

Implementing Health-based Risk Adjustment: Lessons from the Field

David J. Knutson, Jinnet B. Fowles

Preface

The Risk Adjustment Impact Study (RAIS) project convened a panel of implementation experts who have been implementing health-based risk adjustment for different products in different markets (Key Findings). The panel first met in July, 2000 to discuss lessons learned during implementation. This report distills and elaborates the points raised during the initial meeting. Although the report reflects the comments made by panel members, each expert may not necessarily agree with its emphasis. The authors are solely responsible for the organization and interpretation.

The matrix below is the organizing framework for the discussion that follows. The numbers in each cell refer to the sections of the report.

Implementation Stages	Policy and Program Lessons	Technical Lessons	Management Lessons
Original Context	1	2	3
Planning	4	5	6
Simulation	7	8	9
Implementation	10	11	12

1. What were the policy and program issues when health-based risk adjustment was initially considered? (What were the current or anticipated problems that stimulated consideration of health-based risk adjustment?)

Policy and program issues are determined by the purchaser's characteristics, the purchaser's goals, and the original payment context.

Purchaser Characteristics

Purchaser characteristics play an important role in shaping how health-based risk adjustment is implemented. These key dimensions include

- whether the purchaser is in the public or private sector;
- the geographic range of the purchaser (community vs. national);
- whether the purchaser is purchasing for employees or beneficiaries of entitlement programs;
- the size of the purchaser.

The role of each of these characteristics in shaping the implementation of health-based risk adjustment is described in an issue paper on purchaser characteristics on this Web site.

Goals for implementing health-based risk adjustment

The experts identified a number of different rationales of purchasers who have implemented health-based risk adjustment. They include to

- save money;
- address current adverse risk selection and pay fairly;
- be able to include the disabled in managed care;

- address other payment policy issues, such as geographic variation in rates;
- preserve the ability to offer a choice of plans;
- encourage access to specialty providers.

Articulating the specific purchaser’s goals is key for gaining agreement among stakeholders, for guiding planning and implementation decisions, and for assessing progress.

Understanding the original payment context

It is important to understand the payment context that preceded the implementation of health-based risk adjustment. The experts analyzed three antecedent scenarios: fee-for-service, capitation without any risk adjustment, and capitation with demographic risk adjustment.

- If a purchaser is moving from fee-for-service to capitation, then health-based risk adjustment is a reasonable and integral part of the whole reform package.
- If a purchaser already has a capitated product, then plan concerns are focused on revenue changes introduced by the new method, and a phase-in period for health-based risk adjustment may be warranted.
- If a purchaser has a capitated product, and cannot for some reason move immediately to health-based risk adjustment, then implementing demographic risk adjustment is a significant, positive step toward ultimately implementing health-based risk adjustment.
- If a purchaser already has a capitated product that includes demographic risk adjustment, then the move to health-based risk adjustment is more straightforward.

Obtaining necessary authority

In the state of Washington, obtaining legislative authority to assure access to the data was deemed critical to the success of their implementation.

2. What were the technical lessons learned when risk adjustment was initially considered?

There were several initial concerns that purchasers had when they considered initiating health-based risk adjustment, including the availability and completeness of the data, the adequacy of ambulatory diagnosis data, and data confidentiality.

Will the data be available? Will the data be complete? The experts’ strong consensus was that data will become available “just in time.” The important thing is to keep the momentum going.

Are the ambulatory diagnosis data good enough? Yes; they are not perfect, but they are adequate.

Confidentiality, although always a concern, may be a special concern for private employers because, unlike Medicare or Medicaid, private employers do not directly process claims. They also have an employment relationship with covered employees that introduces constraints that a public entitlement program does not share. Employers who are implementing health-based risk adjustment have learned how to use third-party administrators and modes of encryption to assure confidentiality of claims and encounter data.

3. What were the management lessons learned during this initial period?

The principal management needs in this phase are to establish a multidisciplinary team, identify the stakeholders, and hold lots of meetings with food.

The first need is to establish a multidisciplinary team that will probably include specialized experts in health-based risk adjustment as well as internal experts.

Another major management issue in this phase is identifying the stakeholders. Public purchasers are likely to have more stakeholders than private purchasers have. Stakeholders in public programs also have a formal, if not legislated, role in the process. In contrast, in the private sector, the stakeholder role is less formal and not mandated.

4. What were the policy and program lessons learned when planning for risk adjustment?

The experts stressed the importance of monitoring adaptations to the initial health-based risk adjustment model when those adaptations are made for political or practical expediency. These adaptations may undermine the original financial incentives by reducing the number of lives to which health-based risk adjustment is applied or may create a greater lag than was originally planned. These changes may have a detectable effect on the predictive function of the model that may not have been fully anticipated. This inadvertent consequence should not be confused with the result of a deliberate decision to phase in the model (see discussion in Section 10 below).

5. What were the technical lessons learned during the planning phase?

There are many technical issues to attend to during the planning phase. The major issues are choosing a risk assessment grouper, selecting a prediction model, obtaining payment weights for the grouper, testing alternative models, and evaluating lags.

Choosing a grouper

The most important risk adjustment advance is the inclusion of diagnoses in risk assessment. The specific grouper that is chosen is relatively less important. Some purchasers have almost come to a standstill over the issue of model selection.

That said, there is a certain evaluation logic to choosing an appropriate risk assessment grouper.

A purchaser should use a tested risk assessment grouper, rather than develop a new grouper.

The experts would urge purchasers to look for a grouper that has been tested on the same sources of data that are available, e.g., a grouper tested with only hospital diagnoses versus a model that accepts diagnoses from all settings. Some groupers have been validated for different sources of diagnoses than are available to the specific purchaser.

Other criteria to consider when selecting a risk assessment grouper include

- the transparency of the model (e.g., clinically meaningful, logical algorithms);
- good performance for the specific sub-populations that are an important focus of the purchaser's program (e.g., people with disabilities, frail elderly);
- cost.

Selecting a Prediction Model

There are two types of predictive models: prospective and concurrent. Prospective models use diagnoses from one time period to predict an individual's medical expenditures in a future period.

Concurrent models use diagnoses to explain variation in an individual's medical expenses during the same time period in which the diagnoses were recorded.

Characteristics of a prospective model:

- Potentially moves less money around
- Possibly a good phase-in model, because it moves less money around
- Smaller range of payments
- Higher base rate
- Requires at least six months of assessment data to apply a risk score
- High risk plans may never catch up and low-risk plans will continue to be overpaid
- Requires carve-outs or case-specific adjustments, e.g., for births
- Appropriate when turnover in the population is low

Characteristics of a concurrent model:

- Payments more closely reflect the medical costs; therefore capitation payments may fluctuate more than payments from a prospective model
- Wider range of payments
- Lower base rate
- Does not require a minimum enrollment length to apply a risk score; therefore it allows inclusion of a larger proportion of beneficiaries with short eligibility periods

Concurrent models explain much more of the variation in expenditure than predictive models and should be given serious consideration.

Obtaining payment weights for the grouper

Can weights be established without health plan data or historical fee-for-service data in hand? The experts agreed that, yes, weights can be derived from fee-for-service data. The data can be three to four years old; they should be imported from similar populations in other parts of the country (e.g., use commercial data for commercial populations; Medicaid for Medicaid populations; Medicare for Medicare populations). Based on the experience of these experts, it is appropriate, cost saving, and time saving to import weights from generally similar populations as long as there is a similar benefit package (e.g., similar with respect to mental health coverage, pharmaceutical coverage, and large deductibles).

When choosing weights, public programs need to consider each product's history and political context as well as benefit design. For example, different weights may be needed for a state's medical assistance program than for its health benefit program for the medically needy who do not qualify for Medicaid. This need for different weights occurs because the budgeted or average expenditure for the two products is different even though the benefit design is similar. The health-based risk adjusted payment may not fully capture that difference in average per member/per month expenditures because the base rates were established in the context of different policy objectives.

Testing alternative models

Purchasers can bypass testing the predictive performance of alternative models on their own data before the simulation phase. They have the option of picking a tested grouper based on others' evaluations and importing weights without further testing of the general performance of the

model for their own program on their own data. The next step, which requires using program data from the plans, would be to conduct a simulation as described in Section 8 below.

Lags

The reason to be concerned about lags is that lags will reduce the accuracy of the payment. Lags can affect payment in two ways.

- The length of the eligibility requirement will exclude some beneficiaries who do not meet it.
- Some predictive power is lost. It is harder to accurately predict future, more distant, time periods than proximal time periods.

Lags can be caused in three ways. These include (1) the length of the assessment window, (2) the time required for claims run-out, and (3) the time to execute the risk scoring.

The length of the assessment window required by the model: A prospective, individual model requires continuous eligibility throughout the assessment period and for each month that payment is made. A concurrent model does not require a lengthy diagnosis history. Applying an aggregate plan-level risk score to adjust future payment allows one to eliminate the requirement for a specific individual to be continuously enrolled from the assessment period through the lag to the payment period.

The time required for claims run-out: Plans will have unequal claims run-out periods. The purchaser is unlikely to be able to influence this claims lag. It is important to wait sufficient time for all the plans to reach a similar level of data completion, otherwise payment will be biased.

The time to execute the risk scoring and the frequency of risk scoring: Purchasers can control how often and how fast they compute and assign risk scores. Combined with the usual claims run-out lag, the range can be from a six month minimum (as for BHCAG) to up to 24 months (as for WSHCA).

6. What were the management lessons learned during the planning phase?

Three critical management tasks are building effective teams, managing stakeholders, and educating the team and the stakeholders.

Building effective teams

Executive champion: The key to successful implementation is having an executive champion. (There is a problem when they leave.) The executive champion is important because implementing risk adjustment is a lot of work; they have to know that there is a problem to solve (keep an eye on the goal); and they have to keep at it. They have to be able to maintain focus while they are working on many other issues that compete for their attention.

Project manager: Implementation will also require a committed project manager with protected time, protected project status, and resources dedicated to the project. One mechanism that has been successful in achieving these aims has been to have a special project funded by outside funding agencies that requires outside researchers. Another model, used in Maryland, is to have committed resources available from the University to assist with a number of policy issues, including health-based risk adjustment. These approaches allow access to needed technical expertise, which is not yet routinely available.

Other team members: The team will require a range of expertise that includes

- technical expertise in health-based risk adjustment;
- knowledge about the program (coverage and eligibility policy and rate setting);
- knowledge about claims, encounter, and eligibility data;
- expertise in confidentiality;
- expertise in training and education of diverse audiences.

Other supporters: In addition to these team members, it is useful to include other allies who will support the endeavor, including those interested in collecting and analyzing data for quality assurance purposes. They should be involved at this early phase because the health-based risk adjustment planning may affect their work plan. Their involvement also builds a broader political base.

Identifying and managing stakeholders

Stakeholders must be identified. (See Section 3 above.) They will include the purchaser's governance group (executive management), plan representatives (product manager, financial or actuarial consultant, data expert), provider representatives, and, as appropriate, unions and advocacy groups.

A purchaser will need a formal stakeholder plan for dealing with each of these stakeholders. This plan should limit the number of people involved, while making sure that the person representing a group of stakeholders is trusted by that group.

At the outset, most plans think that they have a higher risk population than average. Experience shows that, not uncommonly, many of the participating plans will have risks that are somewhat below average for the pool, while only a few of the plans will have above average risk. For this reason, it is important to lock in the decision rules before conducting the simulation of plan payments.

Educating the team and the stakeholders

The need for education is critical, endemic, and persistent. At the outset, both within the implementation team and in the stakeholder community, no uniform knowledge base will exist, but there will be many misconceptions about health-based risk adjustment.

Education needs to happen early and often, especially before milestones, such as the simulation and implementation. Each stakeholder may require a different approach. The time and resources needed for education should not be underestimated.

It is critical to set up the authorizing environment. Purchasers should define who will make which decisions and the responsibilities of each party. The method should be decided before a simulation is considered.

7. What were the program and policy lessons learned during the simulation phase?

Lessons learned during the simulation phase include the importance of conducting a simulation, establishing goals and methods for the simulation, and anticipating the effects.

Importance of conducting a simulation

The simulation is a watershed event. This is the first time that purchasers and plans will understand quantitatively the relative risk of the plans. It is the first opportunity for the plans to assess the magnitude of the impact of health-based risk assessment on their payments. The new knowledge permanently alters the relation of the plans to the purchaser and the plans to each other.

Establishing goals

Because of the significance of the simulation, it is important to define clearly the goals for the simulation. As described in Section 5 above, purchasers may want to test and compare test risk assessment groupers for predictive accuracy. This goal for simulation is less important, however, than the goals of testing and improving data quality; demonstrating the relative risk of the plans (how the pie will be shared); and, combined with other components of the payment formula such as the base rate, testing the effect on payments to plans.

Establishing rules

Another step in the simulation phase is establishing rules for the simulation, including the methodology, data availability, and dissemination of results. The rules should be determined before simulation results are known. For example, will the results of the simulations be blinded? How are the results of the simulations to be shared with plans? All the participating implementation sites released the scores of all plans; some of them blinded all results except for the plan's own. If the results are shared, will the relative risk scores be shown without the base rates? Does the purchaser test and show the effect of different phase-in strategies?

If the goals are to focus on the risk assessment method and the relative risk score produced, then the analysis and results should be limited to relative risks and data quality and should not include base rate issues.

Anticipating effects

Some effects have been commonly observed from the simulation phase.

- Quantitative results tend to confirm and reinforce previously-held qualitative perceptions of plan relative risk. An expert will probably be able to recognize differences in risk of about 10%. The simulation not only identifies extremes but also provides more quantitative refinement on the relative ranking of the rest of the plans.
- Data completeness and quality will improve, both leading up to the simulation and as a result of feeding the information back to the plans.
- Experience shows that once data from the simulation are released, some plans may no longer participate.
- Data from the simulations will be used in plan negotiations, independent of implementation of health-based risk adjustment.

8. What were the technical lessons learned the simulation phase?

Technical issues at this point revolve around the data.

Data volume: This phase will be more challenging if there has not been some early testing and data handling in the planning phase. The simulation may be the first time the purchaser will be

handling massive amounts of data, especially the encounter data. It is wise to expect a lot of last-minute processing of encounter data.

Analytic skills: Purchasers need to ensure that they have the analytic capability to assign risk scores (either in-house or consultant). They also need someone who understands the program data and program payment, for example, knowing that diagnoses from certain sources will be included or excluded.

Coordination: This phase requires close coordination of disparate activity, for example, making sure that the risk adjustment experts work in close coordination with the actuaries. At this point, some plan actuaries have experience with health-based risk adjustment and others do not.

Electronic edits: The critical data quality issues for risk adjustment are not necessarily those that are captured in a fee-for-service edit system. It will be necessary to selectively bypass some of these fee-for-service edits.

Data testing: Data should be examined for reasonableness. Examining the frequency distributions of various data elements will help identify incomplete encounter data. Although there are not norms, there is some information about what non-contact percentages to expect. Data may be missing because of sub-capitation or because of carve-outs. A common problem is missing mental health provider data for a program that covers mental health services.

The purchaser should return the data that made it through the grouper to the plans. The plans then can go back to identify gaps. Some plans will know significantly more about their data than other plans and will be better at assessing the quality and completeness of their own data. The plan's degree of experience using its own data for activities other than payment usually predicts its depth of knowledge.

Different types of plans have different types of data problems. Staff models that have limited experience with fee-for-service billing will have concerns about data layout for encounters and the bundling of services. Plans whose systems truncate the number of diagnosis codes per record will worry about the number of diagnoses.

Inclusion of base rates: If the purchaser intends to simulate payments to plans, the base rates should be calculated in coordination with the decisions of the risk adjustment system, for example, which people are included in which rate.

9. What were the management lessons learned during simulation/phase-in?

The simulation creates a teachable moment. Attention from the stakeholders is greatest at this time.

Need for information management plan: The purchaser needs to develop an information management plan to communicate the simulation results to key stakeholders. The objective is to release a balanced message, differentiating it from the opinions of plans that might be unhappy with the results.

Need for education: There should be a serious plan and resources committed to intensive education at the beginning. The educational process needs to be done and redone until people stop coming. It is important to educate plan actuaries annually.

Meeting attendance: When data from the simulation are released to the plans, different plan representatives will attend compared with those who attended the earlier planning meetings. Higher level people will attend, and they are likely to be less knowledgeable about the history and technical processes (in spite of any education that the plan account representative may have tried to impart). Purchasers need to expect this shift in meeting membership and plan for it. Attendance should be monitored.

Plans with lower risk scores are likely to send more senior representatives than plans with high risk scores. Meeting conveners should encourage high level representation from all participating plans to provide balance in the meeting discussions.

10. What were the policy and program lessons learned during implementation?

As implementation progresses, purchasers should monitor the effects on both payment and plan participation.

Monitor effects of implementation

It is important to assess whether or not budget objectives (either keeping the pie size constant or even shrinking) are met. With the exception of Maryland (whose payment temporarily increased substantially to plans due to a technical error in implementation), purchasers experienced some modest increase in per member costs. These increases were attributed to changes in the risk pool (e.g., healthier Medicaid members leaving the pool during stronger economic times, or sicker people becoming employed thus raising the risk for commercial pools). In some cases, the base rate was 3 to 4 years old coming from a time when the overall covered population was a lower risk population.

Purchasers should also monitor plan participation to identify any effect of health-based risk adjustment on provider choice, geographic access, and access to specialists. What plans are leaving? Are the providers in those plans available through other participating plans?

In later cycles, as implementation progresses, purchasers should become interested in whether or not the incentives are transferred to capitated providers. Another indication of health-based risk adjustment impact is increased interest of providers in using the risk adjustment information for clinical management.

A purchaser can also use this information to illuminate policy issues, such as being able to compare efficiency of health care organizations.

11. What were the technical lessons learned during implementation?

Technical issues during implementation include data quality, auditing, updating the weights, improving the risk adjustment method, and improving the risk adjustment process.

Data quality, including completeness and accuracy

Data quality can be an issue at the plan level and also at the provider level. Data concerns at the plan level revolve around completeness, while data issues at the provider level include both completeness and accuracy.

At the plan level, the concern is to capture all diagnoses that have already been recorded by the provider. Plans may be missing diagnoses for two reasons.

- They may be missing encounter data from some providers.
- They may be truncating the number of diagnoses per encounter supplied by the provider.

Plans can, and do, resolve plan-level problems with data completeness.

At the provider level, there are three possible activities that can change the number and distribution of diagnoses and can increase the measured risk for a population when, in fact, the underlying morbidity of the population may be stable.

- Diagnostic discovery -- Increased number and severity of diagnoses are reported, all of which are appropriate. The correction of previous underreporting will reduce the problem of lack of persistence of diagnoses and more fairly represent the illness burden of the population.
- Diagnostic creep -- Increased number and severity of diagnoses for cases where the diagnosis is uncertain. This represents an upward bias in response to payment incentives. Many groupers try to minimize this problem by bundling related diagnoses and by excluding ill-defined codes.
- Tentative diagnoses -- Represents a potential source of error, when a diagnosis is appropriately used to justify a diagnostic procedure (rule-out) or to signal the need to treat a person without confirmatory diagnostic tests as if the patient has the disease (presumptive), because delay in treatment is harmful. Here too, the groupers have rules for excluding codes that are highly likely to be tentative.

No purchaser has detected changes in provider coding patterns yet, but it is important to keep looking and to set up monitoring and auditing systems that examine coding practices.

Error and fraud: validation of encounter data

Another potential problem that could occur at either the provider or the plan level is that the diagnoses in the encounter record are not valid. This situation could result from errors in processing or from fraud – the fabrication of diagnoses.

Auditing for plan-level data completeness

There may be a lack of persistence of some diagnoses (diagnoses that would be expected to continue e.g., quadriplegia). This lack of persistence may be the result of incomplete data submission, for example not receiving special authorization to visit specialists. Encounter data can be benchmarked against fee-for-service data, and an underreporting payment adjustment can be built into the phase-in plan. Colorado and Maryland achieved parity with fee-for-service data within a few implementation cycles.

Auditing for provider-level data accuracy

Some purchasers have begun medical record audits and some have not. One strategy is to develop linkages with other measurement activities such as quality assurance. Washington State Health Care Authority audit efforts (using medical records) have been complicated by confidentiality concerns. Others are seeking to automate data quality monitoring through clinical edits and audits of encounter data for illogical combinations or changes in the relationship of diagnoses and services provided.

Updating the weights

The aging of the weights was found to be less important than the experts had anticipated because of the stability among relative values of different risk categories over time. Weights must be updated eventually (after 3-4 years) to account for changes in practice patterns, coding changes, and significant changes in benefit design.

Why move to weights based on encounter data? The major incentive to move to encounter-based weights comes from the gradual erosion of the fee-for-service data over time. As more individuals move from fee-for-service to managed care, the residual fee-for-service population becomes both smaller and increasingly different from the managed care population. In addition, it may be possible to gain some additional validity from encounter-based weights that reflect the clinical and coding practices of a managed care environment.

Using encounter data for weights requires the highest standard for completeness. Although duplications of diagnoses can be tolerated in the risk assessment, duplications of charges are deadly for establishing proper weights. The second issue in developing weights is understanding how to apply charges to encounter data.

Improving the risk adjustment method

There are two objectives in refining or updating the model.

- (1) Improve the performance of the risk assessment measures, including
 - making the model more robust to coding variation, thus reducing error;
 - making the model more predictive by using updated versions of the grouper. (Note that experts are currently trying to improve models for important special populations, e.g., including non-English speaking, frail).
- (2) Improve the prediction model by
 - changing the model from prospective to concurrent;
 - revisiting assumptions about minimum enrollment requirements.

Improving the risk adjustment process

There are two approaches to improving the process:

- Reduce administrative processing time when possible while preserving data completeness
- Move to more frequent updates of the risk scores

12. What were the management lessons learned during implementation?

It is important to develop the internal capacity to implement health-based risk adjustment routinely, rather than to continue to rely on consultants.

Experts reported that plans would challenge payment or identify problems in payment, but in none of those challenges was the health-based risk adjustment itself questioned; it was some other part of the payment application.

Education remains a continuing management issue. Because people shift positions, new people -- both within the purchasing unit as well as with other stakeholders -- need to be educated.

This report is a product of RAIS: The Risk Adjustment Impact Study
conducted by: Health Research Center, Park Nicollet Institute
funded by: The Robert Wood Johnson Foundation

Copyright©, Park Nicollet Institute, 2001