

TECHNICAL PAPERS ON HEALTH AND BEHAVIOR MEASUREMENT

TECHNICAL PAPER 22

Field Test of Medicare Beneficiary Health Status Registry: Major Findings and Recommendations

Charles F. Turner, Barbara H. Forsyth, Sara C. Wheelless, Michael B. Witt, Kirk Pate, Timothy K. Smith, and Frank J. Mierzwa

Reference Citation

Turner, C.F., B.H. Forsyth, S.C. Wheelless, M.B. Witt, K. Pate, T.K. Smith, and F.J. Mierzwa. (1994) Field Test of Medicare Beneficiary Health Status Registry: Major Findings and Recommendations. *Technical Papers on Health and Behavior Measurement*, no. 22. Washington, DC: Research Triangle Institute.

**Field Test of
Medicare Beneficiary Health Status Registry
MAJOR FINDINGS AND RECOMMENDATIONS**

**Charles F. Turner, Barbara H. Forsyth, Sara C. Wheelless, Michael B. Witt,
Kirk Pate, Timothy K. Smith, and Frank J. Mierzwa**

Research Triangle Institute

August 4, 1994

This report was prepared and the research reported herein were performed as RTI Project 4883 under Contract 500-90-0053 between the Research Triangle Institute and the Health Care Finance Agency.

**Field Test of
Medicare Beneficiary Health Status Registry**

MAJOR FINDINGS AND RECOMMENDATIONS

BACKGROUND

The Medicare Beneficiary Health Status Registry (MBHSR) was conceived as a national data resource that would assist DHHS agencies and health researchers in a variety of important policy-related endeavors related to the provision and financing of health care for the elderly in the United States. These included¹:

- measuring access to medical care,
- understanding the patterns of use and need for health services,
- assessing the impact of changes in provision of health care under Medicare,
- providing a basis for the development of improved risk-adjustment procedures in the setting of reimbursement schedules for health care providers;
- assessing the effectiveness and quality of care provided under Medicare.

The designers of the MBHSR anticipate gathering data annually on the health status of large, probability samples of new enrollees in the Medicare system (with sample sizes of approximately 40,000 per year). These data would be linked with administrative and claims data in the Medicare record system and potentially with MBHSR reinterview data in order to provide a unique source of data on the health status and health care utilization of the elderly population. Over time, the accretion of this database would provide a variety of unique analytic possibilities -- some of which can be foreseen now and others of which may become apparent in the future. Among the opportunities that will be afforded by these data are:

- Continuous monitoring of the health status of the elderly population done with a sample size sufficient to allow detection of changes specific to particular socioeconomic, racial, geographic, and other groupings;²
- Evaluation of regional and other variations in the types of medical treatment obtained by Medicare enrollees with equivalent symptomology;
- Assessment of the relative costs of medical care that will be required over time for persons who report different classes of health conditions at Medicare enrollment;
- Development of research methods for both studying the long-term impact of patterns of health care upon subsequent morbidity and mortality and for estimating the relationship between health care status and subsequent utilization of health care services;
- Provision of a sampling frame for subsequent studies of specialized subgroups of the population aged 65 and older, such as those who had particular health problems prior to enrollment in Medicare.

¹As enumerated in Health Care Finance Agency (HCFA), OMB Clearance Package for the Pilot Study of the Medicare Beneficiary Health Status Registry, Epidemiology Branch, Division of Research, HCFA, 1992.

²It should be noted that a sample of 40,000 for any single year might not provide adequate precision for estimates of, for example, the prevalence of a particular health condition in a particular population subgroup. However, the aggregation of MBHSR data over several years would provide the possibility of obtaining the requisite precision.

DATA COLLECTION STRATEGIES

2. Use of mail data collection with telephone follow up for the MBHSR is a viable design option. This data collection strategy yielded response rates of 82 percent for new enrollees and 75 percent for the 76-80 year old cohort when a short-form MBHSR questionnaire was used (Table 1). When a medium length form was used, the corresponding response rates were 80 percent and 79 percent. These response rates *approach* (but do not equal) those obtained in major federal health surveys that use in-person interviewing.⁴
3. Each of the two followup stages of mail data collection substantially increased the obtained response rate (see Table 2). Over both age cohorts, the second stage of mail response increased response rate by 14 percentage points and the third stage increased it by 9 percentage points. We recommend that at least three stages of data collection be attempted by mail before resorting to telephone data collection.
4. While there were only slight (and statistically insignificant) differences in the final response rates between "new enrollees" and the 76-80 year old cohort, the older cohort was slower in their response to successive stages of mail data collection. So, for example, we note that the first of three stages of mail data collection obtained responses from 42 percent of new enrollees but only 33 percent of 76-80 year olds.
5. Followup interviews conducted by personal visit add substantially to the expected cost of the survey. Presently, we estimate that these costs add in excess of \$125 for each case requiring in-person interviewing.⁵ This data collection strategy did, however, yield a noteworthy increment in the overall response rates. In three metropolitan areas where in-person followup was used (Atlanta, Ga.; Richmond, Va.; Raleigh, N.C.), in-person followup of nonrespondents to the telephone phase of data collection increased response rates from 85 to 90 percent for "new enrollees" and from 82 to 89 percent for Medicare enrollees ages 76 to 80.⁶ Given the tradeoff between cost and benefit, we believe that it is reasonable for HCFA to consider dispensing with in-person followup. We would urge, however, that consideration be given to routine use of in-person interviewing for a small (probability) subsample of nonrespondents to monitor whether this "interviewable" component of nonresponse is having a marked affect on key survey estimates.

⁴For example, with in-person interviewing the Medicare Current Beneficiary Survey achieved a response rate of 87 percent in its initial round of recruitment and interviewing. This is seven to eight percentage points higher than the response rates achieved with mail data collection with telephone followup in the MBHSR (using medium length forms).

⁵This estimate represents only the variable costs of the field interview component plus related data receipt and processing costs. It does not take account of other costs that would be incurred, e.g., additional labor by supervisory staff.

⁶The very small additional number of persons interviewed in-person in these three metropolitan areas makes it impossible to reach reliable conclusions about the impact of this stage of data collection on the bias in our survey estimates.

BIAS DUE TO NONRESPONSE

6. Telephone followup is required both to produce an acceptable overall response rate and also to diminish the disproportionate impact of nonresponse among racial and ethnic minorities (see Table 3). Overall, telephone followup increased response rates by 16 percentage points. This effect was most pronounced (28 percentage points) for non-white persons in the sample.
7. Overall rates of response in the Field Test (see Table 3) were found to be somewhat lower than average for:
 - persons residing in the Northeast region (68 percent versus 75 percent for the total sample);
 - nonwhites (70 percent versus 76 percent for whites);
 - persons residing in metropolitan areas (MSA: 72 percent versus 83 percent for non-MSA residents);
8. Assessment of the changes in response distributions that occurred with each stage of data collection indicated that the additional stages of data collection improved coverage of respondents who reported poorer health and those who did not follow some common guidelines for disease prevention and health promotion. Thus respondents in later stages of data collection were:
 - Less likely to view their overall health as good and more likely to report recent declines in their well being;
 - More likely to report health problems that interfered with their social activities; and who
 - Followed fewer suggested disease prevention practices (e.g., having regular PAP smears, participating in an exercise regimen, etc.);

SURVEY LENGTH AND PRENOTIFICATION

9. Of the three lengths of survey forms tested in the Field Test (see Table 4), we believe that it would be inappropriate to use a form as long as the "Long Form" used in the Field Test. Across the two age strata in the study, the long form achieved an overall response (69 percent) that was considerably lower than either the Medium Length or Short Forms (79 and 77 percent). We recommend that HCFA use a form that is not significantly longer than the Medium Length form used in the Field Test.
10. Prenotification, as used in the Field Test, had an insignificant effect upon survey response (see Table 4). We recommend that the MBHSR dispense with prenotification.

ADEQUACY OF QUESTIONNAIRE AND ITEM COMPLETION

11. A review of the questionnaires returned by respondents and those completed during telephone interviews indicated that respondents, in general, provided usable responses to most questions. Defining a "completed questionnaire" as one that contained valid responses to at least 80 percent of 34 key questions for women and 32 key questions for men, we found that 98.6 percent of the new enrollees and 96.5 percent of the 76-80 year old cohort who returned a mail survey, "completed" their questionnaires. Rates were equally high for telephone interviews (98.4 and

99.3 percent), and these rates were stable across racial and gender groups and across different regions of the country.

12. Overall, response rates for individual questions were high. So, for example, questions on past surgeries, participation in disease prevention activities engaged in, nutrition, medical conditions, and disabilities all had average⁷ item response rates of 98 percent or higher. Somewhat lower rates were observed for questions asking about alcohol or tobacco use (<94 percent average item response), socioeconomic status (93 percent), income and assets (88 percent).
13. Neither questionnaire length nor prenotification had a substantial effect upon item nonresponse rates.

EFFECT OF FORM LENGTH AND PRENOTIFICATION ON SUBSTANTIVE RESPONSES

14. Neither form length nor prenotification appeared to have pervasive or noteworthy effects upon the substantive responses given by those respondents who supplied an answer to a given question.⁸

PROXY RESPONSE

15. Our findings indicate that proxy responses will be an important component of the MBHSR, especially for older enrollees. Without such proxy responses, survey estimates would give a biased portrait of the target population as healthier and more active than they actually are. In particular, we found that:
 - Proxy responses were supplied for 15 percent of cases in the 76 to 80 year old sample,⁹ but only two percent of the younger cohort responded by proxy.
 - Proxy response rates were higher among nonwhites than whites (22 versus 14 percent for 76 to 80 year old sample), but there were no significant differences in the proxy response rates for other demographic subgroups.
 - The most common proxy respondents (for 76 to 80 year olds) were: the child (48 percent) or spouse (32 percent) of the target person, other nonrelatives (9 percent), and siblings (3 percent).

⁷That is, averages across questions of a particular type (e.g., questions asking whether respondent had various surgical procedures).

⁸There was one odd exception to this rule: Respondents who were prenotified tended, on average, to report being in a more positive subjective state during the preceding 30 days (i.e., more frequently felt "happy" and "calm"). Since respondents were randomly assigned to receive prenotification or not, this finding cannot readily be attributed to differences in sample composition. (The null hypotheses that random fluctuations in composition of two groups could account for differences can be rejected with $p < .01$.)

⁹That is to say, questionnaires for 15 percent of all persons in the 76 to 80 year old sample were either completed by a person other than the respondent or were completed with assistance from another person. Beneficiaries who completed the interview on their own or who had someone else merely write down answers were considered to have responded for themselves. Beneficiaries for whom someone else completed the questionnaire (or who had proxy involvement other than having someone write down answers) were considered to have responded by proxy.

- Item response rates were generally high for the proxy respondents. Exceptions were noted for the Rose questions asking about the location of chest pain or discomfort and, for males, questions asking about difficulties in urination.
 - Analyses were performed to assess the impact of proxy response on survey estimates for the 76 to 80 year old sample. These analyses indicated that a sample that excludes proxy respondents would underrepresent the less healthy segments of the population in this age group.¹⁰
16. A methodological study conducted during the Field Test compared responses obtained from proxies with the self reports of target respondents.¹¹ Consistency rates of 90 percent or higher were found for questions asking about surgical procedures, medical conditions, and tobacco use.¹² Consistency rates of 70 to 89 percent were found for questions asking about alcohol use, IADLs, medications, and disease prevention activities (80+ percent).¹³ The lowest consistency rates (20 to 30 percent) were obtained for proxy responses to questions about economic status.
17. A large difference was found in the percentage of persons responding by proxy during the mail versus telephone stages of data collection (29 percent versus 10 percent for the 76 to 80 year old sample). It appears that many of the persons who opened the mail for incapacitated target respondents did not realize that they (or another suitable proxy) could complete the questionnaire on behalf of the target respondent.¹⁴ Proxy response to the mailings (for those who need a proxy) could be improved by stressing the possibility of proxy response in the cover letter.

¹⁰Persons in the 76 to 80 year old sample whose reports were provided by proxies appear (based on those proxy reports) to have more medical conditions, to use more medications, to have lower functional status, and to be less able to handle daily tasks of living by themselves, less mentally healthy, and less likely to engage in preventative health measures. In contrast, there were no substantial differences in the reporting of surgical procedures, nutrition, alcohol and tobacco use, sociodemographic characteristics, or economic status.

It is reasonable to assume that most of these findings represent actual differences in the health status of persons who require a proxy to complete their questionnaires (or telephone interviews) versus those who do not require such assistance. It is possible, of course, that some findings may reflect differences in the way proxies (versus the target individuals) respond to the survey. Some survey questions, for example, inquire about subjects that would be poorly known by the proxies. (The lower item response rates obtained from proxies for questions on the location of chest pain and urination problems suggests that this was the case, at least in some instances.)

¹¹Sampled persons who responded for themselves were asked to identify a proxy, and that person was asked to complete a survey questionnaire reporting on the sampled respondent. The consistency of the reports provided by proxies and by the target respondents was assessed for individual items and by content area. Because of a low response rate (60.4 percent for new enrollees and 46.1 percent for older enrollees) and the resulting small number of observations, this analysis focused on questions that were included in all three lengths of the questionnaire.

¹²These consistency rates did not vary across types of proxy (i.e., spouse, child, etc.).

¹³These rates were generally highest when the proxy was the target person's spouse or, to a lesser degree, the target person's child.

¹⁴Indeed, a number of proxies first called RTI to report that the beneficiary was unable to complete the questionnaire rather than completing the questionnaire for the respondent.

18. Some questions are difficult for proxies to answer, as evidenced by the low response and consistency rates obtained for some items in the Field Test. Consideration should be given to the development of a survey instrument that would identify questions that are inappropriate for proxies to answer and that would route proxy respondents around these questions.¹⁵

RELIABILITY OF MEASUREMENTS

19. As one measure of the quality of the data obtained from the prototype MBHSR instrument, we compared the responses given initially by respondents to those obtained during a reinterview conducted, on average, four weeks after the initial interview.¹⁶ Reinterview consistency was measured as percent of agreeing responses across the initial interview and the reinterview. This measure was selected because it is a measure that can be summarized across items with different measurement properties, different response formats, and different numbers of response options. Reinterview consistency rates were not computed for the few continuous variables, such as height and weight, because the large number of response options for continuous variables makes it difficult to interpret their consistency rates.
20. Across all questions in all three form lengths, respondents showed relatively high levels of consistency over time in their responses.¹⁷ Eighty-four percent of the 65 year-olds and 72 percent of the 76 to 80 year-olds answered more than 70 percent of questions identically in both interviews.¹⁸
21. Mean item reinterview consistency rates (see Table 5) varied across content areas and they also varied somewhat within content areas by age cohort. While consistency across interviews was good in many content areas (see Table 5),¹⁹ reinterview consistency rates of less than 50 percent

¹⁵This is clearly a task that will require creativity and careful pilot testing. Among the options that might be considered are variations in shading, color, or typeface that would provide a cue to proxies that a given question is inappropriate for them. (Instructions to proxies might be included in the amended cover letter; see below). Whatever solutions are tested, it will be necessary to insure that they do not confuse the sampled persons who will be responding for themselves.

¹⁶The reinterview study was conducted with a sample of 501 Medicare beneficiaries who lived in the three PSU's (Raleigh, NC; Atlanta, GA; Richmond, VA); reinterviews were completed for 433 of the 501 beneficiaries who were eligible for reinterview.

¹⁷All initial interview and reinterview responses coded as "blank page", "illegible", "multiple response", "legitimate skip", and "variable not defined for preassigned questionnaire length" were eliminated from consistency analyses. Some of these codes may reflect difficulties completing the initial reinterview form. They were eliminated from consistency analyses because they do not necessarily reflect unreliability linked to particular items.

¹⁸Our judgement that this level of consistency was "relatively high" reflects the fact that many items included in the MBHSR (e.g., items on mood) could be expected to show some true temporal change.

¹⁹For the younger cohort, 83 percent of the items on all three questionnaire lengths had 71 to 100 percent of respondents giving identical answers in both interviews. For the older cohort, 67 percent of the items on all three questionnaire lengths had between 71 and 100 percent of respondents giving identical answers in both interviews.

were found for questions on economic status as well for both the Rose items on location of chest pain and the questions on the use of assistive devices. There were also numerous examples of individual questions for which consistency levels suggest that further efforts are warranted to improve future measurement efforts.²⁰ (In interpreting the results presented in Table 5, it should be remembered that for some measurements -- such as the Rand-36 and Mental Health Index [MHI-5] which assess overall physical and mental health -- inconsistency across interviews may accurately reflect true change in the respondent rather than failings of the measurement procedures.)

22. Reinterview consistency rates were effectively independent of questionnaire length. For both age cohorts, reinterview consistency rates for the set of items included on all three questionnaire lengths were roughly equivalent. Across content areas, consistency rates were lowest for the items that appeared only on the long version of the questionnaire. Our analyses suggest that these lower mean consistency rates were due to the nature of the items themselves rather than to questionnaire length²¹
23. There were variations in response consistency within content area categories depending on whether initial and reinterview responses were provided by the beneficiary or by a proxy respondent. However, across content areas, self report and proxy report yielded roughly equal consistency rates for both age cohorts. These results suggest that including proxy reports in the MBHSR will not introduce excessive or systematic measurement error. There are, however, individual items (e.g., items on male urination problems)²² that yield unacceptably low response consistency when answered by proxy respondents.
24. The short screening test for memory and concentration problems used in the Field Test²³ was effective in identifying respondents who were likely to give inconsistent responses. Consideration should be given to adapting this instrument for use as a screening device in telephone followup for the MBHSR.

²⁰Since high consistency rates were obtained for many questions, we believe that the appropriate conclusion from our findings is that some content areas and individual questions present measurement problems not that our elderly respondents were unreliable respondents.

²¹Consistency rates for reinterview respondents answering long questionnaire versions were higher for items also included in the medium and short questionnaire versions than for items included only in the long questionnaire. Furthermore, these respondent consistency rates were roughly equivalent across questionnaire lengths for those items that were included in the medium and short as well as the long questionnaires.

²²Proxy respondents for the older cohort gave consistent responses only 32 percent of the time to the nine questions on male health included on the medium and long questionnaires. In contrast, respondents in the older cohort who answered these questions for themselves gave consistent responses 59 percent of the time.

²³Katzman, Robert, Brown, T., Fuld, P., Peck, A. Schechter, R., Schimmel, H. (1983). Validation of a Short Orientation-Memory-Concentration Test of Cognitive Impairment. *American Journal of Psychiatry*, 140(6):734-739.

25. Analyses of the types of interview-reinterview inconsistencies indicated that 80 percent of inconsistent responses resulted from respondents giving different substantive responses in the two interviews. Fifteen percent of the inconsistencies occurred when respondents answered an item during one interview, but either refused the same item or left it blank during the other interview.
26. During probing of discrepant responses in the reinterview, respondents reported that difficulties understanding questions and difficulties understanding or selecting response options were the most frequent reasons for inconsistent responses.

VALIDITY OF MEASUREMENTS

27. External indicators of the validity of the reports obtained by the prototype MBHSR instrument are available for a subset of respondents (see Table 6).²⁴ For instances in which Medicare records indicated that a surgical procedure had been performed and paid for by Medicare, we found very high rates of consistent reporting by respondents in the MBHSR Field Test. Although the numbers of respondents having specific procedures were small, we found that the survey instrument elicited complete reporting²⁵ of the following surgical procedures as recorded in the Medicare records:
 - Coronary angioplasty (19 respondents)
 - Implantation of a pacemaker (17 respondents)
 - Coronary bypass surgery (31 respondents);
 - Cataract surgery (17 respondents);
 - Mastectomy (7 respondents);
 - Hip replacement (6 respondents).
28. Reporting of other surgical procedures was less complete, but sufficiently high to encourage considerable confidence in the quality of these data. For example,
 - 12 of 16 recorded hysterectomies were accurately reported;
 - 39 of 44 recorded gall bladder removals were accurately reported.
29. Reporting of most medical conditions produced somewhat lower and more variable levels of accurate reporting in the survey (see Table 6). Thus we found that respondents reported between 75 and 86 percent of the following medical conditions coded in their Medicare files: diabetes, asthma, cancer, heart attack.

²⁴In particular, for respondents in the 76-80 year old cohort, it was possible to use Medicare records surgical procedures performed as a partial check on the accuracy of respondents' reporting of these conditions. Obviously, this comparison is not possible for new enrollees, and it does not permit us to assess the accuracy of reporting of medical conditions and procedures that respondents may have had prior to qualifying for Medicare. Indeed, the practical restriction is somewhat more limiting. Records were available to us only for procedures performed since 1983.

²⁵That is to say, all respondents who had received these surgical procedures under Medicare (according to the available Medicare records) reported having the procedure in the survey.

30. For two medical conditions, stroke and chronic obstructive pulmonary disease (COPD), we found substantially lower levels of accurate reporting. In particular,
- Only 24 of 41 respondents treated for stroke (according to Medicare records) reported having had a stroke in response to the survey;
 - Only 38 of 72 respondents treated for COPD (according to Medicare records) reported having COPD in response to the survey.

These two results are atypical, and they may reflect idiosyncrasies in the way in which diagnosis codes are assigned and used in medical record systems or a misunderstanding of the questions or the conditions by the respondents.²⁶

31. Analysis of the internal consistency of responses to the prototype MBHSR questions indicated that relatively high levels of consistent reporting were found in instances where explicit checks were possible.²⁷ For some items, non-trivial numbers of respondents failed to follow instructions to skip over sets of inapplicable questions. The succeeding items, however, were designed to incorporate response categories appropriate to these respondents. As a consequence, a respondent's failure to follow such "skip" instructions typically resulted in no loss of data nor the reporting of inconsistent answers. It appears that this questionnaire design was successful in reducing respondent burden by allowing respondents to skip inapplicable questions while insuring that data integrity was not compromised in instances in which instructions were not followed.

PHASED START-UP OF MBHSR

In light of the demonstrated feasibility of the MBHSR, issues of full-scale implementation merit some discussion. One of the most important issues is the manner in which this large-scale study would be phased in. While the designers of the Medicare Registry have focused attention upon recruiting cohorts of new Medicare enrollees and then following these new enrollees over time, there may be good reason to start the MBHSR using an alternate design.

There is an attractive simplicity to a research design that samples individuals at Medicare enrollment (typically age 65) and then subsequently collects administrative and reinterview data on these same respondents as the cohort ages. Repeated annually (or at other regular intervals), this design could provide a powerful tool for monitoring changes in the health status of the population and in their utilization of health care services. Nonetheless, it is important to recognize that such a design will take many years to provide data upon the older segment of the Medicare population. Since this

²⁶In the case of stroke, it might also reflect the impact of this condition upon the respondent's memory or cognitive functioning.

²⁷The instances consist of cases in which the response to one question (e.g., respondent reporting they had never used alcohol) implied that certain other responses were logically impossible (e.g., Reporting a nonzero amount of alcohol consumed during the prior 30 days).

segment makes greatest use of medical services, some important analyses would have to wait a decade or more before they could be carried out with the MBHSR as presently conceived. This problem will be particularly acute for analyses involving risk factors, functional statuses, medical conditions, and procedures that are most common among the older segments of the Medicare population (e.g., severe functional status impairments).

To address this startup problem, it may be appropriate to consider phasing in the MBHSR as a mixed design that would ultimately evolve into the simpler design originally contemplated by MBHSR's originators. In such a mixed design, an additional sample might be recruited in early years of the MBHSR in order to represent the older segment of the population. For example, if the evidence of the Field Test were taken to demonstrate the viability of the proposed data collection strategies for populations ages 65-80, one might specify two sampling strata: new enrollees and current enrollees ages 66 through 80. In the first year of the survey one might allocate a portion of the target sample of 40,000 to the strata of 66- to 80-year olds.²⁸ In the second year, one might define the second strata to represent persons aged 67 to 80; and in the third year, it might represent persons aged 68 to 80. Such a design would provide interim data monitoring the health status of representative samples of persons who were 65 to 80 years of age at the time the MBHSR survey program commenced. Such a mixed design would also provide a growing sample of these older Medicare recipients who might be subsampled in followup studies.

Over time such a second sampling strata would diminish in size. It would be displaced by the representative samples of new Medicare enrollees (typically age 65) from Year 1 of the MBHSR, Year 2 of the MBHSR, and so forth. By Year 10 of the MBHSR, the Year 1 sample of new enrollees would have reached an average age of 75, thus diminishing the need for continued sampling from the second strata.

The design sketched above is but one of many mixed designs that might be considered. Indeed, we would point out that the sample design of the Field Test which employed a strata of 76-80 year olds, might also be a reasonable candidate to provide data to monitor the health of the older segments of the Medicare population.

Preparatory Research. There is much design work that could be done in this area -- including better specification of the priority of different uses of the resultant data. We believe that further design work and, in particular, consideration of mixed-mode designs during an MBHSR phase-in period, would be a wise investment of resources if the MBHSR is to be implemented. Further design work in preparation for the fielding of the MBHSR will be appropriate in other areas. We note, in particular that proxies will respond for many elderly respondents with serious health problems in the MBHSR. Further methodological research should be conducted to explore ways of enhancing the

²⁸To insure equivalent precision in estimates, one might choose to balance the size of the sample of persons at each specific age in the older cohort so that the expected standard errors for key estimates were equivalent for persons of age 66, 67, 68, etc.

completeness and accuracy of these responses.²⁹ Similarly, the item response rates and consistency analyses conducted with the Field Test data can be used to identify questions that need refinement and further testing prior to use in the MBHSR. Investments in this developmental work can make a substantial contribution to further improving the quality of the final MBHSR instruments.

²⁹We do not now know, for example, how much of nonresponse during the initial stages of mail data collection involves respondents whose proxies did not realize that they could respond for the target person. Similarly, there is important redesign work to be done if the questionnaire is to be appropriate for self-report but also accommodate proxy response for questions to which proxy response is appropriate.

TABLE 1. Response Rate for Combined Mail and Telephone Survey Data Collection by Survey Form Length and Age Cohort

Cohort	Form Length		
	Short	Medium	Long
65-year olds (a)	82.0	79.6	71.9
(s.e.)	(1.7)	(3.2)	(2.4)
76-80 year olds	75.0	78.8	68.2
(s.e.)	(3.7)	(2.3)	(2.9)
Both Cohorts	77.2	79.0	69.4
(s.e.)	(2.3)	(1.9)	(2.0)

(a) Sample strata includes new enrollees to Medicare, the vast majority of whom are enrolled around the time of their 65th birthday.

TABLE 2. Cumulative Survey Response Rate by Stage of Data Collection and Age Cohort

Cohort	Stage of Data Collection			
	Mail-1	Mail-2	Mail-3	Telephone
65-year olds (a)	42.0	54.9	62.4	77.8
(s.e.)	(2.2)	(2.3)	(2.2)	(1.9)
76-80 year olds	33.0	47.5	57.3	74.1
(s.e.)	(1.4)	(1.7)	(1.5)	(2.0)
Both Cohorts	35.8	49.8	58.9	75.2
(s.e.)	(1.2)	(1.4)	(1.4)	(1.5)

(a) Sample strata includes new enrollees to Medicare, the vast majority of whom are enrolled around the time of their 65th birthday.

TABLE 3. Cumulative Response Rates for Mail and Mail Plus Telephone Data Collection for Selected Subgroups of Sample.

	Mail Only	Mail Plus Telephone
Total Sample	58.9 (1.4)	75.2 (1.5)
Gender		
Male	61.0 (2.2)	75.3 (1.9)
Female	57.4 (1.9)	75.2 (1.9)
Race		
White	61.1 (1.4)	75.9 (1.6)
Not White	42.5 (2.8)	70.1 (1.8)
Region		
Northeast	56.6 (4.7)	67.8 (5.2)
North Central	61.1 (1.8)	78.0 (2.3)
South	53.8 (3.3)	77.4 (3.3)
West	64.6 (3.3)	76.9 (3.2)
Metropolitan Area		
In MSA	55.6 (1.9)	71.7 (1.8)
Not in MSA	66.2 (1.9)	83.1 (2.0)

Note: Sample estimates for all subgroups include both 65 and 76-80 year old cohorts.

TABLE 4. Response Rate for Mail Plus Telephone Data Collection By Length of Questionnaire and Prenotification

	%	(s.e.)
Total Sample	75.2	(1.5)
By Questionnaire Length		
Short	77.2	(2.3)
Medium	79.0	(1.9)
Long	69.4	(2.0)
By Prenotification		
Prenotified	74.4	(2.0)
Not Prenotified	76.1	(2.1)

Note: Tabulation includes both age cohorts and response to 3 waves of mail data collection and telephone followup.

TABLE 5. Mean Consistency Rates for Reinterview Sample By Question Content Area, Age Cohort, and Form Length

Content Area	65 Year Olds (a) Forms			76-80 Year Olds Forms		
	S-M-L	M-L	L	S-M-L	M-L	L
Medications (n)	82.1 (2)	76.8 (3)	92.1 (4)	78.8 (2)	70.6 (3)	90.9 (4)
Health (n)	63.1 (1)	83.4 (3)	63.4 (3)	62.3 (1)	83.5 (3)	59.2 (3)
Assistive Devices (n)	-- --	88.6 (3)	27.9 (22)	-- --	88.1 (3)	34.5 (34)
Chest Pain (Rose Items) (n)	-- --	-- --	43.0 (16)	-- --	-- --	34.9 (17)
Female Health (n)	86.5 (3)	97.7 (2)	76.4 (3)	74.2 (3)	90.5 (2)	92.2 (2)
Socioeconomic Status (n)	78.3 (11)	98.5 (1)	-- --	67.7 (11)	98.9 (1)	-- --
Income and Assets (n)	20.8 (6)	62.2 (8)	37.7 (17)	23.0 (6)	67.5 (8)	42.6 (17)
Male Health (n)	97.5 (2)	73.3 (9)	100.0 (1)	93.5 (2)	50.8 (9)	83.2 (1)
Instrumental Activities (n)	81.6 (5)	98.9 (3)	67.4 (18)*	71.3 (5)	86.8 (3)	82.4 (20)*
Activities of Daily Living (n)	-- --	-- --	77.5 (5)	-- --	-- --	79.2 (6)
Mental Health (MHI-5) (n)	68.5 (5)	-- --	-- --	57.9 (5)	-- --	-- --
Tobacco Use (n)	93.6 (2)	64.6 (2)	-- --	92.9 (2)	63.9 (2)	-- --
Alcohol Use (n)	78.5 (1)	-- --	88.5* (4)*	82.6 (1)	-- --	89.4* (4)
CAGE - Alcohol (n)	-- --	-- --	88.5* (4)	-- --	-- --	89.4* (4)
Surgery (n)	96.4 (7)	97.0 (1)	94.7 (2)	87.4 (7)	89.1 (1)	83.7 (2)
Disease Prevention (n)	86.4 (3)	63.4 (3)	-- --	71.0 (3)	63.6 (3)	-- --
Nutrition (n)	-- --	89.7 (3)	91.2 (1)	-- --	83.6 (3)	90.8 (1)
Medical Conditions (n)	96.2 (3)	98.4 (3)	-- --	95.8 (3)	96.1 (3)	-- --
Rand-36 (n)	63.1 (1)	84.3 (6)	92.7 (1)	62.3 (1)	69.0 (6)	69.8 (1)

Note: Entries reflect questions included in three different combinations of Form Lengths: (1) Questions included on all three form lengths (S-M-L); (2) Questions included only on both the medium and long forms (M-L); and (3) Questions that appeared on only the long form (L). Numbers in parentheses in table indicate number of questions included in each analysis

(a) Sample strata includes new enrollees to Medicare, the vast majority of whom are enrolled around the time of their 65th birthday.

TABLE 6. Percent of Respondents with Claims for Particular Procedures or Conditions in Medicare Files who Report Having Had This Procedure or Condition During the Survey (76-80 Year Old Cohort)

	Number in Medicare Records	% Reported in Survey
<u>PROCEDURES</u>		
Coronary Angioplasty	19	100
Pacemaker	17	100
Coronary Bypass	31	100
Cataract Surgery	17	100
Mastectomy	7	100
Hip Replacement	6	100
Hysterectomy	16	74
Gall Bladder Surgery	44	89
<u>CONDITIONS</u>		
Diabetes	96	80
Asthma	19	76
Cancer	73	82
Heart Attacks	117	75
Stroke	41	53
Chronic Obstructive Pulmonary Disease	72	50