibrated. Such recalibration to accommodate coding prevalence changes is what took place with DRGs in the late 1980's and with the Medicare's all encounter HCC model in 2006.

Concern about endogenous signal changes in response to the US Medicare capitation formula in 2004 was the principle reason why Medicare implemented a "FFS correction" factor in the original payment formula. This adjustment factor remained at 5% for three years, but was reduced to 2.9 percent for 2007 suggesting that the impact was relatively transitory and modest. [20]

## *Health plan response to capitation payments: proportional adjustment*

It is a widely held belief that quantities and qualities of health care services will respond to payment incentives. Ellis and McGuire [21] model partially altruistic providers as responding to lump sum payment incentives by reducing quantities of care relative to fee-based services. Newhouse [15] extended this discussion to health plan behavior in response to capitated payments. The classic supply-side response to capitation would be a reduction in spending on all types of services. This possibility is shown in Figure 5, where spending on both diabetics and healthy types is reduced in response to moving from fees to capitated payments. Since the quantities of services are no longer exogenous, we distinguish the

initial fee based quantities $(A^0, B^0)$ from the capitation induced quantities $(A^1, B^1)$. Costs parameters for

each type also change from $(\alpha^0, \beta^0)$ to $(\alpha^1, \beta^1)$.

{Figure 5 goes about here.}


There is a significant empirical literature that has worried about estimating this plan-level change in

spending in response to capitation incentives, building on the approach of Lee [22]. Recent work by

Terza [23] in this area using nonparametric techniques has further extended the technique. The hallmark

of this approach has been to model total health spending, allowing for a uniform proportional or absolute

reduction in spending in response to capitation incentives. Empirical plan selection models that estimate

how total health care spending differs between managed care, indemnity and other types of health plans

work on the assumption of a simple proportional or additive adjustment to total costs. It is not uncommon

to calculate the cost savings of HMOs or other plan types with an additive constant in either a linear or

loglinear model. This corresponds to assuming that all costs are adjusted additively or proportionally for

all types of consumers. See Glied [24] for a discussion of such results for managed care.

If health plans respond to capitation incentives by reducing spending on all types of services uniformly,

then this does not create particular problem for risk adjustment. A uniform reduction can be

accommodated by reducing the payment parameters $(\alpha, \beta)$ proportionally. This sort of correction is

reflected in the US Medicare adjustments to the AAPCC during the 1980s and 1990, where payments

were reduced by 5% to accommodate expected response by capitated health plans (CMS, 2006).

## *Strategic response by health plans to capitation payments.*

Glazer and McGuire [12] highlight that health plans can behave strategically in how generously they

provide each type of specific health care services. Regulators may be able to prohibit dumping of

unprofitable consumers, but they cannot easily prevent health plans and providers from increasing or decreasing the availabilities of certain specialists or types of services. If certain types of consumers are unprofitable because risk adjustment is imperfect, then health plans have an incentive to avoid enrolling or treating these consumers by reducing the provision of the health care services that are most attractive to them. In the present model, rather than reducing spending on all types of services proportionally, health plans will have an incentive to reduce spending on SP because it is used more heavily by high risk types. At the same time, plans will have an incentive to compete to attract low risk types, by oversupplying GP services to the healthy. This process may or may not result in an increase in GP services being made available to high risk types, depending on whether the health plan is able to differentiate the services offered to the two observable types. This subtle issue underlies a key difference between the work of GM [12] and Ma [25]. GM assume that even though plans know the signal of a consumer, they are not able to differentiate the quantity of services (GP and SP) offered to a given signal type. Hence GM assumes plans cannot offer GP services more generously to the S=0 than the S=1 consumers, while Ma assumes that they can. This leads to different conclusions about plan rationing and optimal payments.

{Figure 6 goes about here.}


Three new insights follow from the GM framework. The first is that each consumer type can be thought of as having a set of indifference curves between GP and SP services. For a given level of spending, there will be some optimal combination of GP and SP services that maximizes consumer utility, and those quantities of good will differ by consumer type. It is natural to ask in this framework whether the amounts offered are the utility maximizing quantities under different payment formulas. Indifference curves are shown in Figure 7 illustrating that the initial allocation at $B^0$ need not be welfare maximizing for high risk types: there is no particular force under FFS reimbursement ensuring that quantities provided

are efficient. Indeed we know that in general fully-insured patients will tend to receive too much treatment of both services.

A second insight is that if risk adjustment is perfect, then capitated health plans will compete by offering services to each type where welfare is maximized. (This result is not shown graphically.) With complete information, health plans will wish to maximize consumer welfare for the same reasons that competitive firms do so. As long as health plans are indifferent between enrolling a given consumer, then they will try to do the best thing by them.

The third insight of GM is that if signals are imperfect, then health plans will distort quantities in an inefficient way. As shown in Figure 7, quantities of SP will be greatly reduced (to point $B^2$), to deter high risk types, while quantities of GP will tend to be increased (to $A^2$) to try to attract the low risk types. Notice the key role of consumer choice. Because consumers can be attracted to a better set of services offerings, then health plans will compete in this dimension. If payment formulas are recalibrated using the new service quantities, then conventional risk adjustment formulas will validate the results of this form of health plan competition by reducing payments for S=1 signals, and increasing them for S=0 signals, confirming the service intensity choices of the health plans. Alternatively, if the conventional risk adjustment formula continues to be based on a distinct FFS sample, as is true for the US Medicare program, then the formula may misrepresent the costs of both types when they are enrolled in the capitated plans. Even though the sponsor is breaking even conditional on the signals, they are not maximizing consumer welfare.

## Optimal Risk Adjustment

In order to solve this service distortion, GM broadens the sponsor's objective function. They introduce the concept of "optimal risk adjustment" in which the sponsor's goal is to maximize consumer welfare rather than to just break even. Instead of restricting the sponsor to consider only unbiased risk adjustment

formulas, they allow the sponsor to make payments for some signals that differ from the expected cost of each signal.

The GM optimal risk adjustment solution to this distortion is to use the structure of demand and the exogenous signal reporting process to systematically overpay for signals S=1 while underpaying for S=0. For each observed signals S=1 implying a high risk type, the health plan should be compensated for more than the incremental cost of that consumer, to cover the inferred presence of other high risk types in the same plan. This overpayment for S=1 will make the health plan more willing to enroll high risk types, and to offer more SP services that will attract them.

{Figure 7 goes about here.}

The GM [12] optimal risk adjustment equilibrium for this simplified example is shown in Figure 77. It is useful to contrast the conventional and optimal risk adjustment results. Continuing with the previous example with imperfect signals, consider the case where $\gamma_L = 0$ but $\gamma_H = \frac{1}{2}$. This corresponds to the case where half of all diabetics have a signal reporting their type, but half do not. There are no false positives. Assume that the regulator calculates conventional risk adjustment payments based on unbiased estimates using the imperfect signal S using FFS data. In this case, conventional risk adjustment would pay $\alpha^0 + \beta^0$, for those with S=1, and pay $\alpha^0 + (1-\gamma_H) \beta^0$ for the S=0 types. If the plan simply enrolled a representative mix of high and low risk types, then it would break even. If instead the plan is able to discourage some of the true high risks from enrolling, then for the extra share of its enrollees with S=0, it could receive $\alpha^0 + (1-\gamma_H) \beta^0 > \alpha^0$ for the low risk types it enrolls, making a profit. Hence in order to achieve this, the plan would reduce SP and increase GP spending, offering service mix such as $A^2$ and $B^2$ in Figure 7.

The GM solution to this incentive is to overpay for the high risk types, making them sufficiently attractive as to encourage plans to compete for them. If $\alpha^0$ and $\beta^0$ were the optimal levels of spending to achieve, then the solution would be to pay $\alpha^0$ for the S=0 types, and $\alpha^0 + \beta^0 / \gamma_H = \alpha^0 + 2\beta^0$ for the S=1 types. Notice that by overpaying two times the expected difference in costs for the high signal, the regulator undoes the

disincentive to attract high risk types.  In general, the FFS level of services would not be the desired level, and some alternative level such as $A^3$ and $B^3$ in Figure 7, where $\alpha^3 < \alpha^0$ and $\beta^3 < \beta^0$.

Although attractive conceptually, achieving the first best in practice may be difficult, for reasons that GM and others acknowledge.  Knowing the optimal consumption points $A^3$ and $B^3$ may not be feasible, and the structure of the information set, such as the rates of true and false positive signals, may not be known to the sponsor.  The optimal risk adjustment payment can be very sensitive to the degree of imperfect signaling.  For small $\gamma_H$, it may be necessary to greatly overpay S=1.  In some situations, underpayment for low risks may also be needed, which imposes nonnegativity constraints on payments.  Still, the GM solution points the direction that should be considered by sponsors where health plans compete to avoid high risk types: overpay for high risk type signals and underpay for low risk type signals relative to conventional risk adjustment, in order to encourage desirable competition to attract high risk type consumers. Glazer and McGuire in Chapter XXX of this book contains an important extension to the GM framework that encompasses optimal risk adjustment when there is noncontractible service quality, and not just service noncontractible quantity.

## *Taste and income variation*

Taste or income variation can be introduced into the problem by adding consumer heterogeneity.  Figure 8 illustrates the case where there are two types of healthy enrollees, whose preferred bundles of services are A' and A", and two types of diabetics, preferring B' and B".  Payments based on a binary signal cannot achieve the first best allocations of both GP and SP for all four types of consumers.  If averages for each signal are paid, then health plans will have an incentive to compete to attract consumers with lower tastes for services.

{Figure 8 goes about here.}

18

## Rationale for focusing on for high cost conditions.

8 can also be used to show the implications of adding further types. Suppose that instead of taste variation, there are now two different types of healthy consumers (e.g, healthy and healthy with allergies), with optimal choices A' and A", with still only one diabetic type at B'. (Ignore B" for this discussion.) It could very well be that signals distinguishing B' from A' and A" are much more reliable than those distinguishing A' from A". This could justify developing payment models that recognize distinctions of diabetics B', but ignore minor differences among the healthy. This is the approach that has been taken in the Netherlands and the US Medicare program, both of which have focused on paying more for only the highest cost conditions.

## Optimal risk adjustment in settings with noncompeting health plans

Before continuing on, it is useful to consider the relevance of the GM service competition from the point of view of other countries. Risk adjustment is used in Canada and Europe in settings where health plans do not choose services so as to actively compete for enrollees. In Canada and Scandinavia, budget allocations are primarily geographic, while in Netherlands and Germany, even though there are many competing health plans, selective contracting is rare, so that health plans do not have available tools that enable meaningful service distortion so as to attract profitable enrollees. Hutchison et al. [8] in discussing Canada go so far as to assert "For health care funding allocations to geographically-defined populations, risk adjustment is irrelevant." Is this true?

There are two important ways in which risk adjustment remains important even for Germany, Canada and other regions without meaningful service level competition. One reason is that there may be historical reasons (geography or employment) why populations in some regions or health plans differ that are not captured well by simple risk adjustment models. As documented below under the international review, even geographic areas may vary significantly in their disease patterns and hence costs. Another reason is

that consumers may temporarily or permanently migrate to use a different region or plan. It is not uncommon for consumers to use a hospital or specialist from a different region, or a different health plan, and the internal transfer payments among regions or health plans are tricky to calculate. Individuals migrating across regional borders will tend to differ from the average. For example, chronically ill patients often relocate to close to centers of excellence that specialize in their illness. The effort that hospitals or providers work to attract consumers from neighboring regions is also a strategic decision, possibly influence by the payment system. Recognizing this health distribution and choices being made in the risk adjustment formula may be important.

GM [12] and especially Glazer and McGuire [2] emphasize that optimal risk adjustment is more about recognizing that unbiased payment should not be the primary objective of the sponsor, but rather a more fundamental objective of maximizing consumer welfare should drive payment calculations. Canadians and Europeans may assign more importance to equity than efficiency goals in their objective function, but that does not reduce the desirability of acknowledging the value of a broader objective function than just minimizing individual level variance in plan profitability, or at the other extreme paying regions or health plans without regard to the healthiness of their enrollees.

## Estimation Issues

This section highlights some of the results that bear on estimation and interpretation of risk adjustment models. Van de Ven and Ellis [1] reviewed much of this up through 1999. For the US, more recent reviews and comparisons of different risk adjustment models are contained in Cumming et al [26], in their report to the Society of Actuaries, the work of Thomas, et al, [9, 27] on provider profiling, and the policy review in Newhouse [28]. The chapter by Leger-Thomas in this book on provider payment also has a direct bearing.[29] Rice and Smith [30] conducted a very useful overview of risk adjustment techniques outside of the US. Other articles on other specific countries are discussed further below.

Empirical models of risk adjustment have to address five issues in model estimation. These are:

1. Choice of the objective function to be optimized

2. Choice of risk adjusters (signals to be used for prediction)

3. The conceptual framework, including the classification system and whether to use mutually exclusive rate cells or additive prediction models

4. Econometric specification, including nonlinear estimators and endogenous plan selection corrections

5. Methods of evaluation and prediction

The discussion in this section is organized around coverage of these five topics.

## *Choice of the objective function to be optimized*

The theory section above has emphasized the choice of the objective function for the sponsor, and the consideration of the allowed risk selection problem that is being solved. Very few empirical studies have estimated "optimal risk adjustment" models that reflect any objective function other than being budget neutral and maximizing predictive power (e.g., the R-square). The two notable examples are Glazer and McGuire [31] which estimates a "minimum variance optimal risk adjustment" model using Medicaid data that seeks to maximize the $R^2$ while maintaining the first best incentive for service provision and Shen and Ellis [16] which estimates a "cost minimizing risk adjustment model" which minimizes the financial cost to the sponsor when capitated, risk selecting plans compete against open enrolling, fee-for-service plan, as is true for the US Medicare program.

European countries often pursue equity objectives that are not easily summarized in a welfare function to be maximized. Many of the risk adjustment formulas reviewed in Rice and Smith [30] reflect transfers and subsidies to increase service provisions to underserved populations (low income households, underserved rural areas, and minority status consumers, for example.) These models are implicitly using more complex objectives than maximizing simple predictive power.

Implicit in the choice of most objective functions is the choice of the dependent variable – what variable is being predicted. In the payment literature total annual spending, including inpatient, outpatient and pharmacy spending, is the most common dependent variable, although in some settings visits, admissions, or other measures of utilization are predicted.

## *Choice of risk adjusters*

As better information and larger samples have become available, there has been a progression towards more elaborate information sets being used for empirical risk adjustment models. Many researchers have suggested that diagnoses come as close as any widely available measure of consumer health status, and hence diagnoses from insurance claims are the most widely used set of information beyond demographic variables. Pharmacy information is sometimes used in place of diagnosis information, although prescription practice is perhaps more subject to variation across doctors and over time than diagnostic coding. Self-reported health status is attractive conceptually, although surveys are more expensive to collect than diagnoses and pharmacy use. Another concern is that consumers may only imperfectly differentiate among illness severities or imprecisely recall doctor's advice and diagnoses. Table 1 summarizes the information sets used by twelve risk adjustment models used primarily in the US. [32-46] Each of the widely-used models tends to be given an acronym, and there are typically many variants of each modeling framework. I use only the acronyms here and not elaborate on the details of each model.

The three earliest risk adjustment models (CI, DCGs, and ACGs) emerged in the US in the 1980's and early 90's using either a diagnoses as reported on insurance claims.[32, 33, 34] Pharmacy information started being used in 1992, and has been an active direction for model estimation in recent years.[35,36,37] Some of the recent models (CRGs, ERGs) combine different information sets, such as pharmacy plus diagnoses.[38, 39] Not shown in the table are alternative models using functional status and disability measures (For discussion and evaluation see Pope et al, [40] and models using self reported health status information, such as the Short Form 36 (SF-36) developed a the New England Medical

Center[41]. Predictive models that use lagged spending information are not included, although this framework is often used. Such models are discussed briefly below. The predictive power of various data sets is evaluated below when discussing evaluation tools.

{Table 1 goes about here}

## Conceptual framework and classification systems for empirical models

The risk adjustment models in Table 1 differ not only in the information set used, but also in how information is used for prediction. Two fundamental approaches are used, a rate cell approach, and linear prediction formulas. In the rate cell approach, each consumer is classified into one unique group, with groups chosen so as to most usefully distinguish patient severity or cost. Consumers are typically assigned to one rate cell based on a complex sorting algorithm in which the most complex or highest cost patients are identified first, and remaining consumers are filtered through a succession of screens until lower cost and less severe individuals are eventually exhaustively assigned to a rate category. The most ambition rate cell approach, the CRG framework, uses over 700 rate cells to classify individuals. A rate cell approach allows the cost of each person in a rate cell to be calculated using a simple weighted or unweighted average of the cost of all consumers in that rate cell.

The primary alternative conceptual approach for estimating risk adjustment models is to use linear or nonlinear additive models, in which a long array of binary or possibly continuous signals are used in a regression framework to predict spending. Some of the additive models use over 200 risk adjusters (signals) for prediction. The predicted cost or service utilization of an individual is then the fitted value of the model for that individual.

Rate cell approaches have the advantage of simplicity to estimate, but suffer from being less predictive than additive models. Even with 700 rate cells, there are simply too many combinations of even the 20 most common conditions ($2^{20} = 1{,}048{,}576$) to be able to easily capture all of the diversity of consumer types. A strength of the rate cell approach can be that for certain conditions, a single rate cell can contain a relatively homogeneous set of consumers to price and evaluate. A weakness is that for many conditions (e.g., asthma, diabetes, mental illness) consumers with this condition as well as other more serious ones will be spread out over multiple rate cells. Additive models can accommodate enormous diversity of patient types, while interactions can capture possible nonlinearities.

Even conditional on the same set of information, (diagnoses or drugs) risk adjustment models differ significantly in how they use this information. Most systems impose hierarchies on the information so that more serious conditions take precedence over less serious conditions. Much of the cleverness, clinical coherence and predictive power are related to the choice of these categories.

Once a rate cell or additive approach, and a classification system have been selected, there remain issues about what time period the information will be used. The US Medicare program and many other risk adjusted payment models use a prospective framework, in which information from a base period (usually one year) is used to predict spending from the subsequent period. The main alternative framework is to use concurrent (sometimes called retrospective) information for prediction, where the diagnoses or other information from a year is used to predict spending or utilization from the same year. The principle argument for prospective models is that it is easier for the sponsor, and only predetermined information can be used to influence selection choices. The argument in favor of concurrent models is that they are much more highly predictive than prospective models.

## Econometric specification issues

Risk adjustment models have been the subject of many refinements on econometric techniques. The large number of zero values, the skewed distributions, heteroskedasticity, and sample selection issues all get frequent attention.

The simplest and most common approach is ordinary least squares in which the dependent variable is untransformed spending. This approach has the great advantage of being fast and simple to estimate, interpret, and explain to non-econometricians. It is the approach used by the US Medicare program. [40, 42] Ash et al. [33] established the concept that to get unbiased estimates for consumers where some have partial year eligibility, then the correct thing to do in an OLS setting is to annualize costs by deflating by the fraction of the year eligible, and then to weight the observation by this same fraction. This weighted annualized regression can be shown to generate unbiased means in rate cells, and corresponding linear regression models.

Manning et al., [47] contains the classic discussion of why OLS can be inefficient in small samples. The two-part log-linear model proposed by Duan et al [48] is the classic discussion of this model, and also develops the smearing correction to ensure that model predictions are not too seriously biased. Subsequent work [49, 50, 51] has highlighted the importance of correcting the mean not only for skewness but also for heteroskedasticity. Fishman et al [52] and Basu et al [53] are recent examples contrasting various refined econometric specifications.

{Table 2 goes about here.}

Despite all of these refinements, OLS remains very popular. Why is this? One reason is that with very large sample sizes, the inefficiency of OLS and concerns about overfitting due to a few very high outliers go away. The large sample sizes and large number of parameters also make some of the nonlinear estimators difficult to estimate using conventional statistical packages. The other reason to prefer OLS is

that for large samples findings will be relatively robust to the econometric specification. Table 1

contrasts five econometric specifications on nine alternative sets of regressors, as developed in Ellis and

McGuire [54]. The first column uses the weighted least squares, including partial year eligibles. The

second estimation approach uses simple OLS, and excludes people with fractional years of eligibility.

This approach focuses on people with the most complete information, and uses a sample that is also used

by the remaining three approaches. The third estimation approach is a heteroskedasticity-corrected square

root model as described in Veazie et al.[55]. The square root of actual covered charges is regressed on

the given set of independent variables, and the squared residuals from this regression are then regressed

on the fitted value, called say G. Heterskedasticity-corrected predictions of spending from this model are

the squared fitted predictions, $G_i^2$, plus the predicted variance for each observation $s^2(G_i)$. The fourth

model uses a "two part OLS with smearing" as described in Manning [56], and Buntin and Zaslovsky

[57]. The first stage is a probit model of the probability of any spending, and the second stage is a linear

model of spending among those with positive spending. The fifth specification uses a generalized linear

model (GLM) evaluated by Buntin and Zaslovsky with the linear portion transformed using the log

transformation, and additive errors assumed to be normally distributed, hence $Y = \exp(X\beta) + \varepsilon$. Ellis and

McGuire [54] also tried estimating two part log models and GLM models with alternative link

and distribution functions however these models either did not converge or had very poor or

negative $R^2$ values. Similar problems with these models were found in Veazie et al [55].

Several results are shown from this table. First, all of the various econometric specifications shown

achieve a very similar $R^2$ for each information set. Conventional risk adjustment models using diagnoses

or lagged spending do much better than age and gender alone, but less well than using disaggregated

spending information. Even if sponsors might never want to use lagged spending signals for payment

purposes, health plans that have this information available can use if for prediction and identification of

profitable and unprofitable consumers given the signals used for risk adjustment. A third and final

observation is that there are diminishing returns to adding more information set as explanatory variables, with the highest prospective $R^2$ achieved in this US Medicare sample at only about 17%.

The last econometric specification issue to be discussed is the appropriate correction for sample selection. Terza [23] has been one of many advocates that risk adjustment models should be corrected for selection bias. Since Health maintenance organizations (HMOs) and unmanaged fee-for-service (FFS) plans have different styles of care, they will tend to attract systematically different types of people who may differ in their tastes for those styles. Moreover, it well predicted by the above theory models that there will be behavioral responses by HMO providers to the incentives of capitation that will cause them to offer different quantities of care than FFS providers. The existence of these selection and moral hazard differences is certain, but it is an empirical question whether these differences are large and seriously bias estimates or whether they are small. The magnitude of these differences is examined below.

## *Methods of evaluation and prediction*

A significant literature on how to compare and evaluate risk adjustment models has developed. Cumming et al [26] in their US study for the Society of Actuaries contains a nice discussion of different metrics and illustrates their use. Ash et al [33] used not only individual $R^2$ as a metric, but also defined and used grouped $R^2$, which is to say how well the models perform for mutually exclusive partitions of an entire sample. Mean absolute deviation (MAD) as well as the standard error of the regression are also common. [33, 26] Finally, the literature often reports what are called "predictive ratios" which is the ratio of predicted to actual spending or utilization[33]. An ideal model would have predictive ratios close to one for every group of possible interest.

There is a decided danger of overfitting with econometric models given how skewed both the dependent and the independent variables are. To evaluate the degree of overfitting, researchers often compare the ordinary R2, MAD and predictive ratios not only for the estimation sample, but also with a reserved split

sample not used for model development or estimation.  The validated $R^2$ and other measures of predictive power is often significantly lower, (5-50 % lower), than the development sample $R^2$.

The validated $R^2$ from three different research studies are shown in Figure 9 and Figure 10. The first set of bars are from Berlinguet [58] who studied three diagnosis-based risk adjustment models (ACGs, DCGs, and CRGs) in three provinces of Canada.  The second set of bars are from Wasem et al [59] who report on their comparison of seven diagnosis- and pharmacy-based risk adjustment models in Germany. The final set of bars is from Cumming et al [26] who evaluated seven diverse risk adjustment models for the US Society of Actuaries.  The figures reveal several things. First, there are distinct differences among the risk adjustment models in terms of their predictive power.  Second, the models using combinations of information, such as the CRG the RxGroups_IPHCC and the ERG models tend to do better than models that use only one type of information. Third, concurrent models (Figure 10) do much better than prospective models.

{Figure 9 goes about here.}

{Figure 10 goes about here.}

## An Extended Example

Every country that has implemented more elaborate risk adjustment using diagnoses or pharmacy information started out by using only demographic information such as age and gender. It is useful to see an example of how risk adjustment can improve upon age and gender prediction.  Figure 11 illustrates the pattern of average annual spending by one year age groups from a sample of 14.6 million privately insured individuals in the US.  This figure was generated at DxCG using the MEDSTAT Marketscan data from 2004.  This figure is unusual for the US in that I have pooled together the commercial under age 65

sample with the MEDSTAT Medicare sample, which is a sample of primarily privately-insured individuals over age 65. Some in this Medicare eligible sample are still working, some are retirees over age 65 still covered by their employer, and some are in Medicare supplementary policies. Numerous patterns will strike the reader, but several deserve note. First, costs increase with age, although not monotonically, and a simple linear or quadratic age specification cannot hope to capture the complex shape. Second, women of child bearing age are distinctly more expensive than males of the same age, but still relatively low cost relative to the elderly. Third, males over age 65 are more expensive than females. Finally, there appears to be a tapering off in average spending at about age 85 years.

{Figure 11 goes about here}

{Figure 12 goes about here.}

Figure 12 is similar to Figure 11 except that it focuses on under age 65 population and instead of showing the graph by gender, it shows the curve for five types of health plans. Each of the lines is drawn with at least 100,000 enrollees, so the patterns are highly stable across age groups. The figure suggests that enrollees in each of the five plan types are relatively similar up until age 40, at which time there is a divergence of HMO and point-of-service (POS) with capitation plans from the rest. By age 40 there is about a 20 percent discrepancy between the HMO and POS with capitation plan from the other three plan types. Is this difference due to selection of healthier people conditional on age, to taste differences, or due to supply side moral hazard response to incentives?

{Figure 13 goes about here.}

Figure 13 provides one answer to this question. Instead of plotting actual spending by age and plan type, the figure shows the risk adjusted spending by age and plan type. Risk adjustment in this example was done using the DCG/HCC concurrent risk adjustment model. The five lines are much closer together, differing by less than 5% across plan types. There is still a difference at age 40, but it is much smaller. This modest difference once spending is risk adjusted suggests that most of the 20 percent observed difference is due to selection differences, not taste, or moral hazard. The implied cost savings from the HMO and POS with capitated from the most common plan type of preferred provider organizations (PPOs) is less than five percent, with a modest gradient upward with age.

It is commonplace for others to be critical of risk adjustment model developers because they ignore sample selection when estimating their models. The above graphical results suggest that sample selection corrections will result in relatively modest errors in the estimated models, errors that are much smaller than differences across age, disease, or gender. Adding in a few more interaction terms to more accurately predict spending may be as valuable as incorporating selection effects into the original models.

## Country experience with risk adjustment

This section discusses the experience of six countries that have use risk adjustment in different ways. Table 3 summarizes some of the key features of segments in each of these countries. The systems implemented are by no means uniform, and there are numerous differences in the number of agents, whether competing health plans or regional orientation is used. Discussion of the history and use of risk adjustment in each country is provided below.

{Table 4 goes about here}

## Canada

Canada does not offer multiple competing health plans in any of its provinces, and hence one might expect the provinces to have minimal incentives for selection. Despite this, a modest selection problem still exists, arising from consumer choice of residence, and consumer choice of providers. Canada is also interesting in that its health care system is similar to many other countries with a social insurance program, including countries as diverse as Australia, France, Norway, and Taiwan. Health system financing, publicly-funded services and delivery systems in Canada vary across the provinces. While there is some variation in financing systems across provinces, there are many similarities. I focus here primarily on Alberta.

{Figure 14 goes about here.}

A stylized view of Alberta Canada's health care system using the previously discussed four agents in health care markets is depicted in Figure 14. All residents are automatically covered by the province in which they reside, and payments are collected from all workers and employers through a mandatory social insurance premiums and general income taxes. The province's ministry of health, Alberta Health and Wellness (AHW), pays for most physician and office-based services using a fee schedule managed by the Alberta Medical Association. In addition to using fees, the provincial government also allocates annual budgets to 9 (previously 17) regional health authorities (RHAs). RHAs are responsible for spending on hospitals, continuing care facilities, home care, public health, community rehabilitation, mental health and a portion of cancer care, which together comprises 56 percent of the total AHW spending.[60] Prescription drugs are not included as a benefit for all residents, but are covered for those over age 65 with a co-pay through an arrangement with Alberta Blue Cross. In addition to this publicly funded system, consumers as individuals or with sponsorship from their employers, are allowed to purchase supplementary insurance that covers pharmacy costs and certain uncovered services.

31

Two issues create an interest in risk adjustment in Alberta: in its allocation among the RHA, and in its payments to primary care providers. The allocation of AHW funds among the RHA is complicated by the fact that residents do not choose where to live randomly, and need not necessarily seek care in the RHA in which they live. Patient choice of facility implies a need to ensure fair payments to each facility given the case mix of patients that they manage, treat or refer.

AHW has explored this geographic selection bias, and finds evidence of it [61]. In particular, they find persons with more serious chronic illnesses are more likely to live in urban areas, and that certain diseases vary significantly across RHAs. Table 3 provides evidence by comparing summary statistics across RHA, where we have chosen the highest and lowest rates observed. (Sample sizes in each RHA exceed 20,000 and hence these differences are all statistically significant at high levels of confidence.)

{Table 3 goes about here.}

Geographic variation within Alberta, in terms of urban versus rural, and distances from rural areas to urban hospitals are probably greater than in many other countries (such as the Netherlands) that are more homogeneous in their access. The Alberta experience reminds us that consumers do sort themselves geographically in meaningful ways. We should emphasize that the Alberta regional allocation formula does take into account the age distribution of the population in each RHA. We take the differences in age across the RHAs not as a problem with unpriced risk heterogeneity by itself, only as a signal that consumers are sorting themselves geographically, and that this creates potential selection problems for Alberta.

## Role of risk adjustment in Canada

Only Alberta and Saskatchewan appear to be using population-based capitation formulas in allocating funds to health authorities. RHA budgets in Alberta are currently allocated using demographic risk adjuster information. A rate cell approach is used in which the rate cells are based on age, gender, and

four socioeconomic categories that reflect income and "aboriginal" (i.e., Native American) status.   AHW

has considered methods that more accurately measure health risk in order to improve fairness of

compensation regions and providers for seriously ill, high expenditure patients.  The Calgary Regional

Health Authority in Alberta was one of the participants in the Berlinguet et al [58] study of diagnosis-

based risk adjustment models, funded by the Canadian Health Services Research Foundation (CHSRF),

discussed above.  AHM also conducted its own evaluation of risk adjustment models from 2002-2004 but

has not yet moved beyond using only demographic information for budget allocations to RHA.  Recent

initiatives emphasizing primary care and coordination among diverse providers also highlight the role of

risk adjustment.

A recent review found British Columbia to be the only province using diagnosis-based risk adjustment for

paying physicians, although many alternative payment formulas are being used. [62]. Risk adjustment

was tested as part of a demonstration project on primary care that started in British Columbia in 1998.

Reid et al [63, 64] evaluated diagnosis-based risk adjustment model using the Adjusted Clinical Groups

(ACG) classification system on data from Manitoba and British Columbia. British Columbia uses a

capitation formula: "a physician funding methodology based on the patient population served, as opposed

to the number of services provided". The B.C. Ministry of Health received $9.6 million from the Federal

Health Transition Fund for a primary care demonstration project to explore new and innovative

approaches for paying for primary health care delivery.  The formula uses the ACG case-mix system to

calculate payments that are based on each consumers' age, sex and diagnoses codes.[65]

## USA Medicare

The USA Medicare program was one of the earliest adopters of capitated payments, and provides a useful

case study for other countries. This program covers nearly 40 million individuals who are either aged,

disabled or have end-stage renal disease (ESRD).  Prior to 1985, when "at risk" HMOs were first

permitted, the traditional indemnity Medicare program looked similar to the Canadian system, with a

government sponsor raising revenue from taxes and insurance premiums, fully insuring geographically defined insurance carriers (health plans) who were contracted to pay services mostly on a fee-for-service basis. There was no incentive for traditional Medicare plans to control costs.

Legislative changes to Medicare were adopted in 1985 to encourage cost containment by encouraging competing managed care health plans. These reforms permitted Medicare HMOs to receive capitation funding and be "at risk" for the cost of their enrollees. The Medicare Managed Care program, now called Medicare Advantage (MA) enrolled about 5.6 million (15%) of the Medicare population in 2002, and since then has had a relatively stable enrollment. Health plans participating in this program are closely regulated in terms of the benefits they can offer and premiums they can charge. Open enrollment is required, but MA plans are allowed to compete in many other ways. For example, health plans have the right to market directly to consumers, or to locate their offices in more or less convenient sites. MA plans are also given the right to choose counties to enter or exit from, to choose additional benefits (such as drug coverage) not offered by indemnity Medicare, to choose the enrollee premium, and to selectively contract with providers. From the start, payments to the MA plans by the government were risk adjusted to reflect the county, age, gender, disability, and institutional status of the health plan's enrollees, using a formula called the Adjusted Average Per Capita Cost (AAPCC).

## Early concerns about biased selection

Since its inception, policymakers have been concerned about whether risk adjustment using the "AAPCC," which uses only demographic information, was sufficient to reduce selection incentives and avoid overpayment of managed care plans. Early evidence clearly indicated that the HMOs were attracting healthier than average enrollees even within each rate cell. An important study by Brown et al [66] concluded that rather than saving money, as intended, the Medicare managed care program was actually costing the Medicare program 5.7 percent more than it would have been if Medicare offered FFS alone. The Brown et al analysis analyzed linked survey and claims information and found that HMO

34

consumers were relatively satisfied with their HMOs, there were no clear quality differences, and that HMOs successfully reduced the use of certain resources, such as inpatient days. Several further government and academic studies further documented risk selection problems with the program.

The major concern in the US about managed care plans is that selective contracting permits plans to distort services and provider availability in ways that encourage favorable selection. This level of service distortion is the primary focus of Glazer and McGuire [12, 11] and reflects a concern expressed in Newhouse [15] and elsewhere. Empirical evidence on the nature of this selection is limited. Cao and McGuire [67] use Medicare FFS claims to detect that rates of spending on certain chronic diseases and certain services are higher in the FFS sector when a higher proportion of individuals are enrolled in HMOs, suggesting service and provider type distortions as predicted by the theory. Recent evidence of a different selection activity is provided in Dallek and Dennington [68] who find that Medicare managed care plans had primary care physician turnover rates averaging 14 percent, with rates over 20 percent in five states. These extremely high turnover rates on primary care physicians must disrupt continuity of care, discouraged continued enrollment by those who are more seriously ill.

## Risk adjustment in the US Medicare program

The US Medicare program reacted to the evidence of biased selection in their M+C program in several ways. During the early and mid 1990's the program funded five major studies of different risk adjustment models, using a wide range of approaches that included inpatient diagnoses, all encounter (inpatient and outpatient) diagnoses, survey methods, risk sharing models, and prior year spending models.[69, 70, 71] Legislated changes in 1997 mandated a transition to health status based risk adjustment. Starting in 1998, the Medicare program slowed the rate of increase in payments, largely eliminating the perceived overpayment, and changing the mechanism for calculating each counties average payment. Medicare also began laying the foundation for implementing risk adjustment by requiring that the M+C plans provide inpatient diagnoses starting in 1998, and all encounter diagnoses in 2000.[70, Ch 3] In January 2000, the

Medicare program implemented the Principal Inpatient Diagnostic Cost Group (PIPDCG) model and used it for payment of 10 percent of the total amount to MA plans. Pope, et al [40] provide a careful description of the framework and the effort that went into trying to reduce the incentives to distort treatment and diagnoses and "upcode" patient severity. Rather than jumping from a demographic to a diagnostic based risk adjustment formula in one year, implementation called for the gradual phasing in of the new formula over seven years. While initially only 10% of the capitated payments were based on the diagnoses, in 2007 all of the risk adjusted payment is diagnosis-based.

While initially only inpatient diagnoses were used, since 2004 the payment formula relies on all diagnoses, and hence is called an all encounter model. The Medicare program decided to use a relatively simplified classification system called the CMS-HCC model, a simplified version of the Diagnostic Cost Group/Hierarchical Condition Categories (DCG/HCC).[6, 42, 43,] Kanika Kapur, a researcher at the Rand Corporation summed up the reasons for the choice of this model as "CMS chose the DCG/HCC model for Medicare risk adjustment, largely on the basis of transparency, ease of modification, and good clinical coherence."[72]

The CMS-HCC model is a prospective model, so that demographics and diagnoses from a base year are used to predict payment for a given individual the following year. Payments are calculated prospectively each January, using data from the previous calendar year. Retrospective adjustments are used to reflect late arriving data relevant to payment. To simplify the data burden on health plans, instead of requiring that the full set of all 15,000 ICD-9-CM diagnosis codes be provided, approximately 3,300 valid diagnoses are clustered into 64 disease groups (HCCs). Regressions were used to generate cost weights on each of these disease groups, together with demographic categories, selected disease interactions, and age*disease interactions. These HCC cost weights are cumulative, in that predictions are the sum of the contributions of each factor rather than cost weights on mutually exclusive categories. Five variants of the basic model are used, with separate payment weights for continuing and new Medicare enrollees,

ESRD enrollees, long-term institutionalized, and participants in "specialty organizations" (which also use self-reported frailty measures).

Data from managed care plans paid using the CMS-HCC formula have not yet been available to independent researchers for analysis, so it is early to assess the impact and success of the new risk adjustment formula. The primary concern when it was implemented was that health plans would game the system and increase the coding intensity of enrollees. To offset this, Medicare reduced predicted payments by five percent for each of the first three years. For 2007 they have announced that only a 2.9 percent offset will be used, suggesting that the coding escalation was not as significant as originally feared. For 2007 a newly recalibrated risk adjustment formula is also being used that updates the cost weights using more recent data.

## USA Privately Insured

The USA privately insured population has an extremely complex set of institutions providing health care. No overview can possibly capture its full complexity, although Cutler and Zeckhauser [73] provide a nice summary. Figure 15 illustrates how the market the five classes of agents interact in this market. The sponsor in almost all cases is the employer. In some states, such as New Jersey, the state sponsors individual and small firm coverage by creating a separate insurance pool to subsidize this group. Other states mandates require that all insurers offer individual or small group coverage at a community rate, either independently or as part of a larger pool. This effectively forces health plans to become sponsors for this type of consumer. More commonly, employers in the US get to choose whether to offer insurance or not, and in only a few states is there a tax or penalty to employers who choose not to offer insurance.

{Figure 15 goes about here.}

Employers get to choose from a large number of competing health plans. Long ago, most consumers enrolled in indemnity plans who offered coverage for any services desired by plan enrollees subject to

deductibles and copayments. Most consumers are now in some form of managed care, often called health maintenance organizations (HMOs) in each geographic area. Numerous further refined health plan contracting forms, such as preferred provider organizations (PPOs), exclusive provider organizations (EPOs) and point of service (POS) health plans differ in how much choice the consumer is given, the levels and variation in demand-side cost sharing, and the tightness of the provider panel that is included. Overall, about half of all employees have a choice of health plans through their employment.

The USA is different from Canada and most European countries in that there is much less emphasis on equity goals and weaker efforts to equalize access to health care. Health care is not uniformly viewed as a merit good to which all are entitled with the same level of access. Instead, freedom of choice and honoring individual heterogeneity of tastes is revealed to be highly valued. While there is certainly a great deal of dissatisfaction with the current system and its enormous inequities in access, leaders remain reluctant to introduce major reform.

Formal risk adjustment is used very rarely by private employers [10]. Glazer and McGuire [11] highlight that employers are generally able to negotiate premiums with health plans or have experience rated premiums that implicitly risk adjusts the payments for the riskiness of the enrollees. Also, while formal risk adjustment is only rarely used for payments from employers to health plans, it is widely used for contracting between health plans and provider groups, or for calculating capitated payments to doctors, hospitals and other provider groups.

One new area in which risk adjustment is increasingly being used in commercial health plans is in calculation of provider quality measures and "pay for performance" measures. Increasingly providers are being offered bonuses for holding down hospitalization rates, achieving targets for vaccinations and screening, and holding down costs. Recent experience in the USA with bonus and performance-reward payments systems and find mixed evidence of its effect.[74] Kapur [72] discusses to the need for careful risk adjustment to avoid creating undesirable selection incentives under pay-for-performance programs.

## The Netherlands

The Netherlands health care system has two broad systems: a compulsory insurance system of "sickness funds" with careful sponsorship and regulation by the government, and a voluntary, private insurance system that is much less tightly sponsored and regulated. Eligibility for the two systems is based on income, with income thresholds varying across population subgroups (employees, self-employed, elderly). A person who is eligible for the compulsory system is not eligible for the voluntary system, and vice versa. The structure of the compulsory system is shown in Figure 16. The key decision-making agents in the Dutch health care system are consumers, who enrol in health plans as individuals (plus non-employed partner and non-employed children); employers and a central insurance fund, which serve as sponsors; and 25 non-profit sickness funds (health plans).

{Figure 16 goes about here.}

Since the early 1990s there has been a slow movement toward managed competition in the Netherlands. The key responsibilities of paying and contracting with providers has being gradually transferred from the central government to competing risk-bearing sickness funds. The bulk of the health care financing comes through payroll taxes, levied on earned income, shared between the employee and their employers. A Central Insurance Fund uses prospective formal risk adjustment to reallocate money among the funds. Although in 1992 only age and gender were used to calculate payments, there was a gradual increase in the use of other signals, as described in van de Ven et al, [75,76]. Along with improved risk adjustment, greater risk has also been imposed on the sickness funds. Whereas originally only 3% of the risk was passed on to health plans, in 2003 the sickness funds' financial risk averaged 52%.

The Netherlands is notable for its gradual enhancements to the risk adjustment formula used. In 1992 only age and gender were used. The following new risk-adjusters were added over time: region, being an employee (yes/no), disability (all in 1995), age/disability (in 1997), and Pharmacy-based Cost Groups

(PCGs) (in 2002). In 2004 Diagnostic Costs Groups (DCGs) using only inpatient diagnoses were added. Both systems are relatively simple with only 13 categories of drugs and 9 categories of diagnoses used. Despite this, PCGs and DCGs are used together to predict an individual's subsequent year expenses. They substantially reduce predictable profits and losses, and reduce the incentives for selection. [76]

## *Germany*

Germany has both a social health insurance (GKV) system that covers about 90 percent of the population and a private health insurance system (PKV) that is only available to high income consumers.[77]   Figure 17 highlights the structure of the social health insurance system which is funded primarily by a payroll tax that varies across health plans, called "sickness funds".  Whereas prior to the health insurance reforms in 1996, employers played an important role as sponsors, since then consumers have been allowed to make choices of health plans without regard to their employment.  Consumers make choices as family units, not individuals.  Employers contribute half of the payroll tax payments for health insurance, but are not allowed legally to restrict or influence the choices of their employees.

{Figure 17 goes about here.}

There are over 200 competing sickness funds (health plans), broadly organized into seven classes, each of which has its own history and association.[78]  For the most part the origins of these plans are unimportant for a high level overview.  It is significant that this different history means that each plan tends to have a systematically different group of consumers, so that risk differences across plans are large, even without intentional selection efforts.  There is just beginning to be a small amount of selective contracting by health plans with providers.

Risk adjustment is used in Germany at three stages.  The primary use is by the sponsor (the central insurance fund, to reallocate funds among competing sickness funds.  In addition capitation payments are

used in transferring money to the regional KV associations of physicians. Capitation formulas are also currently used in calculating ceilings on total physician payments which are often binding.

Since 1994 a central insurance fund, that is external to the rest of the financing system and hence cannot endure any deficits, has reallocated money between sickness funds. The central insurance fund acts as if it had received the income-based payments directly from all of the employees and employers, and pays all of the health plans a needs-based amount that is determined by a demographic risk adjuster. Payments for each of 670 risk cells, defined by age, gender, disability pension status, and entitlement to sickness allowances. Buchner and Wasem [79] note that the system equalizes about 92 percent of the income variation, and that about 70 percent of the redistribution is due to morbidity differences across plans.

There is relatively little evidence of selection problems in Germany, perhaps because of its structure. While there are many plans, they mostly have very similar benefits and identical provider networks. Buchner and Wasem highlight that KV local sickness funds remove much of the risk from the sickness funds, and ensure uniform provider incentives. [79] The KV local sickness funds transfer financial risk back onto providers through their ex post pricing system. There is enormous dissatisfaction with this mechanism among physicians, so reforms to this payment approach are a priority.

One empirically validated selection problem in Germany is the difference in characteristics between those choosing private insurance and those choosing the social insurance program. Buchner and Wasem [79] note that among high income workers who have the choice, individuals are much more likely to choose private insurance while families are more likely to choose social insurance. Hence the average number of persons per contract is 1.18 in the private insurance system (which charges a fixed premium) versus 2.22 in the social insurance system (which collects the premium as a percent of income). While segmentation of the wealthy, healthy and single contracts off from the rest of the population is not a major problem for their system, it does reduce the solidarity contribution (cross subsidy) from these low cost individuals to the rest of the population.

The German government passed legislation in 2002 mandating that the government move to a more comprehensive risk adjustment formula in 2009. Legislation enacting this mandate passed in early 2007. The existing payment formula between the Central Insurance Fund sponsor and the sickness funds will be revised to use a diagnosis and pharmaceutical information, where the choice of indicators is constrained by legislation to include only those conditions costing more than 50% above the average. The reform bill also changes how payments to physicians will be made, with a reduced emphasis on local geographic organizations of physicians. Currently these agents receive payments that are calculated using only the characteristics of employees in each region, and even basic demographic information such as the number of children, people with chronic conditions, and the unemployed are not reflected. Risk adjustment in Germany may also start to be used to calculate more refined payments to office-based physicians. As of 2006, physicians received fee-for-service reimbursements subject to fixed caps by specialty on allowed total revenue. Under the recent reforms, payments to primary care providers will change to use a quarterly fee for each patient seen, a step towards capitated payment for office-based care, although currently proposed to be done without any risk adjustment. The new structure of Germany's provider payment system and health plan choice will increase the role of risk adjustment for many uses.

## Lessons from country experience

While summarizing the experience of many diverse countries is difficult, a few themes are worth highlighting. One is that countries differ dramatically in the settings and problems that the face and their reasons for using risk adjustment. It is important to recognize these differences, and adjust the objective functions for optimal risk adjustment appropriately. Achieving equity goals and ensuring efficient payments when the initial levels of per capita spending are inefficient both may require corrections to the formulas generated using conventional risk adjustment.

Every country started out with a relatively simple risk adjustment formula, such as one based on only age and gender. Many have progressed over a period of years moved from a demographic only risk

adjustment to diagnosis (US, Netherlands, and Germany) or pharmacy based groups (Netherlands). Many countries are considering their next steps, but the interest and adoption of capitation-based payment seems to be accelerating in many countries.

The US Medicare program is an example of adjusting the overall level of risk adjusted payments to avoid overpaying its capitation. Both under the demographic and the diagnosis based groups, the predicted spending was reduced by five percent to reflect differences between the capitated and non-capitated spending levels. This is not the same as optimal risk adjustment that might take into account more sophisticated coding creep and other forms of strategic behavior, but it is a step in that direction, and may be all that is feasible given limited information available to policy makers. The research community could help by conducting more research that will inform policy-makers about the nature of strategic behavior and the relevant parameters to the cost and provision process.

## Conclusions

The literature on risk adjustment is diverse and rapidly growing. It is impossible to summarize the extensive literature that has been reviewed here but it may be useful to mention a few common themes and directions for future research. The major new direction for risk adjustment payment models is to incorporate richer and more explicit objective functions than to just maximize the predictive power. Optimal risk adjustment is an important concept, not only for theoretical research, but also for policy applications of risk adjustment. The uses to which a model will be put should have an important role in the design and estimation of the models.

Risk adjustment models and other predictive models are increasingly being used for diverse, non-payment purposes. Controlling for quality[81], detecting performance improvements,[82] ranking and rating providers,[83] measuring selection incentives [54, 84, 85]and identifying and managing patient at greatest risk of health deterioration are all important uses of the frameworks that have been used here. There is

relatively little academic work on either the conceptual basis or empirical development of models for these more practical uses, and it will be interesting to follow the economics and health services research literature that explores these topics over the coming years.

# References

1. Van de Ven, W., Ellis, R. P., (2000) Risk adjustment in competitive health plan markets. In: Culyer, A.J., Newhouse, J.P.(Eds.), Handbook of Health Economics. North-Holland.
2. Glazer, J., McGuire, T.G., (2006) Optimal risk adjustment, in Jones, A.M., (ed.) The Elgar Companion to Health Economics. Edward Elgar Publishing Co., Glos UK.
3. Glazer, J., and McGuire, T.G. (2007) (chapter in this book)
4. Iezzoni, L.I. (2003) Risk adjustment for measuring health outcomes. Fifth edition. Ann Arbor, MI: AcademyHealth HAP.
5. Kronick, R., and Dreyfus, T. (1997) The Challenge of Risk Adjustment for People with Disabilities: Health-Based Payment for Medicaid Programs. Princeton, NJ.Center for Health Care Strategies.
6. Ash, A.S., Ellis, R.P., Pope, G.C., Ayanian, J.Z., Bates, D.W., Burstin, H.,  Iezzoni L.I., McKay, E., and Yu, W. (2000) Using diagnoses to describe populations and predict costs. Health Care Fin Rev, Spring 21(3): 7-28.
7. Hurley J., Hutchison B., Giacomini M., Birch S., Dorland J., Reid R., Pizzoferrato G., (1999) Policy Considerations in Implementing Capitation for Integrated Health Systems. Canadian Health Services Research Foundation, June.
8. Hutchison B., Hurley J., Reid R., Dorland J., Birch S., Giacomini M., Pizzoferrato G., (1999) Capitation Formulae for Integrated Health Systems: A Policy Synthesis. Canadian Health Services Research Foundation, June.
9. Thomas, J. W.,  Grazier, K.L. and Ward K. (2004a) Comparing Accuracy of Risk-Adjustment Methodologies used in Economic Profiling of Physicians. Inquiry  41, 218-231.
10. Keenan, P.S., M.J. Beewkes Buntin, T.G. McGuire, J.P. Newhouse (2001) 'The prevalence of Formal Risk Adjustment.' Inquiry 38,  245-259.
11. Glazer, J., and McGuire, T.G., (2001) Private employers don't need formal risk adjustment. Inquiry 38, 260-269.
12. Glazer, J., McGuire, T.G., (2000) Optimal risk adjustment of health insurance premiums: an application to managed care. Am Econ Rev 90 (4), 1055–1071.
13. Jack, W. (2006) Optimal risk adjustment in a model with adverse selection and spatial competition. *J Health Econ. 25,* 908-926.
14. Biglaiser, G., Ma, CT.A., (2003) Price and quality competition under adverse selection: market organization and efficiency. RAND Journal of Economics 34 (2), 266–286.
15. Newhouse, J., (1996) Reimbursing health plans and health providers: efficiency in production versus selection. J Econ Lit 34, 1236–1263.
16. Shen, Y., Ellis, R.P., (2002) Cost minimizing risk adjustment. J Health Econ 21, 515–530.
17. Kronick, R, Gilmer, T, Dreyfus, T, and Lee, L, (2000) Improving Health-Based Payment for Medicaid Beneficiaries: CDPS Health Care Fin Rev 21, 29-64
18. Ozminkowski, RJ, Smith, MW, Coffey, RM, Mark, TL,., Neslusan, CA,, and Drabek, J (2000) Private Payers Serving Individuals with Disabilities and Chronic Conditions. Report prepared under contract #HHS-100-95-0044 between the U.S. Department of

Health and Human Services (HHS), Office of Disability, Aging and Long-Term Care Policy (DALTCP). at http://aspe.hhs.gov/daltcp/reports/privpay.htm

19. Frank, R., Glazer, J., McGuire, T.G., (2002) Measuring adverse selection in managed health care. J Health Econ 19, 829–854.

20. Centers for Medicare and Medicaid Services (CMS). (2006) Announcement of Calendar Year (CY) 2007 Medicare Advantage Capitation Rates and Medicare Advantage and Part D Payment Policies, April. Downloaded on August 23, 2006. http://www.cms.hhs.gov/MedicareAdvtgSpecRateStats/Downloads/Announcement2007.pdf

21. Ellis, Randall P., and Thomas G. McGuire. (1986) Provider Behavior Under Prospective Reimbursement: Cost Sharing and Supply. J Health Econ 5,129-152.

22. Lee, L.F. (1983)  Generalized Econometric Models with Selectivity. Econometrica 51, 507-512.

23. Terza, J.V. (2002) Alcohol abuse and employment: a second look. J Appl Econometrics 17, 393-404.

24. Glied SA. (2000) Managed Care, in Culyer AJ , Newhouse JP (eds) *The Handbook of Health Economics,*. North Holland: Amsterdam.

25. Ma, Ching-to Albert, (2003) Managed care and shadow price. Health Econ Let 7, 17–20.

26. Cumming, R.B., Knutson, D., Cameron, B.A., Derrick, B. (2002) A Comparative Analysis of Claims-based Methods of Health Risk Assessment for Commercial Populations. Research study sponsored by the US Society of Actuaries.

27. Thomas, J. W.,  Grazier, K.L. and Ward K. (2004b) Economic Profiling of Primary Care Physicians: Consistency among Risk-Adjusted Measures. Health Serv Res 39, 985-1003.

28. Newhouse, J., (2002) Pricing the Priceless: A Health Care Conundrum. MIT Press, Cambridge MA.

29. Leger-Thomas, Pierre (2007) (Chapter in this book.)

30. Rice, N. and Smith, P. (2001), Capitation and risk adjustment in health care financing:an international progress report, Milbank Q 79, 81-113.

31. Glazer, J., McGuire, T.G., (2002) Setting health plan premiums to ensure efficient quality in health care: minimum variance optimal risk adjustment. J Pub Econ 84, 153–175.

32. Charlson ME, Pompei P, Ales KL, MacKenzie CR. (1987) A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chron Dis 40:373-83.

33. Ash, A.S., Porell, F., Gruenberg, L., et al. (1989) Adjusting Medicare capitation payments using prior hospitalization data, Health Care Fin Rev, 10(4):17-29.

34. Weiner JP, Starfield BH, Steinwachs DM, Mumford LM. (1991) Development and application of a population-oriented measure of ambulatory care case mix. Med Care 29, 453-472.

35. Von Korff M, Wagner EH, Saunders K. (1992) A chronic disease score from automated pharmacy data. J Clin Epidemiol 45, 197-203.

36. Zhao, Y,  Ellis, R.P., Ash, A.S., Calabrese, D., Ayanian, J.Z, Slaughter, J.P., Weyuker L., and Bowen, B. (2001) Measuring population health risks using inpatient diagnoses and outpatient pharmacy data. Health Serv Res 26, 180-193.

37. Fishman PA, Goodman MJ, Hornbrook MC, Meenan RT, Bachman DJ, O'Keeffe Rosetti MC. (2003) Risk adjustment using automated ambulatory pharmacy data: the RxRisk model. Med Care 41, 84–99

38. Averill RF, Goldfield NI, Eisenhandler J, Hughes JS, Shafir BV, Gannon DE, et al. (1999) Development and evaluation of clinical risk groups (CRGs). Wallingford, CT: 3M Health Information Systems.
39. Symmetry Health Data Systems, Inc. (2001). Episode risk groups: ERG user's guide. Phoenix, AZ: Symmetry Health Data Systems, Inc.
40. Pope, G.C., R.P. Ellis, A.S. Ash, C.F. Liu, J.Z. Ayanian, D.W. Bates, H. Burstin, L.I. Iezzoni, M.J. Ingber (2000) Principal Inpatient Diagnostic Cost Group Models for Medicare Risk Adjustment Health Care Fin Rev 21, 93-118
41. Ware, J. E and Gandek B. (1998) Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project, J Clin Epi 51, 903-912.
42. Pope, G.C., R.P. Ellis, A.S. Ash, J.Z. Ayanian, D.W. Bates, H. Burstin, L.I. Iezzoni, E. Marcantonio, B. Wu, (2000) Diagnostic cost group hierarchical condition category models for Medicare risk adjustment. Final Report to Health Care Financing Administration, December.
43. Ellis, R.P., G.C. Pope, L.I. Iezzoni, J.Z. Ayanian, D.W. Bates, H. Burstin, and A.S. Ash (1996) Diagnosis-Based Risk Adjustment for Medicare Capitation Payments. Health Care Fin Rev 12, 101-128.
44. Kronick RT, Dreyfus, T, Zhou Z. (1996) Diagnostic risk adjustment for Medicaid: the disability payment system. Health Care Fin Rev 17, 7-33
45. Hornbrook MC, Goodman MJ, Fishman P, Meenan R, Greenlick MR., (1998) Development and estimation of a global risk adjustment model. Abstr Book Assoc Health Serv Res Meet. 15: 227.
46. Carter, G.M., Bell, R.B., Dubois, R.W., et al. (1997) A Clinically Detailed Risk Information System for Cost. DRU-1731-1-HCFA. RAND. Santa Monica, CA.
47. Manning, W.G. et al. (1987) Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment, Am Econ Rev 77, 251-77.
48. Duan, N., et al, 1983. A Comparison of Alternative Models for the Demand for Medical Care, J Bus & Econ Stat 1, pages 115-26.
49. Mullahy, J. (1998) 'Much ado about two: reconsidering retransformation and the two-part model in health econometrics', J Health Econ 17, 247-282.
50. Manning, W.G., A. Basu, and J. Mullahy, (2005) Generalized Modeling Approaches to Risk Adjustment of Skewed Outcomes Data, J Health Econ 24, 465-488.
51. Manning, WG and Mullahy J. (2001) Estimating log models: To transform or not to transform? J Health Econ 20, 461-494.
52. Fishman, P. Sloan, K. Burgess J., Jr, Zhou C. , Wang, L. (2006) Evaluating Alternative Risk Assessment Models: Evidence from the US Veteran Population, Group Health Center for Health Studies working paper, May, 2006, available at: http://www.centerforhealthstudies.org/ctrstaff/fishman.html
53. Basu A, Manning WG and Mullahy J. (2004) Comparing alternative models: Log vs. Cox Proportional Hazard. Health Econ 13, 749-765.
54. Ellis, R.P. and McGuire T.G.. (2007). Predictability and predictiveness in health care spending. J Health Econ 26. 25-48.
55. Veazie PJ, Manning WG, and Kane RL. (2003) Improving risk adjustment for Medicare capitated reimbursement using nonlinear models. Med Care 41, 741-752.
56. Manning, W.G. (1998) 'The logged dependent variable, heteroskedasticity, and the retransformation problem', J Health Econ 17, 283-296.

57. Buntin, M.B., and Zaslavsky, A.M. (2004) Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures J Health Econ 23, 525-542.

58. Berlinguet, M, Preyra C, and Dean, S. (2005) Comparing the value of three main diagnosis based risk adjustment systems (DBRAS). Canadian Health Services Research Foundation (CHSRF) report.

59. Wasem, J, Lauterbach, L.M., and Schrader, W.F. (2006) Klassifikationsmodelle für Versicherte im morbiditätsorientierten Risikostrukturausgleich. (Classification models for risk adjustment in the morbidity-oriented risk structure reconciliation.) Wissenschaftliches Institut der AOK (WIdO). Downloaded on 8/17/2006 at http://wido.de/fileadmin/wido/downloads/pdf_ggw/wido_ggw_aufs1_0205.pdf

60. Alberta Health and Wellness. (2003) Health and Wellness http://www.health.gov.ab.ca/.

61. Chowdhury, T., Fatoo, H., Friesen, D., Yang, Q.and Kramer, G. (2003) An Evaluation of Risk Adjustment Groupers in a Publicly Funded Health System: A Case Study Using Alberta Data, Alberta Health and Wellness. iHEA Poster Presentation, June.

62. Canadian Institute for Health Information. (2005) The Status of Alternative Payment Programs for Physicians in Canada. http://dsp-psd.pwgsc.gc.ca/Collection/H115-13-2003E.pdf

63. Reid, RJ, MacWilliam L., Verhulst L., Roos N, Atkinson N. (2001) Performance of the ACG case-mix system in two Canadian Provinces. Med Care 39, 86-99.

64. Reid, R.J. Roos, N. P., MacWilliam, L., Frohlich, N., and Black C., (2002) Assessing Population Health Care Need Using a Claims-based ACG Morbidity Measure: A Validation Analysis in the Province of Manitoba. Health Services Research 37, 1345–1364.

65. Lu, M, D.G. Moores, Reid, R. Woodhead Lyons S.C. , and Zhang, J.X.. (2002) A Review of Capitation in Alberta: A Report to the APP Subcommittee of Alberta Health & Wellness and the Alberta Medical Association. 1-77.

66. Brown, R., Bergeron, J.W. and Clement, D.G. (1993) "Do health maintenance organizations work for Medicare?". Health Care Fin Rev 15, 7-24.

67. Cao, Z. and T.G. McGuire (2002), Service-level selection by HMOs in Medicare, J Health Econ 22, 915–931.

68. Dallek, G., and Dennington, A. (2002) Physician Withdrawals: A Major Source of Instability in the Medicare+Choice Program New York: The Commonwealth Fund..

69. Ingber, M (1998) The current state of risk adjustment technology for capitation. J Ambul Care Manage 21,1-28

70. Medicare Payment Commission (1998) Report to the Congress: Medicare Payment Policy.

71. Pope, G.C., K.W. Adamache, R.K. Khandker, and E.G. Walsh. (1998) Evaluating Alternative Risk Adjusters for Medicare, Health Care Fin Rev 20, 109–29.

72. Kapur K., (2005) Risk Adjustment Methods and their Relevance to Pay-or-Play Supplement E to the Report: Challenges and Alternatives for Employer Pay-or-Play Program Design: An Implementation and Alternative Scenario Analysis of California's Health Insurance Act of 2003 (SB 2) Report written for the California Health Care Foundation and the California Managed Risk Medical Insurance Board Project Team. www.ihps.org/pubs/2005_Apr_IHPS_SB2_ESup_Risk_Adj.pdf

73. Cutler, D., and Zeckhauser, R. (2000) The Anatomy of Health Insurance in Handbook of Health Economics, Volume I A Culyer, A., and Newhouse J.P., eds., Amsterdam: Elsevier 563-643.

74. Rosenthal, M.B., Frank, R.G., Li, Z., and Epstein, A.M. (2005) Early evidence with pay-for-performance: From concept to practice. JAMA 294, 1788-1793.

75. Van de Ven, W.P.M.M., K. Beck, F. Buchner, et al. (2003a) "Risk adjustment and risk selection on the sickness fund insurance market in five European countries" Health Policy 65, 75-98.

76. Van de Ven WPMM, van Vliet RCJA, Lamers LM. (2004) Health-adjusted premium subsidies in the Netherlands. Health Affairs 23, 45-54.

77. Busse, R, and Riesberg A. (2004) Health care systems in transition: Germany. Copenhagen, WHO Regional Office for Europe on behalf of the European Observatory on Health Systems and Policies.

78. Fuerstenberg, T., Rochell, B., Roeder, N. (2007) Ambulatory and hospital care in the Federal Republic of Germany Am Heart Hosp J 5, 22-26

79. Buchner, F. and Wasem, J., (2003) Needs for further improvement: risk adjustment in the German health insurance system Health Policy 65, 21-35.

80. Alberta Health and Wellness. (2005) 2005-2006 Regional Health Authority Global Funding, Methodology and funding manual. April.

81. Murgolo, M, (2002) MCBS highlights: comparison of Medicare Risk HMO and FFS enrollees. Health Care Fin Rev 24, 177–185.

82. McKillop, I., Pink, G.H., and Johnson, L.M. (2001) The Financial Management of Acute Care in Canada: A Review of Funding, Performance Monitoring and Reporting Practices.

83. Shwartz, M. Ash, A.S., and Pekoz, E. (2006) Risk adjustment and risk-adjusted provider profiles. International Journal of Healthcare Technology and management 7, 15-42.

84. Eggleston, K., (2000) Risk selection and optimal health insurance-provider payment systems. J Risk Insur 67, 175–198.

85. Eggleston, K., and Bir, A., (2007) Measuring Selection Incentives in Managed Care: Evidence from the Massachusetts State Employee Health Insurance Program J Risk Insur
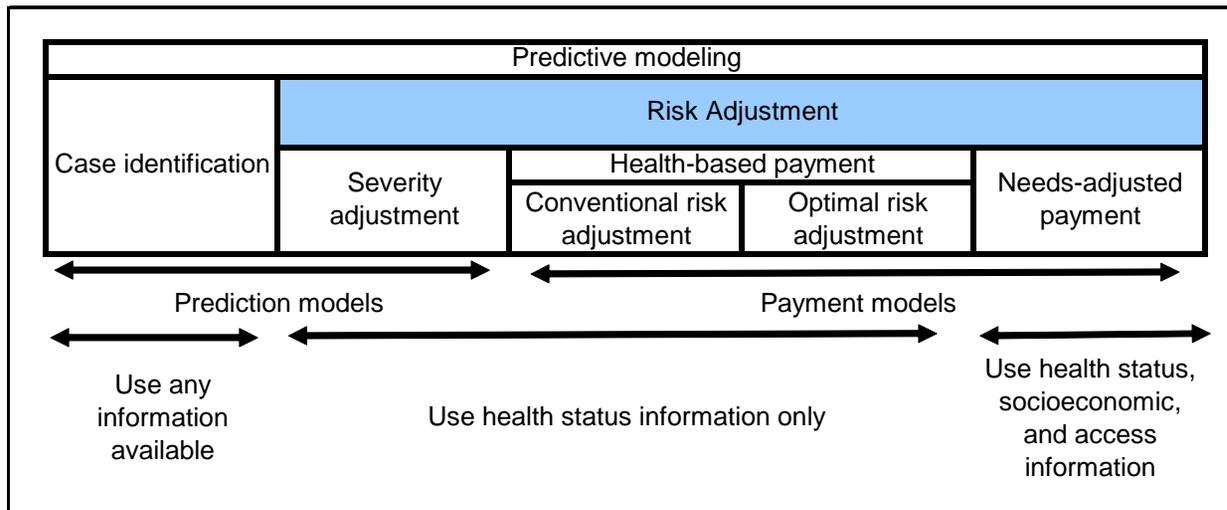
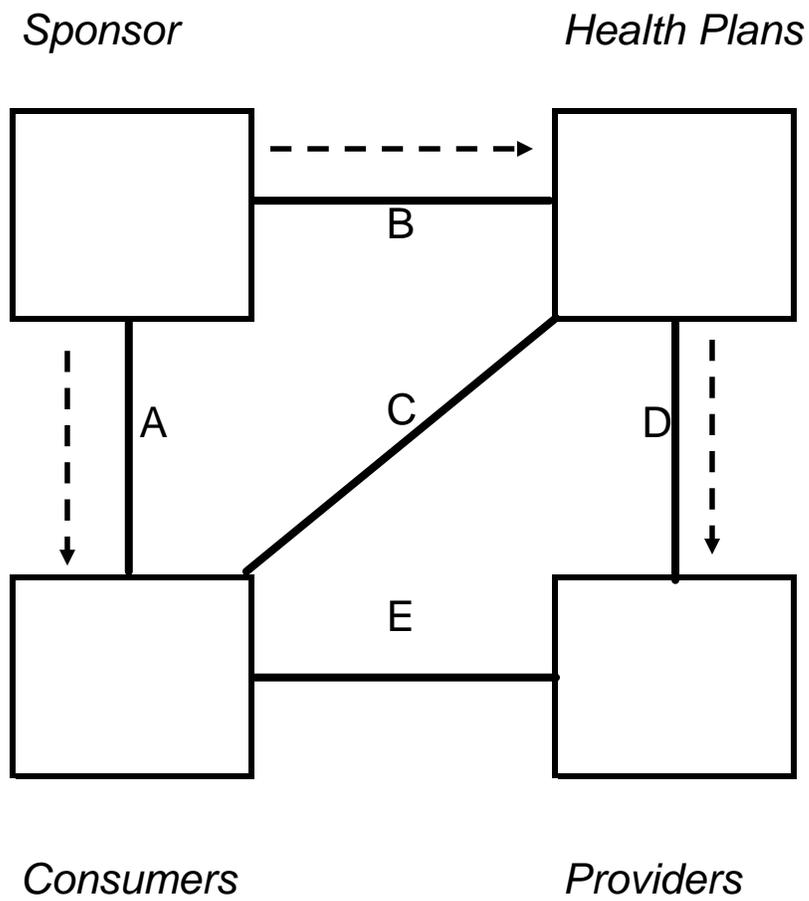**Figure 1: Different terminology used related to risk adjustment**

Sponsor                                    Health Plans

```
┌──────────────┐   - - - - - →    ┌──────────────┐
│              │                  │              │
│              │ ─────────────    │              │
│              │      B           │              │
└──────────────┘                  └──────────────┘
    ┊      │                          │ ┊   │
    ┊      │                    ╲     │ ┊   │
    ┊     A│          C          ╲   D│ ┊   │
    ┊      │                       ╲  │ ┊   │
    ↓      │                        ╲ │ ↓   │
┌──────────────┐        E         ┌──────────────┐
│              │ ─────────────    │              │
│              │                  │              │
│              │                  │              │
└──────────────┘                  └──────────────┘
```

Consumers                              Providers

Figure 2: Four agents, five primary contracting relationships (A-E), and three possible uses of risk adjusted payments

Spending on Specialists

α + β

α+ ß/2

α

Diabetic (high risk)

B

X

A

Healthy (low risk)

α
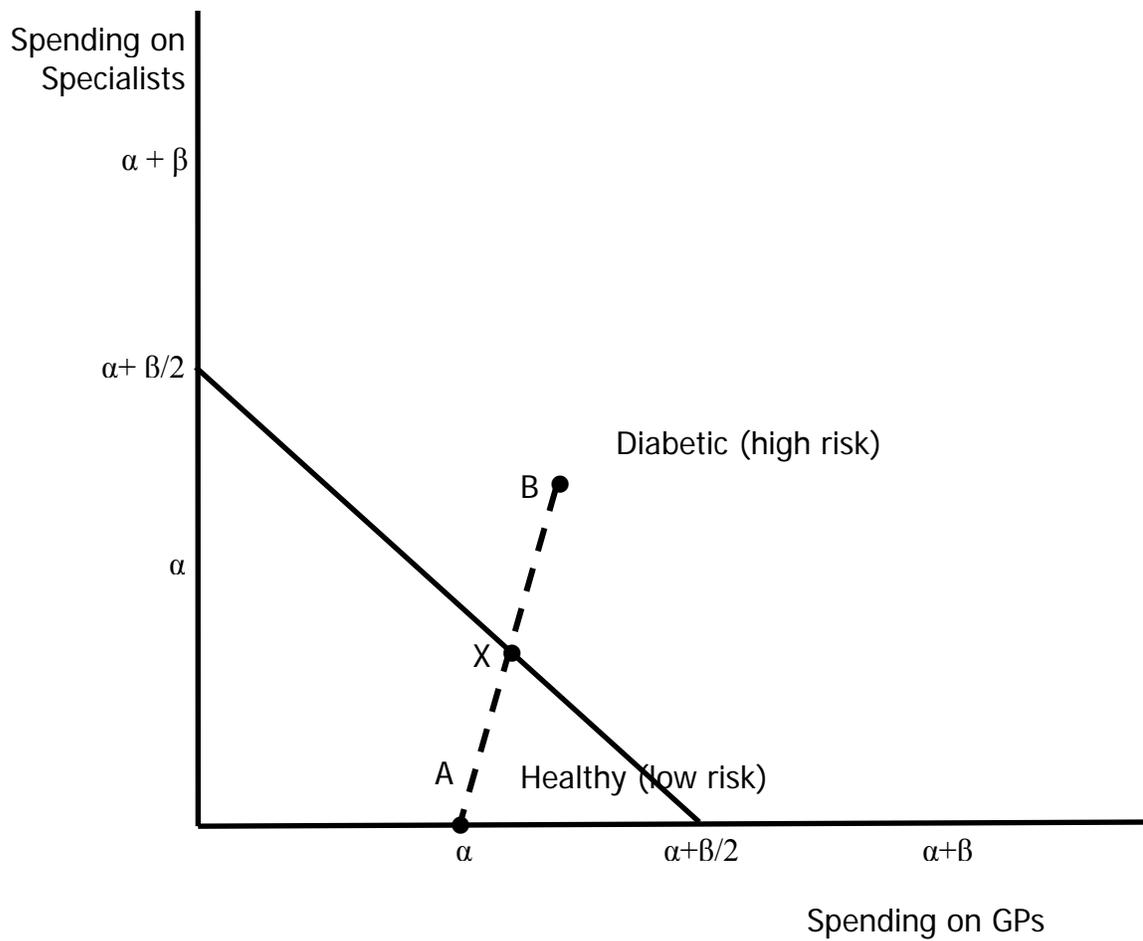
α+ß/2

α+ß

Spending on GPs

**Figure 3: No Risk Adjustment when quantities of each service supplied are exogenous.** Health plan is reimbursed the simple average cost. Average services provided is at X for entire population.
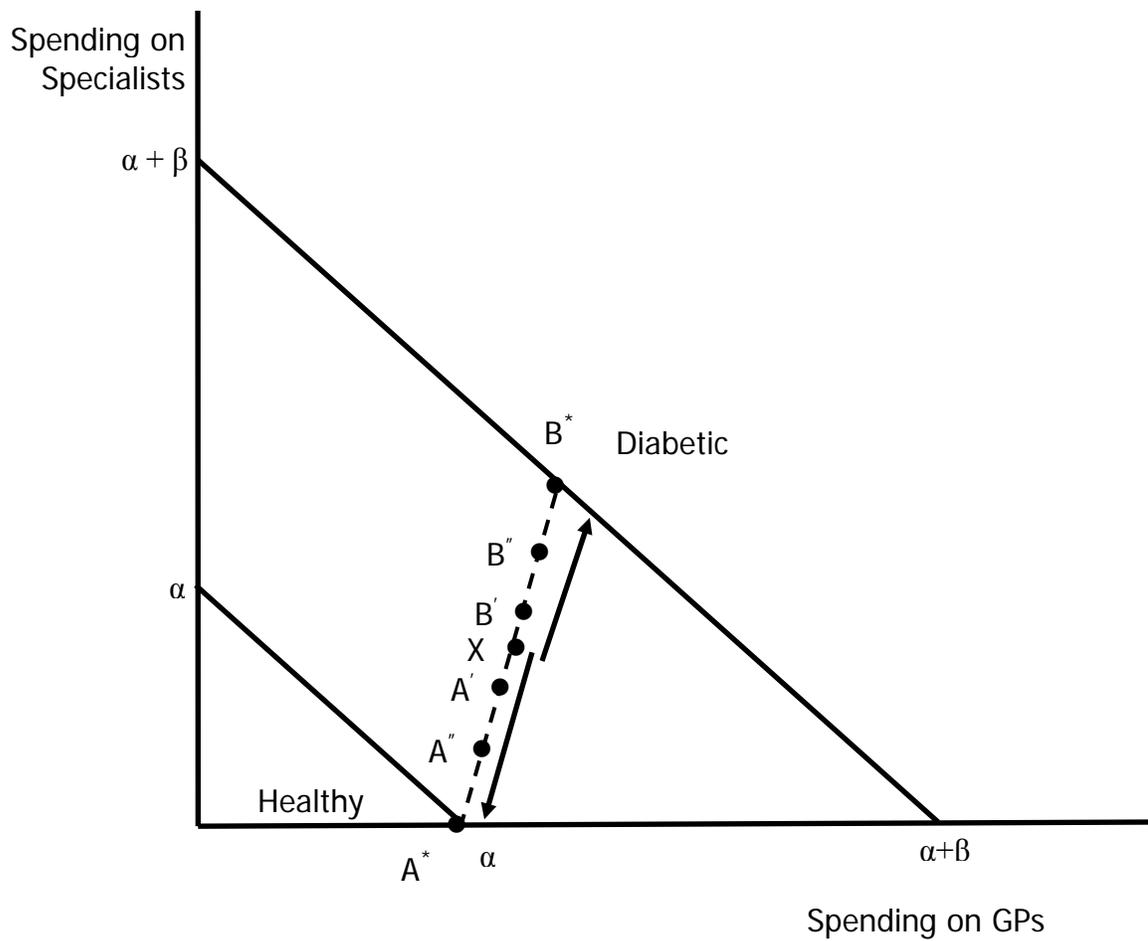
**Figure 4: Perfect and imperfect conventional risk adjustment.** Perfect risk adjustment (A*, B*) is just the limiting case of improving information (X → A' → A" → A* for low risk types and X → B' → B" → B*) so that payments by the sponsor for each type more closely match costs.
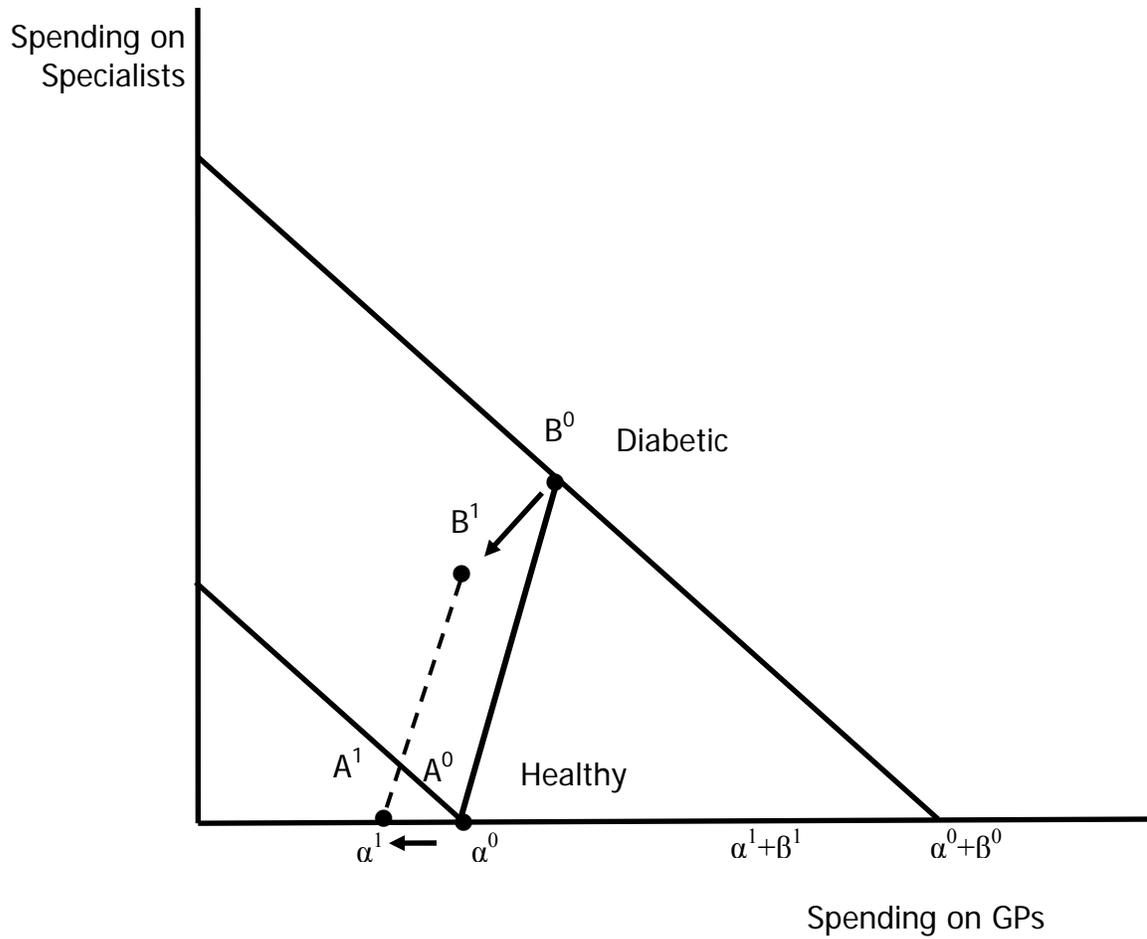
**Figure 5: Capitated quantities of services (A$^1$, B$^1$) will differ from quantities offered under Fee-For-Service (A$^0$, B$^0$) due to supply-side moral hazard.** Simple selection model might have all spending reduced by the same percentage amount. Conventional risk adjustment will yield correct payments if capitated quantities are used.
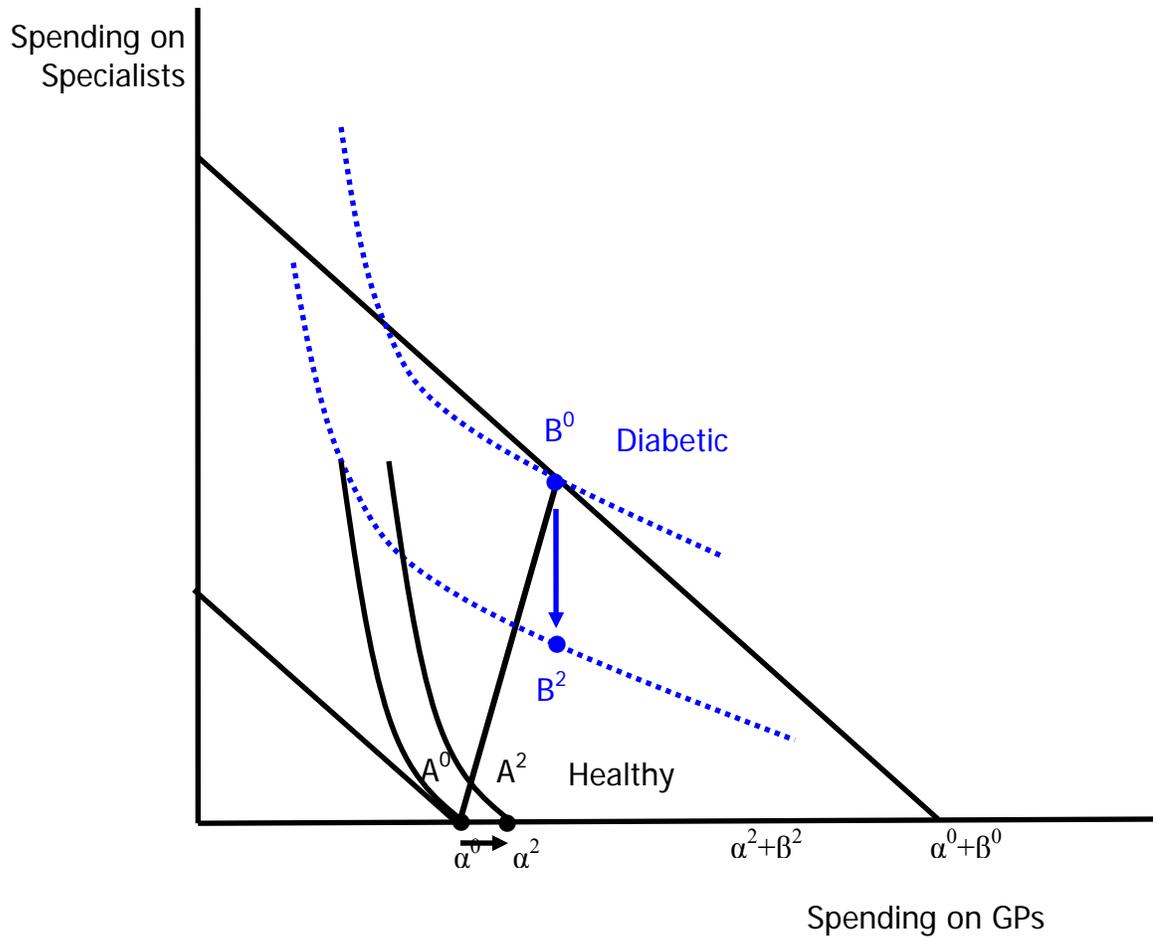
**Figure 6: Conventional Risk Adjustment with quantities of services offered strategically determined.** Unless true types are observable, imperfect signals will tend to cause health plans to oversupply GP services attractive to the healthy (supplying $A^2$ rather than $A^0$) while undersupplying specialist services ($B^2$ versus $B^0$) used by diabetics.
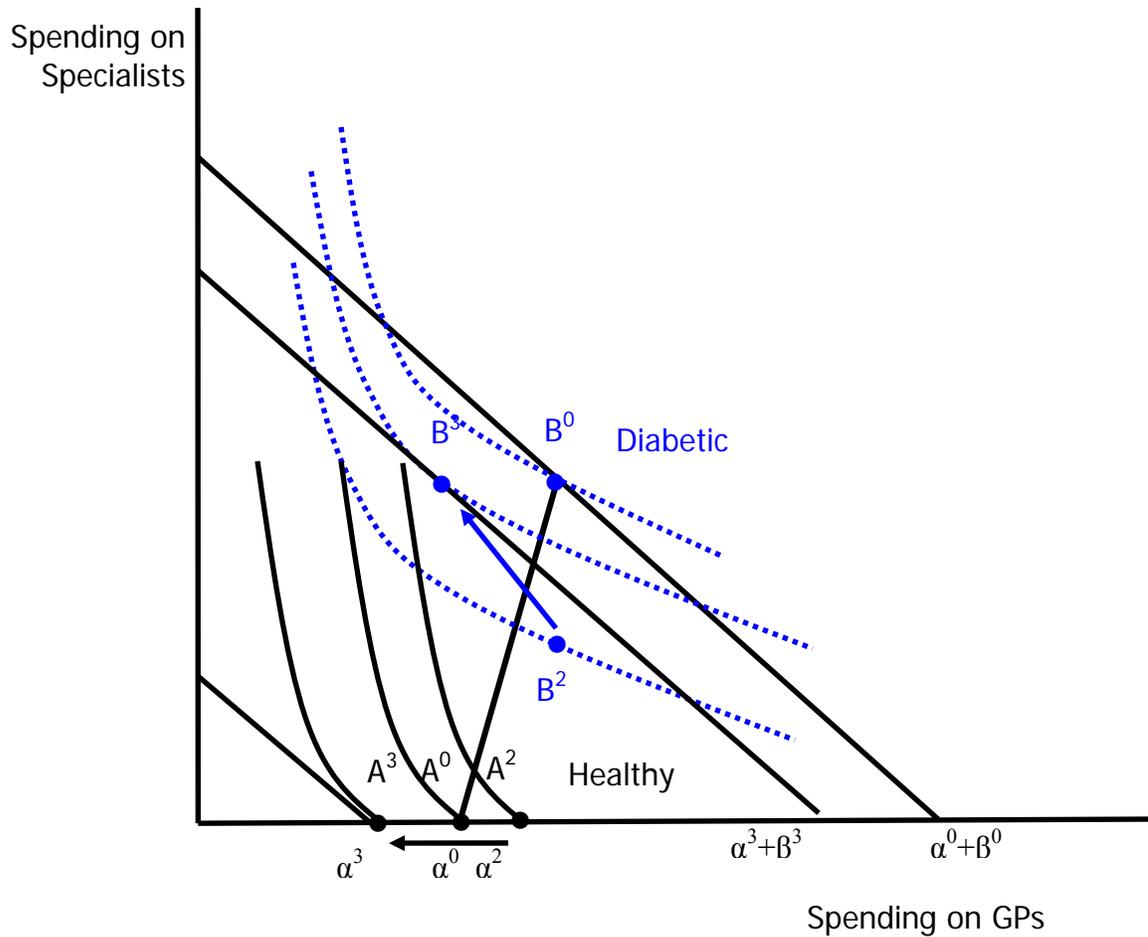
**Figure 7: Optimal Risk Adjustment with quantities of services strategically determined.** By increasing the payments for diabetics to $\beta^3$ which enable $B^3$ rather than $B^2$ and reducing the payment for healthy consumers to enable only $A^3$, health plans can be made indifferent to enrolling healthy or diabetic consumers and induced to provide efficient services to both types.

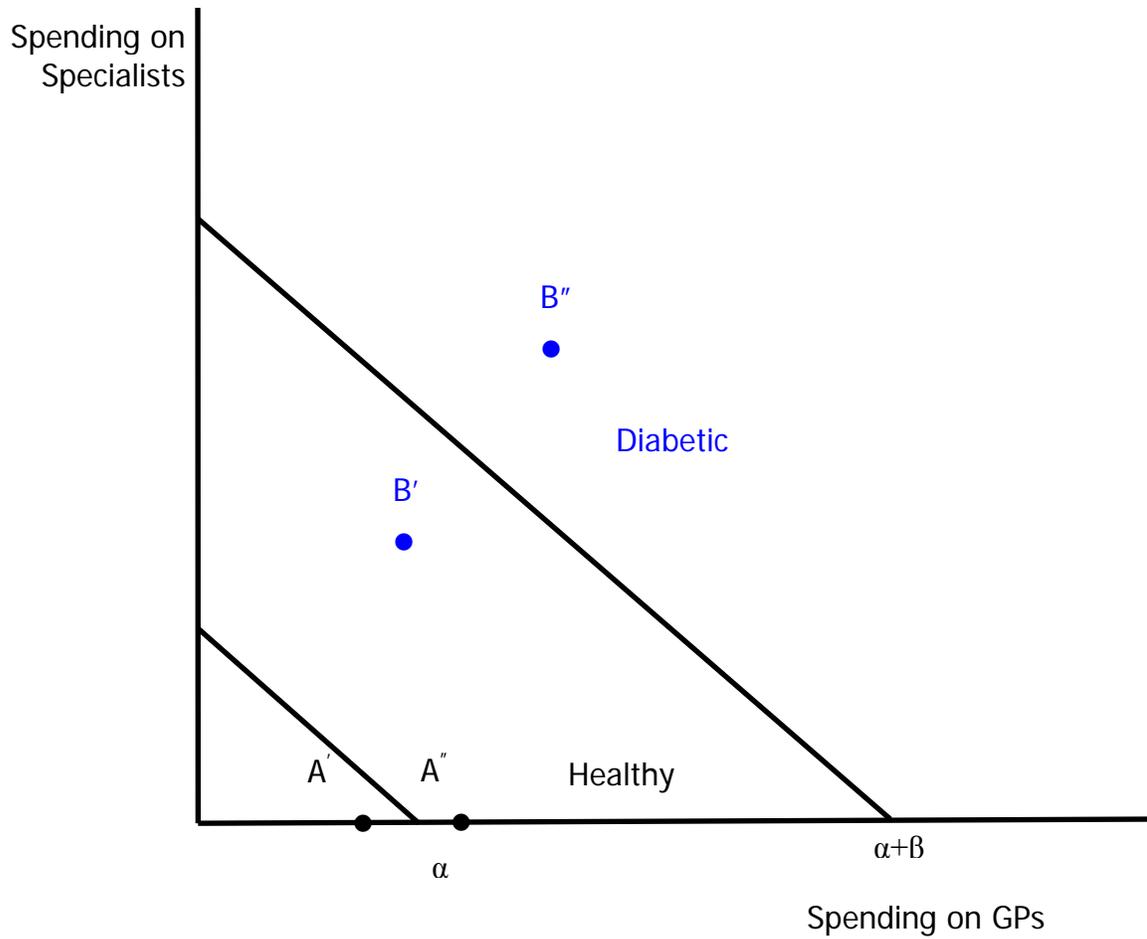**Figure 8: Risk Adjustment with taste or income heterogeneity.** Optimal service choices for two alternative diabetic types are B' and B", while optimal service choices for two healthy types are A' and A". With only two signal values, plans cannot offer efficient quantities of services to all four types. Health plans will have an incentive to compete to be most attractive to the lower cost types, A' and B'.
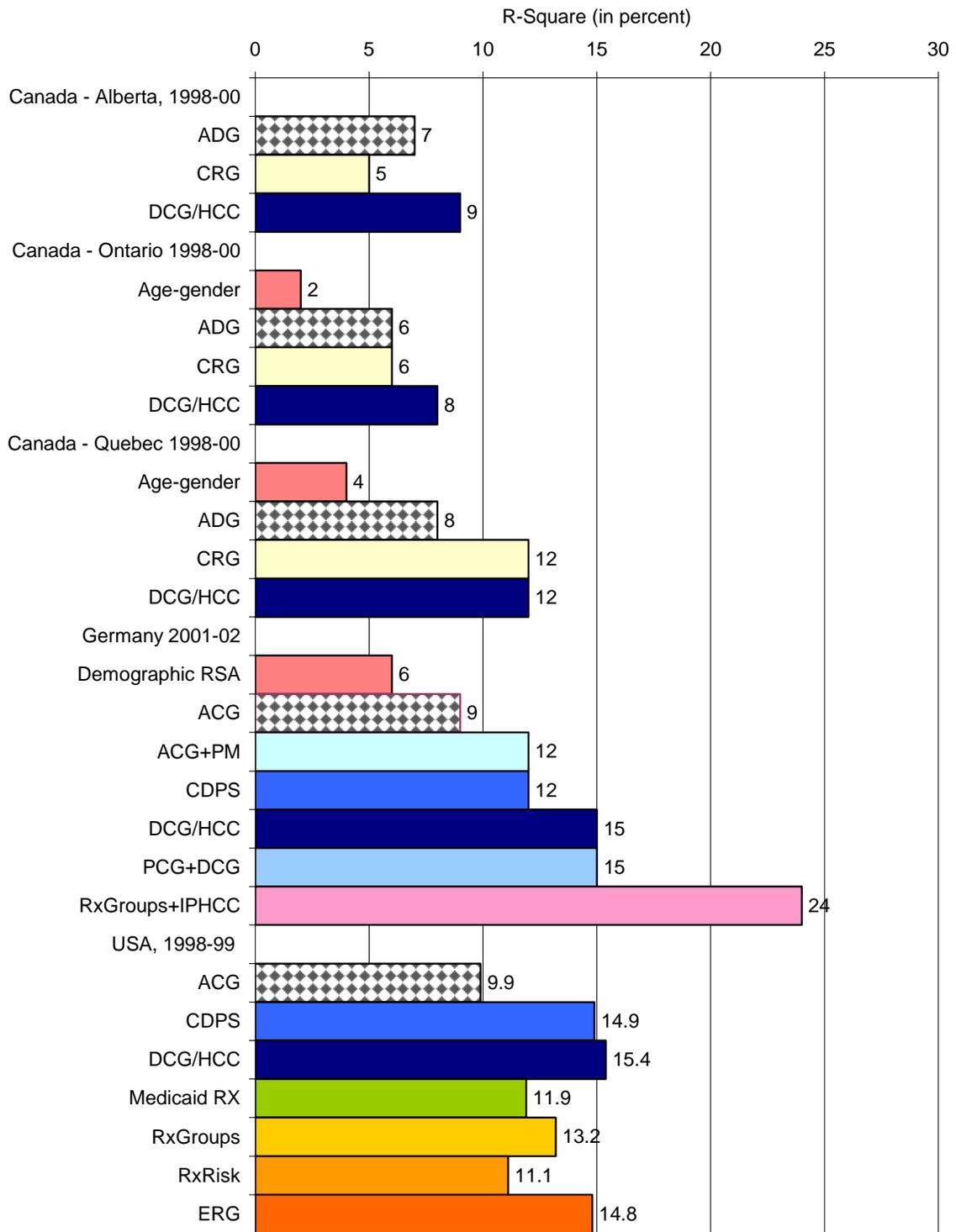
**Figure 9: Predictive power (R-Square) of prospective risk adjustment models in three countries**

Sources: [58, 59, 26]

R-Square (in percent)

**Canada - Alberta, 1998-00**
- ADG: 25
- CRG: 31
- DCG/HCC: 47

**Canada - Ontario 1998-00**
- Age-gender: 2
- ADG: 21
- CRG: 31
- DCG/HCC: 37

**Canada - Quebec 1998-00**
- Age-gender: 4
- ADG: 22
- CRG: 43
- DCG/HCC: 33

**Germany 2001-02**
- Demographic RSA: 7
- ACG: 14
- ACG+PM: 21
- CDPS: 19
- DCG/HCC: 26
- PCG+DCG: 30
- RxGroups+IPHCC: 48

**USA, 1998-99**
- ACG: 28.2
- CDPS: 35.5
- DCG/HCC: 46.6
- Medicaid RX: 24.4
- RxGroups: 27.9
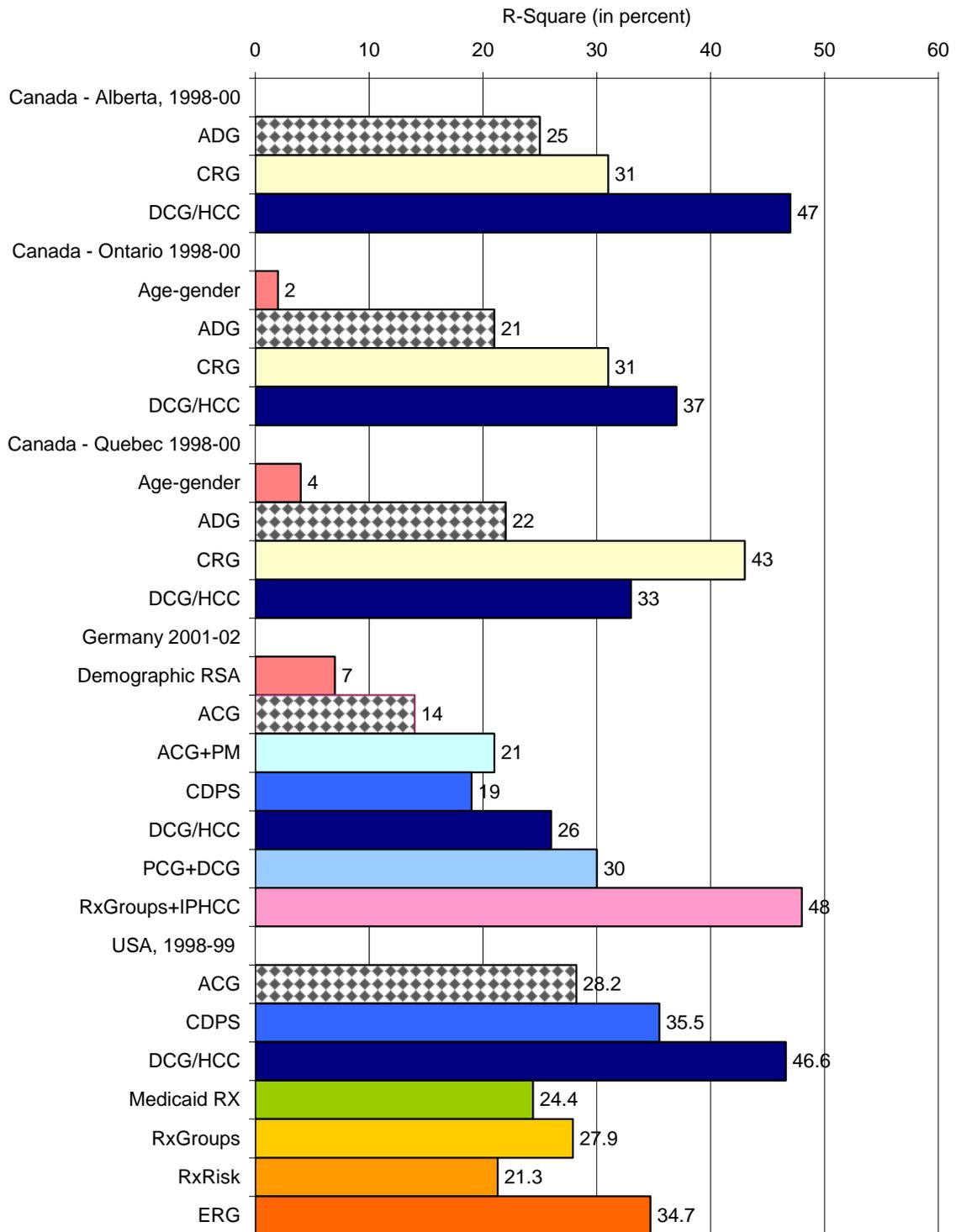- RxRisk: 21.3
- ERG: 34.7

**Figure 10: Predictive power (R-Square) of concurrent risk adjustment models in three countries.**
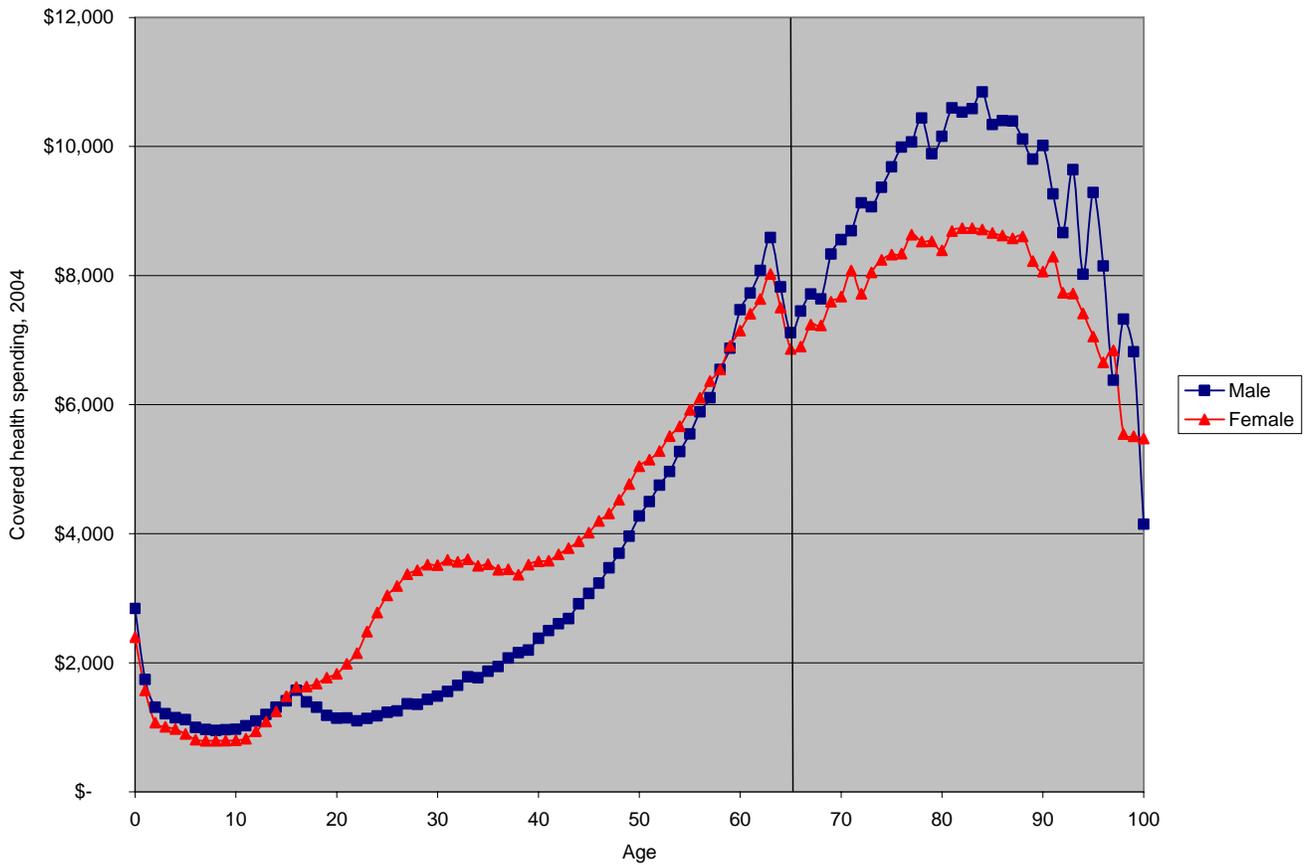
60

Sources: [58, 59, 26]

**Figure 11: US privately-insured health care spending, by age and by gender,**

**2004 MEDSTAT Marketscan data (N=14.6 million)**

Notes: Plots show weighted average annualized health care spending by one-year age intervals, for males and females. Age is beginning of year age. Total health spending includes covered inpatient plus outpatient plus pharmacy covered spending. Sample is merged MEDSTAT commercially insured (under age 65) and MEDSTAT Medicare samples. Discontinuity at age 65 reflects both a sample discontinuity as well as benefit and utilization changes. Source: Authors own calculations at DxCG, Inc.

**Figure 12**: US privately-insured health care spending by age, by health plan type,

**2004 MEDSTAT Marketscan data (N=13.0 million)**

Notes: Plots show weighted average annualized health care spending by one-year age intervals by five major health plan types. Age is beginning-of-year age. Exclusive provider organization (EPA) and missing plan type observations omitted. Total health spending includes covered inpatient plus outpatient plus pharmacy covered spending. Sample is 2004 MEDSTAT commercially-insured (under age 65) sample. Source: Authors own calculations at DxCG, Inc.
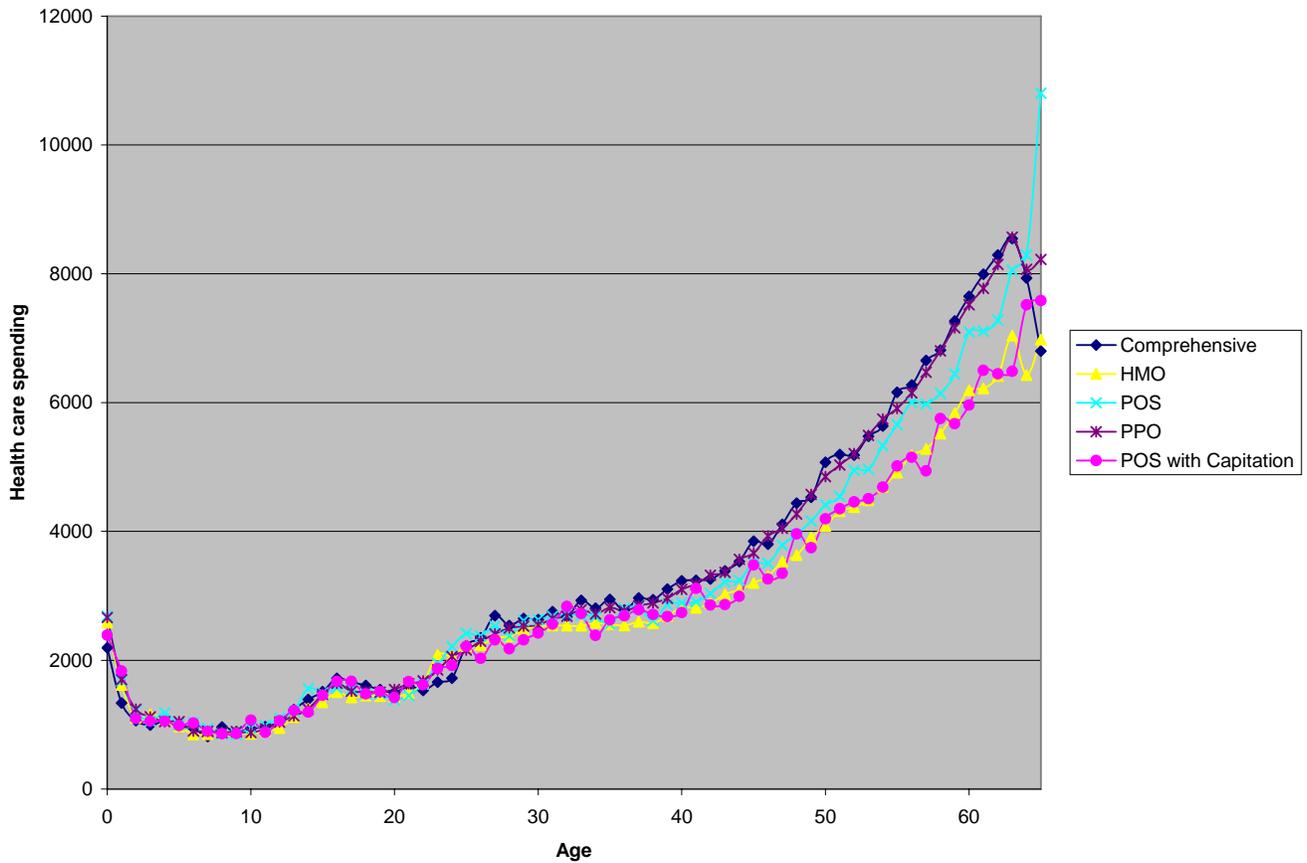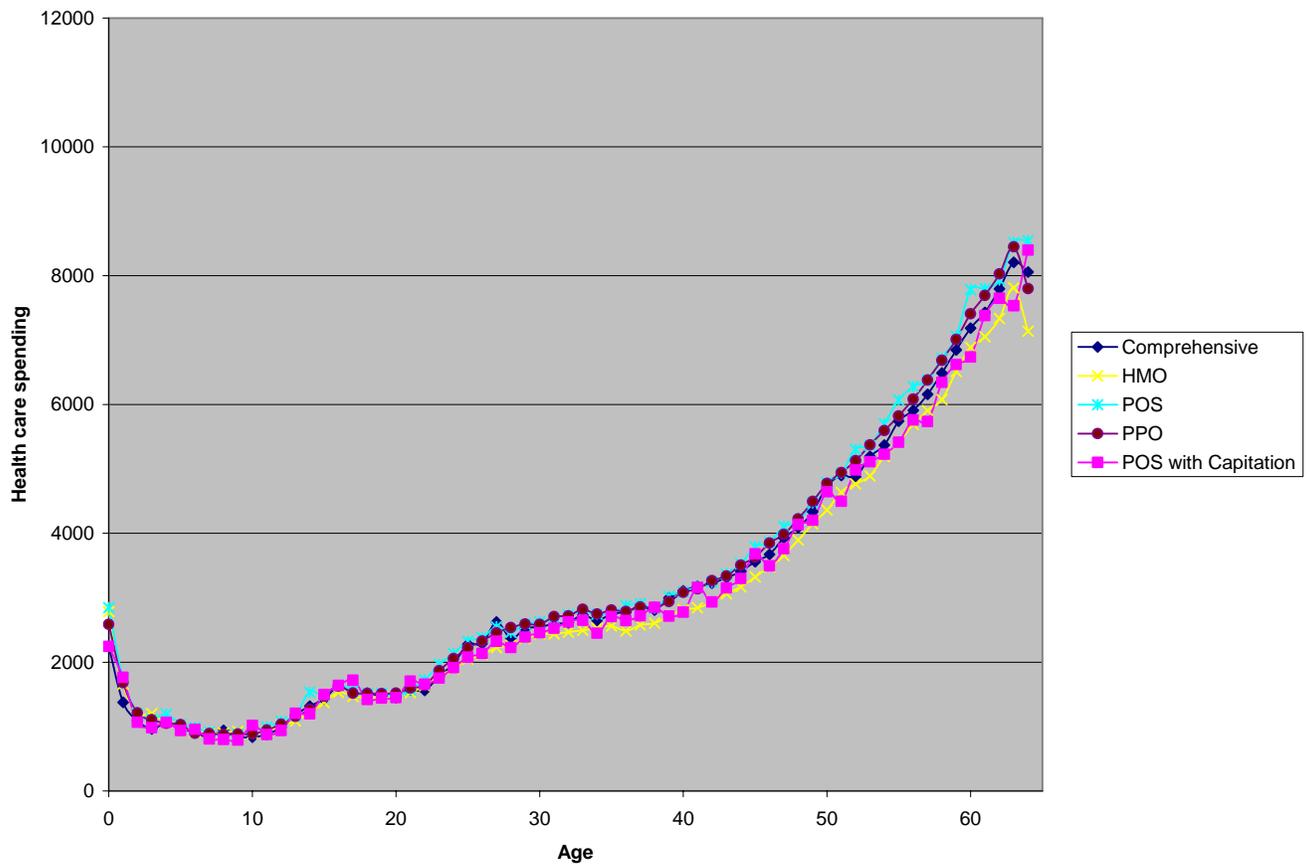
**Figure 13:** Risk-adjusted US privately-insured health care spending, by age and by health plan type, 2004 MEDSTAT Marketscan data (N=13.0 million)

Notes: Plot shows risk-adjusted weighted average annualized health care spending by one-year age intervals by five major health plan types. Average spending at each age was deflated by the ratio of the plan-specific average age-specific relative risk score to the overall age-specific average relative risk score. Risk scores calculated using the concurrent HCC model of DxCG release 6.1 software. Age is beginning-of-year age. Exclusive provider organization (EPA) and missing plan type observations omitted. Total health spending includes covered inpatient plus outpatient plus pharmacy covered spending. Sample is 2004 MEDSTAT commercially-insured (under age 65) sample. Source: Authors own calculations at DxCG, Inc.

*Sponsor*                             "Health Plans"

| Provincial Government | Regional Health Author-ities | Alberta Medical Associa-tion | Alberta Blue Cross | Supple-mentary Plans |
|---|---|---|---|---|

B

A

E

$D_1$      $D_2$      $D_3$

| All Individuals | | Hospitals | Doctors | Drugs |
|---|---|---|---|---|

*Consumers*                           *Providers*
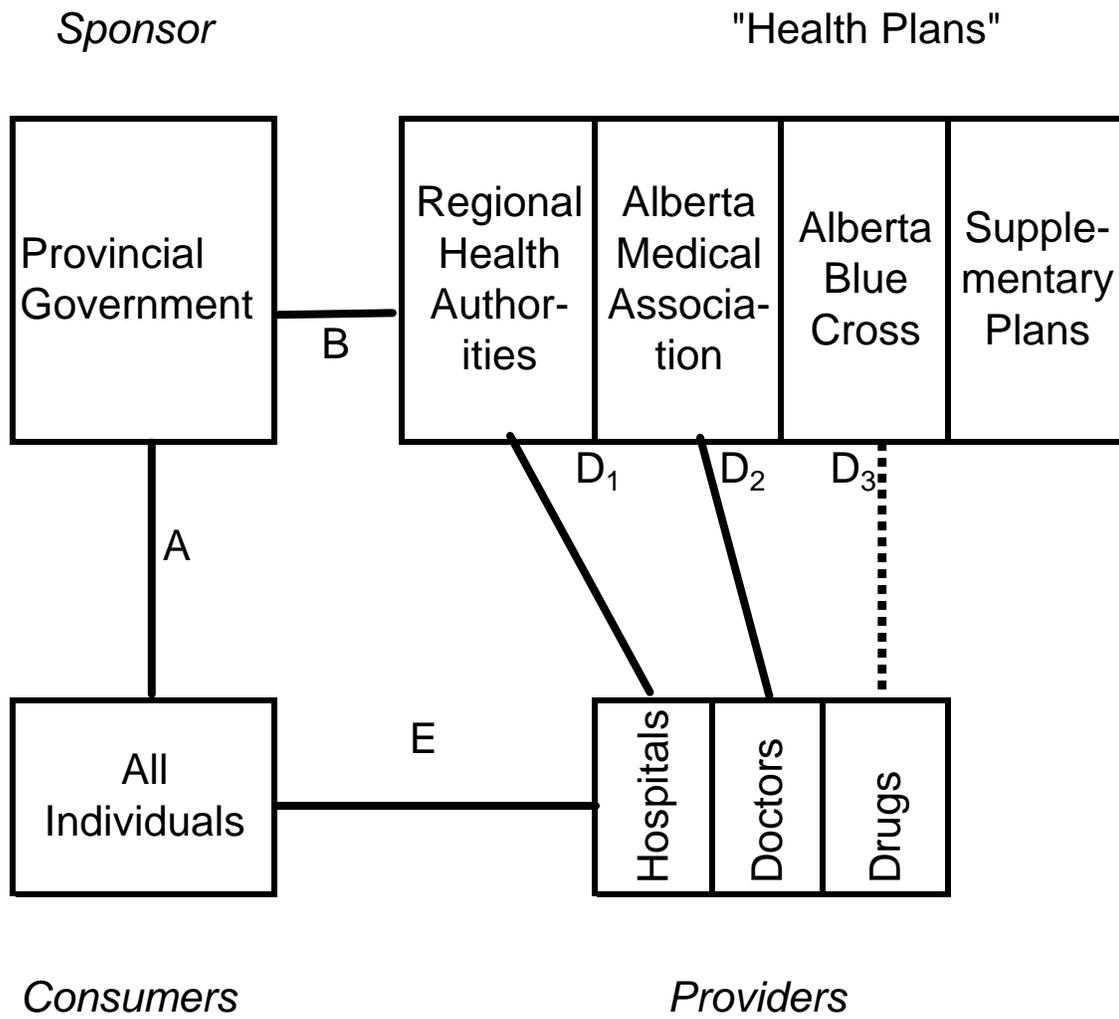
**Figure 14: Canadian Health System. Stylized view of Alberta 2006**

Sponsor

Health Plans

Government

Traditional Indemnity

MEDI-GAP Plans

Managed Care

PPOs

Medicare Enrollees

Consumers

Doctors

Hospitals

Drugs

Providers
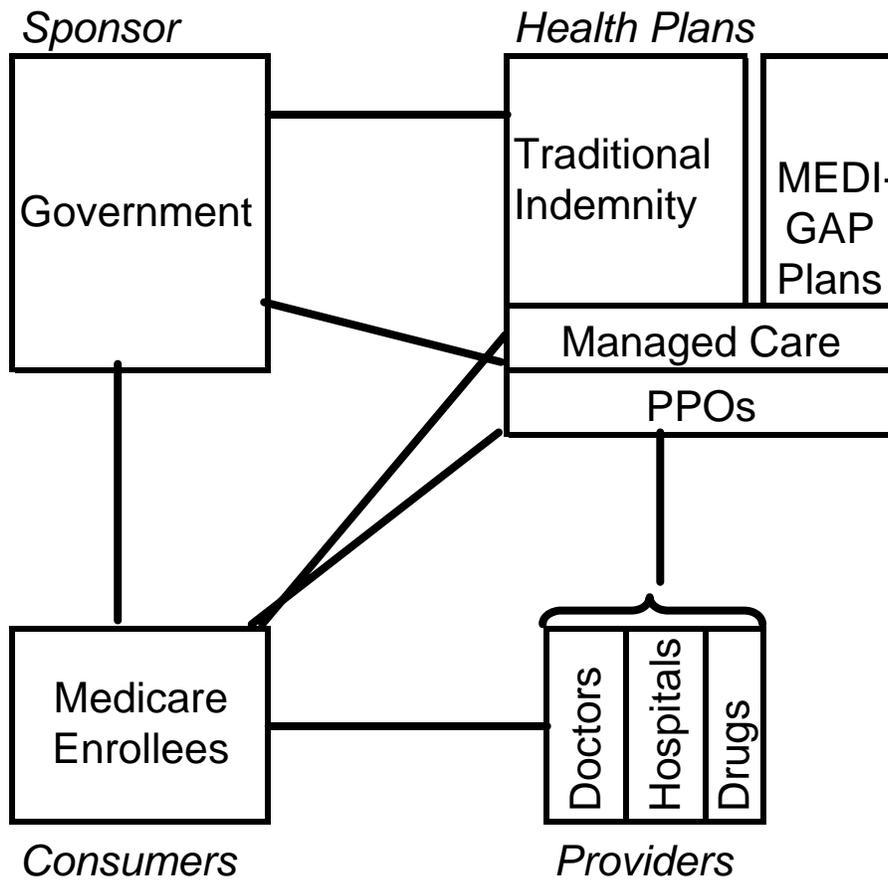
**Figure 15: US Medicare system, 2006**
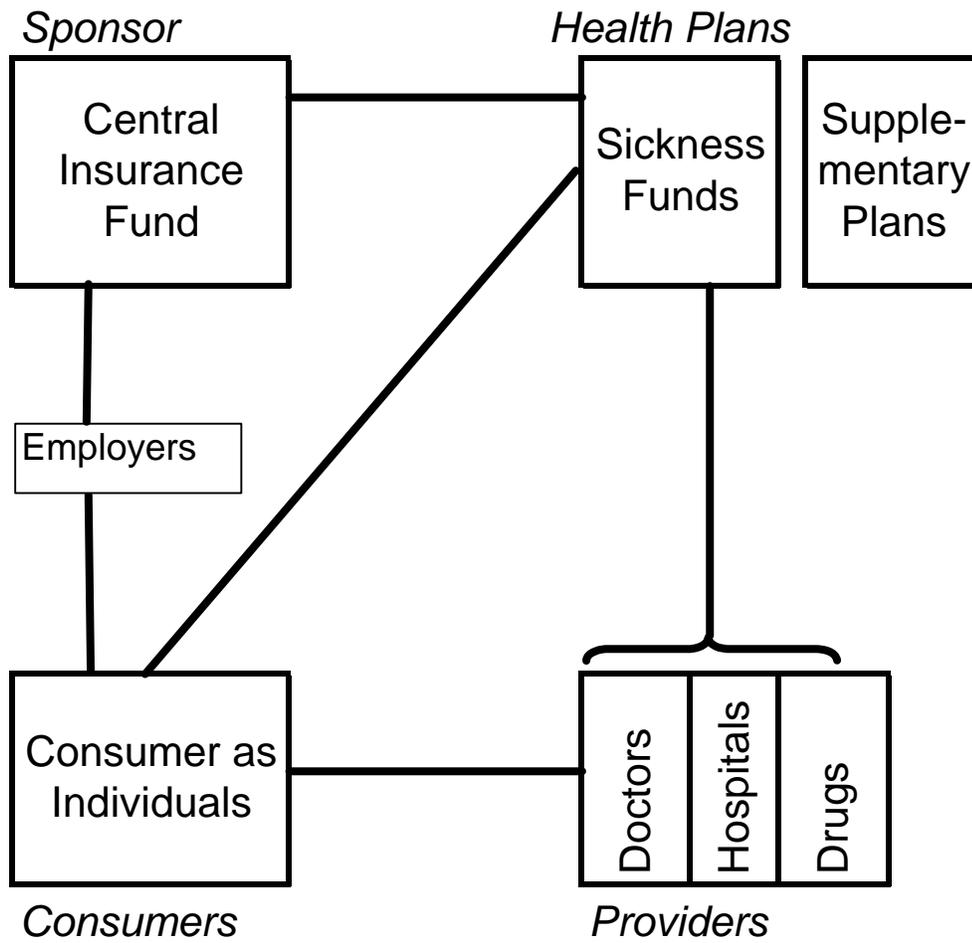
**Figure 16: Netherlands, Mandatory insurance system, 2006.**

*Sponsors* *Health Plans*

Central Insurance Fund

B

282 Sickness Funds and their seven organizations (BKK, AOK, VDAK, IKK, IV, Farmers)

Supple-mentary Plans

Employers

$D_1$

KV

$D_2$

Low and middle Income

High income

Doctors plus others

Hospitals

Drugs

Sick Leave

*Consumers* *Providers, etc.*

**Figure 17: Germany, Statutory Insurance, 2006.**

**Table 1: Overview of major US claims-based risk adjustment models**

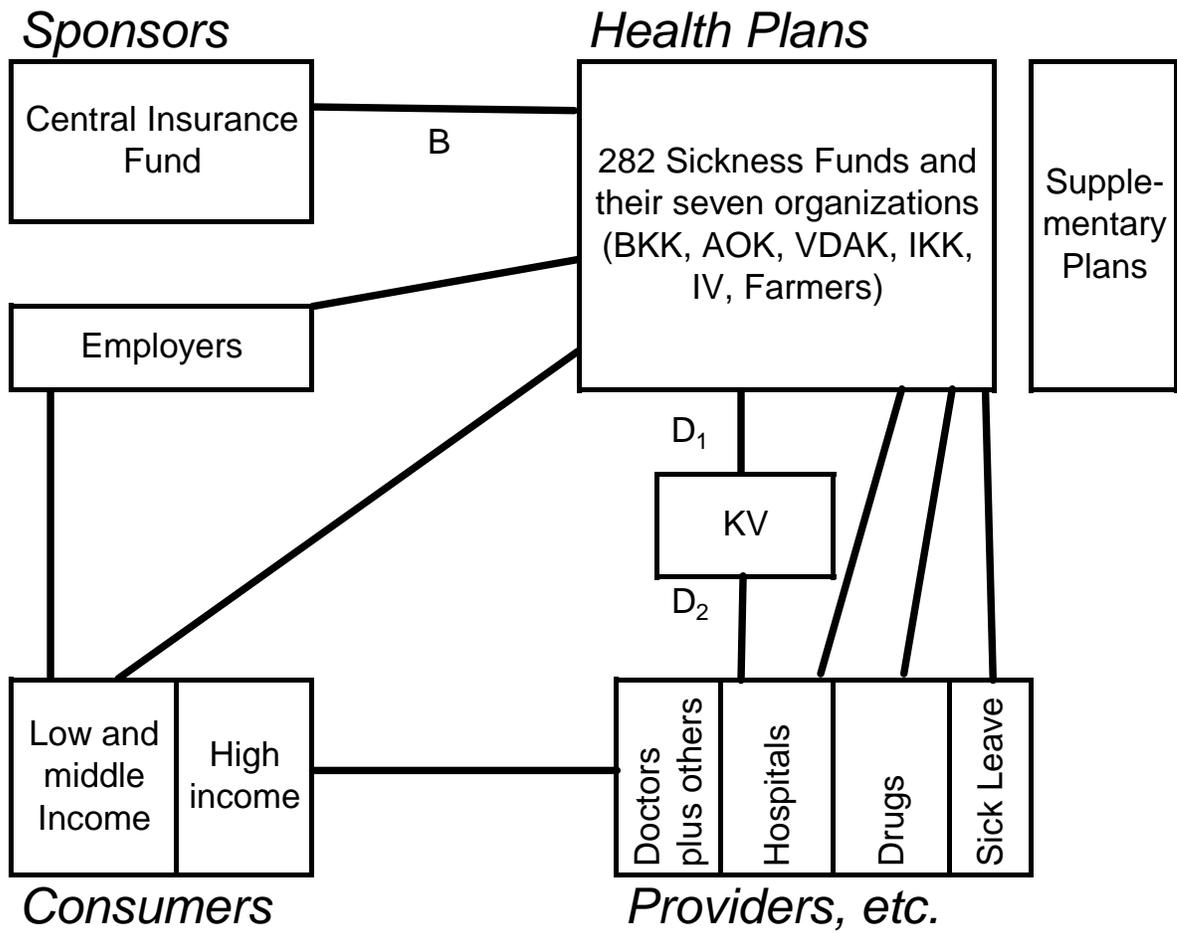| Acronym/key reference | Name | First referenced | Rate Cell or linear regression? | Age/ gender | Inpat. diag-noses | All diag-noses | Phar-macy | Proc codes |
|---|---|---|---|---|---|---|---|---|
| CI[32] | Charleson index | 1987 | regression | X | X | | | |
| DCG[33] | Diagnostic Cost Groups | 1989 | regression | X | X | | | |
| ACG[34] | Adjusted Clinical Groups | 1991 | rate cell | X | X | X | | |
| CDS[35] | Chronic Disease Scores | 1992 | regression | X | | | X | |
| HCC[43] | Hierarchical Condition Categories | 1996 | regression | X | X | X | | |
| CDPS[44] | Chronic and Disability Payment System | 1996 | regression | X | X | X | | |
| GRAM[45] | Global Risk Assessment Model | 1996 | regression | X | X | X | | |
| CD-RISC[46] | Clinically Detailed Risk Indication System for Cost | 1997 | regression | X | X | X | | |
| CRG[38] | Clinical Related Groups | 1999 | rate cell | X | X | X | X | X |
| ERG[39] | Episode Risk Groups | 2001 | rate cell | X | X | X | X | X |
| RxGroups[36] | RxGroups | 2001 | regression | X | | | X | |
| RxRisk[37] | RxRisk | 2003 | regression | X | | | X | |

**Table 2: Predictive power of various information sets and various models**

Dependent Variable:1997 annualized total covered charges

| | Weighted OLS | OLS | Square Root model (hetero-skedasticity-corrected) | Two part linear model | GLM with link = log, dist = normal |
|---|---|---|---|---|---|
| Partial Year Eligibles included? | Yes | No | No | No | No |
| Sample Mean | 6,886 | 5,063 | 5,063 | 5,063 | 5,063 |
| Number of Observations | 1,380,863 | 1,273,471 | 1,273,471 | 1,273,471 | 1,273,471 |
| | $R^2$ | $R^2$ | $R^2$ | $R^2$ | $R^2$ |
| Age and gender only | 0.011 | 0.010 | 0.009 | 0.010 | 0.010 |
| Prior year total covered charges* | 0.089 | 0.096 | 0.113 | 0.120 | 0.105 |
| Diagnoses organized by DCG/HCC* | 0.104 | 0.108 | 0.103 | 0.107 | 0.105 |
| | | | | | |
| Covered charges by  DCG/HCC* | 0.099 | 0.107 | 0.103 | 0.105 | 0.095 |
| Covered charges by Place of Service* | 0.140 | 0.145 | 0.136 | 0.145 | 0.126 |
| Covered charges by Physician Specialty* | 0.142 | 0.152 | 0.143 | 0.152 | 0.131 |
| Covered charges by Type of Service* | 0.150 | 0.155 | 0.146 | 0.154 | 0.134 |
| All of the above except diagnoses* | 0.154 | 0.160 | 0.151 | 0.160 | 0.138 |
| | | | | | |
| "Kitchen sink": All of the above* | 0.169 | 0.171 | 0.161 | 0.169 | 0.147 |

*All Regressions included a constant and 21 age-gender dummy variables

Source: **[54]**, Table 1.

**Table 3**         **Evidence of Geographic Risk Selection in Alberta, Canada**

|  | Highest RHA rate | Lowest RHA rate |
|---|---|---|
| Average Age | 36.7 years | 25.2 years |
| Proportion over age 65 | 14.7% | 2.1% |
| Proportion with pregnancy diagnosis | 3.6% | 1.9% |
| Proportion of diabetics with a chronic complications diagnosis | 13.2% | 8.1% |

Source: [60]

**Table 4**. The practice of risk-adjustment in 6 countries

|  | Canada (Alberta) | Germany | Netherlands | Switzerland | United Kingdom | United States (Medicare) |
|---|---|---|---|---|---|---|
| Current risk-adjusters | age/gender disability income aborigine | age/gender disability | age/gender region pharmacy, inpatient diagnoses | age/gender region | age/gender prior utilization local factors | Age/gender All-encounter diagnoses |
| Rate cell or regression model | 124 rate cells | rate cell | regression model | rate cell | rate cell | regression model |
| Individual or grouped data | individual | individual | individual | individual | group | individual |
| Number of health plans/funding regions | 9 RHA (2006) | 292 Sickness funds (2004) | 25private health insurance funds | 166 sickness funds | 303 Primary Care Trusts | 314 (2006) Medicare Advantage plans |
| Open entry for new health plans? (subject to certain conditions) | No | yes | yes | yes | no | yes |
| Open enrollment every month/…/year | - | year | year | half year | no open enrollment guarantee | monthly |
| Mandatory or voluntary membership | M | V for high income | M | M | M | V |
| Implementation date | 1985 | 1994 | 2004 | 1993 | 1991 | 2004 |