

Contract No.: 500-00-0033/T.O.05
MPR Reference No.: 6138-145

MATHEMATICA
Policy Research, Inc.

**Responses to OMB
Questions for the
Evaluation of the
Medicare Care
Management
Performance (MCMP)
Demonstration**

Volume I

January 8, 2009

*Julita Milliner-Waddell
Stacy Dale
Eric Grau
Lorenzo Moreno*

Submitted to:

U.S. Department of Health and Human Services
Centers for Medicare & Medicaid Services
Office of Research, Development, and Information
C3-23-04 Central Bldg.
7500 Security Blvd.
Baltimore, MD 21244-1850

Project Officer:

Lorraine Johnson

Submitted by:

Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005

Project Director:

Lorenzo Moreno

1. Why is the state (rather than the practice) the most appropriate level of comparison for the matched comparison groups?

States are not being used instead of practices as comparison groups. Practices, which are the unit of *intervention and observation*, are being selected as comparison group from states that resembled the demonstration states on a number of characteristics, as described in detail in Appendix C of the evaluation design report (Moreno et al. 2007), which is included in Volume II of this response. We did match the pool of potential comparison practices to demonstration practices using practice characteristics from the Office Systems Survey and other practice characteristics derived from Medicare enrollment and claims data (Dale and Holt 2008). (See Volume II of this response). The selection of potential comparison practices from a subset of states was to enhance *face validity*—that is, it would not have been very credible that a practice in Massachusetts would be similar to a practice in Arizona, assuming we picked potential comparison practices from the entire nation (excluding the demonstration states). Thus, because the four demonstration sites (states) are likely to differ substantially, the evaluation will estimate impacts separately for each site. Site-level differences may include: physician practice regulations, practice styles, practice settings, adoption of electronic health records, and pay-for-performance penetration. We will also report summary impacts across states to provide an assessment of the overall demonstration effectiveness. Finally, because overall impact estimates may mask important differences within groups, when sample sizes permit, we will estimate impact estimates for subgroups defined by practice features, such as practice affiliation or patient mix, and beneficiary characteristics, such as having a select chronic condition.

It is important to highlight that the analysis will *not* report practice-level impact estimates. As noted, although practices are the unit of intervention, the impact analysis should refer to the overall impact of the intervention by site, or combined sites, rather than by practice. This is in agreement with the standard practice of reporting impacts on the total sample of individuals, rather than on specific individuals who received the intervention and their control or matched comparison group, when the individual is the unit of intervention.

Moreno, Lorenzo, Stacy Dale, Suzanne Felt-Lisk, Leslie Foster, Julita Milliner-Waddell, Eric Grau, Rachel Shapiro, Anne Bloomenthal, and Amy Zambrowski. “Evaluation of the Medicare Care Management Performance Demonstration: Design Report. Final Report.” Princeton, NJ: Mathematica Policy Research, Inc., May 25, 2007.

Stacy Dale, and Jeff Holt. “Selection of MCMP Comparison Practices. Memorandum.” Princeton, NJ: Mathematica Policy Research, Inc., November 12, 2008.

2. Please provide more information about the specific criteria used to identify matched sample states.

Comparison states were selected that were similar to MCMP states in terms of their use of electronic health record or pay-for-performance programs and other key characteristics (such as their ratio of specialists to general practice/family medicine physicians). The methodology for selecting comparison states is described in Appendix C of the evaluation design report (Moreno et al. 2007), which is included in Volume II of this response. The methodology for selecting comparison practices is explained in a memorandum to CMS, which also is included in Volume II (Dale and Holt 2008).

It is important to highlight that at the time we proposed potential comparison states (July 2005), we did not know the number of practices in the Doctor's Office Quality—Information Technology (DOQ-IT) in non-demonstration states. Thus, when we received these counts in early 2008, we realized that the number of DOQ-IT practices available in several potential comparison states was insufficient to find enough matches for the treatment group practices in the four intervention states. Thus, we decided to expand the list of comparison states with the alternate states listed in Appendix C of the design report (Moreno et al. 2007), and Colorado. A summary of the number of potential comparison practices available from the comparison states is reported in Table 1 in Dale and Holt (2008).

Stacy Dale, and Jeff Holt. "Selection of MCMP Comparison Practices. Memorandum." Princeton, NJ: Mathematica Policy Research, Inc., November 12, 2008.

3. Please provide more information about the Office Systems Survey.

The Office Systems Survey (OSS) was originally designed by CMS (Office of Clinical Standards and Quality) for assessing progress of practices served by the Quality Improvement Organizations (QIOs) as part of the Doctor's Office Quality—Information Technology (DOQ-IT) initiative. The OSS was conducted by the Maine Health Information Center on behalf of CMS to all DOQ-IT practices in the nation.

For the MCMP evaluation, CMS agreed to use data collected in summer 2007 as the baseline for treatment group and comparison group practices. Data for 699 treatment group practices and for all DOQ-IT practices in the potential comparison states became available in February 2008.

CMS also decided to collect OSS data in the final year of the MCMP demonstration to be used as the follow-up measurement for both treatment group and comparison group practices. However, CMS decided that Mathematica Policy Research, Inc. (MPR), instead of the Maine Health Information Center, should collect the follow-up data. Furthermore, CMS decided that this data collection would be part of the data collection for the evaluation of the Electronic Health Records Demonstration (EHRD), which MPR also is conducting for CMS. Thus, MPR has prepared a supporting statement for Paperwork Reduction Act submission under the EHRD for the follow-up OSS survey. A copy of the OSS instrument is included as part of Volume II of this response. It is important to highlight that the follow-up OSS survey instrument is comparable to the OSS baseline instrument. MPR will collect data from up to 679 demonstration practices and from up to 548 comparison practices in fall 2009.

The OSS data will allow us to identify whether treatment group practices have changed how they use electronic tools to support quality differently from the comparison practices. The hypothesis is that the demonstration's incentives will increase practices' attention as to how they can use available means, including electronic tools, to improve quality, particularly on the measured dimensions of care. While we expect treatment group practices to be more advanced in their use of electronic tools than the comparison practices, due to longer exposure to assistance from the QIO at the beginning of the MCMP demonstration, it will be of interest for the evaluation to see if they progress measurably over the demonstration period.

In the final year of the demonstration, we plan to produce tables comparing the percentage of the treatment and comparison practices performing health information technology (HIT)-related activities on first measurement, final measurement, and change in these percentages, as well as the percent of practices with advancement in HIT use over the period for each activity included in the instrument. After we have conducted this broad-based analysis, we anticipate following up with additional analysis of two or three content areas that show especially promising results. For a detailed discussion of our analysis plans, see Chapter II of the design report (Moreno et al. 2007).

4. Why does the study use beneficiary self-reports for items that could be collected, perhaps more accurately, in administrative data such as claims data or health records (such as health status, diagnosed chronic conditions, measures etc taken during last health care visit, and flu vaccination over past 2 years)? Will these records be used to verify self-reported information?

Because electronic health records (EHRs) are not mandated by the MCMP demonstration, it will not be feasible to collect measures of health status and use of specific services from them. Furthermore, claims data will only provide comparable measures for both the treatment and comparison group practices for 7 of the 26 quality measures (see Table III. 1 in the evaluation design report). Thus, we have proposed to collect data in the beneficiary survey for an additional 6 quality measures, which will be available for both demonstration and comparison groups, so they can be used to estimate the impacts of the demonstration. Use of the beneficiary survey instrument for collecting data for these six measures ensures that the data are comparable for both study groups. Thus, we will be able to directly assess the demonstration's impacts on 13 of the 26 measures. (We will not ask questions in the survey about the other 13 measures because they are too technical for the beneficiary to know or remember the answer accurately.)

We will also use Medicare claims data to identify Medicare beneficiaries with target and other chronic conditions. Indeed, the demonstration's financial support contractor has already used claims data as part of the assignment of beneficiaries to practices in both the treatment and comparison groups. We will use these measures of health status in our analysis.

For the remaining 13 quality measures that will be extracted from medical records *only for treatment group practices*, we will use these measures for select descriptive and trend analyses, but not for impact estimation because there will be no comparable measure for comparison practices. Furthermore, we do not plan to validate the survey reports against the medical-only-records measures because (1) this validation could only be done for treatment practices, (2) there is no reason to expect that the accuracy of self-reported data will be different for the treatment and comparison groups, (3) the reference and measurement periods for the survey-based and chart-based measures are different, and (4) there are no resources for this comparison, which is likely to be very expensive.

5. For both surveys, are sample cases all being contacted at the same time, or is the survey rolled out over a period of months? If the former, why is the data collection period so lengthy?

We originally envisioned a rolling sample release due to the sites coming into the demonstration over time. However, the delay in the project schedule allowed for all the sites to be recruited by the start of data collection. This allows us to employ a single sample release for both surveys.

The proposed data collection periods of 12 months for the beneficiary survey and 11 months for the physician survey are based on the time we believe it will take to achieve the projected response rates. There are several reasons for needing a long field period for both surveys. Firstly, mail surveys require a longer field period than phone surveys to allow time for the mail to reach sample members and completed questionnaires to be returned before sending additional mailings. Secondly, physicians are a very difficult population to survey and require a great deal of follow-up which takes time (particularly during summer months when many take vacation). Thirdly, in the absence of a monetary incentive to encourage beneficiaries and physicians (especially comparison group practice physicians) to participate, additional follow-up efforts will be needed to reach the targeted response rates. We believe that if the field period is sufficiently long, the survey sufficiently short, and creative contact approaches are used, we can achieve the desired response rates. Nevertheless, we will make every effort to complete data collection in fewer months.

6. What are the planned intervals between mailings in the beneficiary survey (similar to the information provided for the physician survey)?

Several attempts will be made to reach beneficiaries after the initial mailing.

Approximately three weeks following the initial mailing, a reminder postcard will be sent to non-respondents. (This interval allows mail to be forwarded and/or returned if undeliverable.) Then, a second full mailing (letter, FAQs, mail questionnaire, and return envelope) will be sent to remaining non-respondents, approximately four weeks following the first postcard mailing. A second reminder postcard will be sent around week 9 of data collection. A third full mailing, perhaps using priority mail service, will be sent about 3 to 4 weeks later, about week 12 or 13. A third reminder postcard will be mailed to the remaining non-responding sample approximately two weeks later.

In the interim, locating letters will also be sent to alternate addresses for sample members whose mail is returned as undeliverable.

It is important to note that we will be assessing the response to our mail efforts on an ongoing basis and will make mid-stream adjustments as appropriate. For example, if we find that we receive good responses to the reminder postcards and/or our additional full mailings, we may substitute additional reminder postcards and full mailings using regular service before employing the more expensive priority mail option.

We will also be accepting call-ins from sample members from the beginning of data collection. The table below shows the planned data collection activities by week.

Week of Data Collection	Activity
1	Advance letter mailed to beneficiaries
3	First Reminder Postcard Mailed
7	Second full mailing to beneficiaries
9	Second Reminder Postcard Mailed
13	Third full mailing to beneficiaries (this may be a priority mailing)
15	Third reminder postcard mailed
1-15	Telephone call-ins taken
16	Telephone call-outs begin
17-End	Additional Reminder and Specialty (i.e., mailings on request) mailings as needed

7. Confidentiality—In SS A10, please state the statute under which confidentiality is being assured for each survey. The Privacy Act generally does not apply to people in their entrepreneurial capacity; does CMS have a particular reason to think it will apply here? Please also clarify in the cover letters, questionnaire introductions, FAQs, and all other places where confidentiality assurances are made that the data will not be released “except as required by law” or some other similar caveat.

Confidentiality for this project is being assured in accordance with 42 U.S.C. 1306, 20 CFR 401 and 402, 5 U.S.C. 552 (Freedom of Information Act), 5 U.S.C. 552a (Privacy Act of 1974), and OMB Circular No. A-130.

Changes to the advance letters, FAQs, and survey introduction have been made to clarify the limitations of the law. The advance letter for beneficiaries now reads as follows:

The answers you provide will be kept confidential and will not be released, except as required by law. Your information will be used only as part of this evaluation.

The FAQ on confidentiality that will accompany the letter to beneficiaries now reads:

WILL MY INFORMATION BE KEPT CONFIDENTIAL?

Yes. All of the information we collect in the survey will be kept confidential to the extent allowed by law. The information will be used for research purposes only. Your name will never be used in any reports.

The reference to confidentiality in the physician letters now reads as follows:

Your answers will remain completely confidential to the extent allowable by law. Neither your name nor your practices' name will ever be included in any reports prepared as part of this study.

All questionnaire introductions (mail and telephone versions) have been revised to read as follows:

All of your answers will be treated confidentially to the extent allowable by law.

8. Why does the burden table indicate that the cost per response for the beneficiary survey is “NA?” Does CMS assume that no Medicare recipients are working?

We assumed a cost per response for physicians because questionnaires will be mailed to their place of business and we assumed that they would take time out of their work day to complete the survey. In contrast, beneficiaries’ surveys will be mailed to their home. We assumed no cost per response for beneficiaries because even those who are working are not expected to take time out of their workday to complete the survey.

9. Please provide a copy of the first interim evaluation report, due in October 2008.

In summer 2008, CMS decided to drop the first interim evaluation report from the evaluation's Statement of Work as the result of financial constraints and other schedule changes. Instead, CMS asked us to submit the implementation report in lieu of the first interim report. CMS currently is reviewing the draft of this report.

10. Please clarify whether CMS plans to request clearance for the “telephone discussions with highly successful practices.”

CMS does not plan to request clearance for the telephone discussions with highly successful practices because it has only budgeted for contact with no more than nine practices. Furthermore, these discussions will not rely on a structured outline or protocol and will be guided by the particular features of the practices that are determined successful.

11. Please clarify how CMS determines a beneficiary’s medical condition (to be used as a stratification variable).

CMS will identify beneficiary’s chronic conditions from their diagnoses on Medicare claims using an algorithm that the demonstration’s financial support contractor has developed to assign beneficiaries to practices. The details are provided in Wilkin et al. (2007).

Wilkin, John, C. William Wrightson, David Knutson, Erika G. Yoshino, Anahita S. Taylor, and Kerry E. Moroz. “Medicare Care Management Performance Demonstration. Design Report.” Columbia, MD: Actuarial Research Corporation, January 5, 2007.

12. Why does the physician letter get signed by the CPO and the beneficiary letter gets signed by the CIO? Does CMS have cognitive or other results to suggest that these different signatories are more reassuring to these different respondent groups?

The advance letters to beneficiaries, demonstration physicians, and comparison physicians will all be signed by the same officer at CMS—the CMS Privacy Officer. The reference to the CMS Information Officer in Response C.2.a is an error. The sample letters that were provided in Appendices C, H, and I of the original submission correctly showed the CMS Privacy Officer as the signatory.

13. Given that both surveys are expected to have response rates below 80%, please provide the nonresponse bias analysis plan for each, in accordance with OMB statistical standards.

Non-response weights will be calculated using information from the sampling frame as covariates in logistic regression models with a binary indicator of whether the interviewee responded or not as the dependent variable. By choosing covariates that are related both to the outcome variables of interest and to the propensity to respond, non-response bias will be reduced. However, it will not be possible to remove non-response bias entirely. The following describes procedures to investigate non-response bias that is not alleviated by the use of non-response weights. If evidence of such bias is found, further investigation will be required to ascertain the source of the bias, and caution will be needed when reporting and interpreting estimates from the surveys.

We plan to compare respondents and non-respondents on information available from the sampling frame. We will also compare frame values with weighted values from sample respondents, with weights adjusted and unadjusted for non-response. The comparison between sample values using adjusted and unadjusted weights will allow us to (1) see the potential bias with non-respondents removed and no non-response weight adjustment and (2) assess the potential of the non-response bias adjustment to remove any bias or introduce bias. These comparisons will include demographic characteristics of the respondents and non-respondents, as well as membership status (start and stop dates) in HMO, if applicable; Medicare Part A; and Medicare Part B. Although using these variables in the non-response weight adjustment models will alleviate non-response bias, the risk of non-response bias is still increased if response rates differ between subpopulations defined by the different levels of these variables.

In addition, some of the important outcome variables are likely to be strongly correlated with practice-level characteristics. Although the number of beneficiaries and physicians sampled within individual practices will be small, making a practice-level comparison of response rates unrealistic, we will attempt to compare response rates across different types of practices (for example, medium-size practices vs. small practices). We will also compare frame values to sample values using sampling weights adjusted and unadjusted for non-response.

Finally, we will be able to match data from the sampling frame with data from Medicare claims. As indicated in the evaluation design report (Chapter III), data on quality measures (such as whether beneficiaries received appropriate medical tests) can be obtained using information from both the Medicare claims and from the beneficiary survey, and data on continuity of care is available from the Medicare claims, beneficiary survey, and physician survey. We will compare data from Medicare claims, which are available for respondents and non-respondents, with similar items in the beneficiary or physician surveys to determine if unusual response patterns emerge. We will also compare impact estimates for quality measures drawn from the Medicare claims data (for example, whether beneficiaries with diabetes had a dilated retinal exam) for the full sample of beneficiaries (including non-respondents) to impact estimates for the sample of

beneficiaries responding to the survey. The magnitude of the difference between the impact estimates based on the full sample from impact estimates based on the sample of respondents will allow us to assess the degree of non-response bias.

14. We appreciate that CMS is using questions based on previous demonstrations. However, we would like to know whether questions that ask about specific procedures or interventions within the past 12 months to 5 years have been validated, particularly as stand-alone questions (as opposed to being asked in the context of an event history calendar or some other memory aid). If not, we would like to understand CMS's plans for validating in the context of this study.

Most of the questions contained in the beneficiary survey have been used in previous studies as stand-alone items as described below:

- The Social HMO Demonstration, sponsored by CMS and conducted by Mathematica, was the source for questions on chronic health conditions (A2a-A2o); visits to physicians in the past 12 months (B11); and visits to emergency rooms or urgent care centers in past 12 months (B12). Over 200,000 interviews were completed for this study between 1997 and 2008.
- The Evaluation of Programs of Coordinated Care and Disease Management, sponsored by CMS and conducted by Mathematica, was the source for questions on pneumonia vaccination (C1d); lung examination (C1e); foot examination (C1f and C5); provision of materials (C1h); cutting down or quitting drinking (C2c); cutting salt in the diet (C2d); self weighing (C6); doctor's awareness of test results (D3); and satisfaction with various aspects of care (E1). More than 7,000 interviews were conducted for this study between 2003 and 2004.
- The Behavioral Risk Factor Surveillance System (BRFSS), sponsored by the Centers for Disease Control, was the source for questions on usual source of care (B6 and B7); physician's advice regarding increasing exercise (C2a); quitting smoking (C2b); and eating fewer high fat and high cholesterol foods (C2f). The BRFSS is the world's largest, on-going telephone health survey system, tracking health conditions and risk behaviors in the United States yearly since 1984.
- The Picker Ambulatory Care Patient Interview was the source for questions about what to expect in the future (C1i); what to do if symptoms worsened (C1j); and care coordination (D1 and D2). The Picker Ambulatory Care Survey, developed at Beth Israel Hospital (now Beth Israel Deaconess Medical Center), was a spinoff from the Picker Inpatient questionnaire. These questions came from a literature review and focus groups with patients. Many of these questions are now found, in altered form, in the CAHPS Group and Clinician Survey.

In addition to these sources, the MCMP beneficiary survey was pretested with nine Medicare beneficiaries. The questionnaire was completed by mail, and telephone debriefings were conducted following receipt. Pretest participants did not report any problems understanding the questions or providing answers to them during these sessions. Reviews of the completed questionnaires validated such understanding.

15. Please correct the race question on both physician surveys to comply with OMB standards (i.e., instruction should read “please mark 1 or more, delete “other” category, and add final required category—Native Hawaiian or other Pacific Islander).

The race question in the physician surveys has been changed as follows:

Which of the following categories best describes your race?

MARK ONE OR MORE

- 1 American Indian or Alaskan Native
- 2 Asian
- 3 Black or African-American
- 4 Native Hawaiian or Other Pacific Islander
- 5 White

16. Please improve the response to this FAQ, as it is not completely responsive:

Q: What happens if I do not participate in the survey?

A:Your participation is voluntary, but it is also important. Learning about your experiences as a Medicare participant will help CMS improve the services provided by the Medicare program

The response to this FAQ has been changed to read as follows:

Your decision to participate in the survey will not change any of your Medicare benefits or any other benefits you are currently receiving or may qualify for in the future. Your participation in this survey is voluntary, but very important. By participating you will help CMS to improve the services provided to Medicare beneficiaries.