

**Getting to Yes:  
How Encounter Data Become Good Enough for Health-based Risk Adjustment**

Rachel Halpern, David J. Knutson, Jinnat B. Fowles, PhD

Until health care purchasers began implementing health-based risk adjustment and, to a lesser extent, conducting quality assurance activities, health plans had little incentive to collect or submit complete and accurate encounter data. Plans faced with these demands are now developing their capacity to support the requisite data collection. Although there is room for improvement in data collection, the progress of state Medicaid programs and purchasing coalitions in implementing health-based risk adjustment should encourage other purchasers who are considering health-based risk adjustment.

In general, purchasers who incorporate diagnosis-based risk adjustment in their payments to health plans are concerned about two potential problems: (1) collecting complete enough encounter data from the plans to implement risk adjustment, and (2) coping with a potential increase in measured risk. In this paper, we examine the experiences of Maryland, Colorado, and Minnesota Medicaid programs in addressing concerns about the completeness and accuracy of their encounter data.

**Data completeness**

Data completeness is more of an issue for health plans than for providers. The plan's main concern is to be sure to capture all of the diagnoses that have been recorded by the provider. Encounter data submitted by plans to the state, or to any other purchaser, can be incomplete for two reasons.

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

The states of Maryland, Colorado, and Minnesota have had experience with incomplete encounter data. Maryland and Colorado were missing data from providers because of carve-outs or subcapitation arrangements. They each conducted a series of analyses to help identify incomplete data.

## Types of analyses to identify incomplete data from health plans

Maryland requires participating health plans to use the HCFA 1500 and UB92 formats for data uniformity. To identify plans whose data completeness appeared far below average, analysts in Maryland compared each plan's data with an all-plan baseline. Other analyses in Maryland included (1) comparing the utilization in the encounter data to an estimated fee-for-service benchmark; (2) comparing each plan's data with that of the plan with the most complete data; and (3) comparing submitted encounter data with other state data, such as data from external quality review organizations and chart reviews. Maryland Medicaid staff also conducted a study to evaluate the completeness of plans' encounter data by comparing the distributions of Adjusted Clinical Groups (ACGs) and case-mix scores for a cohort of Medicaid recipients who had received their care under fee-for-service between 1995 and 1997 and moved into managed care in 1998 and 1999.

To determine the extent of missing diagnoses from encounter data, Colorado Medicaid also compared data from individuals who moved from fee-for-service Medicaid into an MCO. The state of Colorado's specific concern was with a problem that had been noted even in fee-for-service claims data--that the diagnoses for serious chronic illnesses sometimes were not repeated in the administrative data in subsequent years of enrollment. It would be expected that a chronic disease such as cystic fibrosis would be actively managed and that the relevant diagnosis code would appear at least annually. This problem of non-recurring diagnoses is referred to as a lack of persistence. The lack of persistence was exacerbated for individuals with chronic illnesses who were enrolled in managed care options. The observed level of non-persistence of chronic illness diagnoses has been the basis for a belief that MCOs under-report true illnesses.

## Results of analyses

Consistent results emerged from the analyses of Colorado and Maryland. Both found that the completeness of Medicaid data dropped immediately after the transition from fee-for-service to managed care, as a result of gaps in their encounter data. In both states, the health plans with the least complete data were those that sub-capitated their providers. A crucial consequence of the incomplete data was that health plans' case-mix

scores did not reflect the severity of the illness burdens of the Medicaid beneficiaries enrolled in the plans. A study commissioned by the Colorado Medicaid program in 1998 comparing encounter data with information in medical records revealed that one-quarter to one-third of the care provided to Medicaid recipients was not recorded in the encounter data submitted by providers to health plans. Consequently, health plans submitted incomplete data to the state. In Maryland's comparison of fee-for-service and managed care ACG distributions and case-mix scores, Medicaid recipients' case-mix scores rose slightly between 1995 and 1997 (the fee-for-service period) and then dropped in 1998 and 1999 (the managed care period). The drop in case-mix was 11% for Medicaid adults and children, and 17% for persons with disabilities.

#### Response to results of analyses

Both Colorado and Maryland responded vigorously to finding incomplete data. First, they each adjusted the plans' reimbursement rates to compensate for missing data. Second, they provided extensive technical assistance to plans to assist with their data collection efforts. Because the state invested in the education of health plans regarding risk adjustment data issues, and also provided imminent financial incentives, problems associated with data completeness were addressed and resolved. The completeness of encounter data has increased over the last two years, and case-mix scores are approaching, or have reached, expected values.

The immediacy of financial incentives played a dramatic role in Minnesota. Minnesota Medicaid had required the submission of encounter data from health plans since 1985. Between 1985 and 1999, encounter data were not reliably forthcoming; the state did not pursue the data very enthusiastically; and the health plans saw no consequence in not complying with the submission requirements. In 1999, the year before Minnesota implemented its risk adjustment program, the situation changed dramatically. Health plans realized that the alternative to submitting their encounter data was a potential drop in payment. Minnesota Medicaid, for its part, also had to commit to the data collection process by modifying the submission requirements, excluding some of the more onerous data fields that had previously been specified. At this point, both the

state and the health plans committed themselves to the process, and encounter data began to flow—quickly.

### **Increased numbers of diagnoses being coded**

At the provider level, there are three possible activities that can change the number and distribution of diagnoses. These changes in the way providers record diagnoses will increase the measured risk for a population even though the underlying morbidity of the population remains stable.

- **Diagnostic discovery:** Increased number and severity of diagnoses are reported, all of which are appropriate and based on clear diagnosis indicators. Diagnostic discovery reduces the problem of lack of persistence of diagnoses and more fairly represents the illness burden of the population.
- **Diagnostic creep:** Increased number and severity of diagnoses for cases where the diagnosis is uncertain. This upward bias is a 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. This situation arises when a diagnosis is used appropriately as a rule-out diagnosis to justify a diagnostic procedure. It may also serve as a signal for the need to treat a person as if they have the disease without confirmatory diagnostic tests (presumptive), because delay in treatment is harmful to the patient. Here, too, the groupers have rules for excluding codes that are highly likely to be tentative.

Purchaser concerns about increased coding by providers remain secondary to concerns about incomplete plan-level data. More time and effort is required to influence a provider's coding practice than to produce complete plan-level encounter data. The challenge for purchasers will be to distinguish between diagnostic discovery, diagnostic creep, and a genuine change in case-mix. So far, developing methodologies to make these distinctions has not been a high priority for Colorado, Maryland, or Minnesota. Staff in these programs agree that changes in provider-level coding may become a problem later, but it is not an issue at this stage in the implementation process.

Error and fraud: validation of encounter data

Another problem that could occur at either the provider or plan level is that the diagnoses in the encounter data are not valid. This problem could be the result of errors in processing or be the result of fraud--the fabrication of diagnoses. Neither Maryland nor Colorado has devoted extensive resources to developing methodologies to assess data validity. The 1998 study commissioned for Colorado Medicaid, mentioned previously, was principally a validation study conducted in concert with its routine quality measurement activity. They compared samples of 1997 encounter data with the corresponding medical records. Although data were missing for certain providers, for providers whose data were captured, the situation was encouraging. For these providers, the study found consistency between the number of diagnosis codes recorded in the medical records and the number of diagnoses in the encounter data. The state of Minnesota is developing a request for proposal for a data quality and data integrity function for its Medicaid encounter data.

### **Conclusion**

Risk adjustment cannot be implemented without adequate encounter data. Although data collection and data processing are not trivial tasks, the primary data collection obstacle, to date, appears to be a lack of commitment and resources for such activities from both the states and the participating health plans. One of the most important lessons that has been learned is that commitment, patience, and perseverance must accompany encounter data collection efforts. This is not to deny the efforts required to develop and maintain high-quality encounter data. Nevertheless, these efforts must be preceded by state and health plan commitment to the processes. Once that commitment is established, other obstacles tend to lessen, become negotiable, and are ultimately surmountable. The consensus among states that have implemented health-based risk adjustment is that plans may require a little time (perhaps up to three years) to develop and implement an encounter data collection system that achieves acceptable levels of data completeness.

Encounter data now can be complete and accurate enough for state Medicaid programs and commercial purchasers to implement health-based risk adjustment. This data availability is likely to hold true for Medicare as well. The flow of encounter data has been positively correlated with the level of purchaser commitment toward risk adjustment and with the flow of dollars.