RAND

Adjustment of Consumer Assessments of Health Plans in the Consumer Assessment of Health Plans Study (CAHPS)

Marc N. Elliott, Ron D. Hays, and John Adams

DRU-1832-AHCPR

March 1998

Prepared for the Agency for Health Care Policy and Research

1. ADJUSTMENT OF CONSUMER ASSESSMENTS OF HEALTH PLANS IN THE CONSUMER ASSESSMENT OF HEALTH PLANS STUDY (CAHPS)

Case mix adjustment of consumer ratings can provide more valid plan-to-plan comparisons than unadjusted data by controlling for factors that are related to systematic biases in the way people respond to questions about their health care. Adjusted data are potentially less biased and, therefore, more appropriate for comparing the quality of care delivered. For example, if members of a particular ethnic group are less inclined to complain about bad care and members of this group are disproportionately enrolled in some plans, adjusting for this systematic bias is useful when comparing assessments of different plans.

However, differences in response patterns may reflect real differences in the quality of care as well as systematic biases in reporting and it is difficult to separate out these two factors. It is important to control for exogenous variables that lead to systematic response bias, but to be careful not to control for intervening variables that reflect the quality of the health plan.

Imagine two hypothetical extremes from a randomized experiment in which people are assigned to health plans. In the first situation, there is no response bias, but true quality of care differs on the basis of a variable such as age. In the second case, there is no difference in the quality of care, but there is response bias related to age. The latter case would be ideal for case mix adjustment by age, because the entire effect of such adjustment would be to eliminate response bias without biasing estimates of true quality of care. In the former situation, case mix adjustment would mask actual differences in quality of care associated with age, leaving only differential care by age. To the extent that the truth lies in between these extremes situations, the results will be a combination of the above consequences.

Because of the high stakes associated with how plans are rated by consumers, there is considerable political sensitivity related to the adjustment and presentation of these data. For example, it is possible that some plans will be dissatisfied with and dispute their scores. Those plans whose relative scores dropped after adjustment might be most likely to complain. Similarly, those with the lowest scores without adjustment would be most likely to complain. On the other hand, inadequate adjustment of consumer ratings for case mix differences could reduce the acceptance and usefulness of the results for beneficiaries, health plans, and purchasers.

The Consumer Assessment of Health Plans Study (CAHPS) Implementation Handbook recommends adjusting for age and health status when comparing consumer assessments of health plans. Younger people and those in worse health tend to report more problems and lower satisfaction with health care than do older people and those in better health.

Regression and stratification are perhaps the most appropriate methods used for case mix adjustment. Regression is the most popular method. In this approach, the observed value minus the predicted value for each person in the sample is calculated and the deviation represents the adjusted difference from the overall average rating. It should be borne in mind that any real differences in quality of health care associated with age or health status that hold for all plans will be made invisible by this approach. Unbiased estimates of plan differences in the regression approach are based on the assumption that the regression model is correctly specified, the model is the same for all the plans, and the covariates are measured with negligible error. It is possible to test for the second assumption (testing for interactions between plan and the covariates) and the third assumption (estimating the reliability of covariates). It should be noted, however, that the CAHPS recommended sample sizes will provide power to detect only gross violations of these assumptions.

To the extent that the model differs by plan, there is no meaningful overall plan difference to be estimated. For example, imagine that Plan A has much better ratings for the elderly than Plan B, but they have similar ratings for the young. This would be evident as an interaction between plan and age. It is much more plausible that this interaction reflects differential quality of care by age between plans than an interaction in response bias. Thus, if such a significant interaction occurred, it would be more meaningful to stratify by age and compare Plan A's ratings with Plan B's within age groups.

Any attempt to specify overall plan differences would be tantamount to constructing a weighted average of different effects. While this might ultimately be necessary, one would need to consider carefully what mixing weights proportions to use, as it will affect the relative performance of the two plans.

A second major approach to case mix adjustment is one-way or multi-way stratification by variables such as age or health status. One obvious advantage of this approach in the current context is that it is probably the easiest to explain to a general audience and perhaps the easiest to interpret.

In a sense, you are "closer to the data" than you are with regression. The corresponding disadvantage is that you do not get the "borrowing strength" and "smoothing" that you get with regression. Stratification can result in unstable estimates, particularly with the small cell sizes that are likely to result. Stratification is most straightforward for categorical variables. Continous other variables must be categorized into strata. To the extent that the true relationship between the covariate and response bias is a simple linear relationship, stratification is inferior to regression, as it involves a loss in information by collapsing the covariate into discrete categories. On the other hand, Daley and Shwartz (1994) argue that the relationship between age and some outcomes is unlikely to be linear, and that carefully chosen age strata may result in better models. The same

may be true of health status. While similar results may be obtained in the regression context through the use of dummy variables, stratification is more straightforward and transparent to lay users in this context. Another potential advantage of stratification is that it makes apparent differences by covariates that are hidden in regression. As the discussion of interactions suggested, stratification is also a more useful method when interactions exist between covariates and plan. It should be noted that if more than two stratification variables were used, the advantages of stratification would no longer be available.

One rule of thumb that might be useful is that regression will be more precise if the data is well-behaved (linear relationships, no interactions), but may obscure important information that would be revealed by stratification if the data is "poorly behaved." Additionally, regression will be the only practical approach if a large number of adjustment variables are involved.

Finally, we should consider the special nature of the health status variable, because there exists the possibility of causal relationships between it and plan selection. Consider two models of this relationship. In both models poor health status is associated with low evaluations. In the first model, health status differs among plans because of a priori differential "assignment" to plans (or choice of plans) on the basis of prior health status. This may involve simple geographic convenience, conscious choice, or other factors. The second model is one in which pre-existing health status differences do not exist, but plan choice in fact causes differences in health status to develop.

The first model is the classic case mix adjustment scenario. In the latter case, if plan choice actually causes changes in health status, one would argue against case mix adjustment by health status (at least as measured subsequent to current plan enrollment), because one would be eliminating true information on quality of care (perhaps in addition to eliminating some response bias). In this case, one might

wish to adjust for health status prior to enrollment in current plan (since the causality of the second model would not apply here), but one would not want to adjust for changes in health status since enrollment, or by extension, health status measured after enrollment in the current plan. Obviously, it is possible for a situation to exist that is a combination of the two models. To the extent that the second scenario is true, case mix adjustment by health status is inadvisable because it reflects true differences in quality of care rather than response bias.

BIBLIOGRAPHY

Daley, J. and Shwartz, M. (1994). In Lisa I. Iezzoni (ed) Risk Adjustment for Measuring Health Outcome (P.199-238), Health Administration, Ann Arbor, MI.