



Measuring the Quality of America's Health Care

DATE: September 15, 2006

TO: Sonya Bowen, Chris Haffer, Bill Long

CC: Bill Rogers, Lewis Kazis, Shirley Qian, Alfredo Selim, Kris Spector, Oanh Vuong

FROM: Judy Ng

SUBJECT: Medicare Health Outcomes Survey (HOS) Sample Size Discussion

This memorandum presents a brief background and the key questions discussed during an 8/7/06 group telephone conference regarding HOS sample size issues. It updates an earlier 8/21/06 memorandum on this issue, by including additional analytic documentation by Bill Rogers on sample size recommendations. The goal of the call was to investigate whether and what changes should be made to HOS sample sizes.

A. FINAL RECOMMENDATIONS

Based on further analysis described below, the following are our final recommendations on changing the HOS sample sizes:

- **Increase sample sizes by 100 or 200 to 1,100 or 1,200.**
- **Set plan minimum enrollment sizes at 500.**
- **Survey plans with enrollment sizes of 500-1,000 every 2 year.**

B. BACKGROUND

Currently, the HOS sampling approach involves random selection of 1,000 members per Medicare managed care plan. The sample size of 1,000 was designed to be large enough to allow for attrition due to nonresponse and disenrollment from the plan, and a target of 500 completed surveys. The sample size was determined on the basis of two assumptions: 1) that response rates of 70 percent at baseline and 90 percent at follow-up were achievable (with a 19 percent disenrollment rate between baseline and follow-up data collection), and 2) that baseline and follow-up surveys from approximately 500 enrollees would be sufficient to allow statistically significant comparisons and conclusions. However, the HOS has achieved lower average response rates and higher disenrollment rates than originally assumed.

In addition, while current sampling practices require all Medicare managed care plans to be included in the HOS sampling frame, data collected from plans with an enrollment of less than 100 members are excluded from analysis.

C. HOS SAMPLE SIZE QUESTIONS

The two key questions regarding HOS sample size are as follows:

- 1) Given that the HOS has achieved lower average response rates of 65 percent at baseline and 80 percent at follow-up, and higher disenrollment rates, than originally assumed, what should the increase in sample size be?
- 2) Given that HOS data collected from very small plans are not used, should a minimum plan enrollment size be set; if yes, what should it be?

D. OPTIONS FOR ADDRESSING SAMPLE SIZE QUESTIONS

Several options for addressing these two questions were discussed and are as follows:

a. Question 1 – Sample Size Increase: Increases in sample size were discussed in context of historical assumptions and the ambitions/objectives for the HOS.

In separate analytic documentation (see separate attachment for details), Bill Rogers further explained that the HOS was originally powered to distinguish plans that differ by 2 points on physical and mental component score outcomes (PCS/MCS), the 2 points difference being based on the Medical Outcomes Study. With an analyzable sample size of 500 completed surveys, we could distinguish plans that differ by 2 points with 90% power. However, subsequent results indicate that HOS plans are more similar than the plans in the Medical Outcomes Study. Even if outlier “best” and “worst” HOS plans continue to differ by 2 points in their outcomes, plans in the middle of the distribution are less than 0.5 points within the mean outcomes. On top of this, subsequent events have resulted in analyzable plan sizes of less than 300.

With only 300 follow-ups, we can distinguish plans that differ by a larger point difference, 2.7 points, on PCS or MCS outcomes with 90% power. Our ability to distinguish differences in plans that differ by 2 points has decreased to 70% power. (A difference of 2 points implies effect size differences on the order of 4%). Thus, our current samples are only just sufficient to deduce that positive outliers are plans that do better than the negative outliers. However, smaller differences may also be worth finding. Thus, distinguishing differences in plans by only 1 point may be worth considering.

Three options for increasing the sample size were presented.

- i. Option 1: Double the current sample size.** If the main focus with regard to increasing HOS sample size was to better characterize differences in PCS or

MCS outcomes, achieving this – given current response rates and functional decline rates – would require *at least doubling the current sample size*. Detecting differences of 1 point in PCS or MCS outcomes would require *quadrupling* the analyzable sample sizes to 1,235 completed surveys for 70% power and 2,104 for 90% power. However, if we use the Brazier approach, which analyzes functional status data in different ways, preliminary testing results indicate that analyzable sample sizes of just 700 completed surveys, which requires approximately *doubling* our current sample sizes, could distinguish 1 point differences with 70% power. Bill Rogers suggested a fuller evaluation of the Brazier approach in the future.

- ii. **Option 2: Increase the sample size slightly to 1,100 (or 1,200).** On the other hand, if the main interest in examining plan differences involved focusing on *outlier* plans, then a small increase in sample size would likely be sufficient. Raising the sample size by 10% would decrease the effect size by 5% and raising it 20% would decrease the effect size by 10%. In particular, given the loss of data resulting from the SF-36 to VR-12 transition, and given the HOS is not cheaper to administer despite the shortened length of the VR-12, even a small increase in the sample size would help make up for the loss of data without too much added cost. Bill Rogers estimated that the loss of efficiency due to a shorter survey are in the vicinity of 10-15%, so a 20% sample size increase would be more than enough to make up for this. Thus there is the option of *raising the sample size to 1,100 or 1,200 members per plan*.
- iii. **Option 3: Keep the sample size as is.** Bill Rogers pointed out the changes above won't change the general picture that the plans currently considered to be negative outliers are less than average, while positive outliers are better than average. He stated that what we have in place now is sufficiently powerful to describe HOS differences among outlier plans, so in light of cost, feasibility, or other concerns, another option would be to *keep the sample size as is*.

b. Question 2 – Minimum Plan Enrollment: Based on the above analyses, Bill Rogers recommended the exclusion of plans with enrollment of less than 500 members from the HOS sampling frame. Furthermore, if plans have an enrollment of only 500-1,000 members, he recommended that these plans get surveyed every other year, instead of every year. (Since it is the same population that gets surveyed in these plans each year, not much additional information is obtained by adding, for example, 2007/2008 data to 2006 data). Sonya Bowen also pointed out that surveying these plans every other year may alleviate burden, especially in light of complaints received from plan members who are surveyed every year.