# Evaluation of the Electronic Health Records Demonstration: Final Report

June 15, 2012

Suzanne Felt-Lisk
Lorenzo Moreno
Martha Kovac
Randall Blair
Stacy Dale
Amy Overcash
Rachel Shapiro

With the support of:

Maxwell Benjamin
Mark Brinkley
Barbara Carlson
Mark Flick
Mindy Hu
AnnaMaria McCutcheon
Brenda Natzke
Luis Rodriguez
Amber Tomas
Shinu Verghese
Daniel Weaver
Jelena Zurovac

This page is intentionally left blank.

Contract Number:
HHSM-500-2005-00025I (0006)

Mathematica Reference Number:
06479.481

Submitted to:
U.S. Department of Health and Human Services
Centers for Medicare & Medicaid Services
Room C3-16-07
Mail Stop C3-21-28
7500 Security Blvd.
Baltimore, MD 21244-1850
Project Officer: Curt Mueller

Submitted by:
Mathematica Policy Research
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005
Project Director: Lorenzo Moreno

# Evaluation of the Electronic Health Records Demonstration: Final Report

June 15, 2012

Suzanne Felt-Lisk
Lorenzo Moreno
Martha Kovac
Randall Blair
Stacy Dale
Amy Overcash
Rachel Shapiro

With the support of:

Maxwell Benjamin
Mark Brinkley
Barbara Carlson
Mark Flick
Mindy Hu
AnnaMaria McCutcheon
Brenda Natzke
Luis Rodriguez
Amber Tomas
Shinu Verghese
Daniel Weaver
Jelena Zurovac

This page is intentionally left blank.

## ACKNOWLEDGMENTS

Many individuals were critical to the development of this report. From Mathematica, Leslie Foster offered insightful comments on the Final Report and Randall Brown provided valuable input on our analysis at several stages of this evaluation. Allison Barrett, Grace Ferry, and Christopher Fleming diligently arranged, participated in, and documented the site visits which served as a key data source to this report. Jennifer Schore, the project director for the first two years, provided useful guidance on the design of several components of this evaluation. Cindy George and John Kennedy carefully edited the report, and Cindy McClure, Jill Miller, Jennifer Baskwell, and William Garrett skillfully produced it. From CMS, Lorraine Johnson (the original project officer for the evaluation); Jody Blatt (the project officer for the demonstration); and Curt Mueller (the current project officer for the evaluation) provided excellent suggestions and comments. We also thank Kerry Ketler and Laurie Pekala, both of the Actuarial Research Corporation, for patiently answering our many questions about data files. Finally, we thank the practices who contributed feedback during site visits, and the physicians and office managers who took the time to respond to our surveys.

This page is intentionally left blank.

# CONTENTS

This page is intentionally left blank.

# TABLES

This page is intentionally left blank.

# FIGURE

This page is intentionally left blank.

# ACRONYMS

| | |
|---|---|
| ARF | Area Resource File |
| ARRA | American Recovery and Reinvestment Act of 2009 |
| CAD | Coronary artery disease |
| CCHIT | Certification Commission for Health Information Technology |
| CHF | Congestive heart failure |
| CMS | Centers for Medicare & Medicaid Services |
| EHR | Electronic health record |
| EHRD | Electronic Health Records Demonstration |
| E&M | Evaluation and Management |
| FFS | Fee-for-service |
| HHS | U.S. Department of Health and Human Services |
| HITECH | Health Information Technology for Economic and Clinical Health Act |
| HRSA | Health Resources and Services Administration |
| IT | Information technology |
| MA | Medical assistant |
| MCMP | Medicare Care Management Performance |
| MUA | Medically underserved area |
| MUP | Medically Underserved Population |
| ONC | Office of the National Coordinator |
| OSS | Office Systems Survey |
| REC | Regional Extension Center |
| TIN | Tax identification number |

This page is intentionally left blank.

# EXECUTIVE SUMMARY

This report summarizes the evaluation findings for the Electronic Health Records Demonstration (EHRD), which was funded by the Centers for Medicare & Medicaid Services (CMS), authorized under Section 402 Medicare Waiver Authority, and implemented by CMS. The EHRD was a component of the Federal government's broad health information technology (health IT) strategy, beginning in 2005, to ensure that most Americans have access to secure, interoperable health records by 2014. The demonstration was designed to evaluate whether providing financial incentives increases physician practices' adoption and use of electronic health records (EHRs) and improves the quality of care delivered to chronically ill patients with fee-for-service (FFS) Medicare coverage. To attain these evaluation objectives, the demonstration adopted a stratified, experimental design, in which eligible practices were assigned to a treatment or a control group in equal proportions. The treatment group was eligible to receive the incentive payments, as described below.

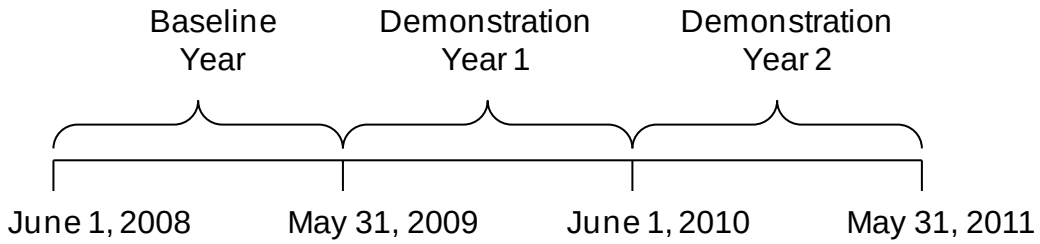
The EHRD expanded upon the Medicare Care Management Performance (MCMP) demonstration and built upon other CMS demonstrations. Unlike the MCMP demonstration, which merely encouraged the use of EHRs, the EHRD required practices to implement a certified EHR and use a core minimum set of functions by the end of year 2 of the demonstration, as measured by the Office Systems Survey (OSS). The main purpose of this report is to describe (1) implementation of the demonstration, and (2) impacts of the demonstration on key outcomes, including EHR adoption, use of specific EHR functions, quality of care, medical services utilization, and cost.

In brief, a large proportion of practices (43 percent) were unable or unwilling to meet program requirements to continue participating beyond year 2, leading CMS to terminate the demonstration early. The evaluation findings about impacts are necessarily incomplete, because of the shorter-than expected time frame. However, they indicate that the demonstration incentives had a strong positive impact on EHR use, despite the attrition. There is also some support for the linkage between EHR use and reduced hospitalizations for chronically ill Medicare beneficiaries, but during this time frame no relationship was found between EHR use or demonstration incentives and cost or other quality-of-care measures.

## A. Demonstration Overview

CMS initially planned to implement the demonstration in 12 sites in two phases one year apart. The agency chose four sites for Phase I: Louisiana, Maryland and the District of Columbia, southwest Pennsylvania, and South Dakota (and some counties in bordering states). Phase II was to have consisted of eight additional sites to start a year later. However, CMS canceled Phase II before it began as a result of the passage of the HITECH Act, which also provides physician financial incentives. Thus, the EHRD demonstration ultimately consisted solely of the four Phase I sites. Figure 1 illustrates the actual demonstration timeline.

Figure 1. Timeline for the Electronic Health Records Demonstration




The demonstration was expected to operate for five years (June 1, 2009–May 31, 2014) but was cancelled at the end of year 2 on August 1, 2011, by CMS. The main reason for the decision to cancel the demonstration early was that practice attrition was substantial at the end of the second year: about 43 percent of all eligible practices in the treatment group had left the demonstration either voluntarily, or more commonly, because they failed to meet program requirements.

The demonstration targeted practices serving at least 50 traditional FFS Medicare beneficiaries with certain chronic conditions for whom the practices provided primary care. Under the original design of the demonstration, primary care providers in practices with 20 or fewer providers (although there were exceptions) were eligible to earn incentive payments for (1) using the minimum functions of a certified EHR (a systems payment); (2) reporting 26 quality measures for congestive heart failure (CHF), coronary artery disease (CAD), diabetes, and preventive health services (a reporting payment); and (3) achieving specified standards on clinical performance measures during the demonstration period (a performance payment). All incentive payments under the demonstration were to be made in addition to normal FFS Medicare payments practices receive for submitted claims. Physicians could have received up to $13,000 and practices up to $65,000 over the first two years of the demonstration. Due to the termination of the demonstration, the reporting and performance payments were never made; CMS made only the systems payment for the first two years of the demonstration in fall 2010 and 2011.

## B. Evaluation Design, Outcome Measures, Data, and Methods

CMS selected Mathematica Policy Research to conduct an independent evaluation of the demonstration. Mathematica developed the evaluation design for the EHRD with the goal of supplying CMS with valid estimates of the incremental effect of offering performance-based financial incentives on a wide variety of outcome measures: quality of care, continuity of care, the use and costs of Medicare services for the chronically ill Medicare beneficiaries served by practices participating in the demonstration, practices' use of health IT, and physician and patient satisfaction. The evaluation also included an implementation analysis to study the implementation of the demonstration and the operational responses of the demonstration practices.

Many aspects of the original demonstration and evaluation implementation plan changed due to the cancellation of Phase II in April 2009 and the termination of the demonstration at the end of year 2. Changes to the original plans are summarized in Table 1.

Table 1. Original Plan Versus Actual Demonstration and Evaluation Timing and Components

| | Original Plan | Actual Timing and Components | Primary Explanation for Change |
|---|---|---|---|
| Demonstration Design | | | |
| Demonstration duration | 5 years (June 1, 2009–May 2014) | 2 years (June 1, 2009–August 1, 2011) | Substantial practice attrition at the end of the second year |
| Number of sites/phases in demonstration | 12 sites (4 sites in Phase I and 8 sites in Phase II ) | 4 sites in Phase I | Phase II was canceled as a result of the passage of the HITECH Act |
| Data Collection | | | |
| OSS administration | Administer survey to treatment group practices annually over the five-year demonstration; survey control group in years 2 and 5 | OSS administered to treatment group practices in years 1 and 2; administered to control group in year 2 | Demonstration was terminated in year 2 |
| OSS content | No questions about ARRA | In year 2 version, asked treatment and control group practices about effect of ARRA funding on their adoption and use of EHR | Federal incentives and resources available to practices for adopting and using EHRs were available as of year 2 of the demonstration |
| Beneficiary survey | Administer mail survey to beneficiaries from treatment and control group practices 36 months after the start of the demonstration (June 2012 for Phase I practices, June 2013 for Phase II) | Not administered | Demonstration was terminated in year 2 |
| Physician survey | Administer telephone survey to physicians from treatment and control group practices 36 months after the start of the demonstration (June 2012 for Phase I practices, June 2013 for Phase II) | Not administered | Demonstration was terminated in year 2 |
| Contacts with participating practices | Site visits in year 1 and follow-up site visits/telephone calls in year 5 | Site visits in year 1 | Demonstration was terminated in year 2 |
| Contacts with withdrawn practices | Telephone calls in years 3 and 5 | Telephone calls in years 1 and 2 | Notable early withdrawal of practices and early termination of the demonstration |
| Number of outcome measures | 26 quality-of-care measures | 5 quality-of-care measures | The demonstration ended prior to the scheduled reporting of quality measures for year 2 |
| Incentive Payment Structure | | | |
| | Practices could receive 3 types of payments:<br>(1) Years 1–5: a systems payment for EHR adoption and use<br>(2) Year 2: a payment for reporting clinical quality measures<br>(3) Years 3–5: payment for performance on clinical quality measures | Years 1–2: practices could receive a systems payment for EHR adoption and use | The demonstration ended prior to the scheduled reporting of quality measures for year 2 |

### Evaluation Design

The EHRD evaluation used a stratified, experimental design to assign 825 eligible physician practices that volunteered for Phase I of the EHRD to treatment and control groups. Practices from Louisiana, Maryland, Pennsylvania, and South Dakota were randomized into treatment and control practices within strata. The stratifying variables were site, practice size, and geographic location of the practice as measured by whether the practice was in a medically underserved area (MUA) or had a medically underserved population (MUP). Randomization was conducted in February 2009, before the demonstration began in June 2009.

### Outcome Measures

For this report, the evaluation examined effects of the EHRD demonstration's incentive payments on the following key outcome and process measures:

- Practices' adoption and use of health IT, as assessed in the site visits and OSS
- Five (of the 26) quality-of-care process measures calculated from claims and proxies for quality of care (such as preventable hospitalizations) available from claims
- Medicare expenditures and service use

### Implementation Analysis

The goal of the implementation analysis was to understand the process through which physician practices changed during the demonstration, and the role played in those changes by both demonstration incentives and factors external to the demonstration. The implementation analysis was based on site visits to 24 practices (16 treatment, 8 control) by a two-person team during May and June 2010. This analysis also used quantitative data to describe participation and the characteristics of participating practices, as well as qualitative data from telephone interviews with practices that did not meet demonstration requirements to continue participating beyond the second year.

### Changes in Outcomes for Medicare Costs, Service Use, and Quality of Care

Impacts for treatment-control differences in health IT use were estimated using regressions that adjusted for practice characteristics, such as practice size and state. All randomized practices (even those that were terminated) that responded to the 2011 OSS were included in the analysis. Observations were weighted to adjust for survey nonresponse and demonstration attrition. Furthermore, a difference-in-differences approach was used to estimate the impact of the demonstration on annualized costs, service use, and other key outcomes for beneficiaries. This approach compared the changes in outcomes over time for the treatment group to the changes in outcomes over time for the control group. The outcomes for beneficiaries assigned to treatment practices (treatment beneficiaries) were compared to the outcomes for beneficiaries assigned to control practices (control beneficiaries) both at baseline (CY 2008–2009) and at follow-up (24 months of the demonstration: June 2009–May 2011). The impact estimation model controlled for several factors often used to risk adjust, such as demographic characteristics and diagnoses.

Office Systems Survey

Measures of practices' health IT use, including their use of EHRs and specific EHR functions, were drawn from the OSS. The OSS also collected information on practice and provider characteristics. In the spring and summer of 2010 and 2011, the web-based OSS was administered on behalf of CMS to eligible practices; the survey was administered to treatment group practices in years 1 and 2, and to the control group practices in year 2. In order to calculate scores for practices' implementation of EHRs for the demonstration, the OSS measured specific EHR functions for practices using an EHR at the time of the survey. These functions fell into five domains: completeness of information, communication about care outside the practice, clinical decision support, increasing patient engagement/adherence, and medication safety.

## C. Summary of Key Findings and Conclusions

The evaluation results showed early impacts of the systems incentive payment on health IT adoption and use and on preventable hospitalizations. However, there were limited impacts across all four states of the incentives on the available quality-of-care process measures, Medicare expenditures, and key measures of Medicare service use (that is, the number of inpatient hospitalizations, emergency room visits, and physician office visits). Results from all components of the evaluation are summarized in Table 2 and are described in more detail in the rest of this section.

Participation. Practices were required to implement and use EHR systems before the end of year 2 to qualify for systems payment. Many practices complied with this requirement; however, 43 percent of the practices randomized to the treatment group had left the program either voluntarily, or more commonly, because they failed to meet program requirements. Thus, an important finding is that many practices were unwilling or unable to fulfill the demonstration requirements.

Site visits and interviews with withdrawn practices suggest two main reasons for the high attrition. First, implementing an EHR is a major, difficult undertaking. Second, many practices lacked some or all of the conditions needed to surmount the difficulties—project management skills; time, labor, and upfront financial resources; and a Medicare FFS caseload large enough to realize sizeable incentive payments. By contrast, practices that met demonstration requirements and continued to participate seemed to be those that had the wherewithal and intention to implement an EHR in the near future anyway, and the financial incentives of the EHRD motivated them to accelerate the process.

Adoption and Use of EHRs. Despite considerable attrition, the analysis of the 2011 OSS data found statistically and substantively significant impacts on several key health IT functions that were reported on both the 2008 demonstration application and the 2011 OSS. This analysis also found that the demonstration had a statistically and substantively significant impact on practices' OSS score as well as on all five OSS domain scores. These findings suggest that the systems payment to practices did incentivize the adoption and use of EHRs during the first two years of the demonstration, as treatment practices had much greater use of EHRs and of specific EHR functions than control practices did. It is notable that these systems payments, independent of the performance payments, resulted in immediate and sizeable impacts on process changes in the practices.

Generally, participating practices had much room for further improvement in EHR use and care management. The large gap between the practices' initial goals for maximizing EHR use and care management for chronically ill beneficiaries and what they reported during site visits validates the goals of the demonstration to incentivize improvements. However, based on the site visits, it also points to the likelihood that stronger financial incentives for performance (particularly related to care management), technical assistance, and a relatively long time frame are probably all needed to achieve goals. This is likely to be particularly true for the typical physician practice that is not as motivated as practices enrolled in demonstration.

Quality of Care. Because of the early termination of the demonstration, the majority of quality measures were not collected by the evaluation team. For the five measures that can be estimated from Medicare claims data, the analysis found no statistically significant impacts in quality measures when pooling claims data across all four demonstration sites. However, the demonstration did have a favorable effect on a common proxy for care quality, preventable hospitalizations. Specifically, the likelihood that beneficiaries with CAD, CHF, or diabetes would have a preventable hospitalization during year 2 fell by 0.5 percentage points for the treatment group relative to the baseline control group mean of 11.8 percent (statistically significant at the 5 percent level). This reduction in preventable hospitalizations was driven by those treatment practices that had the greatest improvement in health IT use. For example, nearly all improvements in preventable hospitalizations related to diabetes were concentrated among the 55 treatment practices that used nearly no health IT functions during baseline, but adopted nearly all diabetes-specific health IT functions by the 2011 OSS. This finding provides supporting evidence that the statistically significant drop in preventable hospitalizations can be attributed to the demonstration's incentives for health IT adoption.

Medicare Expenditures and Service Use. In terms of Medicare service use, the reduction in preventable hospitalizations for the subgroup with CHF, CAD or diabetes only translated to a small, statistically insignificant reduction in the overall hospitalization rate, since preventable hospitalizations are only a fraction of total hospitalizations and because this subgroup only represents half of the total sample. Also, there were no statistically significant treatment control-differences in the number of hospitalizations, emergency room visits, or physician office visits among all practices in the demonstration. Similarly, the demonstration did not appear to affect total Medicare expenditures across all of the sites.[1] Regardless of whether the systems incentive payments were included or excluded from the difference-in-differences estimates, the analysis found no statistically significant impacts of the demonstration on annualized Medicare expenditures during the demonstration's two years. Likewise, there were no statistically significant impacts of the demonstration on selected Part A and Part B expenditures, with one exception (outpatient expenditures increased by $69 during the second year of the demonstration).

[1]Total Medicare costs did not fall as a result of the reduction in preventable hospitalizations in Year 2, likely because any reduction in costs associated with preventable hospitalizations was small in proportion to the large variation in total Medicare costs.

Table 2. Summary of Findings Across All Four Demonstration Sites

| Outcome Domain | Results |
|---|---|
| Implementation Analysis | |
| Participation | Almost half (43 percent) of enrolled practices did not meet minimum EHR adoption and use requirements by the end of year 2 |
| Systems Payments | For practices that met the minimum requirements (57 percent of randomized practices), the average payment was $10,266 in year 1 and $11,064 in year 2 |
| Care Management Processes | No increase in care management for visited practices |
| Treatment-Comparison Analysis of Changes in Claims-Based Outcomes | |
| Use of EHRs | Statistically significant increase (10 to 18 percentage points) on use of EHRs and of specific EHR-functions, and on a summary score (11 percentage points) for calculating the systems payment based on EHR use |
| Quality of Care | No impact on quality measures, but a reduction of 0.5 percentage points in the likelihood that treatment group beneficiaries with CAD, CHF, or diabetes would have a preventable hospitalization during year 2 |
| Medicare Service Use | No impact on number of inpatient hospitalizations, emergency room visits, or physician office visits |
| Total Medicare Expenditures | No impact on annualized Medicare expenditures, with the exception of outpatient expenditures during year 2 |

## D. Limitations and Context to Consider

Although the EHRD evaluation relied on a stratified, experimental design—making it a rigorous study—there were several limitations. First, treatment group practices could have overstated their health IT use because the level of the incentive payment was determined by the level of health IT use they reported in the OSS. Second, because the demonstration ended early, the follow-up period might have been too short for practices to have implemented health IT changes or care management changes that would translate to quality-of-care improvements or to reductions in acute care use and costs. If the demonstration had run for the original five-year term, the lessons learned from the evaluation would have been more reliable—and maybe different—than those drawn from the current analysis. Finally, the exclusion, in error, of a small number of practices originally classified as eligible but later determined to be ineligible after randomization may have introduced selection bias to the OSS intention-to-treat impact estimates.

Aside from technical limitations, one must also consider that the demonstration was conducted during a time of rapid change in health IT policy and in the incentives and resources available for assistance. Efforts that overlapped with demonstration goals had the potential to support and encourage treatment group practices' adoption and use of EHRs, but also could have competed with demonstration activities. For example, in the second half of 2010, when CMS announced that the Medicare Incentive Program would start in 2011, providers could have taken a wait-and-see attitude in anticipation of the potentially larger rewards from the Program. After the Incentive Program started operations in January 2011, there was a four-month overlap between the demonstration and the Program, which could have resulted in demonstration practices changing their behavior despite the systems incentive payments. In fact, a sizeable minority of treatment group practices that responded to the OSS (44 percent) reported that they changed decisions or practice due to the Medicare and Medicaid EHR Incentive Programs by spring 2011. The OSS data showed a similar improvement in use of EHR features incentivized

by the demonstration and the Program alone and by both, which is consistent with both programs improving EHR use over the period.

## E. Lessons from EHRD Relevant to Other Policies and Programs

The following overarching lessons emerged from the EHRD:

1. Efforts with moderate incentive levels can influence use of EHRs, but cannot achieve universal adoption and use in a two-year time frame. Although more than half of practices responded to the financial incentives for implementing and using an EHR system, many practices were not able or willing to do so within the time frame required by the demonstration.
2. Targeting the incentives to individual practitioners instead of practices might be more effective. The site visits found considerable variation within practices in individual practitioners' use of EHRs; often decision making on EHR use occurred at the individual level. However, incentive payments for a practice were often not passed through to individual practitioners, but rather were used for overall support of the practice or its EHR system. Although in the Medicare EHR Incentive Program, eligible professionals who receive the incentive payment can assign it to the practice she belongs to, if any, it remains untested whether payment to the practice or to the individual might be more effective.
3. EHR implementation and use made real progress, but they did not have strong favorable effects on the quality-of-care process measures or on expenditures. The modest decrease in preventable hospitalizations could be attributed to incentives for EHR system use. However, it is unclear whether the systems payments and the performance payments combined would have favorably affected the quality-of-care process measures, service use, or Medicare expenditures (none of which were measurably affected by the systems payment alone during the demonstration's first two years because there was no time to assess whether these payment effected reporting or performance). Findings from the MCMP demonstration suggest that the performance payments had limited favorable impacts on quality of care in selected sites.

The findings and lessons from the evaluation of the EHRD demonstration could have implications for ongoing and future federal initiatives (such as the Medicare EHR Incentive Program, the Comprehensive Primary Care Initiative, the Accountable Care Organization [ACO] initiative, the harmonization of the Physician Quality Reporting System, and others), which use incentive payments to entice practices to meet certain requirements (for example, the "meaningful use" requirements, or the quality-of-care performance targets). One implication of the findings from the EHRD may be that future efforts should set more realistic expectations than CMS had for EHRD. That is, the findings indicate it is not realistic to expect all (or even almost all) targeted small to medium-sized practices to accomplish major changes with incentives that do not nearly cover the costs of purchase and operation of the incentivized health IT systems. Given the difficulties practices experienced as they adopted EHRs, and the low level of care management at present, such efforts should also consider a complementary technical assistance arm to increase the chance that quality as well as EHR use would be improved.

# I. OVERVIEW OF THE EHRD DEMONSTRATION

## A. Objectives

The Electronic Health Record Demonstration (EHRD) was a component of the Federal government's broad health information technology (health IT) strategy, beginning in 2005, to ensure that most Americans have access to secure, interoperable health records by 2014. Authorized under Section 402 of the Medicare Waiver Authority and funded and implemented by the Centers for Medicare & Medicaid Services (CMS), the demonstration was designed to evaluate whether providing financial incentives increases physician practices' adoption and use of electronic health records (EHRs) and improves the quality of care delivered to chronically ill patients with fee-for-service (FFS) Medicare coverage. To attain these evaluation objectives, the demonstration adopted a stratified, experimental design, in which eligible practices were assigned to a treatment or a control group in equal proportions. The treatment group was eligible to receive the incentive payments, as described below.

The EHRD demonstration expanded upon the Medicare Care Management Performance (MCMP) demonstration and built upon other CMS demonstrations. Unlike the MCMP demonstration, which merely encouraged the use of EHRs, the EHRD required practices to implement a certified EHR and use a core minimum set of functions by the end of year 2 of the demonstration, as measured by the Office Systems Survey (OSS).

This report summarizes the evaluation findings from the demonstration. Its main purpose is to describe (1) implementation of the demonstration, and (2) impacts of the demonstration during the first two years on key outcomes, including EHR adoption, use of specific EHR functions, quality of care, medical services utilization, and cost. Chapter I describes the context of the demonstration, including its structure and requirements, and its place in relation to other existing federal and private-sector programs. Chapter II provides an overview of the demonstration, and describes the methodology for analysis presented in the report. Chapter III focuses on practices' participation in the demonstration, factors related to EHR use, and lessons learned from visited practices. Chapter IV describes incentive payments and their relationship to practice characteristics; Chapter V presents the impacts of the demonstration on health IT use, quality of care, and Medicare expenditures and service use. Finally, Chapter VI synthesizes the findings and conclusions, presents the limitations of the evaluation, and describes lessons learned from EHRD relevant to other policies and programs.

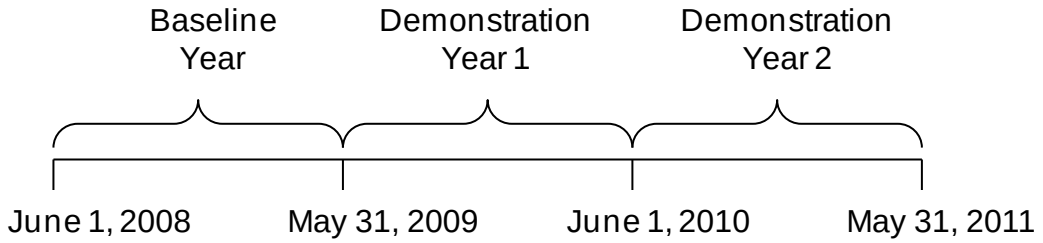
## B. Demonstration Design

### 1. Setting and Duration

CMS initially planned to implement the demonstration in 12 sites in two phases one year apart. The agency chose four sites for Phase I: Louisiana, Maryland and the District of Columbia, southwest Pennsylvania, and South Dakota (and some counties in bordering states). Phase II was to have consisted of eight additional sites to start a year later. However, CMS canceled Phase II before it began as a result of the passage of the Health Information Technology for Economic and Clinical Health (HITECH) Act, which also provides physician financial incentives. Thus, the EHRD demonstration ultimately consisted solely of the four Phase I sites. (See Table II.1 for a summary of all design changes.)

Selection of sites was based on a nationwide competitive process to identify community partners to assist CMS with education, outreach activities, and recruiting physician practices in each site. Recruitment of practices by the community partners began on September 2, 2008, and ended on November 26, 2008. Figure I.1 illustrates the demonstration timeline.

Figure I.1. Time Line for the Electronic Health Records Demonstration




Even after CMS decided to cancel Phase II of the demonstration, it still expected to operate the demonstration for five years (June 1, 2009 – May 2014), with just the Phase I sites. However, CMS later cancelled the demonstration altogether at the end of year 2 on August 1, 2011. The main reason for the early termination was the high rate of practice attrition prior to the 2011 OSS (about 16 percent of all eligible practices in the treatment group). As CMS argued, had the attrition rate continued through the end of the five-year demonstration, the final available sample size would have greatly limited the feasibility of conducting a rigorous and unbiased evaluation of the effectiveness of the financial incentives on EHR use and other outcomes. Table I.1 summarizes practices' participation in the demonstration by the end of the demonstration's second year of implementation.

Table I.1. Summary of Practice Participation in the Demonstration

| | Treatment | Control |
|---|---|---|
| Practices randomized at the start of the demonstration | 412 | 413 |
| Practices eligible for year 2 OSS[a] | 346 | 389 |
| Completed the year 2 OSS | 311[b] | 267 |
| Reported having an EHR | 264[b] | 188 |
| Met minimum requirements for payment | 232 | NA |

Source: Office Systems Survey, conducted in spring and summer 2010 and 2011.

[a] Excludes practices terminated, closed, merged, or withdrawn.

[b] Three practices that were asked to complete the validation survey but did not complete it or failed to provide the requested screenshots are considered to not have completed the OSS (see Appendix F).

EHR = electronic health record; OSS = Office Systems Survey.

## 2. Targeted Practices

The EHRD targeted practices providing primary care to at least 50 traditional FFS Medicare beneficiaries with certain chronic conditions. Under the original design of the demonstration,

primary care providers[2] in practices with 20 or fewer providers (although there were exceptions) were eligible to earn incentive payments for (1) using the minimum functions of a certified EHR; (2) reporting 26 quality measures for congestive heart failure (CHF), coronary artery disease (CAD), diabetes, and preventive health services; and (3) achieving specified standards on clinical performance measures during the demonstration period.

3. Incentive Structure

Treatment group practices were to have had the opportunity to receive three types of payments during the demonstration, a systems payment, a reporting payment, and a performance payment.[3] Because of the early termination of the demonstration, however, they received only systems payments.

Systems payment amounts were determined in large part by practices' self-reports of use of EHR functions, as measured by the OSS. To receive the basic systems payment, practices had to use a minimum set of EHR functions in an EHR system certified either by certification organizations approved by the Office of the National Coordinator for Health Information Technology (ONC) or under the old Certification Commission for Health Information Technology (CCHIT) standards.[4] Practices met the "minimum use requirements" if they used a certified EHR to record visit notes, diagnostic test orders and results, and prescriptions (see Chapter III for more details on this concept). Practices that did not adopt or make minimal use of a certified EHR system by the end of the first year of the demonstration did not receive a payment, but were permitted to remain in the demonstration. Practices that did not take either step by the end of the second year were removed from the demonstration; a total of 79 treatment practices were removed from the demonstration at the end of the second year—of these, 47 practices did not have a certified EHR, and 32 practices did not meet the minimum EHR use requirements. CMS made the system payments for the first two years of the demonstration in fall 2010 and 2011.

4. Beneficiary Assignment

For each demonstration year, beneficiary assignment was a retroactive process. Beneficiaries were assigned to a practice based on services (from Medicare claims data) in the

[2] The following types of providers were eligible to participate in the EHRD if they provide primary care: general practice, family medicine, internal medicine, geriatrics, and such medical subspecialists as cardiologists and endocrinologists and others who completed an internal medicine residency.

[3] Prior to termination of the demonstration, there were plans for practices to receive (1) a payment for reporting specific quality measures for year 2 with additional payment based on the OSS score; and (2) a payment for performance on the same quality measures for years 3 to 5, with additional payment each year based on the OSS score. The demonstration ended prior to the scheduled reporting of quality measures for year 2, and, therefore, this payment was not provided. Practices that met the minimum use requirements for year 2 received the systems payment.

[4] In 2010, new "meaningful use" certification guidelines issued by ONC went into effect that allowed several organizations, including CCHIT, to certify EHRs. CMS followed these guidelines in the two years of the demonstration. Under the demonstration, practices were required to use an EHR that was certified either under the new standards or under the old CCHIT standards.

most recently completed demonstration year. Specifically, each beneficiary with the target chronic conditions (CAD, CHF, and diabetes) or other specific chronic conditions was assigned to the practice providing the plurality of visits for evaluation and management (E&M) services. If two or more practices provided an equal number of visits, the beneficiary was assigned to the practice with the most recent E&M visit. Finally, as described in the demonstration design report (Wilkin et al. 2007), beneficiaries not meeting any of several criteria were excluded (not assigned to a practice). Beneficiaries were excluded if, for more than six months of the demonstration year, they relocated out of the demonstration state or site, received hospice coverage, enrolled in a Medicare coordinated-care plan, or received secondary payer coverage from Medicare due to working aged/disabled status. Beneficiaries who lacked either Part A or Part B coverage for all or part of the demonstration year or were deceased with more than six months remaining in the demonstration year were also excluded.

## C. Context

The demonstration was being conducted during a time of rapid change in health IT policy and the incentives and resources available for assistance. Many of the efforts that overlapped with demonstration goals and could either enhance or compete with demonstration activities were just gearing up, including those that were established under the HITECH act within the American Recovery and Reinvestment Act of 2009 (ARRA). For example, beginning in 2011, eligible professionals could begin receiving payments under either the Medicare or Medicaid EHR Incentive Programs for demonstrating meaningful EHR use, which included meeting required and optional criteria. The criteria overlapped with but also differed from the demonstration's EHR criteria. Further, the Health Information Technology Extension Program funded health IT regional extension centers to provide local technical assistance to support EHR adoption and meaningful use in primary care practices (as well as small rural and critical access hospitals).

Beyond HITECH initiatives, state and local projects had goals that were similar to or overlapping with those of the demonstration. These initiatives, discussed in detail in the implementation report (Felt-Lisk et al. 2011), seemed largely complementary to the demonstration and thus could have enhanced the effectiveness of incentives compared to an environment where supports were less available. However, they were in the early stages; with such a complicated environment, they also had the potential to compete for practices' attention.

# II. EVALUATION DESIGN, DATA, AND METHODS

This chapter describes the goals of the evaluation and the data and methods used in the implementation and impact analyses. Table II.1 summarizes how the original evaluation design was circumscribed by the cancellation of Phase II in April 2009 and the termination of the demonstration at the end of year 2.

CMS selected Mathematica Policy Research as the independent evaluator of the demonstration. Mathematica developed the evaluation design for the EHRD with the goal of supplying CMS with valid estimates of the incremental effect of offering performance-based financial incentives on a wide variety of outcome measures: quality of care, continuity of care, the use and costs of Medicare services for the chronically ill Medicare beneficiaries served by practices participating in the demonstration, practices' use of health IT, and physician and patient satisfaction. The evaluation also included an implementation analysis to study the implementation of the demonstration and the operational responses of the demonstration practices.

Table II.1. Original Plan Versus Actual Demonstration and Evaluation Timing and Components

| | Original Plan | Actual Timing and Components | Primary Explanation for Change |
|---|---|---|---|
| Demonstration Design | | | |
| Demonstration duration | 5 years (June 1, 2009 – May 2014) | 2 years (June 1, 2009 – August 1, 2011) | Substantial practice attrition at the end of the second year |
| Number of sites/phases in demonstration | 12 sites (4 sites in Phase I and 8 sites in Phase II ) | 4 sites in Phase I | Phase II was canceled as a result of the passage of the HITECH Act |
| Data Collection | | | |
| OSS administration | Administer survey to treatment group practices annually over the five-year demonstration; survey control group in years 2 and 5 | OSS administered to treatment group practices in years 1 and 2; administered to control group in year 2 | Demonstration was terminated in year 2 |
| OSS content | No questions about ARRA | In year 2 version, asked treatment and control group practices questions about effect of ARRA funding on their adoption and use of EHRs | Federal incentives and resources available to practices for adopting and using EHRs were available as of year 2 of the demonstration |
| Beneficiary survey | Administer mail survey to beneficiaries from treatment and control group practices 36 months after the start of the demonstration (June 2012 for Phase I practices, June 2013 for Phase II) | Not administered | Demonstration was terminated in year 2 |
| Physician survey | Administer telephone survey to physicians from treatment and control group practices 36 months after the start of the demonstration (June 2012 for Phase I practices, June 2013 for Phase II) | Not administered | Demonstration was terminated in year 2 |

| | Original Plan | Actual Timing and Components | Primary Explanation for Change |
|---|---|---|---|
| Contacts with participating practices | Site visits in year 1 and follow-up site visits/telephone calls in year 5 | Site visits in year 1 | Demonstration was terminated in year 2 |
| Contacts with withdrawn practices | Telephone calls in years 3 and 5 | Telephone calls in years 1 and 2 | Notable early withdrawal of practices and early termination of the demonstration |
| Number of outcome measures | 26 quality-of-care measures | 5 quality-of-care measures | Demonstration was terminated before all measures became available. Two relevant measures could not be constructed reliably. |
| Incentive Payment Structure | Practices could receive 3 types of payments:<br>(1) Years 1 – 5: a systems payment for EHR adoption and use<br>(2) Year 2: a payment for reporting clinical quality measures<br>(3) Years 3 – 5: payment for performance on clinical quality measures | Years 1 – 2: practices could receive a systems payment for EHR adoption and use.[a] | The demonstration ended prior to the scheduled reporting of quality measures for year 2 |

[a]At the end of the year 2 OSS fielding period (June 2011), practices continued to believe they could qualify for the reporting payment for that year. As noted, CMS notified the practices that they were not going to receive the payment due to the demonstration's cancellation until August 2011.

ARRA = American Recovery and Reinvestment Act of 2009; EHR = electronic health record; HITECH = Health Information Technology for Economic and Clinical Health; OSS = Office Systems Survey.

## A. DESIGN OF THE IMPLEMENTATION ANALYSIS

The goal of the implementation analysis was to understand the process through which physician practices changed during the demonstration, and the role played in those changes by both demonstration incentives and factors external to the demonstration. Implementation lessons are critical to understanding how, why, and where the demonstration incentives did (or did not) work. The implementation analysis addressed four broad research questions: (1) What changes did practices make under the demonstration? (2) How did the demonstration influence those changes? (3) How did environmental factors influence those changes? and (4) What actions did practices take with the specific goal of improving care quality and patient health? A detailed description of the data collection methods for site visits with treatment and control practices, the OSS, and telephone discussions with select practices, appears below, along with a brief description of administrative data sources used in the implementation analysis. More details of these interviews are in the evaluation's implementation report (Felt-Lisk et al. 2011).

### 1. Site Visits

For each of the four Phase 1 sites, four demonstration practices and two control practices were selected for case study. Practices were visited by a two-person team during May and June 2010. To select practices for site visits, two steps were followed. First, practices that could be visited within the allotted time frame (one week) based on location were identified. Then, after

reviewing data from the practice applications database (provided by the CMS implementation support contractor for the demonstration) a purposive sample was selected that would provide a mix in terms of urban/rural location, use of health IT, number of providers, number of FFS Medicare beneficiaries, and number of beneficiaries with each chronic condition.

Of the 24 originally selected practices (16 treatment and 8 control), 17 consented to visits and 7 others were selected and recruited to replace those that did not consent to a visit. Individual discussions were usually held with at least two people per practice—a physician and an administrative staff member knowledgeable about the demonstration. When possible, nurses and medical assistants were also interviewed, as well as the medical director and such administrative personnel as the chief information officer and chief financial officer, if applicable. A semi-structured protocol was used during the discussions, which lasted one to two hours per practice.

2. The Office Systems Survey

In the spring and summer of 2010 and 2011, the web-based OSS was administered on behalf of CMS to eligible practices.

The OSS collected information on practice characteristics, provider characteristics, and use of EHRs and other health IT (see Appendix A for OSS design details, including the questionnaires). In order to calculate scores for practices' implementation of EHRs for the demonstration, the OSS measured specific EHR functions for practices currently using an EHR. These functions included prescribing medications, ordering laboratory tests and other procedures, and care management and coordination. All practices that had been randomized to the treatment or control group were asked to participate, with the exception of those that CMS determined had failed to meet terms and conditions of the demonstration.

3. Telephone Contacts to Withdrawn Practices

The team made the first round of calls to withdrawn practices at the end of the demonstration's first year and the second round of calls shortly after the demonstration's termination announcement.

In June and July 2010, 10 practices that had voluntarily withdrawn from the demonstration were contacted for a brief telephone discussion about the reason(s) for their withdrawal. The practices identified for contact were all those that had voluntarily withdrawn and had no other information noted in the database that recorded their withdrawal to indicate a reason. The discussions allowed the team to learn the reason for withdrawal from seven of the practices (the other three did not respond to inquiries). This information was combined with information noted in the demonstration tracking database (explained below), in order to give the broadest possible picture of reasons for withdrawal.

In fall 2011, telephone interviews were conducted with practice representatives about their reasons for withdrawing from the EHRD demonstration. The evaluation team aimed to conduct interviews with a total of 24 practices consisting of 8 practices in each of three groups: (1) those that did not return the OSS, (2) those that did return the OSS but did not implement an EHR by the demonstration deadline, and (3) those that had an EHR and submitted the OSS but did not meet the minimum requirements for continued participation. For each of these three groups, two

practices from each of the four demonstration sites (Louisiana, Maryland, Pennsylvania, and South Dakota) were selected for contact.

In total, full interviews were completed for 20 practices; combined interviews were conducted with corporate-level staff (one person covering three practices and another person covering two practices) for 5 of those practices. Short email responses concerning reasons for withdrawal were collected from another three practices.

4. Administrative Data Sources

The following administrative data sources were used to select practices for site visits, explore practice characteristics, and analyze OSS data for patterns by practice characteristics:

- Demonstration tracking database. The tracking database maintained by the implementation support contractor and provided to Mathematica was used to assess the number of practices participating in the demonstration, and as one source to identify reasons for voluntary withdrawals.
- Demonstration application data. The size of practices' Medicare FFS population was self-reported by practices as part of their application to the demonstration.
- Area Resource File (ARF). The ARF was used to identify urban/rural location based on metropolitan versus non-metropolitan counties (a crude but commonly used method of designating an area as rural).
- HRSA data. Mathematica identified medically underserved area (MUA)/medically underserved population (MUP) status using data from the Health Resources and Services Administration (HRSA). Specifically, each practice's primary location was geocoded and merged with HRSA data by census tract. Addresses for which tracts were not available, and those for which HRSA reported only metropolitan area information, were manually entered into the HRSA website to determine their MUA/P status. Because MUA and MUP are not differentiated in the manual web-based data tool, it was impossible to determine for all practices the MUA status alone. Therefore, it was necessary to use the combined MUA/MUP measure identifying practices that are either in an MUA or an MUP. "MUA" throughout the report means MUA or MUP.
- Medicare claims data. Medicare claims data for the relevant periods were used to calculate the number of assigned Medicare FFS beneficiaries per practice; assignment was determined by the demonstration's implementation contractor.

5. Limitations

Because the evaluation team visited a small proportion of treatment and control group practices, which were not randomly selected, it cannot be assumed that the visited practices represent all demonstration and control group practices. A comparison of visited practices with others on available data showed that they were largely similar; a few exceptions were organizational affiliation (visited treatment group practices were more likely than other treatment group practices to report affiliation with a larger organization) and the number of Medicare beneficiaries served (visited control practices served more of these beneficiaries, compared with other control practices). Further, national generalizations cannot be made: recruitment

experience suggests that the demonstration practices were probably more advanced in their thinking about and use of EHRs than other small practices nationally. In addition, it was noted that the OSS data were self-reported data on EHR use, with simple attestation by respondents of accuracy rather than independent verification.

## B. ANALYSIS OF IMPACTS ON OUTCOMES

This section provides an overview of the evaluation design, expected effects, methods, outcome measures and data sources, and limitations of the evaluation of the EHRD. Further details on the sample, data, outcome measures, and methods are found in Appendix G.

### 1. Design

The EHRD evaluation used a stratified, experimental design to allocate 825 eligible physician practices that volunteered for Phase I of the EHRD into treatment and control groups. If well implemented, this design would ensure that changes in the demonstration outcomes could be attributed to the financial incentive payments and that the impact estimates were unbiased. A stratified design was used to achieve balance on practice characteristics that are important predictors of outcomes, which enhanced the credibility of the design.

Practices from Louisiana, Maryland, Pennsylvania, and South Dakota were randomized into treatment and control practices within strata. The stratifying variables were site, practice size, and geographic location of the practice as measured by whether the practice was in an MUA or had an MUP. Randomization was conducted in February 2009, before the demonstration began in June 2009. According to the proposed plan, the practices were successfully randomized in a way that would minimize chance differences between treatment and control groups on key outcomes during the evaluation. Appendix H includes the randomization plan and the summary of the randomization of practices.

### 2. Expected Effects

Under the demonstration's logic model, available in the evaluation design report (Felt-Lisk et al. 2008), the EHRD intervention was intended to promote high quality of care through practices' use of EHRs and evidence-based guidelines. The financial payments provided to practices were expected to influence the adoption of EHRs and increase the use of specific EHR functions, improvements in practice workflows, and improvement in the continuity of care and care management. These structural and organizational changes were then expected to result in improved quality of care, which could in turn improve practices' financial performance, improve beneficiary health, and reduce Medicare expenditures. It was anticipated that these changes in practice and beneficiary outcomes could then lead to improved physician and beneficiary satisfaction with care.

### 3. Methods

#### a. Practice-Level Impacts

Impacts for treatment-control differences in health IT use were estimated using regressions that adjusted for state, whether the practice was in an MUA, practice size, and the practices' health IT-related use according to the information available on the practices' application for the

demonstration (see Appendix H for the application). All randomized practices (even those that were terminated or withdrew) that responded to the 2011 OSS were included in the analysis. Observations were weighted to adjust for survey nonresponse and demonstration attrition (see Appendix A for details on weights).

b. Beneficiary-Level Impacts

A difference-in-differences approach was used to estimate the impact of the demonstration on annualized costs, service use, and other key outcomes for beneficiaries. With this approach, the outcomes for beneficiaries assigned to treatment practices (treatment beneficiaries) were compared to the outcomes for beneficiaries assigned to control practices (control beneficiaries) both at baseline (CY 2008–2009) and at follow-up (24 months of the demonstration: June 2009–May 2011). Under this approach, the following factors were controlled: demographic characteristics (age, sex, and race) and diagnoses (CAD, CHF, diabetes, and other chronic conditions); practice fixed effects (an indicator variable for each practice that accounts for all practice characteristics that are unchanged over time); and an indicator variable for the follow-up period. Note that it would be inappropriate to control for hierarchical condition category (HCC) scores (often used to risk-adjust) during the follow-up period, since the calculation of these scores takes into consideration certain acute conditions (such as heart attacks and acute diabetes) that the demonstration sought to prevent. However, the model did control for several components of the HCC scores (such as indicators for whether the beneficiary had a targeted condition, as well as the beneficiary's demographic characteristics) that would presumably be unaffected by the demonstration.

4. Outcome Measures and Data Sources

The impact analysis measured the effect of the EHRD demonstration on the following key outcome and process measures:

- Practices' adoption and use of health IT from the OSS
- Five (of the 26) quality-of-care process measures calculated from claims and proxies for quality of care (preventable hospitalizations) available for claims (see Table I.4 for a list of these measures)[5]
- Medicare expenditures and service use

To obtain these measures, the evaluation team relied on the two available data sources: (1) Medicare claims data, and (2) the OSS. Under the original design, a survey of Medicare beneficiaries and physicians would have provided additional data on the quality-of-care measures, beneficiary and physician satisfaction with care, and care coordination. However, these surveys were not administered because the demonstration was terminated.

[5] Only five of the 26 quality-of-care measures were used because the demonstration was terminated before the other measures became available, and two relevant measures could not be constructed reliably due to data availability. See Chapter V Section B for details on the quality-of-care measures.

## 5. Limitations

Although the EHRD evaluation relied on a stratified, experimental design—making it a rigorous study—there are limitations due to (1) key outcome measures construction, (2) termination of the demonstration in year 2, and (3) practice exclusion, attrition, and nonresponse. Chapter VI provides a detailed analysis of these and other limitations.

This page is intentionally left blank.

# III. OVERVIEW OF IMPLEMENTATION AND FINDINGS

This chapter tells the story of the implementation of the demonstration, including the types of practices that participated, the extent to which they were able to meet progressively more demanding demonstration requirements, the evolution of health IT use among participating practices, and care management and quality measurement in the practices. The story draws on several of the data sources described in Chapter II: site visits to four treatment and two control group practices in each of the four sites, telephone contacts with practices that were terminated in summer 2011, OSSs conducted in 2010 and 2011, and data from practice applications to the demonstration in 2008. Most of the chapter focuses on the treatment group; however, it concludes with a summary of findings from the site visits to selected control group practices.

## A. PARTICIPATION

### 1. Participation Requirements

Participation requirements for the demonstration were structured to allow two years for practices enrolled at the start to adopt a certified EHR and begin using a set of EHR functions considered to represent basic use. In year 1, the only participation requirement was to complete the OSS. In year 2, to continue participating beyond the end of the year, practices were required to complete the OSS and conduct the following activities, as attested to on the survey:

- Use an EHR that was certified for at least some of the period beginning June 2009 (demonstration start)
- For at least some patients in the practice, record in the EHR: visit notes, diagnostic test orders (laboratory and imaging orders), diagnostic test results (laboratory and imaging results), and prescriptions

### 2. Recruitment Experience[6]

The community partner in each site served as the recruitment arm for the demonstration and represented diverse collaborations of organizations (such as the South Dakota ehealth Collaborative and the Louisiana Health Care Quality Forum) comprising local stakeholders (such as medical societies or other physician professional organizations), large payers, public entities (such as the Maryland Health Care Commission and the South Dakota Department of Health), and other health care organizations interested in advancing health IT use.

To recruit practices, community partners leveraged hospitals and health systems, and worked with other membership organizations, such as the medical society or the rural health association. The health systems and hospitals were very helpful, according to the community partner staff, and were often very successful in bringing their affiliated practices on board. It may be that health systems and hospitals regarded participation in the demonstration as potentially cost-effective, since they could leverage central health IT support staff and have a

[6] This section is based on telephone contacts with the community partners in spring 2010. Due to staff turnover at the community partner, recruitment details were not available for one site.

larger resource base to cover any up-front costs to make changes to better meet demonstration measures; and at the same time the potential total financial reward to the system would be larger, encompassing several participating practices. The independent practices were much less likely to respond to outreach and more difficult to convince to participate in the demonstration. It appears that the demonstration may have been disproportionately successful in recruiting practices affiliated with hospitals and health systems.

In addition, the demonstration may have been successful primarily in recruiting practices that were further along than many others in their knowledge about and acceptance of EHRs. The community partners reported that the practices that tended to apply were those that already had an EHR, had made the decision to adopt one, or were close to making that decision. However, one community partner representative commented she had believed practices were closer to embracing EHRs than they actually were: "A lot of practices weren't even thinking about it, it wasn't on their radar."

3. Characteristics of Treatment Group Practices

The treatment group practices in each site varied widely in terms of their environment, practice affiliation, and participation in other quality improvement, EHR, and pay-for-performance programs. Most were small to medium-sized practices, and the vast majority had fewer than 10 practitioners (Table III.1). As noted in Chapter II, because the evaluation used an "intent-to-treat" approach, the figures below include data from both participating and nonparticipating treatment group practices that completed the year 1 OSS. (However, only 4 percent of the practices in the data were nonparticipating at the end of the first year of the program.[7])

- Louisiana. Nearly half of Louisiana's treatment group practices were located in an HRSA-designated MUA, far more than in the other sites. Unlike other sites, fully 70 percent of Louisiana's treatment group practices were not participating in any other quality improvement, EHR, or pay-for-performance program at the time of the OSS. In other respects, Louisiana's practices were similar to the average across all sites: 20 percent were rural, about two-thirds were unaffiliated with a larger organization, and most had more than 200 chronically ill FFS Medicare beneficiaries in the practice.
- Maryland and District of Columbia. The Maryland and District of Columbia site's largely urban treatment practices were less often in an MUA than practices in Louisiana or Pennsylvania, but otherwise generally reflected the average characteristics across sites. Nearly three-fourths were unaffiliated with any larger organization, and more than 40 percent were not participating in any other related program. Only 5 of this site's 114 participating treatment practices were located in the District of Columbia.

[7] Fourteen of the 352 practices in Table III.1 were nonparticipating practices. Nonparticipating treatment group practices included in the analysis were those that were randomized to the treatment group but had since voluntarily withdrawn from the demonstration. Treatment group practices that closed or merged with another treatment group practice prior to completing the OSS were not included in the analysis. Participating practices were treatment group practices that had not closed, merged, or withdrawn from the demonstration.

Table III.1. Characteristics of Treatment Group Practices That Completed the 2010 OSS (Percentages)

| Practice Characteristics | All Sites (n=352) | Louisiana (n=79) | Maryland (n=111) | Pennsylvania (n=124) | South Dakota (n=38) |
|---|---|---|---|---|---|
| **Practice Size (total number of providers)[a]** | | | | | |
| 1-2 | 44 | 47 | 42 | 46 | 34 |
| 3-5 | 35 | 33 | 35 | 41 | 18 |
| 6+ | 21 | 19 | 22 | 13 | 47 |
| Percentage in an MUA[b] | 28 | 47 | 17 | 30 | 16 |
| Percentage in a rural area[b] | 16 | 20 | 5 | 6 | 71 |
| **Number of Assigned Medicare FFS Beneficiaries[c]** | | | | | |
| 0-199 | 45 | 39 | 33 | 62 | 34 |
| 200-999 | 47 | 56 | 54 | 35 | 47 |
| 1,000 or more | 8 | 5 | 13 | 3 | 18 |
| **Practice Affiliation[a]** | | | | | |
| Unaffiliated | 57 | 68 | 73 | 31 | 71 |
| Owned by a larger medical group | 5 | 0 | 8 | 5 | 3 |
| Owned by a hospital, hospital system, or integrated delivery system | 32 | 19 | 17 | 57 | 18 |
| Affiliated with an IPA or PHO and not owned by a larger entity | 5 | 9 | 0 | 7 | 5 |
| Other affiliation | 2 | 4 | 2 | 1 | 3 |
| **Participation in Other Quality Improvement, EHR, and Pay-for-Performance Programs[a]** | | | | | |
| No participation | 34 | 70 | 43 | 11 | 18 |
| Private-sector quality improvement program(s) including pay-for-performance | 39 | 3 | 17 | 73 | 58 |
| PQRS | 50 | 22 | 48 | 63 | 66 |
| Other | 17 | 9 | 19 | 21 | 18 |

[a]Source: Office Systems Survey, conducted in spring and summer 2010. Percentages represent practices that responded to the relevant question.

[b]Source: Randomization information (constructed by linking geocoded addresses to data from HRSA's website).

[c]Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment practices at the end of year 1 (June 1, 2009, through May 31, 2010).

IPA = independent practice association; PHO = physician hospital organization.

- Southwestern Pennsylvania. Like Maryland, Pennsylvania's treatment group practices were largely urban, with 30 percent located in MUAs. These practices tended to have relatively lower numbers of Medicare FFS beneficiaries (28 percent had fewer than 200 in total). Smaller practices would be expected to have lower numbers of Medicare beneficiaries and, of the four sites, Pennsylvania had the lowest percentage of practices with six or more physicians. The number of Medicare beneficiaries may be influenced by practice specialty; for example, because they serve children as well as adults, family practitioners may serve fewer Medicare beneficiaries on average than do internists. The number of assigned Medicare beneficiaries may also be influenced by Medicare Advantage participation, since only fee-for-service Medicare beneficiaries were assigned under the demonstration. Pennsylvania practices were far more likely to be owned by a hospital, hospital system, or integrated delivery system than practices in other states (57 percent), and they were more likely to be involved in other related programs; almost three-fourths were participating in private-sector quality improvement or pay-for-performance programs, and nearly two-thirds were in Medicare's Physician Quality Reporting System (PQRS).[8]
- South Dakota. South Dakota had a much larger portion of practices with six or more physicians than other sites (47 percent), and along with that came relatively high Medicare caseloads: more than half the practices reported 1,000 or more Medicare FFS beneficiaries. South Dakota also had by far the most rural environment for the demonstration: 71 percent of its treatment practices were located in rural areas. South Dakota practices were relatively involved with other programs, with 58 percent participating in private-sector quality improvement or pay-for-performance programs and two-thirds participating in PQRS.

4. Attrition Among Originally Randomized Practices

The demonstration began with 412 practices that were randomized into the treatment group. However, by the end of the demonstration's second year, 43 percent of the originally randomized practices were lost to the demonstration (177 practices). Table III.2 shows:

- 66 practices were terminated, closed, merged, or withdrew prior to spring 2011 (row 2)
- 32 did not complete the 2011 OSS, which was required for continued participation (row 7)
- 47 completed the survey but reported they had no EHR (row 6)
- 32 completed the survey and had an EHR but did not meet minimum criteria for EHR use (row 4)

[8] In theory, because PQRS incentivizes reporting of quality data, PQRS practices could be more knowledgeable about their quality measure rates and become interested in using their EHRs to improve such rates.

Although hospitals in the last three categories did not officially withdrew from the program, they did not meet the criteria to continue to participate in the program beyond year 2, so they were terminated by CMS prior to CMS's announcement of the termination of the demonstration.

Table III.2. Summary of EHRD Treatment Practices' Attrition

| Row # | | All Sites | Louisiana | Maryland | Pennsylvania | South Dakota |
|---|---|---|---|---|---|---|
| 1 | Originally randomized to treatment group | 412 | 104 | 127 | 138 | 43 |
| 2 | Terminated, closed, merged, or withdrawn prior to 2011 OSS | 66 | 21 | 18 | 19 | 8 |
| 3 | Eligible for 2011 OSS (not terminated, closed, merged, or withdrawn) | 346 | 83 | 109 | 119 | 35 |
| 4 | Did not complete the OSS | 32 | 14 | 11 | 5 | 2 |
| 5 | Completed the OSS:[a] | 314 | 69 | 98 | 114 | 33 |
| | Of these: | | | | | |
| 6 | Had no EHRr | 47 | 19 | 9 | 15 | 4 |
| 7 | Had EHR but did not meet minimum requirements | 32 | 7 | 7 | 15 | 3 |
| 8 | Total number of 2011 OSS eligibles that did not meet requirements (rows 4+6+7) | 111 | 40 | 27 | 35 | 9 |
| 9 | Total number of 2011 OSS eligibles that met all requirements to participate beyond year 2 (row 3 – row 8) | 235 | 43 | 82 | 84 | 26 |
| 10 | Percentage of 2011 OSS eligibles that did not meet all requirements to continue participation (row 8/row 3) | 32.1 | 48.2 | 24.8 | 29.4 | 25.7 |
| 11 | Percentage of 2011 OSS eligibles that met all requirements to continue participation (100-row 10) | 67.9 | 51.8 | 75.2 | 70.6 | 74.3 |
| 12 | Percentage of originally randomized practices lost from active participation ([rows 8 + 2]/row 1) | 43.0 | 58.7 | 35.4 | 39.1 | 39.5 |
| 13 | Percentage of originally randomized practices that remained active (100-row 12) | 57.0 | 41.3 | 64.6 | 60.9 | 60.5 |

[a]Three practices that were asked to complete the validation survey but did not complete it or failed to provide the requested screenshots are considered to not have completed the OSS (see Appendix F).

In fall 2011, the Mathematica team contacted 24 practices evenly split among the following three categories of practices lost to the demonstration: (1) practices that failed to return the OSS, (2) practices that failed to adopt an EHR, and (3) practices that failed to meet minimum EHR use requirements. Based on interviews with 20 of these practices and emails from 3 more, there were several reasons for their failure to participate.[9]

Practices that did not return the OSS. Eight contacted practices reported the following reasons for failing to return the OSS:

- Lack of awareness or coordination of the survey (four practices). One office manager did not seem to be aware of the OSS, two other practices reported that it was just overlooked before the deadline, and a fourth practice had staff turnover among those responsible for the OSS and demonstration near the deadline for the OSS, and the OSS "fell through the cracks."
- Knowledge they would not meet minimum requirements (three practices). Two of these practices did not implement an EHR, and the third was implementing an EHR but thought that it would not meet the minimum requirements for use of EHR functions. This latter practice was aware of the requirements at the beginning of the demonstration. The obstacles these practices faced included the cost of the EHR (one practice), delays in training and implementation due to lack of planning on their part, as well as a lack of resources to hire temporary clinic staff while the permanent staff trained to use the EHR.
- Organization structure change (one practice). One practice was an independent practice that withdrew from the demonstration at the time of the OSS because it was acquired by a larger organization.

Practices that did not implement an EHR. The eight contacted practices that did not implement an EHR by June 2011 as required by the demonstration reported they enrolled in the demonstration having plans to adopt an EHR by June 2011, but encountered challenges that prevented adoption by that date. Most (six) had selected an EHR but had not yet fully implemented it. One reported fully implementing an EHR and was confused about why the evaluation's information said otherwise.[10] The respondent from the one practice that had not yet selected an EHR commented at length about the barriers to EHR selection and implementation, although it was not clear why his view had changed from the time of initial enrollment. For example:

[9] In addition to these contacts in 2011, in June and July 2010, 10 practices that had voluntarily withdrawn from the demonstration were contacted for a brief telephone discussion about the reason(s) for their withdrawal. The reasons cited by two or more practices included (1) the physician members of the practice did not want to implement an EHR, are advanced in age, and/or will probably retire if/when EHRs are mandatory (two practices); (2) physician "did not agree with" the demonstration, or "had a change of heart" about participating in the demonstration (further details not available) (two practices); (3) physicians at the enrolled practice left the organization, or no primary care physician currently practices at the enrolled practice (two practices); and (4) practice changed ownership (two practices) (Felt-Lisk et al. 2011).

[10] The person from the practice who responded to the OSS had responded erroneously to the question about EHR implementation, indicating the practice had no EHR.

> "I cannot invest a lot in EHRs, as I am a solo practice. The programs [EHRs] are expensive and I cannot decrease my patient load to implement it or I would go bankrupt in months. I cannot invest the money and then find out that it is a bad investment. I am going to let the dust settle first – let others do it first and see how others do it [implement EHRs]. Also, I have heard that a lot of changes are coming in software."

Other practices mentioned the following barriers: the time burden and detailed work to implement the EHR (one practice); the cost of software/hardware (one); delays while waiting for an outside company to make a decision that impacted their EHR selection (one); delay in setting up equipment (one); technical issues with the software (two); challenges recruiting IT and operations staff to set up the software and train staff (two); and being in a network that is implementing the EHR with all of its physicians and being unable to meet the deadline (one).

Practices that had an EHR but did not meet minimum requirements. Most contacted practices in this category were aware of the minimum requirements for EHR use at the beginning of the demonstration and expected to meet the requirements (six of seven). The barriers encountered by these practices included:

- Having a two-part EHR system in which the older part of the system is not certified and the newer part is certified. The practice did not use all the minimally required functions within the part that was certified.
- Lack of in-house IT personnel resources and financial resources to create an interface with the system's laboratory to accomplish laboratory orders electronically, which itself was described as not having robust IT.[11]
- A workflow issue with the practice's system resulting in physicians not embracing the EHR (as the hardware was difficult for them to use and the software had too many "clicks"); this practice also bluntly stated that the physicians in the practice just did not do the work required to be successful in the demonstration.
- Inability to meet the demonstration deadline, although specific plans for rolling out use of EHR functions were proceeding.

5. Demonstration Termination

Due to the rate of attrition described above, and the expectation that some attrition would continue over the remaining three years of the demonstration, CMS expected there would not be enough remaining participants to provide an acceptably strong test of whether the financial incentive worked to improve EHR use, quality of care, and costs of care. Such a test was the purpose of the demonstration, as discussed in Chapter I. Therefore, on August 1, 2011, CMS notified the treatment and control group practices that the demonstration would be discontinued effective August 1, 2011. Practices that met the criteria to receive an incentive for year 2

[11] Although the requirement is for recording laboratory orders in the EHR, not placing them electronically, as a practical matter this practice appeared to believe that recording the orders would only be feasible if they were electronically placed.

received that incentive payment in fall 2011, and the requirement for these practices to also submit quality data in order to receive the incentive payment was waived.

Upon hearing of the discontinuation, CMS received a handful of emails from practices that expressed their disappointment. One commented "your program helped incentivize us to move forward," and another that "it has really made a great impact on our facility, that enhanced patient care and a great financial impact."

## B. EVOLUTION OF HEALTH IT USE

### 1. Reported Changes to Health IT Adoption and Use

Given the EHRD's major goal to increase effective EHR use, ideally, the demonstration would have both increased initial adoption of an EHR system (required, as discussed above), and increased use of EHR systems once they were in place (incentivized by the systems payment). This section describes the trend and pattern in adoption and EHR use, as well as the level of EHR use by the end of the program. The analysis of the finishing level of EHR use provides appropriate context for interpreting the positive trends in adoption and use.

**Adoption.** At the time they applied to the demonstration in fall 2008, 52 percent of treatment group practices did not yet use an EHR (Table III.3). By spring 2011, 37 percent of those without an EHR reported that they had adopted one and begun to use it to some degree. Large gains in EHR adoption were seen across practices with all different characteristics. The rate of adoption was highest for affiliated practices (57 percent) and those located in a rural area (49 percent), which had started out with the greatest need to adopt. The rate of adoption was lowest for practices in Louisiana (29 percent), and those with six or more physicians (28 percent), a group that had relatively lower need to adopt at the start.

**Changes in Use.** Two different analyses of OSS data point to the same result: EHR use increased substantially within the treatment group during the demonstration. The first analysis found change in EHR use for specific queried functions among treatment group practices with data for three points in time, based on a simple yes/no measure of use where any use at all counts as a "yes." The time periods studied were (1) fall 2008, based on practices' EHRD applications; (2) spring 2010, based on their responses to the OSS; and (3) spring 2011, again based on their OSS responses.[12] This method shows the powerful, combined effect of EHR adoption along with practices' beginning to use new functions in EHRs. For example, the increase from 37 to 56 percent of practices using automated patient-specific alerts and reminders comes both from practices that newly adopted an EHR and those that already had one in 2008 but were not using that function (Table III.4). Further observations on the trends shown in Table III.4 include:

[12] The two instruments used slightly different wording for some of these items. The application form is provided in Appendix H and the OSS instrument in Appendix A. Only functions with close wording between the two instruments were included in the table.

Table III.3. Treatment Group Practices' Adoption of EHR Systems During Fall 2008 – Spring 2011 by Practice Characteristics

| | (A) Number of Treatment Group Practices (Eligible and Randomized) | (B) Number of Practices with Information on EHR Use or Non-Use in 2008 and 2011[b] | (C) Number of Practices with No EHR Use in Fall 2008 | (D) Percentage of Practices with No EHR Use in Fall 2008 (Percentage of Column B) | (E) Practices Beginning EHR Use Between Fall 2008 and Spring 2011 (Percentage of Column B) |
|---|---|---|---|---|---|
| All | 412 | 321 | 165 | 51 | 37 |
| Site | | | | | |
| Louisiana | 104 | 70 | 39 | 56 | 29 |
| Maryland | 127 | 100 | 44 | 44 | 34 |
| Pennsylvania | 138 | 116 | 61 | 53 | 42 |
| South Dakota | 43 | 35 | 21 | 60 | 43 |
| Practice Size (total number of providers) | | | | | |
| 1-2 | 215 | 146 | 85 | 58 | 40 |
| 3-5 | 122 | 115 | 55 | 48 | 37 |
| 6+ | 75 | 60 | 26 | 42 | 28 |
| Practice Affiliation | | | | | |
| Unaffiliated | NA | 263 | 127 | 48 | 32 |
| Affiliated[a] | NA | 58 | 38 | 66 | 57 |
| Located in a Rural Area | | | | | |
| Yes | 69 | 53 | 32 | 60 | 49 |
| No | 343 | 268 | 133 | 50 | 34 |
| Located in an MUA | | | | | |
| Yes | 120 | 89 | 50 | 56 | 41 |
| No | 292 | 232 | 115 | 50 | 35 |
| Participating in Another EHR, Quality Improvement, or Quality Reporting Program | | | | | |
| Yes | NA | 307 | 155 | 50 | 37 |
| No | NA | 14 | 10 | 71 | 36 |

Sources: Office Systems Survey, conducted in spring and summer 2011, and demonstration application data, fall 2008.

[a]Owned by a hospital, hospital system, or larger medical group, or affiliated with a larger medical group, independent practice association, physician hospital organization, or other entity, according to the 2011 OSS. Most of these practices are owned by a larger organization (Felt-Lisk et al. 2011).

[b]Three practices responded to both 2010 and 2011 OSS, but did not provide information about EHR use, so could not be included here.

EHR = electronic health record; MUA = medically underserved area; NA = not applicable.

Table III.4. Percentage of Treatment Practices Using Selected Functions Queried at Time of EHRD Application

| EHR/Health IT Function | Treatment Group - Fall 2008 (Application) | Treatment Group - Spring 2010 | Treatment Group - Spring 2011 |
|---|---|---|---|
| Number of practices[a] | 324 | 324 | 324 |
| Percent of practices with an EHR | 62 | NA | NA |
| Any EHR/health IT use | 48 | 81 | 90 |
| Electronic patient visit notes | 47 | 65 | 82 |
| Electronic patient problem lists [core MU] | 46 | 66 | 83 |
| Automated patient-specific alerts and reminders | 37 | 46 | 61 |
| Patient email | 9 | 19 | 29 |
| Patient-specific educational materials [menu set MU] | 37 | 36 | 56 |
| Laboratory tests: online order entry | 30 | 26 | 34 |
| Radiology tests: online order entry | 17 | 14 | 19 |
| Printing and/or faxing Rx | 55 | 71 | 81 |
| On-line Rx transmission to pharmacy [core MU] | 35 | 64 | 86 |

Source: Office Systems Survey conducted in spring and summer 2010 and 2011, and demonstration application data, fall 2008.

[a]Excludes practices that closed or were terminated by CMS. Includes only practices with health IT use data in 2008 and 2011. Includes all practices that indicated any use of the functions listed and those that responded to the question about whether they used an EHR, including practices that indicated that they did not use an EHR but they did use a stand-alone registry (2 practices), stand-alone e-prescribing system (31 practices), or both (1 practices) in 2010 or 2011.

MU = meaningful use; NA = not applicable.

- Use of all the functions queried in 2008 increased during the period.
- Large increases in use of electronic patient visit notes and problem lists to relatively high levels of use (from 46 percent to 82 and 83 percent, respectively). This is not surprising, given the robust rates of system adoption discussed above, since those are likely some of the first functions to be used after an EHR is installed.
- E-prescribing became commonplace over the period, with online prescription ordering increasing from 35 percent of practices to 86 percent. Medicare's E-Prescribing Incentive Program provided incentive payments beginning in 2009 to encourage e-prescribing, and may have been an important driver of this increase.
- Use of automated patient-specific alerts and reminders, and use of the system to generate patient-specific educational materials both increased dramatically in the treatment group, with use of automated reminders growing from 37 percent to 61 percent of practices, and system use for educational materials growing from 37 to 56 percent in 2011.
- Use of patient email, online ordering of radiology tests, and online ordering of laboratory tests remained relatively infrequent throughout the period (34 percent of

practices or fewer used these functions in 2011), although use of patient email climbed substantially from 9 to 29 percent.

The second analysis examined change in a composite score of EHR use between 2010 and 2011, created from practices' OSS responses to questions about use of 53 functions, among practices that used EHRs in both periods. Treatment group practices that used an EHR in both years 1 and 2 of the demonstration reported increased use in the second year by an average of 9 points on a 100 point scale (Table III.5). This method considered both new use of queried functions and use of functions that were previously used for more of the practice's patients (Appendix E provides details about the scoring; Appendix A provides the OSS instrument). Table III.5 also shows:

- Overall, 77 percent of year 1 EHR users reported improving their use by year 2 (that is, they reported using more of the specific functions queried, and/or using the functions for more of their patients).
- Above-average EHR use score increases for practices in Maryland, practices with one or two physicians or 6 or more physicians, and those that were located in an urban area (average of 10 to 13 point increase).
- Use score increases for practices in Louisiana and Pennsylvania were less than the average increase by 6 and 7 points respectively).[13]

2. Level of EHR Use in 2011

Although the changes in EHR use during the demonstration period are substantial, the average treatment group practice still has a long way to go to maximize use of its EHR to improve care. The OSS asked about a fairly comprehensive set of 53 EHR functions that are logically connected to improved care (although many are not yet empirically proven and reported in the literature to be associated with better care). If the ideal circumstance is practices using all these functions for three-fourths or more of their patients, which would be reflected in a Total OSS Score of 100, then the current shortfall is substantial, since the average treatment group practice Total OSS Score in 2011 was only 64.

In addition to calculating a Total OSS Score, composite scores were calculated for five domains of the OSS: completeness of information, communication about care outside the practice, clinical decision support, increasing patient engagement, and medication safety. While the elements of "meaningful use" were developed independently of these domains, many of the "meaningful use" elements fall into one of these domains. The many similarities between the "meaningful use" elements (Stage 1) and the EHRD OSS items, and their differences, are listed in Felt-Lisk et al. 2011, Appendix B. Table III.6 shows the average scores on each domain.[14] The main findings are as follows:

[13] These states did not have substantially different OSS scores in 2010 from other states and were not the states with the highest 2010 means.

[14] For details about the descriptive statistics of the OSS scores used for calculating the systems payment, see Felt-Lisk and Verghese (2011).

- Most practices in the demonstration that used an EHR used it in a fairly robust manner on the completeness of information and medication safety domains (mean scores 80 and 79, respectively).
- Communication about care outside the practice and decision support domains was used to a considerably lesser extent as measured by the OSS (both mean scores 58).
- Use of functions related to patient engagement lagged the other domains considerably with a mean score of only 40 percent.

Table III.5. Changes in Total OSS Score Among EHR Users in the Treatment Group During 2010-2011, by Practice Characteristics

| | (1) Number of Treatment Group Practices with OSS Responses Indicating EHR Use in Both 2010 and 2011 | (2) Mean Total OSS Score in 2011[a] | (3) Mean Change in Total OSS Score 2010-2011 Among EHR Users During Both Periods | (4) Percentage of 2010 EHR Users That Improved OSS Score in 2011 Among EHR Users During Both Periods |
|---|---|---|---|---|
| All | 220 | 65 | 9 | 77 |
| Site | | | | |
| Louisiana | 41 | 59 | 7 | 59 |
| Maryland | 74 | 69 | 13 | 93 |
| Pennsylvania | 82 | 64 | 6 | 66 |
| South Dakota | 23 | 69 | 10 | 100 |
| Number of Physicians | | | | |
| 1-2 | 92 | 63 | 8 | 70 |
| 3-5 | 80 | 66 | 9 | 80 |
| 6 + | 48 | 70 | 10 | 85 |
| Owned by/Affiliated with Larger Organization | | | | |
| Affiliated | 41 | 71 | 9 | 78 |
| Unaffiliated | 179 | 64 | 9 | 77 |
| Located in a Rural Area | | | | |
| Yes | 35 | 66 | 12 | 83 |
| No | 185 | 65 | 8 | 76 |

Source: Office Systems Survey (OSS) conducted in spring and summer 2010 and 2011.

Note: Table is restricted to treatment group practices not closed or terminated by CMS that indicated they used an EHR in both 2010 and 2011. The 34 practices with stand-alone registries and/or stand-alone e-prescribing systems are excluded from this table, as were practices that failed to respond to the question directly asking if they used an EHR. Three practices that responded to both surveys discontinued EHR use between 2010 and 2011; their OSS score was not calculated in 2011 so they are excluded from the table.

[a]Medians for each row were between 0 and 3 points higher than the means.

Table III.6. Mean 2011 Office Systems Survey Scores by Domain, Among EHR Users

| OSS Domain | All Sites | Louisiana | Maryland | Pennsylvania | South Dakota |
|---|---|---|---|---|---|
| Completeness of information | 80 | 80 | 81 | 79 | 83 |
| Communication about care outside the practice | 58 | 51 | 56 | 61 | 64 |
| Clinical decision support | 58 | 50 | 63 | 55 | 65 |
| Increasing patient engagement/adherence | 40 | 30 | 41 | 43 | 40 |
| Medication safety | 79 | 80 | 80 | 77 | 79 |
| Total OSS score | 64 | 59 | 65 | 64 | 67 |
| Sample Size | 267 | 50 | 89 | 99 | 29 |

Source: Office Systems Survey (OSS), conducted in spring and summer 2011.

Note: OSS Domain Scores are unweighted, on a 100-point scale.

### 3. Influence of the Beginning of the Medicare and Medicaid EHR Incentive Programs

A majority of EHRD treatment group practices (56 percent) had not changed their decisions or practice due to the ARRA "meaningful use" incentives and penalties as of spring 2011, when the OSS was administered (Table III.7). Figures were similar for control group practices, data shown in Appendix B, Table B.8. These questions were added to the 2011 OSS at the request of CMS. The data were not collected on the 2010 OSS because final program rules were not yet in place. However, a sizable minority of practices (44 percent) had changed their decisions or practice, based on their awareness of the incentives or penalties or both. It appears that both the incentives and penalties components of the law influenced many practices. When practices did not make any changes due to the incentives or penalties, it was generally a deliberate decision, since only 2 percent reported they both did not make changes and were unaware of the incentives and penalties.

Findings by practice characteristics (see Appendix B, Table B.7) included:

- Affiliation mattered. Practices that were affiliated with or owned by a larger entity were less likely than others to report that they had changed decisions or the pace of change in response to learning of the ARRA incentives (in both the treatment and control groups). Note that this may reflect the decision-making process and information flow within the larger organization; if the larger organization governs EHR use in affiliated practices, then even if the incentive program influenced the larger organization's decisions, the smaller practices might not know of this influence. Of those that reported making changes, affiliated practices in the treatment group (more so than in the control group) were more likely to report that the type of response was to accelerate adoption of an EHR system rather than to accelerate use.

Table III.7. Percentage of Treatment Group Practices That Changed Decisions or Practice Due to ARRA "Meaningful Use" Incentives and Penalties

| Item | Percentage of Treatment Group Practices (n=319) |
|---|---|
| Changed Decisions or Practice Due to ARRA in Response to: | |
| Total | 44 |
| Both incentives and penalties | 22 |
| Incentives (only) | 14 |
| Penalties (only) | 7 |
| Did Not Change: | |
| Total | 56 |
| Aware of both incentives and penalties | 50 |
| Aware of incentives but not penalties | 2 |
| Aware of penalties but not incentives | 3 |
| Unaware of both incentives and penalties | 2 |

Source: Office Systems Survey, conducted in spring and summer 2011.

Note: Five practices that did not answer the questions were omitted from the table. Amounts may not add to totals due to rounding.

- Wide variation by state. Louisiana and Maryland practices were more likely to report making changes in response to the incentive program than others, with 46 and 49 percent of treatment group practices reporting such changes, respectively, compared to 21 and 34 percent in Pennsylvania and South Dakota. This pattern was not seen in the control group, however (see Appendix B, Table B.7).
- Large difference in awareness of ARRA incentives between the treatment and control group practices in Louisiana. Only 6 percent of treatment group practices in Louisiana were unaware of the ARRA incentives, whereas 24 percent of the control group practices reported being unaware. This raises the question of whether the treatment group received more information about the incentive program.
- Rural practices were responsive to penalties. Rural practices were somewhat more likely to say they were influenced by the ARRA penalties than non-rural practices (38 percent of rural treatment group practices were influenced versus 28 percent of non-rural practices).
- Size did not seem to matter. The pattern of response did not differ among practices with one to two providers, with three to five providers, and with six or more providers.

Most practices expected to meet Stage 1 Medicare "meaningful use" criteria in either 2011 or 2012 (67 percent of treatment group practices). Eighteen percent of treatment group practices reported already meeting the criteria, and only 5 percent said they would not meet the Stage 1

criteria "in the foreseeable future."[15] Practices with six or more providers appeared prepared to meet the criteria earlier than smaller practices: 72 percent of treatment group practices with six or more providers said they either already met the requirements or could meet them in 2011, versus 58 percent of practices with one to two providers.

Over 40 percent of treatment group practices believed that the payments they could receive from "meaningful use" were substantially larger than the EHRD incentive payments. Twenty-eight percent said the two sources of incentives were "about the same," while only 5 percent thought the EHRD incentive payments were substantially larger. Nearly a quarter of practices had not estimated EHRD payments, and so did not express a view on the comparative benefits between the two programs.

Despite the perception among many of greater incentive under "meaningful use," the data do not show practices losing focus between 2010 and 2011 on EHR functions incentivized by EHRD but not by the Medicare and Medicaid EHR Incentive Programs. To analyze this important issue, we first identified OSS items that overlapped with "meaningful use" elements in the core set (required of eligible practitioners to receive an incentive) and menu set (optional) (Stage 1) (flagged in Table B.6 in Appendix B). One might expect practices to focus most intensively on improving their use of EHR functions that were incentivize by both programs, and that this pattern would be most apparent for functions that were part of the "meaningful use" core set and EHRD. Conversely, one might expect practices to focus less on EHR functions that were incentivized only in EHRD. Table III.8 shows that practices did not follow this expected pattern: Change in average use of functions incentivized by both programs (top row) moved from 2.6 to 2.8 on a 4-point scale (change of .24), while use of functions incentivized only by EHRD (bottom row) increased more, rising from 1.58 to 1.97 (change of .39).. Appendix B, Table B.6 shows the trend for each function queried in the OSS and how the functions were classified to develop the function group results displayed in the table below.

As of spring 2011, few practices reported having received grants or subsidies, bonuses or incentives, or loans related to the purchase or use of an EHR system outside of EHRD or the incentive program. Only 7 percent reported receiving a grant or subsidy, 4 percent a bonus or incentive, and 3 percent a loan. In terms of their future expectations, only about 7 percent were expecting a bonus or incentive outside of EHRD or "meaningful use," while fewer than 1 percent expected to receive a grant or subsidy, or loan (Appendix B, Table B.7).

4. Changes to Practice Workflow and Staffing Due to Implementation of EHRs

Practices that implemented EHR systems within the three years prior to the 2010 site visits reported a few changes to practice operations, including workflow and staffing. The changes that occurred were related to office routines, such as improved documentation, more efficient operations, and improved patient care. The EHRs were not yet being used to directly influence care coordination.

[15] Only a very small percentage said they would meet the criteria in 2013-2015—4 percent of treatment practices.

Table III.8. Changes in Use of Functions Incentivized by EHRD Only vs. EHRD and Medicare and Medicaid EHR Incentive Programs (Treatment Group Practices with OSS Responses in Both 2010 and 2011)

| | Mean Item Score, 2011 (0-4, 4 is use for 3/4 or more patients)[a] | Change in Mean from 2010 | Number of Responses[b] |
|---|---|---|---|
| EHR Functions Incentivized in EHRD AND: | | | |
| Core Set Stage 1 Medicare & Medicaid EHR Incentive Programs | 2.84 | 0.24 | 287 |
| Menu Set Stage 1 Medicare & Medicaid EHR Incentive Programs | 1.50 | 0.34 | 287 |
| Likely related to Core Set requirement to implement a clinical decision support rule | 2.50 | 0.50 | 286 |
| EHR Functions incentivized in EHRD but not incentivized by Stage 1 Medicare & Medicaid EHR Incentive Programs | 1.97 | 0.39 | 287 |

Source: Office Systems Survey, conducted in spring and summer 2010 and 2011.

Note: EHRD items queried on the OSS were often not exactly the same as the Medicare and Medicaid EHR Incentive Programs measures; some judgment was required to classify the items most related. Details about the comparability between specific measures were provided in Appendix B of Felt-Lisk et al. (2011). Details on the trends in specific functions and which functions were classified into which rows are provided in Appendix B, Table B.6.

[a]The mean item score does not represent an exact measurement of the extent of use, as it is the average response regarding the range of proportions of patients for whom the function was used. Specifically, values of 0 to 4 for each item represent the proportion of the practice's patients for whom the EHR function was used over the past month: 0 = None; 1 = Some but less than 1/4; 2 = 1/4 or more, but less than 1/2; 3 = 1/2 or more but less than 3/4; and 4 = 3/4 or more.

[b]Only practices with complete data for all functions within a group were included in calculating the mean item score for the group.

Improved documentation processes. Interviewed staff at treatment practices were quick to mention that they were better able to document clinical notes, test results, and prescriptions with the EHR than with paper charts. Not only were electronic charts more legible, they were also more complete, and physicians were able to easily review patients' charts prior to visits. Four of the 16 practices reported that medical assistants (MAs) were now responsible for documenting patients' vital, laboratory, and imaging results, either at the time of the visit (for example, for blood pressure results) or after (for test results). One large treatment practice (14 physicians, 53 total employees) reported it had not only improved the volume of documentation, but also standardized the recording of medical, social, and family histories, as recommended by an EHR advisory committee composed of physicians, nurse practitioners, and administrative staff. Electronic dictation also played a role in improved documentation; two treatment practices lauded the ability of their EHR systems to accept electronic dictation through Dragon NaturallySpeaking, which also made physicians more amenable to using the EHR systems and fully documenting their clinical notes. However, although the documentation was visibly improved, staff at five treatment practices noted that physicians spent more time documenting notes than with paper, either during the patient visit (which took time away from face-to-face interactions with patients) or at the end of the day. In addition, a physician at one treatment

practice questioned the volume of documentation, stating that if, for example, if he is screening a patient for diabetes during an office visit, he should not have to review some of the general health maintenance items, as they are not all relevant to the visit.

More efficient practice due to elimination of paper charts. Several interviewed staff reported there noticeable improvements in efficiency at their practices. Specifically, one treatment practice noted that the administrative burden was dramatically reduced now that there was no need to pull paper charts and print results to put into the charts. In addition, two treatment practices reported that patients received answers to their questions more promptly because administrative staff could (1) look up information in the electronic system (without pulling a paper chart) and (2) contact the physician via electronic message. An office manager at a solo treatment practice noted that the physician now had about 10 additional minutes during each visit to spend with the patient (from 20 minutes to 30) because he could enter all his notes in the EHR during instead of after the visit. This is remarkable, but it is not typical of impacts of the EHR implementation reported by other practices; instead, it reflects the physician's relatively inefficient style of documentation before the shift, and a good fit of the system to his typing ability.

Integrating EHR into established systems. One treatment practice associated with a large medical group was benefiting from support of the medical group's IT department in the ongoing implementation of its EHR system. The IT department was supplementing the medical group's home-grown system (which contained an e-prescribing system, a clinical data port, a patient portal, and a link to hospital medical records and tests) with the EHR in such a way that its electronic capabilities (for care management, for example) would be seen as an extension to those already in existence. The EHR was being customized so it could be used at all of the medical group's locations.

Change in number of staff. During the initial implementation phase of its EHR, a large treatment practice augmented its staff (up to 53 employees) with several temporary employees charged with scanning and entering data from the paper charts into the electronic system. Two of these temporary staff were later hired as MAs and entered patient data. Another treatment practice (with three physicians) was hoping to decrease the number of staff once operations became more efficient.

5. Factors Influencing Change in Health IT Use

The major factor that seemed to be influencing demonstration practices to adopt EHRs was the national trend toward adoption—the perception that this is how business will need be done in the future. Three practices also decided to adopt EHRs to promote interoperability with the associated hospital system and, if associated with a larger group, other practices. When asked what factors had been helpful in acquiring and implementing their EHR, practices often pointed to strong training or strong advisory groups or resources. Strong training included high-quality vendor support or, in one instance, an IT supervisor who was a former high school teacher, while strong advisory groups included steering committees that were formed at the practice or larger medical group level. In one case, a consulting firm was hired to manage the transition, to avoid overloading existing administrative staff. A user group for the practice's EHR system was cited as another helpful resource. Finally, one medical group that owned a visited practice identified specific staff and physicians who received additional compensation for being the point person for questions on EHR use. The same group used data to identify physicians not using the EHR well,

and proactively provided personal assistance. This group also tied a small portion of physician salaries to EHR use.

The demonstration itself had some influence on 5 of the 16 visited treatment group practices' EHR adoption or use by May/June 2010, according to the site visit interviews. The influence was to speed up or heighten attention to adopting an EHR or, in one case, to provide a structure (through the OSS) for planning enhancements to EHR use.

Physicians and staff of visited practices commented on several barriers to adoption and use of EHRs. Although staff members at practices with EHRs were committed to using their EHRs, they noted several hurdles to their effective use, including the time and labor necessary to implement the systems, the complexity of the electronic systems, lack of interoperability with other systems, and insufficient technical support. Those without EHRs were hindered by the expected high costs of adoption, both in terms of money and labor, although three of the four visited practices without EHRs were committed to adopting an EHR within a year. The fourth was a solo practitioner in a rural area who was experiencing a declining revenue stream and found that the available incentives would not cover the system costs.

## C. CARE MANAGEMENT AND QUALITY MEASUREMENT

### 1. Care Management

The site visit findings confirmed that the demonstration's goal of improving practices' care management was indeed a useful goal. However, possibly because of its early emphasis on reward for health IT use only, the demonstration was not influencing care management in the visited practices at the time of the site visits.

The site visits found that there was room for much growth in practices' care management. The good news is that all but 2 of the 16 visited treatment group practices articulated one or more care management activities, defined as routines designed to improve patient care. However, all—including two practices that were considerably more advanced than the others—had a long way to go to achieve the advanced primary care medical home ideals beginning to be embodied by the "meaningful use" criteria and laid out in medical home joint principles endorsed by major physician organizations.

At the time of the site visits in spring 2010, 6 of the 16 practices were using front desk and medical assistant staff to help update patient information and identify and/or fill gaps in missing preventive services. Five discussed using flags, alerts, and reminders, including some that were using paper charts rather than EHRs. Three routinely gave patients individualized guidance. A few practices did other things: one had about 20 standing orders in place (whereby care could be given or referred under specific circumstances without the physician initiating the order), and the practice believed this had boosted its quality performance and improved the consistency of patient care for the targeted services, such as pneumonia vaccine.

In terms of care beyond the patient visit, 7 of the 16 practices were calling or sending letters to at least some patients who were identified as needing a service; 3 had a process in place to obtain information from a patient's other clinicians; and others have a case manager made available by the local hospital (1 practice) or a patient portal that they populate with key

information after each visit to support patients and family members in accurately remembering care instructions and self-management guidance (1 practice).

2. Quality Measurement

As originally designed, the demonstration would have provided practices information about their quality of care relative to that of other practices starting in the second year of the demonstration. The site visits suggested this would have been a wholly new phenomenon for 5 of the 16 practices, and would have provided potentially useful new information for at least four more.

Eleven of the 16 visited treatment group practices were engaged to some degree with quality measurement outside the demonstration, while a handful (five practices, including at least one in each site) reported not having seen any quality measures calculated for their practice (other than the summary report from the demonstration). Four practices periodically saw payer-specific data from insurers with whom they contract—such as an annual HEDIS report, or quality measures computed from claims or enhanced claims (for example, claims plus laboratory) data—but that was the extent of their involvement in quality measurement. In contrast, five practices were seeing quality measures data that was generated for the whole of their practice (some of them also received quality reports from payers). These efforts included:

- Two practices' larger owner organizations ran "dashboard reports" monthly or quarterly from their EHR system for all the practices in the organization (including those that were visited) with measures that overlap the demonstration measures.
- One practice's owner organization used a contractor to collect data from paper charts for diabetes measures for all the owned practices (this system did not have an EHR).
- One practice participated in Minnesota Community Measurement, a public reporting initiative.
- In one practice, the physician who was interviewed reported running graphs and analyzing data "every couple of weeks" as a hobby.

Another practice participated with its Medicare Quality Improvement Organization (QIO) in its "core prevention" initiative, an effort under the QIO program's 9th Scope of Work, and so saw flu and pneumonia vaccination rates for its Medicare population. Finally, one practice said it tracked patient health indicators for diabetes as well as receiving reports from local payers, but more information was not available.

## D. FINDINGS FROM SITE VISITS TO CONTROL GROUP PRACTICES

In addition to visiting 16 treatment group practices, the Mathematica team visited 8 control group practices and conducted discussions similar to those described above. There was no apparent difference between the two groups of visited practices in the progress of EHR adoption and use, the plans for the future, or the care management practices as of the time of the visits in May/June 2010. This section concludes with observations from the control group visits.

- Influences on EHR adoption. Control practices, like treatment practices, were influenced in their adoption of EHRs by the nationwide move toward health IT; two control practices were aware of the Medicare and Medicaid EHR Incentive Programs.

Technical assistance also played a large role for four of the control practices, which utilized the same types of support (IT consultant, vendor, corporate IT department, training) reported by treatment practices.

- Workflow changes with EHR implementation. The control group reported similar workflow changes as they implemented their EHRs, such as an improvement in documentation due to the EHR templates. As with four of the treatment practices, MAs at one control practice had more responsibility for documentation of patient vital and test results in the EHR. None of the control practices had experienced or expected any staff increases or reductions as a result of the electronic systems.
- Barriers to EHR adoption and use. Similar barriers to effective adoption and use of EHRs were reported, including the labor and monetary cost of implementing and sustaining the system, the burden of transitioning from paper to electronic charts, and the complexity of the templates. One control practice was frustrated by the lack of response from vendors to problems occurring after normal business hours, as there had been several instances when the EHR system's database was inaccessible (including during the site visit), and the problem was not fixed until the next business day.
- Level of EHR use. Like their treatment counterparts, control practices reported varying levels of use of the minimum EHR functions. At the time of the site visits, the five control practices with EHRs reported that they recorded clinical notes and prescriptions electronically, but only some could record and receive diagnostic tests through their EHRs (three practices could receive laboratory orders; four could receive imaging results).
- Plans for the future. In the future, control practices expect to: (1) create patient portals (two practices); (2) improve interoperability with other systems (one practice); (3) print educational materials for patients (one practice); and (4) use an integrated disease registry (one practice). As with the visited treatment practices, control practices were investigating ways to customize their EHRs to help with care management, including improved interoperability with hospitals and laboratory and imaging vendors to enable greater awareness among physicians and practice staff of test results and reminders for test orders.
- Care management. Like the treatment group, most visited control group practices articulated at least one care management activity (six of the eight practices). One that did not was an independent three-physician practice not using an EHR, which indicates that lack of care management is not strictly limited to solo practitioners (though the other practice that did no care management was, in fact, a solo practice).
- Quality measurement. The only notable difference between the visited control group practices' experience with quality measurement and that of the treatment group was that proportionately fewer of the visited control group practices were engaged in quality measurement (4 of 8 control practices, versus 11 of 16 treatment practices). As with the treatment group, several control group practices reported receiving quality measure reports from payers on those specific populations, while several others saw quality reports generated for their whole practice by larger organizations that owned them or through their EHR.

# IV. INCENTIVE PAYMENTS AND THEIR RELATIONSHIP TO PRACTICE CHARACTERISTICS

This chapter describes average incentive payments received per practice in years 1 and 2 of the demonstration, as well as the association between average payment and practice characteristics.

## A. Average Incentive Payments During Each Year

While the number of treatment practices participating in the demonstration and responding to the OSS decreased slightly over time, with 338 practices responding to the OSS in year 1 (93 percent of the 363 eligible practices) and 311 responding in year 2 (90 percent of the 346 eligible practices), the number of practices meeting minimum requirements to receive an incentive payment increased from 198 in year 1 to 231 in year 2 (Table IV.1). In percentage terms, over the course of the demonstration, the percentage of practices that received an incentive payment (out of those that responded to the OSS) increased from 59 percent in year 1 to 74 percent in year 2; likewise, the average incentive payment received by this group of practices increased from $6,014 to $8,218. The average incentive payment per practice that met minimum requirements increased slightly, from $10,266 in year 1 to $11,064 in year 2. Few practices that responded to the OSS in both years received an incentive payment in year 1 but not in year 2 (3 percent); practices that responded to the OSS in both years more often received an incentive payment in year 2, but not in year 1 (17 percent).

## B. Practice Characteristics and Incentive Payments

Because certain types of practices might respond differently to the demonstration incentives, the analysis also examined whether there were significant variations by practice characteristics in incentive payments per practice in year 1 and year 2, for those practices that met minimum requirements to receive payment.

Incentive payments were associated significantly with practice size and site. As expected, larger practices received higher payments in both years, in terms of number of physicians (Table IV.2). While larger practices received higher payments under the demonstration, this finding was partially attributable to the payment structure (specifically, practices were paid on a per-beneficiary basis, though there was a per-physician cap on payments). Incentive payments were also associated significantly with the site in which the practice was located. One possible explanation for this finding is the variation in average practice size by site. South Dakota had a much larger proportion (47 percent) of practices with six or more physicians compared to other sites, and along with that came relatively high Medicare caseloads: more than half the practices reported 1,000 or more Medicare FFS beneficiaries; as shown below, practices in that state received the largest payments per practice, on average, in both years. Another factor that could explain these results is that, because South Dakota had the smallest sample size (only 37 practices were participating at the end of year 1, compared to at least 86 in other states), the estimates are less precise than for those of other states.

Table IV.1. Incentive Payments Received by Treatment Group Practices During the Demonstration

| | Year 1 | Year 2 | Cumulative |
|---|---|---|---|
| 1. Average incentive payment per practice that responded to the OSS | $6,014 | $8,218 | $7,070 |
| 2. Percentage of practices that responded to the OSS and met the minimum requirements to receive payment | 58.6 | 74.3 | NA |
| 3. Total number of practices that responded to the OSS[a] | 338 | 311 | NA |
| 4. Total incentive payments across all practices that met the minimum requirements in OSS | $2,032,569 | $2,555,748 | $4,588,317 |
| 5. Average incentive payment per practice that met the minimum requirements in OSS | $10,266 | $11,064 | $10,695 |
| 6. Total number of practices that met minimum requirements in OSS | 198 | 231[b] | NA |
| 7. Percentage of practices that responded to the OSS in years 1 and 2 and received incentive payments in year 1 but not in year 2 | 3.3 | NA | NA |
| 8. Percentage of practices that responded to the OSS in years 1 and 2 and received incentive payments in year 2 but not in year 1 | NA | 17.0 | NA |

Source: Office Systems Survey, conducted in spring and summer 2010 and 2011; payment data provided by EHRD's implementation support contractor for treatment group practices that submitted OSS data in 2010 and/or 2011.

[a]Includes eligible and participating treatment practices that completed the OSS.

[b] In 2011, three practices that were asked to complete the validation survey but did not complete it or failed to provide the requested screenshots are considered to not have completed the OSS (see Appendix F).

NA = not applicable; OSS = Office Systems Survey.

There was no statistically significant association between average payment per practice and other practice characteristics, including practice affiliation, location in a rural/non-rural area, location in a medically underserved area, or participation in another EHR, quality improvement, or quality reporting program.[16]

Expanding the sample to include all practices that responded to the OSS (N = 338 and 311 in years 1 and 2, respectively), yielded similar results—larger practices tended to receive larger payments, and practice site was significantly associated with average payment (see Appendix I). However, in year 1, participation in another EHR, quality improvement, or quality reporting program was also significantly associated with average payment; the average payment for participating practices was $2,841 more than the average for other practices (p = .001).

[16] Because data on the number of beneficiaries included in the calculation of the incentive payments was not available, the percentage maximum potential payment was not calculated.

Table IV.2. Average Incentive Payments During the First and Second Years of the Demonstration, by Practice Characteristics

| | Average Payment (Dollars) in Year 1 | Average Payment (Dollars) in Year 2 |
|---|---|---|
| Site | | |
| Louisiana | $10,487 | $10,755 |
| Maryland | $12,216 | $13,006 |
| Pennsylvania | $7,046 | $8,088 |
| South Dakota | $14,597 | $15,261 |
| p-value | 0.000*** | 0.000*** |
| Practice Size | | |
| 1-2 physicians | $4,864 | $4,856 |
| 3-5 physicians | $9,999 | $11,017 |
| 6 or more physicians | $18,914 | $19,697 |
| p-value | 0.000*** | 0.000*** |
| Practice Affiliation | | |
| Unaffiliated | $11,062 | $10,703 |
| Affiliated[a] | $9,402 | $11,344 |
| p-value | 0.138 | 0.548 |
| Located in a Rural Area | | |
| Yes | $12,232 | $12,142 |
| No | $9,928 | $10,831 |
| p-value | 0.145 | 0.343 |
| Located in a Medically Underserved Area | | |
| Yes | $9,019 | $9,995 |
| No | $10,770 | $11,474 |
| p-value | 0.157 | 0.210 |
| Participating in Another EHR, Quality Improvement, or Quality Reporting Program | | |
| Yes | $10,559 | $10,967 |
| No | $9,503 | $11,905 |
| p-value | 0.399 | 0.588 |
| Number of Practices[b,c] | 198 | 231 |

Sources: Office Systems Survey, conducted in spring and summer 2010 and 2011; payment data provided by EHRD's implementation support contractor for treatment group practices that submitted OSS data in 2010 and/or 2011; baseline characteristics from practice applications and HRSA's Area Resources File.

Notes: The p-values from testing the equality of means across binary variables are from t-tests. The p-values from testing the equality of means across practice characteristics for a variable with multiple (>2) categories (practice size) are from the F-test of an analysis of variance.

[a]Owned by a hospital, hospital system, or larger medical group, or affiliated with a larger medical group, independent practice association, physician hospital organization, or other entity.

[b]Includes eligible and participating treatment practices that completed the Office Systems Survey.

[c] In 2011, three practices that were asked to complete the validation survey but did not complete it or failed to provide the requested screenshots are considered to not have completed the OSS (see Appendix F).

***Statistically significant at 1 percent level; **Statistically significant at 5 percent level; *Statistically significant at 10 percent level

This page is intentionally left blank.

# V. IMPACTS OF THE DEMONSTRATION ON KEY OUTCOMES DURING YEARS 1 AND 2

This chapter examines the impacts of EHRD on health IT use, quality of care, and Medicare service use and expenditures. As outlined in Chapter II, the demonstration's financial payments were intended to increase practices' use of EHRs and evidence-based guidelines, which could in turn improve quality of care, as defined by more effective disease management, case management, or administration of recommended preventive care. This improved quality of care could generate improved beneficiary health outcomes and reduce net Medicare expenditures through reduced emergency room, inpatient, and outpatient utilization.

As a result of the demonstration's incentive payments, treatment group practices were more likely than control group practices to report using EHRs; they were also more likely to report using EHRs for specific tasks, such as communicating with outside providers, supporting decision making, increasing patient engagement, and encouraging medication safety (Tables V.1 and V.2). Across all four sites, the demonstration appeared to decrease the likelihood of preventable hospitalizations related to CAD, CHF, or diabetes during year 2; these reductions were concentrated in practices that used EHRs before the demonstration. However, there were no significant impacts across participating sites on the five available claims-based quality measures the EHRD was intended to affect. Moreover, there is no evidence that Medicare costs and service use fell during the demonstration period. Table V.3 summarizes the main findings from the difference-in-differences analysis of these claims-based outcomes; detailed impact tables are provided in Appendix J.

## A. Impacts on Health IT Use in Years 1 and 2

To determine the impact of the demonstration's financial incentives on health IT use, treatment and control practices' responses to the 2011 OSS were compared. This analysis used data from the 324 treatment practices (or 80 percent of the 405 eligible,[17] randomized treatment practices) and 268 control practices (or 65 percent of the eligible, randomized control practices) that completed the follow-up OSS.[18] First, this section examines treatment-control differences in selected measures of health IT use, gleaned from the OSS. Second, it presents treatment-control differences in overall OSS scores and the five OSS domain scores. Overall OSS scores provide a general measure of practices' general health IT use, whereas domain scores measure practices' health IT use related to specific objectives, such as increasing patient engagement or ensuring medical safety. These analyses suggest that the demonstration had large positive impacts on practices' use of health IT. Moreover, an analysis of nonresponse bias suggests that these

[17] CMS determined that 7 of the 412 randomized treatment practices and one of the 413 randomized control practices were ineligible before the demonstration.

[18] These comparisons are made under an intent-to-treat analysis approach, in which all treatment practices that completed the relevant survey are compared to all control practices that completed the relevant survey. This stands in contrast to a treatment-on-the-treated approach, which would estimate the impact of the demonstration on only those treatment practices that actually adopted an EHR.

Table V.1. Impacts of EHRD on Health IT Use, by Function

| EHR/Health IT Function | Treatment Group Adjusted Mean | Control Group Adjusted Mean | Impact | p-value |
|---|---|---|---|---|
| Any EHR/Health IT Use | 89.8 | 80.2 | 9.6 | <0.01 |
| Electronic Patient Visit Notes | 83.8 | 68.6 | 15.2 | <0.01 |
| Electronic Patient Problem Lists [MU-C] | 84.5 | 70.4 | 14.1 | <0.01 |
| Automated Patient-Specific Alerts and Reminders | 63.1 | 45.3 | 17.9 | <0.01 |
| Electronic Disease-Specific Patient Registries | 70.9 | 53.2 | 17.7 | <0.01 |
| Patients' Email | 30.8 | 29.8 | 1.0 | 0.80 |
| Patient-Specific Educational Materials [MU-M] | 58.6 | 42.2 | 16.4 | <0.01 |
| Online Referrals to Other Providers | 70.2 | 57.7 | 12.5 | <0.01 |
| Laboratory Tests: | | | | |
| Online order entry | 35.7 | 35.1 | 0.6 | 0.87 |
| Online results viewing | 68.3 | 58.5 | 9.8 | 0.01 |
| Radiology Tests: | | | | |
| Online order entry | 19.5 | 22.4 | -2.9 | 0.39 |
| Online results viewing (reports and/or digital films) | 46.5 | 40.5 | 6.1 | 0.12 |
| E-Prescribing: | | | | |
| Printing and/or faxing Rx | 82.2 | 69.1 | 13.1 | <0.01 |
| Online Rx transmission to pharmacy [MU-C] | 86.8 | 71.8 | 14.9 | <0.01 |
| Number of Practices (Weighted) | 405 | 412 | | |
| Number of Practices (Unweighted) | 324 | 268 | | |

Sources: Office Systems Survey, conducted in spring and summer of 2011 and data drawn from the applications practices submitted to EHRD in 2008.

Notes: Reported means are regression-adjusted. Regressions control for state, MUA, practice size, and health IT-related variables practices reported on the application to the demonstration. Observations for treatment and control group practices are adjusted for nonresponse to the 2011 OSS and for demonstration attrition. The weighted sample reflects all randomized practices, except for seven treatment practices and one control practices that were determined by CMS to be ineligible before the demonstration. Eighty percent (324 of 405) of eligible randomized treatment practices and 65 percent (268 of 412) of eligible randomized control group practices responded to the year 2 OSS.

MU-C indicates a function related to a Stage 1 meaningful use core set item; MU-M indicates a function related to a Stage 1 meaningful use menu set item. See Chapter III for additional information on meaningful use items.

CMS = Centers for Medicare & Medicaid Services; EHRD = Electronic Health Records Demonstration; MUA = medically underserved area; OSS = Office Systems Survey; Rx = prescription.

Table V.2. Impacts of EHRD on OSS Scores, by Domain

| OSS Score (Means) | Treatment Group Adjusted Mean | Control Group Adjusted Mean | Difference | p-value |
|---|---|---|---|---|
| Overall OSS score | 54.35 | 42.81 | 11.54 | <0.01 |
| OSS Score Domains | | | | |
| 1. Completeness of information in the EHR | 11.71 | 9.31 | 2.40 | <0.01 |
| 2. Communication of care outside the practice | 10.93 | 8.99 | 1.94 | <0.01 |
| 3. Clinical decision support | 10.82 | 8.51 | 2.30 | <0.01 |
| 4. Increasing patient engagement | 5.81 | 4.37 | 1.45 | <0.01 |
| 5. Medication safety | 14.68 | 11.31 | 3.37 | <0.01 |
| Number of Practices (Weighted) | 405 | 412 | | |
| Number of Practices (Unweighted) | 324 | 268 | | |

Sources: Office Systems Survey, conducted in spring and summer 2011, and data drawn from applications practices submitted by practices to EHRD in 2008.

Notes: Reported means are regression-adjusted. Regressions control for state, MUA, practice size, and health IT-related variables practices reported on the application to the demonstration. Observations for treatment and control group practices are adjusted for nonresponse to the 2011 OSS and for demonstration attrition. The weighted sample reflects all randomized practices, except for seven treatment practices and one control practices that were determined by CMS to be ineligible before the demonstration. Eighty percent (324 of 405) of eligible randomized treatment practices and 65 percent (268 of 412) of eligible randomized control group practices responded to the year 2 OSS.

CMS = Centers for Medicare & Medicaid Services; EHRD = Electronic Health Records Demonstration; MUA = medically underserved area; OSS = Office Systems Survey.

impacts would have been even larger if more control group practices had responded to the follow-up survey.[19]

[19] Specifically, although about half of 2011 OSS respondents reported having EHRs at baseline, only about 43 percent of the full sample (including OSS nonrespondents) reported having EHRs at baseline. This suggests that practices without EHRs (at baseline and follow-up) were less likely than those with EHRs to respond to the follow-up OSS. Although regressions are weighted to account for the fact that practices without EHRs at baseline were less likely to complete the OSS than those with EHRs, it is likely that nonresponse bias would overstate practices' EHR use, particularly for the control group relative to the treatment group (because practices in the control group were less likely than those in the treatment group to respond to the follow-up OSS). The end-result of this nonresponse bias would be that impact estimates of the EHRD on health IT use would be artificially low.

1. Impacts on Selected Measures of Health IT Use

Using data from all treatment and control practices that completed the 2011 OSS, this analysis found statistically significant impacts of the demonstration on several key health IT functions measured on both the 2008 demonstration application and the 2011 OSS (Table V.1). After controlling for practice characteristics and baseline health IT use, treatment group practices were 10 to 18 percentage points more likely than control group practices to report the following functions: using an EHR, making electronic patient visit notes, keeping electronic patient problem lists, using automated patient-specific alerts and reminders, using electronic disease-specific patient registries, disseminating patient-specific educational materials, making online referrals to other providers, viewing lab tests online, printing and faxing prescriptions, and digitally transmitting prescriptions to pharmacies. In particular, there are large treatment-control differences for automated patient-specific alerts and reminders, and for electronic disease-specific patient registries (18 percentage point treatment-control difference in both cases). These treatment-control differences were similar in magnitude and statistical significance regardless of the use of baseline controls or the application of nonresponse weights. However, the demonstration had no statistically significant impacts on practices' likelihood to report using patients' email, ordering laboratory tests, ordering radiology tests online, or viewing radiology results online.

2. Impacts on Health IT Performance Score

In addition, this analysis found that the demonstration had a statistically significant and positive impact on practices' overall OSS scores as well as all five OSS domain scores (Table V.2). After controlling for practice characteristics and baseline health IT use, treatment group practices' overall OSS scores were more than 11 points higher than those of control group practices, on average (54 of 100 for treatment versus 43 of 100 for control group practices).[20] In addition, treatment group practices' scores on all five domains were at least 1.5 points higher than control group practices' (with a maximum score of between 17 and 22 points in each domain). There were notably large impacts on the completeness of information in the EHR and medication safety domains (2.4 and 3.4 points, respectively). In analyses that limited the sample to EHR users (excluding the 96 practices without an EHR), positive impacts on health IT use were present regarding the completeness of information and on medication safety; however, there were no significant treatment-control differences in communication of care outside the practice, clinical decision support, or increasing patient engagement.

[20] Please see Appendix E for a summary of the methodology used for calculating the OSS scores. Chapter III provides a descriptive analysis of this measure for treatment group practices.

In addition to treatment practices' higher average scores for all five domains, treatment practices also had significantly higher scores than control practices on most functions within each domain. Treatment practices were at least 10 percentage points more likely to report using 17 of the 18 functions related to completeness of health IT information (such as using allergy lists, developing problem lists, and so on) (Appendix Table J.1). For example, 86 percent of treatment practices, compared to 72 percent of control practices, reported using EHRs to create, update, store, and display allergy lists for patients. Similarly, treatment practices were significantly more likely than control practices to report using EHRs to conduct (1) 12 of 19 functions related to communication of care, (2) 16 of 17 functions related to clinical decision support, (3) 23 of 28 functions related to patient engagement, and (4) 8 of 9 functions related to medication safety. Nearly all of these statistically significant differences were sizeable, with the treatment group being 6 to 18 percentage points more likely to use each function than the control group. Conversely, there were no functions that the control group was significantly more likely to use than the treatment group.

Sensitivity tests confirmed that the results were similar in regressions that did not use baseline control variables and in regressions that did not use nonresponse weights. In short, the demonstration appeared to have a positive impact on practice's use of health IT across a variety of outcome measures and regression specifications.

## B. Impacts on Quality of Care

In this section, Medicare claims data are analyzed to determine whether the demonstration's system incentive payments had an impact on practices' quality of care, as measured by the five claims-based quality measures available for this analysis. The demonstration's impact on preventable hospitalizations is also analyzed, as these measures are commonly used as proxies for quality of care. The impacts are based on difference-in-differences estimates, which compare changes in outcomes over time for the treatment group to those for the control group, while controlling for beneficiary and practice characteristics (as explained in Chapter II).

### 1. Impacts on Quality Measures

As shown in Table V.3, this analysis found no statistically significant impacts in quality measures when pooling claims data across all four demonstration sites. Within particular sites, however, there were a few statistically significant impact estimates. For example, in Maryland, relative to control group beneficiaries, treatment group beneficiaries experienced a favorable 2-percentage point increase in urine tests for protein in year 2, and an unfavorable 2-percentage point reduction in screening for breast cancer in year 1 (Appendix Table J.2). In addition, treatment group beneficiaries in South Dakota were slightly more likely than control group beneficiaries to have a blood test for HbA1c in year 2. These state-by-state results indicate that there were three statistically significant impacts (two negative and one positive) at the 10 percent level among the five different measures that were possible within each state during each year (a total

Table V.3. Summary of EHRD Impacts on Claims-Based Outcome Measures Between Baseline, Year 1, and Year 2 of the EHRD Evaluation

| | All Sites | | LA | | MD | | PA | | SD | |
|---|---|---|---|---|---|---|---|---|---|---|
| Outcome | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 |
| **Quality-of-Care Measures** | | | | | | | | | | |
| Among Beneficiaries with CAD: | | | | | | | | | | |
| Cholesterol/lipid test | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns |
| Among Beneficiaries with Diabetes: | | | | | | | | | | |
| HbA1c test | ns | ns | ns | ns | ns | ns | ns | ns | ns | Incr |
| Cholesterol/lipid test | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns |
| Urine test | ns | ns | ns | ns | ns | Incr | ns | ns | ns | ns |
| Among Female Beneficiaries Ages 40–69: | | | | | | | | | | |
| Breast cancer screening | ns | ns | ns | ns | Decr | ns | ns | ns | ns | ns |
| **Preventable Hospitalizations** | | | | | | | | | | |
| Any Cardiac Hospitalization (among beneficiaries with CAD) | ns | ns | ns | Decr | ns | ns | ns | ns | ns | ns |
| Any Hospitalization Related to CHF | ns | ns | ns | ns | ns | ns | Incr | ns | Decr | ns |
| Any Hospitalization Related to Diabetes | ns | ns | ns | Decr | ns | ns | ns | ns | Decr | ns |
| Any Hospitalization Related to CAD, CHF, or Diabetes | ns | Decr | ns | Decr | ns | ns | ns | ns | Decr | ns |
| **Medicare Expenditures** | | | | | | | | | | |
| Total Expenditure | ns | ns | Incr | ns | ns | ns | ns | ns | ns | ns |
| Part A Services | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns |
| Inpatient Hospitalizations | ns | ns | Inc | ns | ns | ns | ns | ns | ns | ns |
| Part B Services | ns | ns | Incr | Incr | ns | ns | ns | ns | ns | ns |
| Physician Services | ns | ns | Incr | Incr | ns | ns | ns | ns | ns | ns |
| Outpatient Services | ns | ns | ns | Incr | ns | ns | ns | ns | ns | ns |
| Home Health Services | ns | ns | Incr | ns | ns | ns | ns | ns | ns | ns |

*Table V.3 (continued)*

| Outcome | All Sites | | LA | | MD | | PA | | SD | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 |
| Medicare Service Use | | | | | | | | | | |
| Inpatient Hospitalizations | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns |
| ER Visits | ns | ns | ns | ns | ns | ns | ns | ns | ns | Incr |
| Physician Office Visits | ns | ns | Incr | Incr | ns | ns | ns | ns | ns | ns |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment and control group practices at the end of the baseline period (2008); at the end of year 1 (June 1, 2009, through May 31, 2010), and/or at the end of year 2 (June 1, 2010, through May 31, 2011).

Notes: All estimates were obtained from difference-in-differences models that control for beneficiary demographic characteristics, diagnoses, reason for Medicare eligibility, practice-fixed effects, and indicators for whether the observation is in the year 1 or the year 2 period (described in Appendix G). "Incr" implies an increase (and "Decr" implies a decrease) from baseline for the treatment group relative to the control group in year 1 or year 2 of the demonstration that was statistically significant at the 10 percent level. For example, across all sites between baseline and year 2, outpatient services increased for the treatment group relative to the control group, as shown in the second column of the sixth row in the Medicare expenditures panel. Increases imply favorable outcomes for quality measures, but increases imply unfavorable outcomes for preventable hospitalizations, expenditures, and service use.

Decr= Decrease; EHRD = Electronic Health Records Demonstration; ER = emergency room; HbA1c = hemoglobin A1c; Incr = Increase; LA = Louisiana; MA = Massachusetts; MD = Maryland; PA = Pennsylvania; SD = South Dakota.

ns = not statistically significant at the 10 percent level.

of 40 significance tests).[21] This small number of statistically significant differences could reflect the potential of statistical tests to generate some statistically significant differences when a large array of variables is considered, even if these differences are due to chance.[22] (See Appendix Table J.2 for detailed treatment-control comparisons.)

In addition, one subgroup analysis explored a potential correlation between quality of care and health IT adoption. Compared to practices with little or no improvement in health IT related to CAD and diabetes, treatment group practices with large improvements in health IT use related to CAD and diabetes experienced a larger increase in the average percentage of beneficiaries who received blood tests related to these conditions (see Appendix Table J.3); however, these differences were not statistically significant.[23]

2. Impacts on Preventable Hospitalizations

For preventable hospitalizations, across all sites, the demonstration had a favorable impact on a composite measure for whether beneficiaries with any of the three target conditions (CAD, CHF, or diabetes) had a preventable hospitalization related to their chronic condition in year 2. Specifically, the likelihood of any preventable hospitalizations fell by 0.5 percentage points (or 4 percent) relative to the baseline control group mean of 11.8 percent ($p = 0.05$, Appendix Table J.4). The demonstration's effects on preventable hospitalizations were also favorable and statistically significant in South Dakota in year 1 (for CHF- and diabetes-related hospitalizations as well as hospitalizations related to all three target conditions, $p = 0.04$ and 0.05, respectively), and in Louisiana in year 2 (for CAD- and diabetes-related hospitalizations as well as hospitalizations related to all three target conditions, $p = 0.03$ and 0.01, respectively). These findings regarding the likelihood of a preventable hospitalization largely mirror treatment-control differences in the number of preventable hospitalizations among beneficiaries with these target conditions (Appendix Table J.5). However, there are more statistically significant treatment-control differences in the likelihood of preventable hospitalizations—as opposed to the average number of each type of hospitalization—because likelihood measures reflect up to four types of hospitalizations related to each targeted condition.

Several subgroup analyses were conducted to examine whether these estimates differed by practice characteristics (for example, number of physicians, caseload, or prior use of EHRs) or

[21] A 10 percent level of significance was used (instead of the more conventional 5 percent) to ensure that potentially important site-specific estimates for changes in outcomes were identified. With a lower significance level, these estimates might not have been identified. Statistical tests were two-tailed, in which the hypothesis of no impact was rejected if the value of the test statistic was either sufficiently small or sufficiently large.

[22] For example, the Bonferroni adjustment for multiple comparisons would imply that a p-value should be less than 0.02 for any of five related outcomes (such as the five quality measures examined here) to avoid falsely concluding at the 10 percent level that the demonstration improved quality during a particular year in a specific state. Because none of the treatment-control differences in quality measures had p-values below 0.02, there is no solid evidence that these statistical differences are not due to chance.

[23] For example, the largest increases in blood tests for cholesterol or lipids (among beneficiaries with CAD) and blood tests for HbA1c (among beneficiaries with diabetes) were detected among 55 treatment practices that used nearly no health IT functions during baseline, but adopted nearly all diabetes-specific health IT functions by the 2011 OSS.

by beneficiary characteristics (entitled to Medicare as a result of age versus disability; see Appendix Table J.6 for a subgroup analysis of preventable hospitalizations related to CAD, diabetes, or CHF). These analyses suggest that reductions in the likelihood of preventable hospitalizations were concentrated in practices with one or two physicians, caseloads of more than 100 beneficiaries per physician, and that used EHRs before the demonstration.

Finally, a correlation analysis assessed whether improvements in EHR use tailored to beneficiaries with CAD, diabetes, and CHF were associated with reductions in preventable hospitalizations related to CAD, diabetes, and CHF, respectively. Interestingly, reductions in preventable hospitalizations related to CAD and diabetes were concentrated in the treatment group practices that had the greatest improvements in EHR use between baseline and year 2 (see Appendix Table J.7). For example, nearly all improvements in preventable hospitalizations related to diabetes were concentrated among 55 treatment practices that used nearly no health IT functions during baseline, but adopted nearly all diabetes-specific health IT functions by the 2011 OSS. On average, these high-adopting practices experienced reductions of 2 percentage points in the likelihood of diabetes-related hospitalizations over the full demonstration period. Similarly, 45 treatment practices that used nearly no health IT functions during baseline, but adopted nearly all CAD-specific health IT functions by the 2011 OSS, displayed larger decreases in preventable hospitalization rates than treatment practices with lower adoption rates. On average, these high-adopting practices experienced reductions of 2 percentage points in the likelihood of CAD-related hospitalizations over the full demonstration period.

## C. Impacts on Medicare Expenditures and Service Use

This section analyzes the demonstration's impact on total Medicare expenditures, expenditures for particular Medicare services, and use of Medicare-covered services. These expenditure analyses help determine whether, over its two years of duration, the demonstration appeared to reduce Medicare costs through reduced emergency room, inpatient, and outpatient utilization. All of the impacts are based on difference-in-differences estimates that compare changes in outcomes for the treatment and control groups, after adjusting for beneficiary and practice characteristics. Under the study's intent-to-treat approach, impact estimates are based on Medicare records for all relevant FFS beneficiaries from all practices randomly assigned to the demonstration's treatment and control groups.

### 1. Impacts on Total Medicare Expenditures

Under an analysis that pooled all sites in the demonstration, the demonstration did not appear to affect total Medicare expenditures (Table V.4). Regardless of the inclusion or exclusion of the system incentive payments from calculations, the analysis found no statistically significant impacts on annualized Medicare expenditures during the demonstration's two years.[24] (Note that all confidence intervals include a difference of $0 between treatment and control groups.) Subgroup analyses of total Medicare expenditures across all sites revealed that regardless of the number of physicians, caseload, or prior use of EHRs, the demonstration had no

[24] Chapter IV provides a summary of payments to treatment group practices that met minimum requirements in years 1 and 2.

statistically significant impact on total expenditures (Appendix Table J.8). As further discussed below, total Medicare costs did not fall as a result of the reduction in preventable hospitalizations in Year 2, likely because any reduction in costs associated with the lower preventable hospitalization rate was small in proportion to the large variation in total Medicare costs.

Site-level analyses replicated these expenditure results for the pooled sample, with one exception. In Louisiana, total Medicare expenditures per beneficiary were more than $500 higher for treatment group practices than control group practices in the first year of the demonstration (see Appendix Table J.9).

2. Impacts on Expenditures for Particular Medicare Services

This section examines EHRD impacts on selected annual Medicare Part A and Part B expenditures. With the exception of outpatient services during the second year of the demonstration (impact of an additional $69 billed per patient, significant at the 10 percent level), there are no statistically significant impacts on Medicare expenditures per beneficiary in the pooled analysis of all sites during the two years of the demonstration (Table V.3).

However, in Louisiana, treatment group practices had Part B expenditures that were $180 and $240 more per beneficiary than control group practices in year 1 and year 2, respectively. (See Appendix Table J.9 for detailed estimates.) These differences were driven by statistically significant impacts on physician services (both years) and outpatient service expenditures (year 2 only). During year 1, treatment group practices in Louisiana also had higher inpatient and home health services expenditures (differences of $292 and $118, respectively). Other statistically significant impacts of the demonstration are found in hospice expenditure in South Dakota and Maryland, and durable medical equipment in Pennsylvania (not shown). However, these treatment-control differences are small in magnitude (less than $40 per patient in either direction).

Regarding impacts in each demonstration site, there was a statistically significant (and unfavorable) increase in emergency room visits among treatment group beneficiaries in South Dakota during year 2.[25] However, this increase of 0.03 visits per beneficiary (per year) was small in magnitude (Appendix Table J.10). In Louisiana, however, treatment group practices experienced an increase in the number of physician visits in both years of between 0.3 and 0.4 additional visits per beneficiary (per year). This substantial and statistically significant change is consistent with the higher expenditures in physician services among treatment group practices in Louisiana discussed previously.

[25] Given the small sample size of 87 randomized practices in South Dakota, impact estimates for this site are not likely to be as precise as estimates for larger sites in the demonstration.

Table V.4. Impacts of the EHRD on Total Medicare Expenditures During Years 1 and 2, by Site

| | Control Group Mean Annual Medicare Expenditures (Dollars) | Difference-in-Differences Estimate (Dollars) | 95 Percent Confidence Interval (Dollars) | p-value |
|---|---|---|---|---|
| Excluding Year 1 Incentive Payment | 11,480 | 80 | -170 to 330 | 0.53 |
| Including Year 1 Incentive Payment | 11,480 | 94 | -156 to 344 | 0.46 |
| Excluding Year 2 Incentive Payment | 11,743 | 129 | -147 to 405 | 0.36 |
| Including Year 2 Incentive Payment | 11,743 | 148 | -129 to 423 | 0.30 |
| **Number of Observations[a]** | **94,267** | **800,524** | | |
| **Number of Practices[b]** | **410** | **822** | | |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment and control group practices at the end of the baseline period (2008); at the end of year 1 (June 1, 2009, through May 31, 2010), and/or at the end of year 2 (June 1, 2010, through May 31, 2011).

Notes: The demonstration's effect on total Medicare expenditures per beneficiary (excluding or including incentive payments) was estimated using a difference-in-differences regression model that controlled for beneficiary demographic characteristics, diagnoses, reason for Medicare eligibility, practice-fixed effects, and indicators for whether the observation is in the year 1 or the year 2 period. Incentive payment per beneficiary in year 1 (or year 2) was calculated by dividing the total payment received by a practice during year 1 (or year 2) by the number of beneficiaries assigned to that practice during year 1 (or year 2). Standard errors are robust to the clustering of beneficiaries within practices.

[a]The number of observations for the treatment and control groups is the number of beneficiaries assigned to treatment and control practices at the end of year 2. The number of observations for the difference-in-differences estimate is equal to the sum of the number of beneficiaries (with nonmissing control variables) in the treatment and control groups across baseline, year 1, and year 2 of the demonstration.

[b]The number of practices in the analysis is all randomized practices that had assigned beneficiaries in the baseline period, year 1, or year 2. Only 3 of the original 825 randomized practices had no assigned beneficiaries in any of these periods, and were thus excluded from the analysis.

EHRD = Electronic Health Records Demonstration.

### 3. Impacts on Use of Medicare-Covered Services

Consistent with expenditure results, there were no statistically significant treatment-control differences in the number of inpatient hospitalizations, emergency room visits, or physician office visits in the pooled analysis of all sites in the demonstration (Table V.3 and Appendix Table J.10). This finding was consistent across all subgroups related to the number of physicians, caseloads, or prior EHR use. (See Appendix Table J.11 for a subgroup analysis of impacts on inpatient hospital stays.)

Regarding treatment-control differences in utilization rates—or beneficiaries' likelihood to use a service during the year—there were no statistically significant impacts across all sites. The reduction in the preventable hospitalization rate for the subgroup with CHF, CAD or diabetes

only translated to a small, statistically insignificant reduction in the overall hospitalization rate. Specifically, in Year 2, while there was a 4 percent reduction in the likelihood of preventable hospitalizations among those with CAD, CHF or diabetes, the reduction in the likelihood of any hospitalization among this population was smaller--half of a percentage point on a mean of 35 percent, which translates to a 1.4 percent, statistically insignificant, reduction (not shown). This is likely because preventable hospitalizations are only a fraction of all hospitalizations; while 11.8 percent of those with CAD, CHF or diabetes had a preventable hospitalization, 35 percent of this subgroup had any hospitalization. Moreover, this subgroup (those with CAD, CHF, or diabetes) only represents 53 percent of the total sample, so the reduction in the overall hospitalization rate for the full sample was less than a third of a percentage point (or about 1 percent) and statistically insignificant.

Within sites, however, there was a statistically significant 1-percentage point decrease in the probability of emergency room visits among treatment group beneficiaries relative to their control group counterparts in Pennsylvania, but only in year 1 (Appendix Table J.12). Also in Pennsylvania, beneficiaries in treatment group practices were about 1 percentage point less likely than beneficiaries in control group practices to make outpatient visits during both years. In Louisiana, consistent with the home health expenditure results described earlier, patients in treatment group practices were 1 percentage point more likely than those in control group practices to use home health services in year 1. All other significant treatment-control differences were smaller than 1 percentage point in either direction.

Subgroup analyses for Medicare service use outcomes did not reveal any clear patterns by practice or beneficiary characteristics (Appendix Table J.11); moreover, none of the results suggested that expenditures declined for a particular subgroup. Similarly, various sensitivity tests (summarized in Appendix Table J.13), such as estimating the model without practice-fixed effects, trimming extreme values of expenditures, and including individual-fixed effects, confirm the findings reported here. Taken together, these results suggest that Medicare costs and service use did not decline among treatment group beneficiaries relative to the control group over the course of the demonstration.

# VI. CONCLUSION

The goals of the demonstration were to use financial incentives to encourage the implementation and use of certified EHR systems among primary care physicians in targeted practices. In addition to these systems payments, the EHRD also intended to use financial performance incentives to improve the quality of care to eligible fee-for-service Medicare beneficiaries. This chapter examines whether the demonstration was on target to meet these goals, though only a limited assessment could be conducted because the demonstration was terminated by CMS in August 2011, after only two years. In addition, the chapter presents a context for interpreting the results, and describes the limitations of the evaluation and lessons learned.

## A. Summary and Discussion of Results

Participation. Practices were required to implement and use EHR systems before the end of year 2 to qualify for systems payment. Many practices complied with this requirement; however, a sizeable portion (43 percent) of the practices randomized to the treatment group had left the program either voluntarily, or more commonly, because they failed to meet program requirements. Thus, the first important finding is that many practices were unwilling or unable to fulfill the demonstration requirements.

Site visits and interviews with withdrawn practices suggest two main reasons for the high attrition. First, implementing an EHR is a major, difficult undertaking. Second, many practices seemed to lack some or all of the conditions needed to surmount the difficulties—project management skills, time and labor and upfront financial resources, and a Medicare FFS caseload large enough to realize sizable incentive payments. By contrast, practices that met demonstration requirements and continued to participate seemed to be those that had the wherewithal and intention to implement an EHR in the near future anyway, and the financial incentives of the EHRD motivated them to accelerate the process.

EHR Adoption and Use. Despite considerable attrition, the analysis of the 2011 OSS data found statistically and substantively significant impacts on several key health IT functions that were reported on both the 2008 demonstration application and the 2011 OSS. This analysis also found that the demonstration had a statistically and substantively significant impact on practices' OSS score as well as all five OSS domain scores. These findings suggest that the systems payment to practices did incentivize the adoption and use of EHRs during the first two years of the demonstration as treatment practices made much greater use of EHRs and of specific EHR functions than control practices did. It is notable that these systems payments, independent of the performance payments, resulted in immediate and sizeable impacts on process changes in the practices.

Generally, participating practices had much room for further improvement in EHR use and care management, validating the goals of the demonstration. The large gap between the practices' initial goals for maximizing EHR use and care management for chronically ill beneficiaries and what they reported during site visits validates the goals of the demonstration to incentivize improvements. However, based on the site visits, it also points to the likelihood that stronger financial incentives for performance (particularly related to care management), technical

assistance, and a relatively long time frame are probably all needed to achieve goals. This is likely to be particularly true for the typical physician practice that is not as motivated as practices enrolled in demonstration.

Quality of Care. Because of the early termination of the demonstration, the majority of quality measures were not collected by the evaluation team (see Chapter I and Appendix G). For the five measures that can be estimated from Medicare claims data, the analysis found no statistically significant impacts in quality measures when pooling claims data across all four demonstration sites. However, the demonstration did have a favorable effect on a common proxy for care quality, preventable hospitalizations. Specifically, the demonstration significantly reduced the likelihood that beneficiaries with CAD, CHF, or diabetes would have a preventable hospitalization during year 2. This reduction in preventable hospitalizations was driven by those treatment practices that had the greatest improvement in health IT use, strengthening the likelihood that this finding is attributed to the demonstration's incentives for health IT.

Expenditures and Service Use. The demonstration did not appear to affect total Medicare expenditures across all sites of the demonstration. Regardless of whether the system incentive payments were included in or excluded from the difference-in-differences estimates, the analysis found no statistically significant impacts of the demonstration on annualized Medicare expenditures during the demonstration's two years. Likewise, there were no statistically significant impacts of the demonstration on selected Part A and Part B expenditures, with one exception (outpatient expenditures increased by $69 during the second year of the demonstration). Finally, consistent with expenditure results, there were no statistically significant treatment control-differences in the number of inpatient hospitalizations, emergency room visits, or physician office visits among all practices in the demonstration.

## B. Interpreting the Demonstration Results Within the Context of Other Programs

The demonstration results must be interpreted not only in light of the early termination of the demonstration, but also in light of the rapid, concurrent changes in health IT policy and the incentives and resources available for assistance.

Efforts that overlapped with demonstration goals had the potential to support and encourage treatment group practices' adoption and use of EHRs, but also could have competed with demonstration activities. These efforts included those that were established under the HITECH act within ARRA. For example, beginning in 2011, eligible providers could begin receiving payments under the Medicare EHR Incentive Program for demonstrating meaningful EHR use, which included meeting a core set of required criteria and several criteria providers could choose from a menu. Although the meaningful use concept and the demonstration were aligned in their focus on EHRs being used in ways that promote quality care, and much of the Incentive Program criteria overlapped with the demonstration's criteria, there were some differences.

In the second half of 2010, when CMS announced that the Medicare Incentive Program would start in 2011, providers could have taken a wait-and-see attitude in anticipation of the potentially larger rewards from the Program (see Chapter III). After the Incentive Program started operations in January 2011, there was a four-month overlap between the demonstration and the Program, which could have resulted in demonstration practices changing their behavior despite the system incentive payments. For instance, as noted in Chapter III, a sizeable minority

of treatment group practices that responded to the OSS (over 40 percent) reported changing decisions or practice due to the Medicare and Medicaid EHR Incentive Programs by spring 2011.

Further, the Incentive Program offered a potentially more attractive payment structure—payments were offered to providers rather than to practices as a whole. In this context, the large positive impact of the system incentive payments on EHR use is remarkable because the behavior of practices in the control group did not seem to change as the result of the Program or the overlap between this program and the demonstration. However, it is unclear whether the Program would have had as much influence on EHR adoption and use in an environment unaccompanied by additional EHR-related incentives.

Beyond HITECH initiatives, state and local projects also had goals that were similar to or overlapped with those of the demonstration. Although these initiatives, summarized in Chapter I, could be interpreted as complementary efforts, those in the early stages of development seemed to contribute to a complicated environment that competed for practices' attention.

In sum, the evaluation's mixed findings (that is, the demonstration had favorable impacts on EHR use and on preventable hospitalizations even though more than two-fifths of practices were terminated) have important implications for policy. These findings could be used as a benchmark for interpreting the progress of the EHR Medicare Incentive Program, despite the fact that the demonstration was not designed to assess that Program.

## C. Limitations of the Evaluation

Although the EHRD evaluation relied on a stratified, experimental design—making it a rigorous study—it had the following limitations:

Key outcome measures construction. Treatment group practices could have overstated their health IT use because the level of the incentive payment was determined by the level of health IT use they reported in the OSS. Although an OSS validation analysis found no major discrepancies between the responses to the OSS and the validation survey, the validation survey was only a crude check on practices' reporting of their health IT use. Also, the OSS impact estimates rely on year 2 data only (because OSS data were not available for the control group for year 1), so they cannot determine whether the behavior of control group practices in year 1 was different than the behavior in year 2 and how year 1 and year 2 estimates would relate to one another.

Termination of the demonstration in year 2. Because the demonstration ended early, the follow-up period might be too short for practices to have implemented health IT changes or care management changes that would translate to quality-of-care improvements or to reductions in acute care use and costs. Indeed, the phased-in incentive structure of the original demonstration design was premised on the expectation that practice behavior change would occur over time, not all at once. The analysis includes only a limited number of quality-of-care measures because 21 of the 26 measures targeted by the demonstration were impossible to collect or construct given the circumscribed conditions of the evaluation once the demonstration was cancelled. Therefore, the evaluation cannot fully assess whether there was a relationship between health IT and quality. In addition, because of the limited availability of quality measures, the evaluation could

not extensively investigate why the demonstration affected preventable hospitalizations but did not seem to affect other quality-of-care measures.

Practice exclusion, attrition, and nonresponse. After randomization, the implementation support contractor identified seven treatment practices and one control practice that originally were classified as eligible but later were determined to be ineligible. The exclusion of these practices may have introduced selection bias to the OSS intention-to-treat impact estimates because this analysis must include all eligible, randomized practices; fortunately, the number of excluded practices was small, so this limitation is not a major concern. Because of differential response rates between the treatment and control groups in the OSS, the comparison between treatment and control group practices could also be unreliable. Nonresponse analytic weights were calculated to minimize this bias, although there is no guarantee that these adjustments controlled for all biases in the analysis.

## D. Lessons from EHRD Relevant to Other Policies and Programs

This evaluation provides some evidence about the health IT experience of a limited sample of small to medium-sized primary care practices serving Medicare fee-for-service beneficiaries. Because of the demonstration's termination, this evidence needs to be interpreted cautiously. If the demonstration had run for the original five-year term, the lessons learned from the evaluation would have been more reliable, and maybe different, than those drawn from the current analysis.

Although some policymakers may view the early impact of the system incentive payment on health IT adoption and use and on preventable hospitalizations favorably, other policymakers will be disappointed with the limited impact of the incentives on the available quality measures, Medicare expenditures, and the number of inpatient hospitalizations, emergency room visits, or physician office visits among all practices in the demonstration. The following overarching lessons—as well as unanswered questions—emerged from the EHRD:

1. Efforts with moderate incentive levels can influence use of EHRs, but cannot achieve universal adoption and use in a two-year time frame. While over half of practices responded to the financial incentives for implementing and using an EHR system, many practices were not able or willing to do so within the time frame required by the demonstration. Their decision to not respond to the incentives raises the important question of whether the incentives should have been larger, particularly larger than those offered by the Medicare EHR Incentive Program.
2. Targeting the incentives to individual practitioners instead of practices might be more effective. The site visits found considerable variation within practices in individual practitioners' use of EHRs; often decision making on EHR use was at the individual level. However, incentive payments for a practice were often not passed through to individual practitioners, but rather were used for overall support of the practice or its EHR system. Although in the Medicare EHR Incentive Program, eligible professionals who receive the incentive payment can assign it to the practice she belongs, if any, it remains untested whether payment to the practice or to the individual might be more effective.
3. Progress of EHR implementation and use is real, but it did not have strong favorable effects on the quality-of-care process measures or on expenditures. The modest

improvement in preventable hospitalizations are likely to be attributable to incentives for EHR system use. However, it is unclear whether the systems payments and the performance payments combined would have favorably affected the quality-of-care process measures, service use, or Medicare expenditures (none of which were measurably affected by the systems payment alone during the demonstration's first two years because there was no time to assess whether these payment effected reporting or performance). Findings from the MCMP demonstration suggest that the performance payments had limited favorable impacts on quality of care in selected sites.

In addition to those lessons, the following questions remain:

- What differentiates practices that are more or less doing better in adopting and using EHRs?
- How can policy tools be employed to reinforce drivers of success and shift more practices to succeed?
- Will incentives be large enough under current public- and private-sector medical homes initiatives, such as CMS's Comprehensive Primary Care (CPC) initiative or the Accountable Care Organization (ACO) initiative, to induce practices to make more transformational changes in their care management processes, including the use of EHRs for this purpose?
- Will CMS's more timely input to practices on the quality of the care they provide to patients, such as that contemplated under a harmonized Physician Quality Reporting System (PQRS)/EHR Incentive Program quality-measure reporting system, convince practices of the importance of reporting data from "meaningful use" certified products?
- Will Regional Extension Center assistance to practices frustrated with the implementation and use of EHRs become an important complement to payments to eligible providers by the EHR Incentive Program?

The findings and lessons from the evaluation of EHRD could have implications for ongoing and future federal initiatives (such as the Medicare EHR Incentive Program, CPCI, ACO, the harmonization of PQRS, and others) that use incentive payments to entice practices to meet certain requirements (for example, the "meaningful use" requirements).

One implication of lessons learned from the EHRD may be that future efforts should set realistic expectations based on these findings. That is, the findings indicate it is not realistic to expect all (or even almost all) targeted small to medium-sized practices to accomplish major changes with incentives that do not nearly cover the costs of purchase and operation of the incentivized activity. Given the difficulties practices experienced as they adopted EHRs, and the low level of care management at present, such efforts should also consider complementary technical assistance to increase the chance that quality as well as EHR use would be improved.

This page is intentionally left blank.

# REFERENCES

Dale, Stacy, Randall Blair, and Lorenzo Moreno. "Revised MDDs for EHRD after Accounting for Practice Attrition." Memorandum to CMS. Princeton, NJ: Mathematica Policy Research, June 22, 2011.

Esposito, Dominick, Randall Brown, Arnold Chen, Jennifer Schore, and Rachel Shapiro. "Impacts of a Disease Management Program for Dually Eligible Beneficiaries." Health Care Financing Review, vol. 30, no. 1, fall 2008, pp. 27-45.

Felt-Lisk, Sue, and Shinu Verghese. "EHRD Demonstration Year 2 OSS Score Data Transfer." Memorandum to CMS. Washington, DC: Mathematica Policy Research, August 31, 2011.

Felt-Lisk, Sue. "Revised OSS Scoring Plan." Memorandum submitted to CMS (EHRD-052), June 10, 2009 (revised on July 28, 2010 and April 25, 2011).

Felt-Lisk, Suzanne, Rachel Shapiro, Christopher Fleming, Brenda Natzke, Allison Barrett, and Grace Anglin. "Electronic Health Records Demonstration Evaluation: Implementation Report 2010." Princeton, NJ: Mathematica Policy Research, January 20, 2011.

Felt-Lisk, Suzanne, Lorenzo Moreno, Rachel Shapiro, Audra Wenzlow, Martha Kovac, and Jennifer Schore. "Evaluation of the Electronic Health Records Demonstration: Evaluation Design." Princeton, NJ: Mathematica Policy Research, December 23, 2008.

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 to Centers for Medicare & Medicaid Services. Princeton, NJ: Mathematica Policy Research, May 25, 2007

Trisolini, Michael, Musetta Y. Leung, Sherry Grund, Mary Schrader, and Francis Landiza. Medicare Care Management Performance Demonstration Quality Measurement and Reporting Specifications. Waltham, MA: RTI International, August 3, 2007.

Wilkin, John C., Kerry E. Moroz, Erika G. Yoshino, and Laurie E. Pekala. "Electronic Health Records Demonstration Waiver Cost Estimate." Columbia, MD: Actuarial Research Corporation, December 13, 2007.

This page is intentionally left blank.

# APPENDIX A

# OSS DESIGN AND DATA COLLECTION

This page is intentionally left blank.

This appendix provides details on the design and conduct of the annual Office Systems Survey (OSS). It covers the instrument design and pretest, sample design and release, the survey fielding process, final response rates, and survey weighting and nonresponse.

## A. Instrument Design and Pretest

The OSS measured practices' use of electronic health record (EHR) systems and their specific functions, such as prescribing medications, ordering laboratory and radiology tests, and care management and coordination. Treatment group practices were required to complete an annual OSS for each demonstration year. Control group practices were asked to complete the survey at the end of the second and fifth demonstration years, although the demonstration was terminated after the first round of data collection for the control group. A validation of 25 percent of the responding practices was conducted to validate practice's responses to the OSS each year (see Appendix H for details on the validation survey and findings).

Treatment group practices were required to participate in the OSS in order to receive financial systems payments. For this reason, treatment practices had a strong motivation to participate in the OSS and a survey incentive payment was not needed to ensure a high response rate. Control group practices, on the other hand, received no demonstration payment for adoption and use of an EHR, so they had no clear incentive to complete the OSS. Therefore a $50 incentive was offered to control group practices to ensure a comparably high response rate to the survey.

The OSS was designed as a web-based survey instrument. The OSS instrument drew heavily upon a similar OSS instrument that was designed and administered by another organization in 2007 under the Medicare Care Management Performance (MCMP) demonstration evaluation. However questions were added to capture EHR functions and use in greater detail. The OSS contained six main sections:

- Section 1: General Practice Information. This section collects information on the enrolled practices, including name, address, telephone number, affiliation with an independent practice association (IPA) or physician hospital organization (PHO), and participation in any other quality reporting or improvement initiatives.
- Section 2: Provider Profile. This section asks about the providers enrolled in the demonstration, such as their name, specialty, credentials, languages spoken, primary practice location, and provider and Medicare billing numbers (provider identification number).
- Section 3: Use/Planned Use of EHRs, E-Patient Registry, or E-Prescribing Systems. This section collects information about the various types of electronic systems (EHRs, patient registries, and prescribing) currently in use or planned for use in the practice, and the number of providers who use these systems.
- Section 4: EHR, Patient Registry, and Prescribing System Functions. This section collects information about the various functions that practices use for each of the systems identified in Section 3, and the proportion of patients for which they use each function. Functions are organized under five domains: (1) completeness of information, (2) communication of care outside the practice, (3) clinical decision

support, (4) use of the system to increase patient engagement/adherence, and (5) medication safety.

- Section 5: Receipt of ARRA funding. In the second year of the OSS (2011), a section was included to gather information about practices' receipt of ARRA funding and how it may have affected their adoption and use of EHRs.
- Section 6: Data Attestation. This section requests that the respondent confirm that the responses are a correct assessment of the practice, that the survey responses are accurate, and that they may be subject to validation.

The questionnaire was pretested with eight practices (as per OMB requirements). Interview completion times ranged from 60 to 105 minutes, with an average length of 83 minutes. Practices were mailed a paper-and-pencil questionnaire and instructed to fill it out, keep track of how long it took to complete, and fax it back to Mathematica. A telephone debriefing was conducted with practice respondents to assess their cognitive understanding of key terms and to identify any difficulties in answering questions or navigating the instrument. Respondents were also asked whether they needed help from other practice staff to answer any questions and the level of effort it took those additional staff to assist in answering questions. (A copy of the instrument is included at the end of this appendix.)

## B. Sample Design and Release

A total of 825 practices voluntarily enrolled in the demonstration across the four CMS selected sites: Louisiana; Maryland/Washington, DC; Pennsylvania; and South Dakota (412 treatment group practices and 413 control group practices). All were small to medium-sized practices (20 or fewer physicians, although there were exceptions) that provide primary care to at least 50 fee-for-service Medicare beneficiaries with congestive heart failure, coronary artery disease, diabetes, or other chronic diseases. Seven treatment practices and one control practice failed to meet the terms and conditions of the demonstration and were determined ineligible by CMS prior to the start of the demonstration. This resulted in a total of 405 eligible treatment and 412 eligible control practices at the start of the demonstration.

The OSS was administered to treatment practices in years 1 and 2 of the demonstration (2010 and 2011), and to control group practices in year 2 of the demonstration. In year 1, a total of 405 treatment practices were released and attempted for surveying. In year 2, only 362 treatment group and 392 control group practices were released and attempted for surveying. This was because, by year 2 of the OSS fielding, 43 treatment practices and 20 control practices had closed, merged with another practice, actively refused to participate in the year 1 OSS and requested not to be contacted again, or were ineligible because they failed to meet the terms and conditions of the demonstration.

## C. Survey Fielding and Response Rates

Each yearly round of the OSS was administered over a 9- to 10-week period beginning in April and ending in June.[1] Multiple attempts were made to encourage practices to complete the web-based survey. A personalized advance letter was mailed to practices just prior to the start of data collection. The letter was printed on CMS letterhead and signed by the CMS Privacy Officer. It described the survey; provided the survey web address, a secure login, and personal identification number to access the web-based survey; and was accompanied by a fact sheet with answers to commonly asked questions about the study and the survey. (Copies of the letters and fact sheets are included at the end of this appendix.) The letter also provided a toll-free number to call if a practice had a question or preferred to complete a paper-and-pencil questionnaire. In addition, a toll-free help desk and general email address were established to assist practices in completing the online survey.

An initial email was then sent to all practices at the start of the survey fielding period. The email contained the same information as the advance letter, including a hyperlink to the website and the practice's secure login and password. Multiple follow-up emails and letters were sent to nonresponsive practices encouraging them to respond at roughly 10- to 14-day intervals throughout the fielding period. Email reminders encouraging response increased during the final weeks of the fielding period at roughly five- to seven-day intervals. Practices that requested a paper-and-pencil questionnaire were mailed one along with a personalized letter, fact sheet, and postage-paid return envelope.

In 2010, a total of 405 treatment practices were sent letters and emails requesting their participation, including 367 practices still participating in the demonstration and 38 practices that had voluntarily withdrawn from the demonstration. A total of 355 treatment practices responded to the OSS and 353 completed it (the other two only partially completed it). An unweighted final response rate of 87 percent was achieved (92 percent among participating demonstration practices and 37 percent among withdrawn practices).

In 2011, a total of 754 practices (362 treatment and 392 control) were sent letters and emails requesting their participation, including 348 treatment practices still participating in the demonstration and 14 treatment practices that had voluntarily withdrawn from the demonstration. A total of 316 treatment practices responded to the OSS and 314 completed it (the other two only partially completed it). A total of 270 control practices responded and 267 completed it (the other three only partially completed it). An unweighted final response rate of 87 percent among those that were attempted was achieved for treatment practices (90 percent among participating treatment practices, 38 percent among withdrawn treatment practices), and 68 percent was achieved for control group practices. The response rate for eligible treatment and control practices combined, including 60 that closed or terminated, was 71 percent.

[1] The year 1 OSS was fielded over a 9-week period, from April 18 until June 21, 2010. The year 2 OSS was fielded over a 10-week period, from April 1 until June 10, 2011.

Table A.1 presents the number of completed OSS surveys and the unweighted response rate among those that were attempted for the demonstration by survey fielding year, research status, and site. Table A.2 presents the final disposition of eligible practices for the demonstration by survey fielding year and research status.

Table A.1. Number of Practices That Completed the OSS and Unweighted Response Rates Among Those Attempted by Survey Fielding Year, Research Status, and Site

| Site | 2010 Treatment Group Respondents | 2010 Control Group Respondents | 2011 Treatment Group Respondents | 2011 Control Group Respondents |
|---|---|---|---|---|
| Louisiana | 79 (77)[a] | NA | 69 (66) | 49 (49) |
| Maryland/DC | 111 (87) | NA | 98 (77) | 74 (58) |
| Pennsylvania | 125 (90) | NA | 114 (83) | 108 (76) |
| South Dakota | 38 (100) | NA | 33 (77) | 36 (82) |
| Total | 353 (87) | NA | 314 (87) | 267 (68) |
| Sample Size[b] | 405 | NA | 362 | 392 |

Source: Office Systems Survey (OSS), conducted in spring and summer 2010 and 2011.

[a]In parentheses, the response rate among attempted cases.
[b]Excluding practices that closed, merged, or were terminated by CMS.

Table A.2 presents the final disposition of eligible practices for the demonstration by survey fielding year and research status.

Table A.2 Final Sample Disposition of Eligible Practices by Survey Fielding Year and Research Status

| Final Sample Disposition | 2010 Treatment Group Respondents | 2010 Control Group Respondents | 2011 Treatment Group Respondents | 2011 Control Group Respondents |
|---|---|---|---|---|
| Participating Practices | | | | |
| Complete | 339 | NA | 314 | 267 |
| Partial Complete | 2 | NA | 2 | 3 |
| Refusal/No Response | 19 | NA | 26 | 112 |
| Merged | 2 | NA | 3 | 0 |
| Closed | 5 | NA | 4 | 10 |
| Withdrawn Practices | | | | |
| Complete | 14 | NA | 5 | NA |
| Partial | 0 | NA | 3 | NA |
| Refusal/No Response | 24 | NA | 5 | NA |
| Total Eligible Sample[a] | 405 | NA | 362 | 392 |

Source: Office Systems Survey, conducted in spring and summer 2010 and 2011.

[a]Excluding practices that closed, merged or were terminated by CMS prior to the survey fielding period.

## D. Survey Weighting and Nonresponse Adjustment

The target population for the study is all eligible practices assigned to either treatment or control groups. Because the OSS was intended to be delivered to all practices in the target population, no adjustment was needed for different probabilities of sample selection. There were three main reasons that practices did not complete the OSS: (1) they withdrew from the study, (2) they closed or merged with another practice before the OSS was fielded, or (3) they were approached but did not respond to the OSS. Because reasons for dropping out or not responding were found to be related to treatment status and other practice characteristics, the responding practices are no longer representative of the target population. Thus, weights were constructed to control as much as possible for the resulting bias. Weights were constructed in several steps, after excluding the eight practices (seven treatment practices and one control practice) that were determined to be ineligible due to not meeting the study requirements.

First, weights were adjusted to account for mergers so that the practice remaining after the merge represented both practices initially in the sample. There were three such practices, all in the treatment group, that merged with other treatment group practices.

Second, the propensity for a practice to drop out of the study (because they closed or otherwise terminated participation in the study) was modeled using logistic regression, and the weight of each practice was multiplied by the inverse predicted propensity to remain in the study. We used a stepwise procedure to determine the logistic regression model, which considered including variables for randomization group, state, whether the practice had an EHR at baseline, the size of the practice, size squared, an indicator of practice size, whether the practice was in an urban area, whether the practice was in a medically underserved area, and the interaction of randomization group with state, EHR at baseline, and the indicator of size. The final model included the following predictive variables: randomization group, state, EHR at baseline, indicator of practice size, and the interaction of randomization group with state. This adjusts the weights for the attrition of 60 practices (20 control practices and 40 treatment practices), for which the year 2 OSS was not attempted.

Third, the propensity of practices to respond to the OSS, if approached, was modeled using a separate logistic regression, and the weights of practices were multiplied by the inverse response propensity in order to adjust for nonresponse. Once again, we used a stepwise procedure to determine the logistic regression model using the same set of practice characteristics as described above for the attrition model. The final model included the following predictive variables: randomization group, whether the practice had an EHR at baseline, the indicator of size, state, and the interaction of randomization group with EHR at baseline. Separate adjustments were needed for attrition and nonresponse because the relationships between prediction variables and the two types of non-completes were different.

For the 581 practices that completed the year 2 OSS, we then multiplied the three weighting factors to adjust for merging, attrition, and nonresponse. The final weighting steps consisted of a ratio adjustment to ensure the sum of weights within each randomization group was equal to the number of eligible practices initially assigned (405 treatment and 412 control). Large (outlier) weights were trimmed, and a final ratio adjustment was made.

This page is intentionally left blank.

# APPENDIX B

# SUPPLEMENTAL TABLES TO CHAPTER III

This page is intentionally left blank.

# TABLES

This page is intentionally left blank.

# FIGURES

This page is intentionally left blank.

Table B.1. Adoption of EHR Systems During Fall 2008 – Spring 2011 by Practice Characteristics (Percentages, Unless Otherwise Noted)

| | (B) | (C) | (C1) | (D) | (E) | (F) | (G) | (G1) | (H) | (I) |
|---|---|---|---|---|---|---|---|---|---|---|
| | Number of TG Practices with Information on EHR Use in 2008 and 2011 | Percentage of Column B TG Practices with No EHR Use in Fall 2008 | Number of Column B TG Practices with No EHR Use in Fall 2008 | Percentage of Column B TG Practices Beginning EHR Use Between Fall 2008 and Spring 2011 | Number of CG Practices (Eligible and Randomized) | Number of CG Practices with Information on EHR Use in 2008 and 2011 | Percentage of CG Practices with No EHR Use in Fall 2008 | Number of CG Practices with No EHR Use in Fall 2008 | Percentage of CG Practices Beginning EHR Use Between Fall 2008 and Spring 2011 | Difference Between TG and CG in Percentage Newly Using EHRs (D–H) |
| All | 321 | 51.4 | 165 | 36.8 | 263 | 263 | 51.3 | 135 | 24.0 | 12.8 |
| Site | | | | | | | | | | |
| Louisiana | 70 | 55.7 | 39 | 28.6 | 49 | 49 | 53.1 | 26 | 14.3 | 14.3 |
| Maryland | 100 | 44.0 | 44 | 34.0 | 71 | 71 | 36.6 | 26 | 16.9 | 17.1 |
| Pennsylvania | 116 | 52.6 | 61 | 42.2 | 109 | 109 | 56.0 | 61 | 34.9 | 7.4 |
| South Dakota | 35 | 60.0 | 21 | 42.9 | 34 | 34 | 64.7 | 22 | 17.6 | 25.2 |
| Practice Size (total number of providers) | | | | | | | | | | |
| 1-2 | 146 | 58.2 | 85 | 39.7 | 124 | 124 | 58.1 | 72 | 26.6 | 13.1 |
| 3-5 | 115 | 47.8 | 55 | 37.4 | 91 | 91 | 49.5 | 45 | 24.2 | 13.2 |
| 6+ | 60 | 41.7 | 25 | 28.3 | 48 | 48 | 37.5 | 18 | 16.7 | 11.7 |
| Practice Affiliation | | | | | | | | | | |
| Unaffiliated | 263 | 48.3 | 127 | 32.3 | 218 | 218 | 52.3 | 114 | 22.5 | 9.8 |
| Affiliated[a] | 58 | 65.5 | 38 | 56.9 | 45 | 45 | 46.7 | 21 | 31.1 | 25.8 |
| Located in a Rural Area | | | | | | | | | | |
| Yes | 53 | 60.4 | 32 | 49.1 | 40 | 40 | 67.5 | 27 | 32.5 | 16.6 |
| No | 268 | 49.6 | 133 | 34.3 | 223 | 223 | 48.4 | 108 | 22.4 | 11.9 |
| Located in an MUA | | | | | | | | | | |
| Yes | 89 | 56.2 | 50 | 40.4 | 77 | 77 | 40.3 | 31 | 19.5 | 21.0 |
| No | 232 | 49.6 | 115 | 35.3 | 186 | 186 | 55.9 | 104 | 25.8 | 9.5 |
| Participating in Another EHR, Quality Improvement, or Quality Reporting Program | | | | | | | | | | |
| Yes | 307 | 50.5 | 155 | 36.8 | N/A | N/A | N/A | N/A | N/A | N/A |
| No | 14 | 71.4 | 10 | 35.7 | N/A | N/A | N/A | N/A | N/A | N/A |

Source: Office Systems Survey conducted in spring and summer 2011 and demonstration application data, fall 2008.

[a]Owned by a hospital, hospital system, or larger medical group, or affiliated with a larger medical group, IPA, PHO, or other entity. Most of these practices are owned by a larger organization (see Felt-Lisk et al. 2011).

CG = control group; EHR = electronic health record; MUA = medically underserved area; TG = treatment group.

Table B.2. Progress in Health IT Use – Percentage of Treatment and Control Group Practices Using Selected Functions Queried at Time of Application to Participate in the Demonstration (Percentages)

| EHR/Health IT Function | Treatment – Fall 2008 (Application) | Treatment – Spring 2010 | Treatment – Spring 2011 | Control – Fall 2008 (Application) | Control – Spring 2011[b] |
|---|---|---|---|---|---|
| Number of practices responding to 2011 OSS[a] | 324 | 324 | 324 | 268 | 268 |
| Number of practices with EHR use data | 201 | 263 | 290 | 172 | 214 |
| Any EHR/health IT use | 48.5 | 81.2 | 89.5 | 48.1 | 79.5 |
| Electronic patient visit notes | 46.6 | 64.8 | 82.4 | 48.1 | 67.9 |
| Electronic patient problem lists [MU-C] | 46.3 | 66.4 | 83.3 | 47.4 | 69.4 |
| Automated patient-specific alerts and reminders | 37.0 | 46.0 | 60.8 | 32.8 | 46.3 |
| Electronic disease-specific patient registries | 15.4 | 50.3 | 68.5 | 20.9 | 53.7 |
| Patient email | 9.3 | 19.4 | 29.3 | 7.1 | 30.6 |
| Patient-specific educational materials [MU-M] | 37.3 | 36.1 | 55.9 | 36.9 | 43.7 |
| Laboratory tests: online order entry | 29.9 | 25.6 | 34.0 | 35.1 | 37.3 |
| Radiology tests: online order entry | 16.7 | 13.6 | 19.1 | 17.2 | 22.8 |
| Printing and/or faxing Rx | 54.6 | 71.0 | 81.5 | 51.9 | 67.9 |
| Online Rx transmission to pharmacy [MU-C] | 34.6 | 63.9 | 85.8 | 26.9 | 70.9 |

Source: Office Systems Survey (OSS) conducted in spring and summer 2011, and demonstration application data, fall 2008.

[a]Excludes practices that did not respond to the 2011 OSS as well as those that closed or were terminated by CMS.

[b]Control group was not surveyed in spring 2010.

EHR = electronic health record; IT = information technology; MU-C=Stage 1 meaningful use core set item.

MU-M=Stage 1 meaningful use menu set item.

Table B.3. Changes in EHR Use Among Users of EHRs During 2010–2011, by Practice Characteristics

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | Number of TG Practices with OSS Responses Indicating EHR Use in Both 2010 and 2011 | TG - Mean Total OSS Score Among 2011 Users | TG - Mean Change in Total OSS Score 2010–2011 (EHR Users, Both Periods) | TG - Percentage of 2010 Users that Improved OSS Score in 2011 (EHR Users, Both Periods) | Number of CG Practices with OSS Responses Indicating EHR Use in 2011 | CG-Mean Total OSS Score Among 2011 Users | Difference TG – CG Total OSS Scores in 2011 (Column 2 - Column 5) |
| All | 220 | 65.4 | 9.0 | 77.2 | 188 | 59.9 | 5.5 |
| Site | | | | | | | |
| Louisiana | 41 | 59.4 | 7.2 | 58.5 | 30 | 55.6 | 3.8 |
| Maryland | 74 | 68.9 | 13.1 | 93.2 | 54 | 58.6 | 10.3 |
| Pennsylvania | 82 | 64.2 | 6.0 | 65.9 | 86 | 61.5 | 2.9 |
| South Dakota | 23 | 68.9 | 9.7 | 100.0 | 18 | 63.3 | 5.6 |
| Number of Physicians | | | | | | | |
| 1-2 | 51 | 63.0 | 8.3 | 70.3 | 51 | 56.5 | 6.6 |
| 3-5 | 73 | 63.1 | 9.2 | 80.0 | 55 | 57.2 | 5.9 |
| 6+ | 85 | 68.5 | 9.8 | 85.4 | 70 | 63.5 | 5.0 |
| Owned/Affiliated with Larger Organization | | | | | | | |
| Yes | 41 | 70.6 | 8.9 | 78.0 | 38 | 70.4 | 0.2 |
| No | 179 | 64.2 | 9.0 | 77.0 | 150 | 57.2 | 7.0 |
| Rural | | | | | | | |
| Yes | 35 | 65.6 | 12.1 | 82.9 | 26 | 54.7 | 10.8 |
| No | 185 | 65.4 | 8.4 | 76.1 | 162 | 60.7 | 4.7 |
| Participation in Another Quality or EHR Initiative (2010) | | | | | | | |
| Yes | 217 | 65.3 | 8.8 | 76.9 | N/A | N/A | N/A |
| No | 3 | 71.3 | 19.0 | 100.0 | N/A | N/A | N/A |

Source: Office Systems Survey (OSS) conducted in spring and summer 2011.

Note: Table is restricted to treatment group practices not closed or terminated by CMS that indicated they used an EHR in both 2010 and 2011 (treatment group), or 2010 (control group, as they were not surveyed in 2010). The 34 practices with stand-alone registries and/or stand-alone e-prescribing systems were excluded from this table, as were practices that failed to respond to the question directly asking if they used an EHR. Three practices that responded to both surveys discontinued EHR use between 2010 and 2011. They are excluded from the table because their OSS score was not calculated in 2011.

CG = control group; EHR = electronic health record; TG = treatment group.

Table B.4. Mean OSS Scores for Each Domain, by Practice Characteristics, for Treatment and Control Group EHR Users, 2011

| | TG – Total OSS Score | CG – Total OSS Score |
|---|---|---|
| All | 63.5 | 59.9 |
| Site | | |
| Louisiana | 58.8 | 55.6 |
| Maryland | 64.9 | 58.6 |
| Pennsylvania | 63.5 | 61.5 |
| South Dakota | 67.0 | 63.3 |
| Number of Physicians | | |
| 1-2 | 61.4 | 55.7 |
| 3-5 | 62.9 | 64.1 |
| 6+ | 69.0 | 61.6 |
| Owned/Affiliated with Larger Organization | | |
| Yes | 68.1 | 70.4 |
| No | 62.3 | 57.2 |
| Rural | | |
| Yes | 63.2 | 60.7 |
| No | 64.7 | 54.7 |
| Participation in Another Quality or EHR Initiative (2010) | | |
| Yes | 63.3 | N/A |
| No | 69.7 | N/A |

Source: Office Systems Survey (OSS) conducted in spring and summer 2011.

Note: The table includes the 268 treatment group and 188 control group practices who responded to the 2011 OSS and indicated that they used an EHR.

CG = control group; EHR = electronic health record; TG = treatment group.

Table B.5. Changes in Use of Functions Incentivized by EHRD Only vs. EHRD and Medicare and Medicaid EHR Incentive Program (Treatment Group Practices with OSS Responses in Both 2010 and 2011)

| | Mean Item Score, 2011 (0-4, 4 Is Use for 3/4 or more Patients)[a] | Change in Mean from 2010 | Number of Responses |
|---|---|---|---|
| EHR Functions Incentivized in EHRD AND: | | | |
| Core Set Stage 1 Medicare & Medicaid EHR Incentive Program | 2.84 | 0.24 | 286 |
| Menu Set Stage 1 Medicare & Medicaid EHR Incentive Program | 1.50 | 0.34 | 286 |
| Likely related to Core Set requirement to implement a clinical decision support rule | 2.50 | 0.50 | 285 |
| EHR Functions Incentivized in EHRD but Not Incentivized by Stage 1 Medicare & Medicaid EHR Incentive Program | 1.97 | 0.39 | 286 |

Source: Office Systems Survey (OSS) conducted in spring and summer 2011.

Note: EHRD items queried on the OSS were often not exactly the same as the Medicare and Medicaid EHR Incentive Program measures; some judgment was required to classify the items most related. More specifics about the comparability between specific measures were provided in Appendix B of Felt-Lisk et al. 2010. Details on the trends in specific functions and which functions were classified into which rows are provided in Appendix B, Table B.6.

[a]The mean item score does not represent an exact measurement of the extent of use, as it is the average response regarding the range of proportions of patients for whom the function was used. Specifically, values of 0 to 4 for each item represent the proportion of the practice's patients for whom the EHR function was used over the past month: 0 = None; 1 = Some but less than 1/4; 2 = 1/4 or more, but less than 1/2; 3 = 1/2 or more but less than 3/4; and 4 = 3/4 or more.

EHR = electronic health record.

Table B.6. Trend in Treatment Group Practices' Responses to Function-Specific Questions (Percentages, Unless Otherwise Noted)

Domain 1. Completeness of Information

| Functions | Percentage of Practices with Some Records Transitioned/ Some Paper Charts Pulled, 2011 | Percentage Beginning Transition/ Beginning to Stop Paper Charts During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Applies for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| 4.1a. Paper records that have been transitioned to the EHR system. By "transitioned" we mean either scanned documents in full into the EHR or keyed in data items by hand (such as patient demographics, medical history, blood pressure readings, test results) | 90.9 | 21.2 | 3.1 | 31.2 | 286 |
| 4.1b. Paper charts that were pulled for scheduled patient visits over the past month | 58.0 | 16.4 | 1.4 | -26.6 | 286 |

| | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| 4.1d. Clinical notes for individual patients<br>Refers to using the electronic system to create, update, store and display clinical notes. | 93.4 | 20.2 | 3.5 | 31.3 | 286 |
| 4.1e Allergy lists for individual patients<br>Refers to using the electronic system to create, update, store and display a list of medications or other agents (food, environmental) to which patient has a known allergy or adverse reaction.MU-C | 96.2 | 17.8 | 3.6 | 21.5 | 286 |
| 4.1f. Problem or diagnosis lists for individual patients<br>Refers to using the electronic system to create, update, store and display a list of problems or diagnoses for a patient.MU-C | 94.4 | 19.9 | 3.6 | 36.3 | 286 |

| | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| 4.1g. Patient demographics (for example, age or sex)<br>Methods of entry include direct keyboard entry (typing); entering notes/data using templates, forms or drop-down menus; or dictation with the voice transcribed manually or via voice recognition into text that is later integrated into the system. MU-C | 96.2 | 18.5 | 3.7 | 20.2 | 286 |
| 4.1h. Patient medical histories | 93.7 | 19.2 | 3.5 | 33.9 | 286 |
| 4.1i. Recording (or entering) laboratory orders into electronic system<br>Methods of entry include direct keyboard entry (typing); entering notes/data using templates, forms or drop-down menus; or dictation with the voice transcribed manually or via voice recognition into text that is later integrated into the system.<br>Includes orders for lab tests conducted by external providers and the practice itself. | 90.2 | 25.0 | 3.4 | 65.5 | 286 |
| 4.1j. Receiving laboratory results by fax or mail and scanning paper versions into electronic system<br>Refers to converting the image or text from paper into a digital image or text that is saved in the electronic system.<br>Includes results from lab tests conducted by external providers and the practice itself. | 88.5 | 25.3 | 2.5 | 42.1 | 286 |
| 4.1k. Reviewing laboratory test results electronically<br>Refers to (1) system tracking that results have been received and (2) physician examining screens with displays of results stored in the system. | 87.7 | 23.6 | 3.1 | 27.9 | 285 |
| 4.1l. Recording (or entering) imaging orders into electronic system<br>Methods of entry include direct keyboard entry (typing); entering notes/data using templates, forms or drop-down menus; or dictation with the voice transcribed manually or via voice recognition into text that is later integrated into the system.<br>Includes orders for imaging conducted by external providers and the practice itself. | 86.4 | 26.0 | 3.1 | 70.7 | 286 |

| | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| 4.1m. Receiving imaging results by fax or mail and scanning paper versions into electronic system<br>Refers to converting the image or text from paper into a digital image or text that is saved in the electronic system.<br>Includes results from imaging conducted by external providers and the practice itself. | 86.7 | 24.0 | 2.6 | 27.7 | 286 |
| 4.1n. Reviewing imaging results electronically<br>Refers to (1) system tracking that results have been received and (2) physician examining screens with displays of results stored in the system. | 82.2 | 31.8 | 2.7 | 57.1 | 286 |
| 4.1o1. Recording that instructions or educational information were given to diabetes patients | 98.5 | 30.0 | 3.1 | 24.9 | 200 |
| 4.1o2. Recording that instructions or educational information were given to coronary artery disease patients | 98.7 | 34.4 | 2.9 | 29.4 | 152 |
| 4.1o3. Recording that instructions or educational information were given to congestive heart failure patients | 96.5 | 36.1 | 2.8 | 19.0 | 144 |
| 4.1o4. Recording that instructions or educational information were given to preventive care patients | 96.8 | 29.8 | 3.1 | 21.0 | 190 |
| 4.1p. Recording (or entering) prescription medications (new prescriptions and refills) into electronic system<br>Methods of entry include direct keyboard entry<br>(typing); entering notes/data using templates, forms or drop-down menus; or dictation with the voice transcribed manually or via voice recognition into text that is later integrated into the system. | 98.6 | 16.1 | 3.8 | 15.1 | 286 |

Domain 2. Communication of Care Outside the Practice

| Functions | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| Laboratory Orders | | | | | |
| 4.2a. Print and fax laboratory orders to facilities outside the practice<br>Order is first printed and then sent over a telephone line using a stand-alone fax machine. | 68.9 | 28.4 | 1.4 | 47.1 | 286 |
| 4.2b. Fax laboratory orders electronically from system, or order electronically through a portal maintained by facilities outside the practice<br>Order is generated electronically, using a macro or template, and faxed directly through the electronic system to the laboratory or ordered directly without using any paper or a stand-alone fax machine. | 39.9 | 19.2 | 0.9 | 23.1 | 286 |
| 4.2c. Transmit laboratory orders electronically directly from system to facilities outside the practice that have the capability to receive such transmissions MU-C<br>Order is sent as machine-readable data. | 38.5 | 15.8 | 1.2 | 12.3 | 286 |
| Imaging Orders | | | | | |
| 4.2d. Print and fax imaging orders to facilities outside the practice<br>Order is first printed and then sent over a telephone line using a stand-alone fax machine. | 77.6 | 27.1 | 1.8 | 40.9 | 286 |
| 4.2e. Fax imaging orders electronically from system, or order electronically through a portal maintained by facilities outside the practice<br>Order is generated electronically, using a macro or template, and faxed directly through the electronic system to the imaging facility without using any paper or a stand-alone fax machine. | 30.4 | 17.8 | 0.7 | 16.7 | 286 |
| 4.2f. Transmit imaging orders electronically directly from system to facilities outside the practice that have the capability to receive such transmissions MU-C<br>Order is sent as machine-readable data. | 21.7 | 12.3 | 0.6 | 16.6 | 286 |

| Functions | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| Laboratory Results | | | | | |
| 4.2g. Transfer electronic laboratory results (received in non-machine readable form, such as an e-fax) directly into system<br>Refers to saving or attaching an electronic submission, such as an e-fax, that is not electronically searchable in the EHR system. (An e-fax is a transmission of the image of a document directly from a computer or multi-purpose printer without the use of stand-alone fax equipment to generate the paper-based image.) | 29.4 | 14.4 | 0.6 | 9.7 | 286 |
| 4.2h. Enter laboratory results manually into electronic system in a searchable field (whether received by fax, mail or phone)<br>Methods of entry include direct keyboard entry (typing); entering notes/data using templates, forms or drop-down menus; or dictation with the voice transcribed manually or via voice recognition into text that is later integrated into the electronic system and is searchable.MU-M | 68.9 | 27.4 | 1.2 | 36.9 | 286 |
| 4.2i. Receive electronically transmitted laboratory results directly into system from facilities that have the capability to send such transmissions<br>Results are received electronically and do not need to be manually uploaded or posted into the system. MU-M | 75.9 | 21.2 | 2.7 | 33.1 | 286 |

| Functions | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| Imaging Results | | | | | |
| 4.2j. Transfer electronic imaging results (received in non-machine readable form, such as an e-fax) directly into system<br>Refers to saving or attaching an electronic submission, such as an e-fax, that is not electronically searchable into the EHR system. (An e-fax is a transmission of the image of a document directly from a computer or multi-purpose printer without the use of stand-alone fax equipment to generate the paper-based image.) | 29.0 | 15.4 | 0.7 | 9.5 | 286 |
| 4.2k. Enter imaging results manually into electronic system in a searchable field (whether received by fax, mail or phone)<br>Methods of entry include direct keyboard entry (typing); entering notes/data using templates, forms or drop-down menus; or dictation with the voice transcribed manually or via voice recognition into text that is later integrated into the electronic system and is searchable. | 58.0 | 27.1 | 1.4 | 41.5 | 286 |
| 4.2l. Receive electronically transmitted imaging results directly into system from facilities that have the capability to send such transmissions<br>Results are received electronically and do not need to be manually uploaded or posted into the system.<br>MU-C | 51.0 | 17.8 | 1.6 | 26.6 | 286 |

| Functions | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| Referral and Consultation Requests | | | | | |
| 4.2m. Enter requests for referrals to or consultation with other providers (for example, specialists, sub-specialists, physical therapy, speech therapy, nutritionists)<br>Refers to recording physician or patient requests for referral/ consultation, scheduling the referral/consultation, and tracking results of referral/ consultation. | 77.9 | 24.3 | 2.6 | 65.9 | 285 |
| Sharing Information with Other Providers | | | | | |
| 4.2n. Transmit medication lists or other medical information to other providers (for example, hospitals, home health agencies, or other physicians) MU-C | 65.4 | 20.5 | 1.9 | 10.6 | 286 |
| 4.2o. Transmit laboratory results to other providers (for example, hospitals, home health agencies, or other physicians)<br>Results are sent as machine-readable data. | 47.9 | 19.5 | 1.4 | 16.3 | 286 |
| 4.2p. Transmit imaging results to other providers (for example, hospitals, home health agencies, or other physicians)<br>Results are sent as machine-readable data. | 44.1 | 18.2 | 1.3 | 20.2 | 286 |
| 4.2q. Receive electronically transmitted reports directly into system, such as discharge summaries, from hospitals or other facilities that have the capability to send such transmissions | 52.1 | 23.6 | 1.7 | 48.1 | 286 |

| Functions | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| Prescription Orders | | | | | |
| 4.2r. Print prescriptions (new prescriptions and refills) on a computer printer and fax to pharmacy or hand to patient | 83.9 | 17.8 | 1.1 | -6.4 | 286 |
| 4.2s. Fax prescription orders (new prescriptions and refills) electronically from electronic system<br>The prescription is faxed without using any paper or a stand-alone fax machine. MU-C | 73.4 | 29.1 | 1.0 | -26.2 | 286 |
| 4.2t. Transmit prescription orders (new prescriptions and refills) electronically directly from system to pharmacies that have the capability to receive such transmissions MU-C<br>The prescription is sent and received without relying on a stand-alone fax machine at either the provider's office or the pharmacy. | 97.2 | 25.7 | 3.4 | 81.0 | 286 |

Domain 3. Clinical Decision Support

| Functions | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| 4.3a. Enter information from clinical notes into documentation templates<br>Documentation templates are preset formats that determine what information will be displayed on each page and how it will be displayed. Templates usually allow information to be displayed as discrete data elements (that is, each element of data is stored in its own field or box). For example, the clinical notes page can have separate boxes for entry of notes or data about a patient's height, weight, blood pressure, or other vital signs.<br>Methods of entry include direct keyboard entry (typing); entering notes/data using templates, forms or drop-down menus; or dictation with the voice transcribed manually or via voice recognition into text that is later integrated into the system. | 88.4 | 20.6 | 3.1 | 27.9 | 285 |
| 4.3b. View graphs of patient height or weight data over time MU-C | 82.8 | 22.0 | 2.8 | 59.5 | 285 |
| 4.3c. View graphs of patient vital signs data over time (such as blood pressure or heart rate) MU-C | 82.8 | 21.6 | 2.8 | 47.1 | 285 |
| 4.3d. Flag incomplete or overdue test results MU-CDS | 70.9 | 21.6 | 2.1 | 48.4 | 285 |
| 4.3e. Highlight out of range test levels MU-CDS<br>Refers to system comparing test results with guidelines or provider-determined goals for this patient | 75.8 | 21.3 | 2.8 | 34.4 | 285 |
| 4.3f. View graphs of laboratory or other test results over time for individual patients | 69.8 | 23.4 | 2.2 | 52.0 | 285 |
| 4.3g. Prompt clinicians to order necessary tests, studies, or other services MU-CDS | 70.2 | 22.7 | 2.2 | 38.3 | 285 |
| 4.3h1. Review and act on reminders at the time of a patient encounter regarding interventions, screening, or follow-up office visits recommended by evidence-based practice guidelines for diabetes patients MU-CDS | 93.0 | 27.2 | 3.2 | 23.8 | 200 |

| Functions | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| 4.3h2. Review and act on reminders at the time of a patient encounter regarding interventions, screening, or follow-up office visits recommended by evidence-based practice guidelines for coronary artery disease patients MU-CDS | 96.0 | 32.1 | 3.3 | 70.4 | 151 |
| 4.3h3. Review and act on reminders at the time of a patient encounter regarding interventions, screening, or follow-up office visits recommended by evidence-based practice guidelines for congestive heart failure patients MU-CDS | 95.8 | 36.7 | 3.2 | 73.6 | 144 |
| 4.3h4. Review and act on reminders at the time of a patient encounter regarding interventions, screening, or follow-up office visits recommended by evidence-based practice guidelines for preventive care patients MU-CDS | 97.4 | 29.8 | 3.3 | 12.5 | 190 |
| 4.3i. Reference information on medications being prescribed<br>Electronic system displays information about medications stored in its e-prescribing module/ subsystem or offers providers links to Internet websites with such information. | 85.6 | 26.8 | 2.7 | 40.1 | 284 |
| 4.3j. Reference guidelines and evidence-based recommendations when prescribing medication for a patient<br>Electronic system links to published diagnosis-specific guidelines or recommendations that includes appropriate medications for that diagnosis | 66.5 | 21.6 | 2.0 | 34.8 | 284 |
| 4.3k. Search for or generate a list of patients requiring a specific intervention (such as an immunization) MU-M | 59.2 | 24.1 | 0.8 | 22.7 | 284 |
| 4.3l. Search for or generate a list of patients on a specific medication (or on a specific dose of medication) MU-M | 60.6 | 22.0 | 0.7 | 16.5 | 284 |
| 4.3m. Search for or generate a list of patients who are due for a lab or other test in a specific time interval MU-M | 52.5 | 22.0 | 0.7 | 21.6 | 284 |

| Functions | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| 4.3n. Search for or generate a list of patients who fit a set of criteria, such as age, diagnosis, and clinical indicator value. For example, age less than 76, diagnosed with diabetes, and has an HbA1c greater than 9 percent. MU-M | 62.0 | 26.1 | 0.8 | 21.3 | 284 |

Domain 4. Use of the System to Increase Patient Engagement/Adherence

| Functions | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| 4.4a. Manage telephone calls<br>Refers to bringing up a patient's record whenever the patient calls or is called by the office and noting reason for the call. | 90.5 | 19.2 | 3.4 | 33.5 | 285 |
| 4.4b. Exchange secure messages with patients | 33.3 | 15.8 | 0.6 | 21.7 | 285 |
| 4.4c. Allow patients to view their medical records online MU-M | 30.5 | 17.2 | 0.6 | 32.2 | 285 |
| 4.4d. Allow patients to provide information online to update their records | 30.5 | 18.9 | 0.5 | 33.6 | 285 |
| 4.4e. Allow patients to request appointments online | 33.3 | 19.2 | 0.7 | 39.3 | 285 |
| 4.4f. Allow patients to request referrals online | 22.1 | 14.4 | 0.4 | 23.9 | 285 |
| 4.4g1. Produce hard-copy or electronic reminders for diabetes patients about needed tests, studies, or other services (for example, immunizations) MU-M | 78.5 | 32.9 | 2.2 | 27.5 | 200 |
| 4.4g2. Produce hard-copy or electronic reminders for coronary artery disease patients about needed tests, studies, or other services (for example, immunizations) MU-M | 75.3 | 34.0 | 2.0 | 27.3 | 150 |
| 4.4g3. Produce hard-copy or electronic reminders for congestive heart failure patients about needed tests, studies, or other services (for example, immunizations) MU-M | 76.2 | 36.7 | 2.0 | 27.7 | 143 |
| 4.4g4. Produce hard-copy or electronic reminders for preventive care patients about needed tests, studies, or other services (for example, immunizations) MU-M | 76.7 | 30.2 | 2.1 | 17.4 | 189 |
| 4.4h1. Generate written or electronic educational information to help diabetes patients understand their condition or medication MU-M | 85.9 | 35.2 | 2.4 | 49.7 | 199 |
| 4.4h2. Generate written or electronic educational information to help coronary artery disease patients understand their condition or medication MU-M | 82.7 | 37.0 | 2.1 | 59.0 | 150 |
| 4.4h3. Generate written or electronic educational information to help congestive heart failure patients understand their condition or medication MU-M | 82.5 | 38.6 | 2.1 | 49.1 | 143 |

| Functions | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| 4.4h4. Generate written or electronic educational information to help preventive care patients understand their condition or medication MU-M | 85.2 | 36.6 | 2.3 | 55.5 | 189 |
| 4.4i1. Create written care plans (personalized to patient's condition or age/gender for preventive care) to help guide diabetes patients in self-management | 67.8 | 41.3 | 1.7 | 87.9 | 199 |
| 4.4i2. Create written care plans (personalized to patient's condition or age/gender for preventive care) to help guide coronary artery disease patients in self-management | 68.7 | 49.4 | 1.5 | 93.9 | 150 |
| 4.4i3. Create written care plans (personalized to patient's condition or age/gender for preventive care) to help guide congestive heart failure patients in self-management | 64.3 | 45.6 | 1.4 | 86.2 | 143 |
| 4.4i4. Create written care plans (personalized to patient's condition or age/gender for preventive care) to help guide preventive care patients in self-management | 63.0 | 42.4 | 1.4 | 92.9 | 189 |
| 4.4j1. Prompt provider to review patient self-management plan (or patient-specific preventive care plan) with the diabetes patient during a visit | 64.3 | 36.2 | 1.8 | 62.0 | 199 |
| 4.4j2. Prompt provider to review patient self-management plan (or patient-specific preventive care plan) with the coronary artery disease patient during a visit | 68.7 | 45.7 | 1.8 | 79.4 | 150 |
| 4.4j3. Prompt provider to review patient self-management plan (or patient-specific preventive care plan) with the congestive heart failure patient during a visit | 65.7 | 44.3 | 1.8 | 73.4 | 143 |
| 4.4j4. Prompt provider to review patient self-management plan (or patient-specific preventive care plan) with the preventive care patient during a visit | 61.4 | 36.6 | 1.7 | 70.0 | 189 |
| 4.4k1. Modify self-management plan (or patient-specific preventive care plan) as needed following a diabetes patient visit | 62.3 | 34.3 | 1.8 | 65.3 | 199 |

| Functions | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| 4.4k2. Modify self-management plan (or patient-specific preventive care plan) as needed following a coronary artery disease patient visit | 66.0 | 45.1 | 1.7 | 87.2 | 150 |
| 4.4k3. Modify self-management plan (or patient-specific preventive care plan) as needed following a congestive heart failure patient visit | 63.6 | 44.3 | 1.8 | 81.4 | 143 |
| 4.4k4. Modify self-management plan (or patient-specific preventive care plan) as needed following a preventive care patient visit | 61.4 | 38.5 | 1.8 | 94.7 | 189 |
| 4.4l. Identify generic or less expensive brand alternatives at the time of prescription entry<br>Electronic system includes formularies that identify generic or less expensive alternatives to selected medication or offers providers links to Internet websites with such information. | 88.0 | 24.7 | 3.1 | 50.3 | 284 |
| 4.4m. Reference drug formularies of the patient's health plans/ pharmacy benefit manager to recommend preferred drugs at time of prescribing MU-M<br>Preferred drugs refer to medicines that receive maximum coverage under the patient's health plan. | 83.8 | 26.8 | 2.8 | 44.7 | 284 |

Domain 5. Medication Safety

| Functions | Percentage of Practices Using this Function in 2011 | Percentage that Began Using Function During 2010–2011 | Mean Item Score, 2011 (0-4, 4 Is Use for All/Nearly All Patient Visits) | Percentage Change in Mean from 2010 | Number of Responses |
|---|---|---|---|---|---|
| 4.5a. Maintain medication list for individual patients<br>Refers to using the electronic system to create, update, store and display a list of all medications (prescription and non-prescription) that the patient is taking. MU-C | 98.2 | 15.5 | 3.8 | 14.4 | 284 |
| 4.5b. Generate new prescriptions (that is, system prompts for common prescription details including medication type and name, strength, dosage, and quantity) MU-C | 98.6 | 15.8 | 3.8 | 17.6 | 284 |
| 4.5c. Generate prescription refills (that is, system allows provider to reorder a prior prescription by revising original details associated with it, rather than requiring re-entry) MU-C | 97.9 | 14.8 | 3.7 | 13.4 | 284 |
| 4.5d. Select individual medication for prescription (for example, from a drop-down list in the electronic system) MU-C | 97.9 | 15.5 | 3.7 | 18.4 | 284 |
| 4.5e. Calculate appropriate dose and frequency, or suggest administration route based on patient parameters such as age, weight, or functional limitations | 70.1 | 30.6 | 2.3 | 50.6 | 284 |
| 4.5f. Screen prescriptions for drug allergies against the patient's allergy information MU-C | 96.5 | 17.5 | 3.6 | 15.8 | 284 |
| 4.5g. Screen new prescriptions for drug-drug interactions against the patient's list of current medications MU-C | 96.8 | 19.2 | 3.6 | 31.2 | 284 |
| 4.5h. Check for drug-laboratory interaction<br>Such as to alert provider that patient is due for a certain laboratory or other diagnostic study to monitor for therapeutic or adverse effects of the medication or to alert provider that patient is at increased risk for adverse effects.<br>Electronic system may either store this information or link to Internet websites with such information. | 41.2 | 19.9 | 1.2 | 19.5 | 284 |
| 4.5i. Check for drug-disease interaction<br>Electronic system may either store this information or link to Internet websites with such information. | 55.3 | 29.6 | 1.9 | 61.3 | 284 |

Source: Office Systems Survey (OSS) conducted in spring and summer 2011.

Notes: N = 277 practices. The sample for all function-specific questions includes all practices who completed an OSS and implemented some sort of an electronic tool (an EHR, an electronic patient registry, or an electronic prescribing system) by the end of demonstration year 2.

MU-C: Also in “meaningful use” core set

MU-M: Also in “meaningful use” menu set

Table B.7. Treatment Group Practices' Responses to ARRA Incentives Questions by Practice Characteristics (Percentages, Unless Otherwise Noted)

| | All | 1-2 Providers | 3-5 Providers | 6+ Providers | Rural | Non-rural | Affiliated | Unaffiliated | LA | MD | PA | SD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Responding Practices | 319 | 144 | 115 | 60 | 53 | 266 | 58 | 261 | 70 | 100 | 114 | 350 |
| 5.1. Did the announcement of ARRA funding change the decision to adopt an EHR system or change the pace of changes in use? (N = 319 ) | | | | | | | | | | | | |
| Yes | 36.7 | 35.4 | 39.1 | 35.0 | 39.6 | 36.1 | 24.1 | 39.5 | 45.7 | 49.0 | 21.1 | 34.3 |
| No (but aware) | 58.6 | 59.0 | 59.1 | 56.7 | 56.6 | 59.0 | 72.4 | 55.6 | 48.6 | 47.0 | 75.4 | 57.1 |
| No, not aware | 4.7 | 5.6 | 1.7 | 8.3 | 3.8 | 4.9 | 3.4 | 5.0 | 5.7 | 4.0 | 3.5 | 8.6 |
| 5.1a. If yes, how? (N = 120) | | | | | | | | | | | | |
| Accelerated adoption | 55.8 | 61.5 | 52.2 | 50.0 | 57.1 | 55.6 | 78.6 | 52.8 | 65.6 | 49.0 | 44.4 | 83.3 |
| Accelerated use | 42.5 | 36.5 | 45.7 | 50.0 | 42.9 | 42.4 | 21.4 | 45.3 | 34.4 | 49.0 | 51.9 | 16.7 |
| Delayed adoption | 1.7 | 1.9 | 2.2 | 0 | 0 | 2.0 | 0 | 1.9 | 0 | 2.0 | 3.7 | 0 |
| Delayed use | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5.2. Did the policy of penalties starting 2015 influence the decision to adopt an EHR system or change the pace of planned changes in use? (N = 319 ) | | | | | | | | | | | | |
| Yes | 29.5 | 30.6 | 27.8 | 30.0 | 37.7 | 27.8 | 27.6 | 29.9 | 32.9 | 35.0 | 20.2 | 37.1 |
| No (but aware) | 66.8 | 65.3 | 69.6 | 65.0 | 56.6 | 68.8 | 70.7 | 65.9 | 58.6 | 63.0 | 77.2 | 60.0 |
| No (not aware) | 3.8 | 4.2 | 2.6 | 5.0 | 5.7 | 3.4 | 1.7 | 4.2 | 8.6 | 2.0 | 2.6 | 2.9 |
| 5.3. When are most or all MDs in the practice expected to meet Stage 1 Medicare meaningful use criteria? (N = 314 ) | | | | | | | | | | | | |
| Already met these | 17.8 | 19.7 | 15.8 | 17.2 | 26.9 | 16.0 | 12.1 | 19.1 | 34.8 | 15.0 | 12.3 | 11.8 |
| 2011 | 43.6 | 38.0 | 44.7 | 55.2 | 28.8 | 46.6 | 50.0 | 42.2 | 33.3 | 49.0 | 51.8 | 20.6 |
| 2012 | 22.9 | 21.8 | 23.7 | 24.1 | 28.8 | 21.8 | 32.8 | 20.7 | 21.2 | 25.0 | 12.3 | 55.9 |
| 2013 | 3.2 | 4.2 | 3.5 | 0 | 5.8 | 2.7 | 3.4 | 3.1 | 1.5 | 2.0 | 6.1 | 0 |
| 2014 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2015 | 0.3 | 0.7 | 0 | 0 | 0 | 0.4 | 0 | 0.4 | 0 | 0 | 0.9 | 0 |

| | All | 1-2 Providers | 3-5 Providers | 6+ Providers | Rural | Non-rural | Affiliated | Unaffiliated | LA | MD | PA | SD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Other year | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Not in foreseeable future | 5.4 | 4.9 | 8.8 | 0 | 3.8 | 5.7 | 0 | 6.6 | 1.5 | 1.0 | 12.3 | 2.9 |
| Don't know, not familiar enough to estimate | 6.7 | 10.6 | 3.5 | 3.4 | 5.8 | 6.9 | 1.7 | 7.8 | 7.6 | 8.0 | 4.4 | 8.8 |
| 5.4. Average estimate of Medicare or Medicaid incentive money that all the practices' physicians combined expect to receive over all the years it is available (N = 159) | | | | | | | | | | | | |
| Average (mean) estimate per physician | 162,167.3 | 253,848.5 | 147,325.4 | 36,118.9 | 179,839.5 | 158,552.5 | 92,097.2 | 180,518.9 | 430,069.7 | 32,475.8 | 177,756.2 | 124,198.8 |
| Percentage of practices unable to estimate | 49.8 | 53.5 | 55.7 | 30.0 | 43.4 | 51.1 | 43.1 | 51.3 | 62.9 | 48.0 | 51.8 | 22.9 |
| 5.5. How does the anticipated Medicare or Medicaid meaningful use incentive money compare to the anticipated funding by the EHRD over the course of the demonstration? (N = 160) | | | | | | | | | | | | |
| About the same | 28.1 | 26.9 | 33.3 | 23.8 | 35.7 | 26.5 | 21.2 | 29.9 | 42.3 | 33.3 | 20.0 | 20.0 |
| Meaningful use dollars are substantially larger | 43.1 | 32.8 | 49.0 | 52.4 | 35.7 | 44.7 | 51.5 | 40.9 | 23.1 | 31.5 | 60.0 | 52.0 |
| EHRD dollars are substantially larger | 5.0 | 4.5 | 7.8 | 2.4 | 7.1 | 4.5 | 3.0 | 5.5 | 3.8 | 7.4 | 3.6 | 4.0 |
| Have not estimated EHRD dollars | 23.8 | 35.8 | 9.8 | 21.4 | 21.4 | 24.2 | 24.2 | 23.6 | 30.8 | 27.8 | 16.4 | 24.0 |
| Have not estimated meaningful use dollars | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| | All | 1-2 Providers | 3-5 Providers | 6+ Providers | Rural | Non-rural | Affiliated | Unaffiliated | LA | MD | PA | SD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5.6. Other than through EHRD or meaningful use, funding received from other sources for purchase or use of an EHR system since June 2009 (N = 319 ) | | | | | | | | | | | | |
| % with grant or subsidy | 6.9 | 7.6 | 6.1 | 6.7 | 7.5 | 6.8 | 5.2 | 7.3 | 2.9 | 3.0 | 14.9 | 0 |
| % with bonus or incentive | 3.8 | 4.2 | 4.3 | 1.7 | 0 | 4.5 | 3.4 | 3.8 | 4.3 | 4.0 | 4.4 | 0 |
| % with loan | 2.5 | 0.7 | 4.3 | 3.3 | 3.8 | 2.3 | 1.7 | 2.7 | 0 | 6.0 | 1.8 | 0 |
| No | 86.8 | 87.5 | 85.2 | 88.3 | 88.7 | 86.5 | 89.7 | 86.2 | 92.9 | 87.0 | 78.9 | 100.0 |
| 5.6a. Average (mean) amount received among those with other sources of support (N = 32 ) | | | | | | | | | | | | |
| Grant or subsidy | 131,473.7 | 93,000.0 | 236,500.0 | 82,000.0 | 343,000.0 | 112,812.5 | 94,000.0 | 151,071.4 | | 50,000.0 | 149,437.5 | |
| Bonus or incentive | 30,133.3 | 30,500.0 | 19,933.3 | 60,000.0 | | 30,133.3 | 11,000.0 | 33,960.0 | 19,933.3 | 55,000.0 | 11,000.0 | |
| Loan | 128,581.0 | 44,000.0 | 54,229.6 | 356,750.0 | 253,750.0 | 86,858.0 | 80,000.0 | 135,521.1 | | 153,665.2 | 53,328.5 | |
| 5.7. Other than through EHRD or meaningful use, funding expected to be received from other sources for purchase or use of an EHR system between now and 2016 (N = 316) | | | | | | | | | | | | |
| % expecting grant or subsidy | 0.9 | 0.7 | 0.9 | 1.7 | 0 | 1.1 | 1.8 | 0.8 | 2.9 | 0 | 0.9 | 0 |
| % expecting bonus or incentive | 6.6 | 6.3 | 7.9 | 5.0 | 1.9 | 7.6 | 3.5 | 7.3 | 4.3 | 9.1 | 7.1 | 2.9 |
| % expecting loan | 0.3 | 0.7 | 0 | 0 | 0 | 0.4 | 0 | 0.4 | 1.4 | 0 | 0 | 0 |
| No | 92.1 | 92.3 | 91.2 | 93.3 | 98.1 | 90.9 | 94.7 | 91.5 | 91.4 | 90.9 | 92.0 | 97.1 |
| 5.7a. Average (mean) amount expected among those expecting other sources of support (N = 13 ) | | | | | | | | | | | | |
| Grant or subsidy | 2,011,333.3 | 6,000,000.0 | 10,000.0 | 24,000.0 | | 2,011,333.3 | 10,000.0 | 3,012,000.0 | 3,005,000.0 | | 24,000.0 | |

| | All | 1-2 Providers | 3-5 Providers | 6+ Providers | Rural | Non-rural | Affiliated | Unaffiliated | LA | MD | PA | SD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bonus or incentive | 121,888.9 | 17,333.3 | 29,000.0 | 900,000.0 | | 121,888.9 | 1,000.0 | 156,428.6 | 30,000.0 | 226,250.0 | 51,000.0 | |
| Loan | 45,000.0 | 45,000.0 | | | | 45,000.0 | | 45,000.0 | 45,000.0 | | | |

Source: Office Systems Survey conducted in spring and summer 2011.

EHR = electronic health record; LA = Louisiana; MD = Maryland; PA = Pennsylvania; SD = South Dakota.

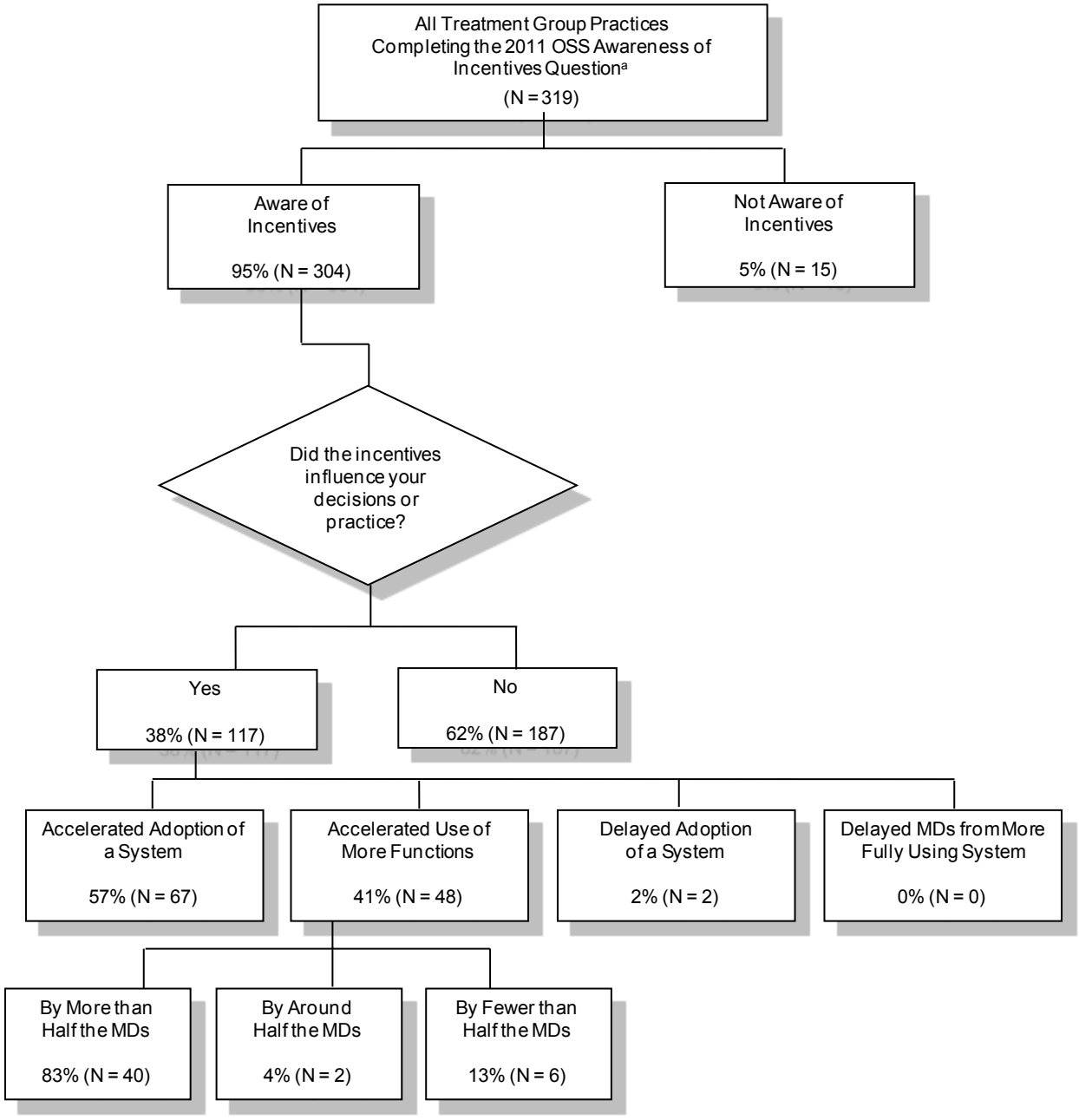
Table B.8. Control Group Practices' Responses to ARRA Incentives Questions by Practice Characteristics (Percentages)

| | All | 1-2 Providers | 3-5 Providers | 6+ Providers | Rural | Non-rural | Affiliated | Unaffiliated | LA | MD | PA | SD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Responding Practices | 255 | 119 | 91 | 45 | 38 | 217 | 44 | 211 | 47 | 67 | 107 | 34 |
| 5.1. Did the announcement of ARRA funding change the decision to adopt an EHR system or change the pace of changes in use? (N = 253) | | | | | | | | | | | | |
| Yes | 31.2 | 31.4 | 31.9 | 29.5 | 35.1 | 30.6 | 20.5 | 33.5 | 32.6 | 37.3 | 23.6 | 41.2 |
| No (but aware) | 59.3 | 55.9 | 60.4 | 65.9 | 56.8 | 59.7 | 68.2 | 57.4 | 43.5 | 56.7 | 68.9 | 55.9 |
| No, not aware | 9.5 | 12.7 | 7.7 | 4.5 | 8.1 | 9.7 | 11.4 | 9.1 | 23.9 | 6.0 | 7.5 | 2.9 |
| 5.1a. If yes, how? (N = 80) | | | | | | | | | | | | |
| Accelerated adoption | 57.5 | 60.5 | 55.2 | 53.8 | 76.9 | 53.7 | 55.6 | 57.7 | 46.7 | 52.0 | 53.8 | 85.7 |
| Accelerated use | 40.0 | 36.8 | 44.8 | 38.5 | 15.4 | 44.8 | 44.4 | 39.4 | 46.7 | 48.0 | 46.2 | 7.1 |
| Delayed adoption | 1.3 | 2.6 | 0 | 0 | 0 | 1.5 | 0 | 1.4 | 6.7 | 0 | 0 | 0 |
| Delayed use | 1.3 | 0 | 0 | 7.7 | 7.7 | 0 | 0 | 1.4 | 0 | 0 | 0 | 7.1 |
| 5.2. Did the policy of penalties starting 2015 influence the decision to adopt an EHR system or change the pace of planned changes in use? (N = 256) | | | | | | | | | | | | |
| Yes | 34.8 | 35.8 | 33.0 | 35.6 | 47.4 | 32.6 | 29.5 | 35.8 | 54.2 | 22.4 | 29.9 | 47.1 |
| No (but aware) | 60.5 | 56.7 | 64.8 | 62.2 | 42.1 | 63.8 | 65.9 | 59.4 | 31.3 | 74.6 | 67.3 | 52.9 |
| No (not aware) | 4.3 | 6.7 | 2.2 | 2.2 | 10.5 | 3.2 | 4.5 | 4.2 | 12.5 | 3.0 | 2.8 | 0 |

Source: Office Systems Survey conducted in spring and summer 2011.

EHR = electronic health record;; LA = Louisiana; MD = Maryland; PA = Pennsylvania; SD = South Dakota.

Figure B.1. Awareness and Influence of ARRA "Meaningful Use" Incentives on Adoption and Use of an Electronic Health Record (EHR) System in the EHRD Treatment Group



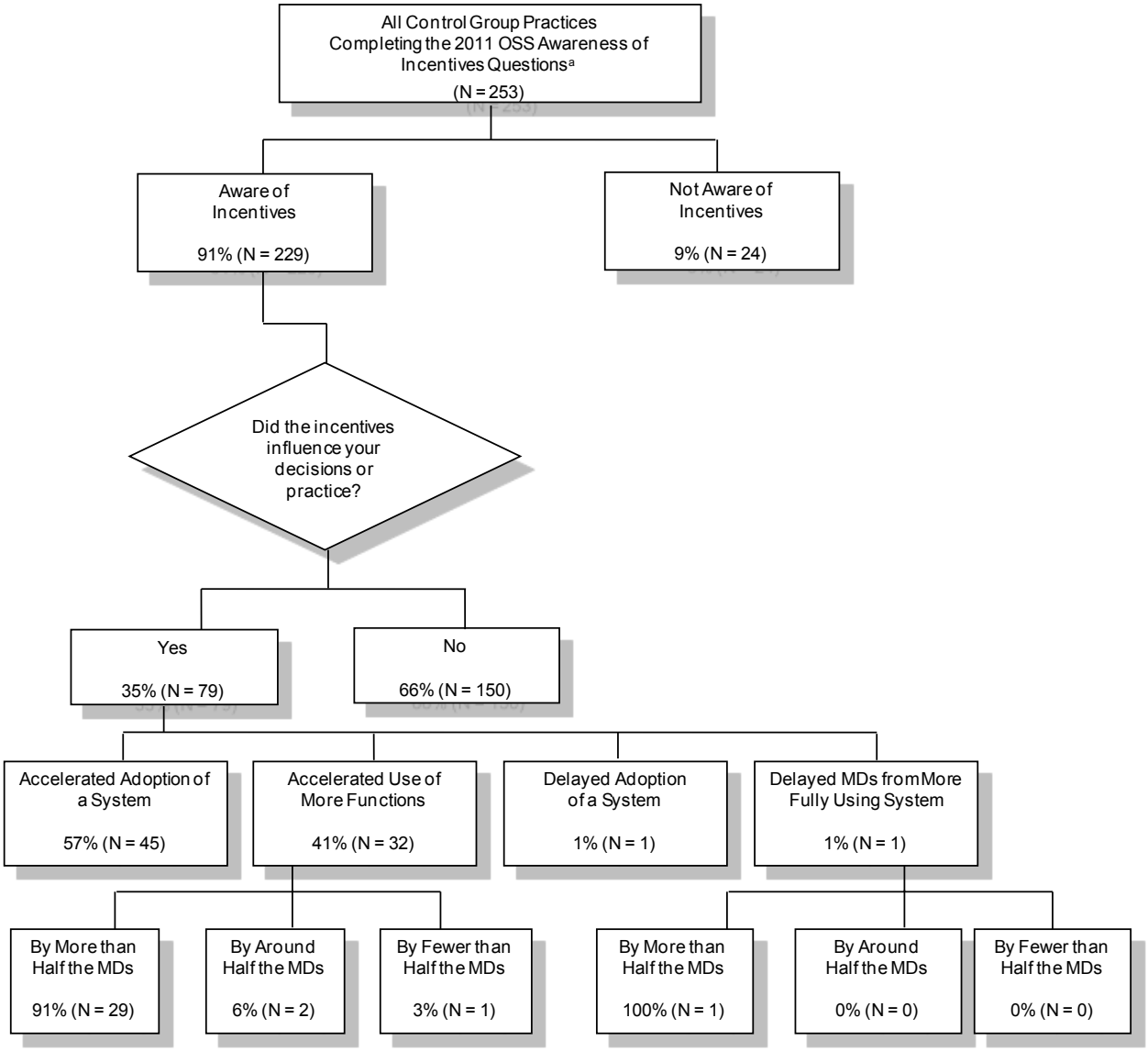

Source: Office Systems Survey conducted in spring and summer 2011.

Note: Percentages may not add to 100 percent due to rounding.

[a]Five treatment group practices responded to the OSS but did not answer questions related to meaningful use.

MD = Medical Doctor.

Figure B.2. Awareness and Influence of ARRA "Meaningful Use" Incentives on Adoption and Use of an Electronic Health Record (EHR) System in the EHRD Control Group



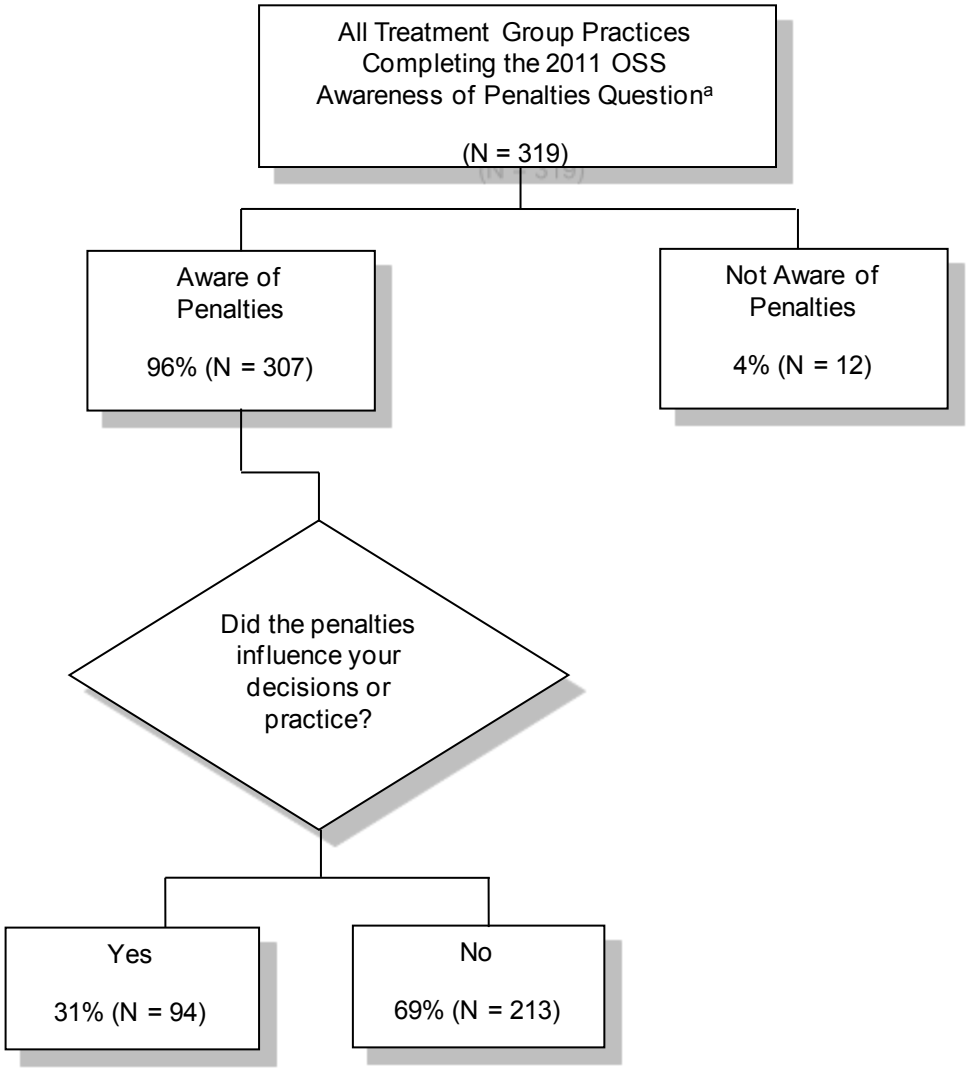

Source: Office Systems Survey (OSS) conducted in spring and summer 2011.

Note: Percentages may not add to 100 percent due to rounding.

[a]Fifteen additional control group practices responded to the OSS but did not answer questions related to meaningful use.

MD = Medical Doctor.

Figure B.3. Awareness and Influence of 2015 Penalties for Lack of Meaningful Use in the EHRD Treatment Group

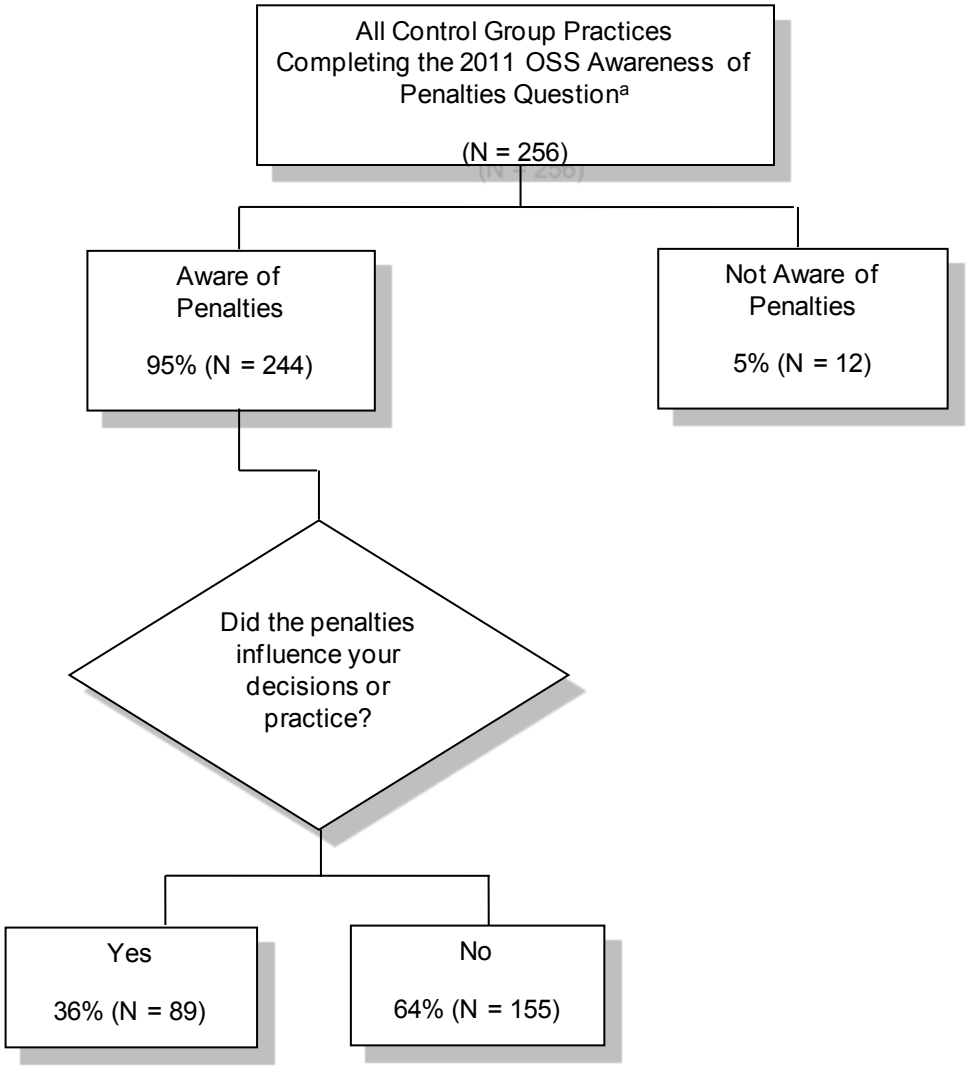



Source: Office Systems Survey (OSS) conducted in spring and summer 2011.

Note: Percentages may not add to 100 percent due to rounding.

[a]Five treatment group practices responded to the OSS but did not answer questions related to meaningful use.

Figure B.4. Awareness and Influence of 2015 Penalties for Lack of Meaningful Use in the EHRD Control Group




Source: Office Systems Survey (OSS) conducted in spring and summer 2011.

Note: Percentages may not add to 100 percent due to rounding.

[a]Twelve control group practices responded to the OSS but did not answer questions related to meaningful use.

This page is intentionally left blank.

# APPENDIX C

# CHARACTERISTICS OF VISITED TREATMENT AND CONTROL GROUP PRACTICES COMPARED WITH OTHER TREATMENT AND CONTROL PRACTICES

This page is intentionally left blank.

Table C.1: Characteristics of Visited Treatment and Control Group Practices Compared with Other Treatment and Control Group Practices (Percentages, Unless Otherwise Noted)

| Practice Characteristic | Number of Visited Treatment Practices | Percentage of Visited Treatment Practices | Number of Treatment Practices Not Visited | Percentage of Treatment Practices Not Visited | Number of Visited Control Practices | Percentage of Visited Control Practices | Number of Control Practices Not Visited | Percentage of Control Practices Not Visited |
|---|---|---|---|---|---|---|---|---|
| Practice Size (total number of providers) | | | | | | | | |
| 1-2 | 5 | 33.3 | 133 | 43.0 | 6 | 100.0 | 261 | 99.6 |
| 3-5 | 5 | 33.3 | 104 | 33.7 | 0 | 0.0 | 0 | 0.0 |
| 6+ | 5 | 33.3 | 72 | 23.3 | 0 | 0.0 | 1 | 0.4 |
| Percentage in an MUA[a] | 4 | 26.7 | 86 | 27.8 | 1 | 16.7 | 76 | 29.0 |
| Percentage in a Rural Area[b] | 3 | 20.0 | 51 | 16.5 | 1 | 16.7 | 40 | 15.3 |
| Practice Affiliation | | | | | | | | |
| Unaffiliated | 7 | 46.7 | 138 | 44.7 | 3 | 50.0 | 99 | 37.8 |
| Owned by or affiliated with a larger organization | 8 | 53.3 | 171 | 55.3 | 3 | 50.0 | 163 | 62.2 |
| Participation in Other Quality Improvement, EHR, and Pay-for-Performance Programs | | | | | | | | |
| No participation | 3 | 20.0 | 48 | 15.5 | 1 | 16.7 | 59 | 22.5 |
| At least some participation | 12 | 80.0 | 261 | 84.5 | 5 | 83.3 | 203 | 77.5 |

Source: Office Systems Survey (OSS), conducted in spring and summer 2011.

Note: Table includes only practices that responded to the 2011 OSS (15 of the 16 visited treatment group practices, and 6 of the 8 visited control group practices responded to the survey). Percentages indicate practices that responded to the relevant question.

[a] Randomization information (done by linking geocoded addresses to data from HRSA website).

[b] HRSA Area Resource File was used to identify urban and rural counties.

EHR = electronic health record; HRSA = Health Resources and Services Administration; MUA = medically underserved area.

This page is intentionally left blank.

# APPENDIX D

# CHARACTERISTICS OF CONTACTED PRACTICES FAILING PROGRAM REQUIREMENTS AT THE END OF DEMONSTRATION YEAR 2, COMPARED WITH OTHER PRACTICES FAILING THESE PROGRAM REQUIREMENTS

This page is intentionally left blank.

Table D.1: Characteristics of Contacted Practices Failing Program Requirements at the End of Demonstration Year 2, Compared with Other Practices Failing These Program Requirements (Percentages, Unless Otherwise Noted)

| Practice Characteristic | Number of Contacted Practices – Failing Program Requirements End of Year 2 | Percentage of Contacted Practices – Failing Program Requirements End of Year 2 | Number of Practices Not Contacted – Failing Program Requirements End of Year 2 | Percentage of Practices Not Contacted – Failing Program Requirements End of Year 2 |
|---|---|---|---|---|
| Practice Size (total number of providers) | | | | |
| 1-2 | 6 | 30.0 | 35 | 49.3 |
| 3-5 | 8 | 40.0 | 26 | 36.6 |
| 6+ | 6 | 30.0 | 10 | 14.1 |
| Percentage in an MUA[a] | 1 | 4.3 | 28 | 32.6 |
| Percentage in a Rural Area[b] | 6 | 26.1 | 7 | 8.1 |
| Practice Affiliation[a] | | | | |
| Unaffiliated | 10 | 50.0 | 44 | 62.0 |
| Owned by or affiliated with a larger organization | 10 | 50.0 | 27 | 38.0 |
| Participation in Other Quality Improvement, EHR, and Pay-for-Performance Programs | | | | |
| No participation | 7 | 35.0 | 35 | 49.3 |
| At least some participation | 13 | 65.0 | 36 | 50.7 |

Source: Office Systems Survey (OSS), conducted in spring and summer 2011.

Note: includes the 23 practices that failed to meet program requirements at the end of demonstration Year 2 and were contacted, and 86 such practices that were not contacted. This group includes practices that reported not using an EHR, practices that did not complete the 2011 OSS, and practices that submitted the OSS and had an EHR but otherwise failed to meet program requirements (described in Chapter I). The characteristics based on OSS data are limited to the 20 contacted and 71 non-contacted practices that responded to the applicable questions. Percentages indicate practices that responded to the relevant question.

[a]Randomization information (done by linking geocoded addresses to data from HRSA website).

[b]HRSA Area Resource File was used to identify urban and rural counties.

EHR = electronic health record; HRSA = Health Resources and Services Administration; MUA = medically underserved area

.

This page is intentionally left blank.

# APPENDIX E

# EHR PERFORMANCE SCORE CALCULATION

This page is intentionally left blank.

MATHEMATICA
Policy Research, Inc.

600 Maryland Ave. S.W., Suite 550
Washington, DC 20024-2512
Telephone (202) 484-9220
Fax (202) 863-1763
www.mathematica-mpr.com

MEMORANDUM

TO: Lorraine Johnson

FROM: Sue Felt-Lisk[1]

DATE: 6/10/2009[2]
EHRD–052

SUBJECT: Revised OSS Scoring Plan

---

Physician practices assigned to the treatment group of the Centers for Medicare & Medicaid Services' (CMS) Electronic Health Records Demonstration (EHRD) will receive payments for their use of EHR systems based on their responses to the Office Systems Survey (OSS). Practice responses to the OSS in year 2 pertaining to the minimum requirements (Section C) will also determine their eligibility to continue participating in the demonstration in years 3 and beyond. This memorandum describes the plan for scoring Office Systems Survey responses for the purposes of calculating those payments.

The next two sections provide background and explain the method used to develop the OSS scoring plan (which includes assigning individual questions to broader domains). Section C explains the method for determining whether practices pass the minimum requirement to qualify for an incentive (years 1 and 2), and to continue participating in the demonstration (after year 2). Section D explains how questions and domains are scored, and Section E describes how the OSS summary score is built from the domain scores. The scoring plan ends with a description of how payment is determined, in Section F. Appendix A provides details about scoring registry and e-prescribing functions for practices with stand-alone systems.

## A. BACKGROUND

The EHRD, which is authorized under Section 402 Medicare Waiver Authority, is being implemented by CMS. It expands upon the Medicare Care Management Performance (MCMP) demonstration as well as building upon other CMS demonstrations. Specifically, the EHRD tests whether performance-based financial incentives increase physician practices' adoption and use of electronic health records (EHRs) and improve the quality of care practices deliver to chronically ill patients with fee-for-service Medicare coverage.

[1] Input and comments throughout the development of this plan from Lorenzo Moreno and Jennifer Schore, and our physician researcher colleagues Mai Pham, Anne O'Malley, and Arnold Chen are gratefully acknowledged. In addition, this plan reflects the decisions and guidance of both CMS and Assistant Secretary for Planning and Evaluation (ASPE) staff and CMS contractors per several telephone conferences during August through December 2008 and related emails.

[2] Section 2 was revised to clarify that question 4.1c (an informational item, not an EHR function) is excluded from scoring, on July 28, 2010. Introductory material to the scoring plan and Section C was revised to emphasize that all minimum requirements must be met for continued participation in the demonstration after year 2, as has always been referenced in Section A (April 25, 2011).

Under the demonstration, treatment group practices will have the opportunity to receive three types of payments. The first, called the *systems* payment, (up to $5,000 per physician to a maximum of $25,000 per practice) is based on use of an EHR. It is available in all five years of the demonstration. The systems payment will be based on the practice's use of a minimum set of functions in an EHR system that was certified under the old Certification Commission for Healthcare Information Technology (CCHIT) certification standards or by another authorized certification body under the new "meaningful use standards", as measured by responses to an annually administered practice survey (called the Office Systems Survey or OSS). Additional payment will be provided for use of more sophisticated EHR functions. Practices that have not adopted minimal use of the EHR system by the end of the first year will not receive payment, but may remain in the demonstration.

In year 2, practices have the opportunity to receive systems payments and a second type of payment (up to $3,000 per physician to a maximum of $15,000 per practice) for reporting on specific clinical quality measures. (Practices that have not adopted minimal use of their EHR system by the end of the second year will be removed from the demonstration.)

In years 3 to 5, practices will have the opportunity to receive systems payments and a third type of payment (up to $10,000 per physician to a maximum of $50,000 per practice), a quality payment, for performance on specific clinical quality measures. The financial payments will be in addition to the normal fee-for-service Medicare payment practices receive for services delivered. Physicians could receive up to $58,000 per provider, up to a maximum of $290,000 per practice over the five years of the demonstration.[3]

The EHR Demonstration summary issued by CMS (June 10, 2008) states that practices will receive up to $45 per beneficiary (for beneficiaries with chronic conditions assigned to the practice) based on their performance on the Office Systems Survey.[4] For example, a single overall score on the survey will be used to calculate the percentage of the $45 per beneficiary that the practice will receive. So a practice with two or more physicians that scores 60 percent on the survey and has 200 beneficiaries with chronic conditions assigned to it would receive $5400 (200 x $45 x 60 percent).

## B. METHOD

To develop options for scoring the OSS, we first explored whether existing scoring mechanisms might serve as models. We reviewed the scoring of the DOQ-IT version of the OSS

[3] John C. Wilkin, Kerry E. Moroz. Erika G. Yoshino, and Laurie E. Pekala. "Electronic Health Records Demonstration Waiver Cost Estimate." Columbia, MD: Actuarial Research Corporation, December 13, 2007.

[4] "Electronic Health Records (EHR) Demonstration: Demonstration Summary," CMS. Dated June 10, 2008.

that was used to gauge DOQ-IT program progress. We concluded that a more comprehensive scoring mechanism is required for calculating EHR demonstration payments, that is, one that takes into account use of all or nearly all the EHR functions queried on the OSS. We spoke with key NCQA staff regarding scoring for the Physician Practice Connections, but their objectives are very different from those of the EHR demonstration in that they aim to determine if a practice has desirable care management practices in place, regardless of whether they are electronic. For example, a practice can get a high score on the Physician Practice Connections instrument without having an EHR.

The scoring plan described in this memo relies on the following principles:

- The plan should be kept as simple as possible. The plan should not vary the scoring method by demonstration year; rather it must remain constant.
- The plan should recognize that early in the demonstration some practices will be new to the use of EHRs, but that over the demonstration period, use of EHRs could increase substantially.
- An overall summary score should be built up from domain scores on a relatively small number of domains that are conceptually distinct and would be perceived as relevant and meaningful to providers and CMS.

An initial draft of this plan (dated August 15, 2008) was reviewed and discussed with CMS and CMS partner staff through four telephone conferences (held during August through December 2008). Decisions made during those discussions are reflected in the text that follows.

We considered using the MCMP OSS data to support factor analysis during the development of the plan, but decided against it. Factor analysis relies on linear regression methods to identify groups of questions in a survey whose responses tend to be highly correlated and, therefore, can be grouped into a single factor or domain. Factor analysis aims at identifying the most important domains in a dataset. The main reasons for deciding not to use factor analysis to set the domains include technical considerations, limitations on available data, and potentially greater difficulty to explain the rationale. From a technical perspective, factor analysis can result in any number of domains, and the domains may or may not be ones that CMS, physician practices, or other interested parties would view as logical, despite their statistical basis. We also considered limitations in available data. Specifically, the only data potentially available for factor analysis is based on an older version of the OSS used for the MCMP, therefore the questions added to the OSS for the EHRD would not correspond to the domains that could be identified from the older OSS. Finally, explaining that the domains are based on statistical analysis to audiences that are not familiar with factor analysis may not be as satisfying as reviewing domains that have intuitive face validity and are grounded in medical practice.

C. MINIMUM REQUIREMENT FOR SYSTEM PAYMENT (YEARS 1 AND 2) AND PARTICIPATION (AFTER YEAR 2)

To receive a systems payment, practices must have implemented a certified EHR (that was certified under the old CCHIT certification standards or the new "meaningful use standards") and be using it for the following minimum core functions: recording of patient visit notes, recording of diagnostic test orders and results, and recording of prescriptions. Practices that are not using a certified EHR to perform the minimum core functions by the end of year 2 will be dropped from participation in the demonstration and will not be eligible to receive any incentive payments. To pass the minimum requirement to qualify for any incentive and for continued participation in the demonstration after year 2, all of the following question responses are required:

1. Certified EHR: yes to OSS question 3.5 [Does the EHR system have a valid certification for some or all of the period from June 1, 2009 to the present?]

2. Recording of visit notes: non-zero response (that is, a response other than "none") to question 4.1d [Clinical notes for individual patients]

3. Recording of diagnostic test orders: non-zero responses to 4.1i [Recording (or entering) laboratory orders into electronic system] AND 4.1l [Recording (or entering) imaging orders into electronic system]

4. Recording of diagnostic test results: non-zero response to *any* of 4.1j, 4.2g, 4.2h and 4.2i (pertaining to laboratory results) AND non-zero response to *any* of 4.1m, 4.2j, 4.2k, and 4.2l (the parallel questions pertaining to imaging results). For reference, the 4.1j, and 4.2g through 4.2i for laboratory are:

    - 4.1j: Receiving laboratory results by fax or mail and scanning paper versions into electronic systems

    - 4.2g: Transfer electronic laboratory results (received in non-machine readable form, such as an e-fax) directly into system

    - 4.2h: Enter laboratory results manually into electronic system in a searchable field (whether received by fax, mail or phone)

    - 4.2i: Receive electronically transmitted laboratory results directly into system from facilities that have the capability to send such transmissions

5. Recording of prescriptions: Non-zero response to 4.1p [Recording (or entering) prescription medications (new prescriptions and refills) into electronic system].

If a practice passes this minimum requirement, an overall OSS score is calculated to determine the size of the per-patient payment to the practice.

## D. QUESTION AND DOMAIN SCORING

### 1. Question Scoring

Fifty-three EHR functions are scored through response to questions on the OSS. Most questions are scored on a 0 to 4 (5-point) scale. The response choices for most items translate directly into their score, with 0 less desirable, representing no use of a function, and 4 indicating the function is used for "3/4 or more" patients. One question (4.1b–proportion of paper charts pulled) requires scoring in reverse of the response choices, because a better score on this question is lower. For the items pertaining to report generation, we will recode the responses on a 3-point scale such that 0 [Not used during last year]=0, 1 [As-needed basis at least once]=2, and 2 [Regularly for full practice]=4.

#### Hierarchical Item Sets

The OSS contains five sets of hierarchical items that are each scored as a set. A hierarchical item set consists of several consecutive questions in the OSS that represent progressively more advanced ways of using the EHR, so that as a practice advances in its use, it will indicate less use of the less advanced process and more use of the more advanced process.

The questions associated with these sets pertain to ordering laboratory tests, ordering radiology tests, receiving laboratory results, receiving radiology results, and prescription ordering. We will score each set of questions (representing a single function) together such that the result is a score between 0 and 4 just as with the other items. However, additional steps are necessary to arrive at the score for the set.

1. We will weight the response to the most advanced method most heavily in the score; specifically the lowest-level question response will be multiplied by 1; the middle-level response will be multiplied by 2; and the most advanced level question response will be multiplied by 3.

2. The products of the responses times their weight (1,2, or 3) will be summed and divided by 12, then multiplied by 4. In mathematical terms this is written [((Q1X1)+(Q2X2)+(Q3X3))/12] X 4. Twelve is the appropriate denominator since we would want a practice that responded with the highest response (3/4 or more) to the most advanced function to receive the maximum points. The multiplication times 4 is in order to rescale the result to a 0 to 4 scale, similar to most of the other OSS items. However, because of the exact categorical response boundaries of the individual items, the result can exceed 4, therefore capping is sometimes necessary.

3. Apply caps as follows, based on the response to the highest-level question in the hierarchy: cap to 4.0 if the highest-level question equals 4; cap to 3.75 if the highest-level question equals 3, and cap to 3.5 if the highest-level question equals less than 3.

This method of capping reserves the top score (4) for practices that use the most advanced level function for 75 percent or more of their patients.

An example of scoring a hierarchical set is shown in the table below.

| Ordering Laboratory Tests (Hierarchical Items) | Example Practice Response to proportion of patients for which function used | Response is Multiplied by a Weighting Factor |
|---|---|---|
| Print and fax laboratory orders | 2 [1/4 or more but less than 1/2] | 2 X 1 = 1 |
| Fax laboratory orders electronically from system | 3 [1/2 or more but less than 3/4] | 3 X 2 = 6 |
| Transmit laboratory orders directly from system to facilities that have the capability to receive such transmissions | 2 [ 1/4 or more but less than 1/2]] | 2 X 3 = 6 |
| Sum of Weighted Responses Divided by 12 and Multiplied by 4 | | 13/12 = 1.08 X 4 = 4.3 |
| Cap Applied Based on Response to Highest-level Item (Transmit laboratory orders directly…) | | 3.5 |

*Condition-Specific Items*

Seven items pertaining to care management are asked on a condition-specific basis for diabetes, CAD, CHF, and prevention. This is because practices tend to begin using these functions as part of their attempt to improve quality on specific conditions, rather than all at once. For each of these items, a total score will be developed across the conditions. The total will be divided by 16, which is the total possible points since each of the 4 condition-specific items represents 4 possible points. Then the total will be rescaled to a 0 to 4 point scale similar to the other items by multiplying the percentage of possible points achieved by 4. The table below provides an example.[5]

[5] The other six condition-specific items in the OSS are (1) produce hard-copy or electronic reminders for patients about needed tests, studies, or other services; (2) generate written or electronic educational information to help patients understand their medical condition or medication; (3) record that instructions or educational information were given to patient; (4) create written care plans to help guide patients in self-management; (5) prompt provider to review patient self-management plan with the patient during a visit; and (6) modify self-management plan as needed following a patient visit.

| 4.3h Review and act on reminders at the time of a patient encounter… | Response |
|---|---|
| Diabetes | 4 [3/4 or more] |
| CAD | 0 [None] |
| CHF | 2 [1/4 or more, but less than 1/2] |
| Prevention | 2 [1/4 or more, but less than 1/2] |
| Total Score: | 8/16 = .50 |
| Rescaled Score: | .50 x 4 = 2 |

2. Exclusions from Scoring

Items 3.5 and 3.5a (certified EHR as of June 1, 2009 or later) will not be included in the scoring. Nearly all functions queried on the OSS are scored, except the following that pertain to minimum requirements, and three additional items:

### Minimum Requirement Items

4.1d (maintain clinical notes)

4.1i (record or enter laboratory orders)

4.1l (record or enter imaging orders)

4.1j (receive laboratory results by fax or mail and scan paper versions into electronic system)

4.1m (receive imaging results by fax or mail and scan paper versions into electronic system), and

4.1p (record or enter prescription medications (new prescriptions and refills) into electronic system)

### Other Items

4.4e (allow patients to request appointments online)

4.4f (allow patients to request referrals online)

4.1c (method used to transition paper records to the EHR system)

The reason for excluding 4.4e and 4.4f is that they pertain to the interaction of the patient with the practice management system rather than the EHR system, and to date they have no known implications for quality improvement or savings. (Progress on these items will still be tracked in

the evaluation.) Item 4.1c is not scored because it is informational in nature—it does not indicate use of an EHR function. As noted above, the minimum requirement for receiving laboratory results may be met by 4.1j (which is never scored), or by any of 4.2g, 4.2h, or 4.2i (which are always scored). Similarly, the minimum requirement for receiving imaging results may be met by either 4.1m (which is never scored), or by any of 4.2j, 4.2k, or 4.2l (which are always scored).

3. Recodes to 0 if Stand-Alone Registry or E-Prescribing System Is Not Linked to EHR

Condition-specific "registry" items are asked of practices that indicate that they have a stand-alone registry or who use their EHR to identify patients with specific diagnoses or medications; identify patients overdue for specific therapies; facilitate prompt ordering of specific laboratory tests or recommended drugs; and facilitate prompt communication with patients requiring follow-up. Similarly, items about e-prescribing are asked of practices that indicated they have a stand-alone e-prescribing system or who e-prescribe through their EHR. However, when a practice has a stand-alone registry or stand-alone e-prescribing system, their responses to the related questions will only be counted in their OSS score if they also indicated the stand-alone systems they use are linked to their EHR. If the stand-alone systems are not linked, the practice's responses to the condition-specific function items or e-prescribing items will be recoded to 0 for purposes of scoring. This implements a CMS policy decision that these questions should only be counted if the stand-alone systems are linked in some way to the practice's EHR. Appendix A displays the threshold questions and the implications of various responses to them for asking and scoring the registry and e-prescribing items.

4. Domain Definitions and Scoring

Five domains were defined to represent the objectives of the functions queried in the OSS:

1. Completeness of Information
2. Communication About Care Outside the Practice
3. Clinical Decision Support
4. Increasing Patient Engagement/Adherence
5. Medication Safety

Each of these objectives is intuitively tied to care improvements. Some functions could contribute to more than one of the five objectives represented by the domains. However, in order to keep the scoring approach simple we included each question in only one domain representing its predominant objective. The predominant objective for each function was based on consensus among CMS and ASPE staff and CMS contractors. Use of a single predominant domain per question avoids complexity in understanding how any given function contributes to the score, and allows the questions in the OSS to be ordered by domain without repetition.

The steps to score each domain are (1) sum the points for each question within the domain, and (2) calculate the percentage of possible points achieved in each domain.

Note that the method for scoring each domain gives each function within the domain equal weight because all functions are scored on a 0 to 4 scale. The number of scored functions per domain ranges from 9 to 14.

## E. OSS SUMMARY SCORE

The OSS summary score will be calculated by multiplying each domain score by its weight, and summing the products. Domain weights were decided by CMS through consensus among involved CMS and ASPE staff, after considering input from MPR and ARC. The domain weighting scheme gives three domains slightly higher weights based upon CMS' understanding from a literature review conducted by ARC that at present, in general, evidence suggests the potential for savings from use of EHR functions related to electronic laboratory and radiology ordering, clinical decision support and medication safety checks (domains 2, 3 and 5).

Three additional points will be added to the summary score each year the practice uses a system with current certification. This is hoped to encourage practices to upgrade their systems as certifications expire, while allowing practices that do not choose to make such an upgrade to remain in the demonstration. The total score is capped at 100.

The example below assumes item points within each domain have been assigned, hierarchical and condition-specific items have been recoded and rescaled as described above, stand-alone system items have been recoded where applicable, and the sum has been calculated for each domain.

| A<br>Domain | B<br>Number of Items | C<br>Sum of Item Points/Possible Points | D<br>Domain Score (of a possible 100) (Column C x 100) | E<br>Domain Weight | F<br>Product (Column D X E) |
|---|---|---|---|---|---|
| 1. Completeness of Information | 9 | 34/36 | 94.4 | .17 | 16.0 |
| 2. Communication About Care Outside the Practice | 10 | 0/40 | 0.0 | .22 | 0.0 |
| 3. Clinical Decision Support | 14 | 53/56 | 94.6 | .22 | 20.8 |
| 4. Increasing Patient Engagement/Adherence | 11 | 20/44 | 45.5 | .17 | 7.7 |
| 5. Medication Safety | 9 | 15/36 | 41.7 | .22 | 9.2 |
| Total OSS Score (Rounded up to Nearest Integer) | | | | | 54* |
| Add 3 Points for Current Certification | | | | | 57* |

*Maximum score is 100.

F. DETERMINATION OF PAYMENT

CMS has decided to provide a minimum payment of $13.50 per beneficiary with chronic illness assigned to the practice for practices that meet the minimum criteria. (The maximum payment, as noted, is $45 per beneficiary.) This recognizes that the OSS contains an extensive list of EHR functions related to care improvement, a list that not every EHR may be capable of at the start of the demonstration, and that is unlikely to be implemented by a practice all at once or quickly. Without attention to this in the scoring or payment methodologies, practices may be discouraged from enrolling in the demonstration, feeling they would not have a reasonable chance to get more than a very small portion of the systems payment in the first two years. CMS set the minimum at 30 percent of the $45 total available amount. The OSS score will be applied to determine how much of the remaining $31.50 the practices will receive. The examples below illustrate how payment is calculated.

| | Practice #: | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| Minimum Criteria: | Met | Not Met | Met | Met |
| OSS Score: | 25 | 50 | 75 | 100 |
| EHR Certification: | | | | |
| Ever | Yes | Yes | Yes | Yes |
| Current | Yes | No | No | Yes |
| Adjusted OSS Score | 28 | na | na | 100* |
| Minimum Payment | $13.50 | $ 0.00 | $13.50 | $13.50 |
| OSS Score-Based Payment | $8.82 | na | $23.63 | $31.50 |
| Total Payment Per Beneficiary with Chronic Illness[6] | $22.32 | $0.00 | $37.13 | $45.00 |

*Maximum score is 100.

cc: Jennifer Schore, Lorenzo Moreno, Rachel Shapiro

[6] Actual payment per beneficiary may be lower if the physician or practice runs up against the demonstration caps on total physician or practice revenue from the demonstration.

APPENDIX A: OSS ITEMS RECODED TO ZERO IF PRACTICE USES STAND-ALONE REGISTRY OR E-PRESCRIBING NOT LINKED TO AN EHR

TABLE A.1

THRESHOLD QUESTIONS FOR REGISTRY AND E-PRESCRIBING ITEMS

| Registry: | Practice Response: | Implication: |
|---|---|---|
| 3.9a : Has your practice at this location implemented an EHR (rather than a stand-alone patient registry) to perform registry functions, such as tracking patients who have a specific chronic illness, or receive preventive care (that is, immunizations, mammography and other cancer screening) for at least one condition? (By "implemented" we mean an EHR has been purchased, installed, and tested, and is currently being used.) | Yes | The registry questions in Table A.2 are asked and scored |
| | No | Continue to item 3.9b |
| 3.9b: Has your practice at this location implemented a stand-alone patient registry to track patients who have a specific chronic illness, or receive preventive care (that is, immunizations, mammography and other cancer screening) for at least one condition? (By "implemented" we mean a registry has been purchased, installed, and tested, and is currently being used.) | Yes | Continue to item 3.9c |
| | No | The registry questions in Table A.2 are not asked, and thus receive a "0" score |
| 3.9c: Is this stand-alone patient registry linked with your EHR system? That is, do you electronically update the registry from the EHR system? | Yes | The registry questions in Table A.2 are asked and scored |
| | No | The registry questions in Table A.2 are asked, but the responses are recoded to "0" for purposes of payment |
| E-Prescribing | | |
| 3.15a: Has your practice at this location implemented an EHR to generate prescriptions? (By "implemented" we mean an EHR has been purchased, installed, and tested, and is currently being used.) | Yes | The e-prescribing questions in Table A.2 are asked and scored |
| | No | Continue to item 3.15b |
| 3.15b: Has your practice at this location implemented a stand-alone electronic prescribing system to generate prescriptions? (By "implemented" we mean an electronic prescribing system has been purchased, installed, and tested, and is currently being used.) | Yes | Continue to item 3.15c |
| | No | The e-prescribing questions in Table A.2 are not asked, and thus receive a "0" score |
| 3.15c: Is this stand-alone prescription system linked with your EHR system? That is, do you electronically update the prescription system from the EHR system? | Yes | The e-prescribing questions in Table A.2 are asked and scored |
| | No | The e-prescribing questions in Table A.2 are asked, but the responses are recoded to "0" for purposes of payment |

TABLE A.2

REGISTRY AND E-PRESCRIBING ITEMS WHOSE SCORING MAY BE AFFECTED BY THE THRESHOLD QUESTIONS LISTED IN TABLE A.1

| |
|---|
| Registry-Related Items |
| 4.1o. Record that instructions or educational information were given to patient |
| 4.3h. Review and act on reminders at the time of a patient encounter regarding interventions, screening, or follow-up office visits recommended by evidence-based practice guidelines |
| 4.4g. Produce hard copy or electronic reminders for patients about needed tests, studies, or other services (for example, immunizations) |
| 4.4h. Generate written or electronic educational information to help patients understand their condition or medication |
| 4.4i. Create written care plans (personalized to patient's condition or age/gender for preventive care) to help guide patients in self-management |
| 4.4j. Prompt provider to review patient self-management plan (or patient-specific preventive care plan) with the patient during a visit |
| 4.4k. Modify self-management plan (or patient specific preventive care plan) as needed following a patient visit |
| E-Prescribing Items |
| 4.1p. Recording (or entering) prescription medications (new prescriptions and refills) into electronic system [Minimum requirement, no score for payment purposes—minimum requirement not met if e-prescribing system not linked to EHR] |
| 4.2r. Print prescriptions (new prescriptions and refills) on a computer printer and fax to pharmacy or hand to patient |
| 4.2s. Fax prescription orders (new prescriptions and refills) electronically from electronic system |
| 4.2t. Transmit prescription orders (new prescriptions and refills) electronically directly from system to pharmacies that have the capability to receive such transmissions |
| 4.3i. Reference information on medications being prescribed |
| 4.3j. Reference guidelines and evidence-based recommendations when prescribing medication for a patient |
| 4.4l. Identify generic or less expensive brand alternatives at the time of prescription entry |
| 4.4m. Reference drug formularies of the patient's health plans/ pharmacy benefit manager to recommend preferred drugs at time of prescribing |
| 4.5b. Generate new prescriptions (that is, system prompts for common prescription details including medication type and name, strength, dosage, and quantity) |
| 4.5c. Generate prescription refills (that is, system allows provider to reorder a prior prescription by revising original details associated with it, rather than requiring re-entry) |
| 4.5d. Select individual medication for prescription (for example, from a drop-down list in the electronic system) |
| 4.5e. Calculate appropriate dose and frequency, or suggest administration route based on patient parameters such as age, weight, or functional limitations |
| 4.5f. Screen prescriptions for drug allergies against the patient's allergy information |
| 4.5g. Screen new prescriptions for drug-drug interactions against the patient's list of current medications |
| 4.5h. Check for drug-laboratory interaction |
| 4.5i. Check for drug-disease interaction |

# APPENDIX F

# OSS VALIDATION SURVEY AND FINDINGS

This page is intentionally left blank.

This appendix summarizes the validation study of the Office Systems Survey (OSS) for years 1 and 2 of the Electronic Health Records Demonstration (EHRD) evaluation.[2] For the year 1 data, the analysis was limited to the treatment group because the year 1 OSS did not collect data from the control group. The control group was included in the year 2 analysis as a face validity check of the treatment group practices' responses, not as a formal test of differences between the two groups. The purpose of the survey was to confirm practices' self-reported use of various electronic health record (EHR) functions in the OSS. Sections A and B describe the background and rationale for validating the OSS and summarize the validation survey instrument, respectively. Data collection is summarized in Section C, and the analysis plan is described in Section D. Sections E and F summarize findings and discuss their implications, respectively.

As described in detail below, our analysis did not find major discrepancies between the responses to the OSS and the validation survey. This encouraging finding suggests that treatment group and control group practices reported data on EHR use accurately and reliably.

## A. Background and Rationale for Validating the OSS

The EHRD was authorized under Section 402 Medicare Waiver Authority and was implemented by CMS. It expanded upon the Medicare Care Management Performance demonstration and was designed to test whether financial incentives (1) increased physician practices' adoption and use of EHRs, and (2) improved the quality of care practices deliver to chronically ill patients with fee-for-service Medicare coverage. The demonstration was expected to run for five years, but was cancelled on August 1, 2011, by CMS because practice attrition was substantial at the end of the second year of the demonstration; this greatly limited the feasibility of conducting a rigorous and unbiased evaluation of the demonstration's effectiveness by the end of the five-year demonstration. The year 1 OSS was administered only to treatment group practices; the year 2 OSS was administered to both control and treatment group practices.

Under the demonstration, treatment group practices had the opportunity to receive two types of payments in its first two years, a systems payment and a reporting payment.[3] The systems payment was an important underlying reason for this validation analysis, since payment amounts were determined in large part by practices' self-reports of use of EHR functions, as measured by the OSS. To receive the basic systems payment, practices had to use a minimum set of EHR functions in an EHR system certified either by certification organizations approved by the Office of the National Coordinator (ONC) or under the old Certification Commission for Health

[2] The OSS was fielded twice to treatment group (demonstration) practices, and once to control group practices. CMS terminated the demonstration effective August 1, 2011, due to a high rate of attrition among demonstration practices.

[3] Prior to termination of the demonstration, there were plans for practices to receive (1) a payment for reporting specific quality measures for year 2 with additional payment based on the OSS score; and (2) a payment for performance on the same quality measures for Years 3 to 5, with additional payment each year based on the OSS score. The demonstration ended prior to the scheduled reporting of quality measures for year 2, and, therefore, this payment was not provided. Practices that met the minimum use requirements for year 2 received the systems payment.

Information Technology (CCHIT) standards.[4] Practices met the minimum use requirements if they used a certified EHR to record visit notes, diagnostic test orders and results, and prescriptions. Practices that did not adopt or implement minimal use of a certified EHR system by the end of the first year of the demonstration did not receive a payment, but were permitted to remain in the demonstration. Practices that did not take either step by the end of the second year were removed from the demonstration.[5] Under the original demonstration design, beyond the basic systems payment, practices could have earned further systems payments for use of more sophisticated EHR functions, as measured by the OSS. The reporting payment would have been available for reporting on specific clinical quality measures for year 2. All incentive payments under the demonstration were to be made in addition to normal fee-for-service Medicare payments practices receive for submitted claims. Physicians could have received up to $13,000 and practices up to $65,000 over the first two years of the demonstration (Wilkin et al. 2007). Due to the termination of the demonstration, CMS only made the system payment for the first two years of the demonstration in fall 2010 and 2011.

At the start of the demonstration in June 2009, there were 412 treatment group practices and 413 control group practices enrolled. By the end of the EHRD's second year of implementation, 346 treatment and 389 control practices were participating in the demonstration, of which (1) 264 treatment and 188 control practices had completed the OSS and reported having an EHR, (2) 82 treatment and 201 control practices did not have an EHR, and (3) 35 treatment and 122 control practices did not complete the OSS. Of the 264 treatment practices that completed the OSS and had an EHR, 232 treatment practices met the minimum requirements to qualify for the systems payment, as described in the OSS scoring plan (Felt-Lisk 2009).[6]

Treatment practices could potentially inflate their OSS responses in order to receive more payment. Hence, in the first and second years of the demonstration, a validation of practices' responses was conducted to determine whether treatment practices were actually using the minimum functions as indicated in the OSS responses. In the second year of the demonstration, treatment practices' responses also were compared to those of control practices, which received no incentive payments and hence had no reason to inflate their responses. The comparison of treatment and control practices' responses could be envisioned as a face validity check of the treatment group practices' responses, not as a formal test of differences between the two groups. Treatment and control practices that were eligible for the validation survey were defined as those that had adopted and were currently using an EHR system and those that had completed the full survey, including complete data in Section 4 of the OSS instrument.

[4] In 2010, new "meaningful use" certification guidelines issued by ONC went into effect that allowed several organizations, including CCHIT, to certify EHRs. CMS followed these guidelines in the two years of the demonstration. Under the demonstration, practices were required to use an EHR that was certified either under the new standards or under the old CCHIT standards.

[5] A total of 79 treatment practices were removed from the demonstration at the end of the second year. Of these, 47 practices did not have a certified EHR, and 32 practices did not meet the minimum EHR use requirements.

[6] Practices included in the validation survey did not have to meet the minimum requirements; the only requirements for inclusion in the survey were that the practice completed the OSS and reported having an EHR.

## B. Validation Survey

Each year, after the OSS fielding period was ended, the evaluation team randomly selected a sample of practices (approximately one-quarter of treatment practices in year 1, and approximately one-quarter of treatment practices and half of control group practices in year 2) that completed the OSS and met the criteria for inclusion in the validation sample and asked them to complete a validation questionnaire.[7] This questionnaire asked for details about the use of specific EHR functions for small samples of patients, such as electronically recording clinical notes, laboratory orders, and results. The questionnaire was designed to collect information that could be compared with practices' reported EHR function use in the OSS.[8] This included information on practices' use of the full range of EHR functions, from use of basic functions meeting minimum demonstration requirements to use of advanced functions. In the validation survey, practices reported on several specific functions used in the past two weeks. A description of the key EHR topic areas addressed in the validation survey follows below.

Maintaining Clinical Notes. Practices were asked to select three dates in the two weeks prior to completion of the survey during which physicians at the practice saw more than five patients. Practices were then asked to verify whether or not at least 75 percent of the patients seen by a physician on each of those dates had an electronic clinical note.

Laboratory Results and Orders. Practices were asked to select the first three patients with encounters in the two weeks prior to completion of the survey for whom laboratory results were reported. Practices then documented, among these patients, the number for whom the laboratory results were received electronically.[9] They were then asked to document how the laboratory results were received by the electronic system—that is, whether the results were received by fax or mail, scanned, entered manually (keyboard entry), received by e-fax and transferred into the system, or received directly electronically into the system. The practices also answered the same set of questions for laboratory orders for these same three patients.[10]

Imaging Results and Orders. Practices were asked to select the first three patients with encounters during the two weeks prior to completion of the survey for whom imaging results

[7] In year 2, a larger sample of control group practices was randomly selected because a lower response rate was expected among control group practices than among treatment group practices.

[8] Mark Leavitt, CCHIT's former executive director, provided guidance on the development of the OSS validation instrument.

[9] In the analysis of results from the year 1 validation survey, evidence was found that practices completed the validation survey under the assumption that printing and faxing or receiving and scanning results/orders did not constitute electronic functions; however, these methods are indeed included in the definition of electronic receipt of orders and results. In response to this apparent misinterpretation by practices to the definition of electronic receipt of orders and results in the year 1 validation survey, the year 2 validation survey was revised to specify that electronic receipt means being received by fax or mail and scanned or entered manually (keyboard entry), received by e-fax and transferred into the system, or received directly electronically into the practice's system. This is the same definition that was used in the year 2 OSS.

[10] The response categories for how laboratory orders were sent were revised to the following for the year 2 validations survey: mailed, scanned, printed and faxed to lab or handed to patient, transferred directly via e-fax, or transmitted directly into the EHR.

(including radiology, MRI, and CAT scans) were reported. Practices followed the same procedure as with the questions related to laboratory results and orders.

Prescription Medication Orders. Practices were asked to select three patients with encounters during the two weeks prior to completion of the survey for whom prescription medications were refilled. Practices documented how many of these patients had the physician's refill order electronically documented in the system, as well as how the prescription orders were sent (printed and faxed or handed to patient, faxed electronically, or transmitted directly via the EHR).

For validation purposes, practices were required to fax to Mathematica screen shots of the EHR confirming each response to the validation survey, with de-identified patient information. In addition, practices provided the three dates selected for the Maintaining Clinical Notes questions (Questions 1a–1c) and partial patient identifiers (last four digits of patients' social security numbers) for all patients for whom data were reported for the survey. This information would allow CMS to verify that the submitted data were not fabricated if CMS were to conduct a complete audit of the practices' EHR records. At the end of the survey, practices were required to certify that their responses were accurate to the best of their knowledge and that they understood that the results could be subject to verification.

## C. Data Collection

Practices were eligible to be included in the survey sample if they met the following criteria: (1) they had adopted and were using an EHR system at the time they completed the OSS; and (2) they had completed the full OSS, including complete and scorable data in Section 4 of the instrument. The year 1 validation survey was administered to a stratified, random sample of 237 eligible practices between June 21 and July 13, 2010. The year 2 validation survey was administered to a stratified, random sample of 212 eligible treatment group practices and 205 eligible control practices between June 9 and July 21, 2011. Treatment practices were not included in the year 2 sample if they responded to the year 1 validation survey, with the exception of four treatment practices[11] because CMS wanted to have all treatment group practices respond to the validation survey at least once over the original duration of the demonstration.

In preparation for random selection, practices were first stratified by practice size and by site. There were three categories for practice size: 1 to 2 physicians, 3 to 5 physicians, and 6 or more. There were four sites: Louisiana, Maryland, Pennsylvania, and South Dakota. This resulted in 12 strata. In year 1, drawing a stratified, random sample of 25 percent of the eligible treatment practices[12] yielded 64 practices to solicit for the validation survey. Drawing a

[11] One practice had partially completed the year 1 validation survey. Another practice was allowed to revise its year 1 OSS responses after completing the year 1 validation survey. Two practices were requested by CMS to ensure that they had accurately implemented and used their EHRs.

[12] In October 2008, Mathematica proposed to CMS to administer the validation questionnaire to all treatment group practices. In January 2009, in response to requests from CMS, the plan was revised to propose sampling 25 percent of all treatment group practices. In response to subsequent requests from CMS and the feasibility of

stratified, random sample of 25 percent of the eligible treatment practices and 50 percent of the eligible control practices yielded 61 treatment practices and 105 control practices to solicit for the validation survey.[13] To ensure successful completion of the validation survey, reminder e-mails were sent (three in year 1 and four in year 2) and all nonrespondents were telephoned to encourage them to complete the survey. Mathematica staff successfully administered an online version of the validation survey to 62 treatment practices in year 1,[14] and 58 treatment practices and 69 control practices in year 2.[15]

## D. Analysis Plan

The analysis was twofold: (1) choosing relevant OSS questions and converting response categories of the OSS and the validation survey to a common scale, and (2) assigning discrepancy scores according to the magnitude of the discrepancy between the OSS response to a specific question (or group of questions) and the validation survey response to the equivalent question. Given that a relatively small number of practices was sampled and validation survey responses were based on having practices sample only three patients or three office working days, no statistical techniques, such as calculating confidence intervals around estimates or conducting formal hypothesis tests, could be applied. Instead, an agreement (or concordance) analysis was used, which, although simple, is very powerful for purposes of identifying systematic discrepancies between the responses of the same individual to two different measurement instruments. The analysis is described in more detail below.

### 1. Choosing Relevant OSS Questions

In both year 1 and year 2, practices' validation survey responses were compared to their OSS responses for six key minimum EHR functions: (1) clinical notes, (2) laboratory results, (3) laboratory orders, (4) imaging results, (5) imaging orders, and (6) prescription refill orders. For each of the six functions, the proportions of patients reported in the OSS were compared with the information on the different sets of three patients collected in the validation survey. The specific OSS and validation survey questions used for each of the six comparisons is presented in Table F.1.

---

*(continued)*

sampling OSS respondents, the design memo was further adjusted to sample 25 percent of treatment group practices with a complete OSS and scorable data.

[13] Twenty-five percent of eligible treatment practices and 50 percent of eligible control practices were sampled from each of the 12 strata, which resulted in samples with slightly more than 25 percent of the treatment group and 50 percent of the control group. A larger proportion of control group practices than treatment group practices were sampled because a much lower response rate was expected for the control group than the treatment group, as the control group did not receive any incentives under the demonstration.

[14] Two treatment practices were not included in the year 1 analysis. One treatment practice only partially completed the year 1 validation questionnaire and did not provide any screen shots, and a second practice was no longer in operation when the validation questionnaire was administered.

[15] Three additional practices (1 treatment and 2 control) partially completed the year 2 validation survey and were not included in the year 2 analysis.

Table F.1. OSS and Validation Survey Questions Used for the Validation Analysis

| OSS Question Number and Description | Validation Survey Question Number and Description |
|---|---|
| **Maintaining Clinical Notes for 75 Percent or More of Patients** | |
| 4.1d. Clinical notes prepared for individual patients | 1. Presence of an electronic clinical note for 75 percent or more of every patient seen during each date. |
| **Laboratory Results Received Electronically** | |
| 4.1j. Receiving laboratory results by fax or mail and scanning paper versions into electronic system<br>4.1k. Reviewing laboratory test results electronically | 2a. For how many of these patients was the laboratory result received electronically in the practice's system? |
| **Laboratory Orders Sent Electronically**[a] | |
| 4.2a. Print and fax laboratory orders to facilities outside the practice<br>4.2b. Fax laboratory orders electronically from system, or order electronically through a portal maintained by facilities outside the practice<br>4.2c. Transmit laboratory orders electronically directly from system to facilities outside the practice that have the capability to receive such transmissions | 2d. For how many of these patients was the laboratory order sent electronically? |
| **Imaging Results Received Electronically** | |
| 4.1m. Receiving imaging results by fax or mail and scanning paper versions into electronic system<br>4.1n. Reviewing imaging results electronically | 3a. For how many of these patients was the imaging result received electronically in the practice's system? |
| **Imaging Orders Sent Electronically**[b] | |
| 4.2d. Print and fax imaging orders to facilities outside the practice<br>4.2e. Fax imaging orders electronically from system, or order electronically through a portal maintained by facilities outside the practice<br>4.2f. Transmit imaging orders electronically directly from system to facilities outside the practice that have the capability to receive such transmissions | 3d. For how many of these patients was the imaging order sent electronically? |
| **Prescription Orders Sent Electronically** | |
| 4.2r. Print prescriptions on a computer printer and fax to pharmacy or hand to patient | 4ba. Prescription orders printed and faxed to pharmacy or handled by patient |
| 4.2s. Fax prescription orders (new prescriptions and refills) electronically from electronic system | 4bb. Prescription orders faxed electronically |
| 4.2t. Transmit prescription orders electronically directly from system to pharmacies that have the capability to receive such transmissions | 4bc. Prescription orders transmitted directly (electronically) |

[a] Practices' responses to the validation survey were updated to reflect that laboratory orders were submitted electronically for those practices that said in the validation survey that they did not submit laboratory orders electronically (validation survey question 2d = 0), but in later questions, responded in the affirmative to (1) printing and faxing (validation survey question 2ec = 1, 2, or 3), (2) faxing electronically (validation survey question 2ed = 1, 2, or 3), or (3) transmitting electronically directly from the system (validation survey question 2ee = 1, 2, or 3).

[b] Practices' responses to the validation survey were updated to reflect that imaging orders were submitted electronically for those practices that said in the validation survey that they did not submit imaging orders electronically (validation survey question 3d=0), but in later questions, responded in the affirmative to (1) printing and faxing (validation survey question 3ec = 1, 2, or 3), (2) faxing electronically (validation survey question 3ed = 1, 2, or 3), or (3) transmitting electronically directly from the system (validation survey question 3ee = 1, 2, or 3).

The selection of relevant OSS questions for conversion of the response categories faced several challenges relating to practices' understanding of the term "electronic." In completing the year 1 OSS, some practices appeared to have questions about what constituted electronic entry for laboratory and/or imaging orders. Specifically, after receiving their OSS summary reports in late July 2010, some practices had questions regarding the definition of recording orders and results for two OSS questions (OSS question 4.1.i, which asked for the proportion of patients for whom laboratory orders were recorded or entered into the electronic system; and OSS question 4.1.l, which asked for the proportion of patients for whom imaging orders were recorded or entered into the electronic system). CMS clarified the definition to include scanning as a valid form of entering orders into the electronic system. This update had implications for practices' OSS scores and incentive payments—30 practices failed to meet the minimum function requirements for receiving a systems payment only because of their response to one or both of these questions. CMS gave these practices the opportunity to update their responses to these questions only. The practices resubmitted their changes in revised form on August 27, 2010, to CMS.

Given the practices' confusion regarding the two OSS questions described above, and because the year 1 validation survey did not provide a definition of the term "electronic,"[16] it was hypothesized that practices completed the year 1 OSS and validation survey under the assumption that printing and faxing or receiving and scanning results/orders do not constitute electronic functions. The year 1 validation survey data provided evidence for this, as several practices reported that results and orders were not sent "electronically" (validation questions 2d and 3d), but reported in follow-up questions that these orders were (1) printed and faxed (validation survey questions 2ec and 3ec), (2) faxed electronically (validation survey questions 2ed and 3ed), or (3) transmitted electronically directly from the system (validation survey questions 2ee and 3ee). In the validation survey that practices accessed online, practices that had responded that the laboratory or imaging results were not received "electronically" (for any of the three patients who had results reported to the practice) were automatically skipped out of questions that asked for specifics on how these orders were electronically received.[17] Hence, the extent to which practices received results by fax/mail and scanned/entered the results manually (keyboard entry), received results by electronic fax and transferred them into the system, or received results directly electronically into the practice's electronic system could not be determined.

To address these challenges and the practices' confusion, answers were corrected among those practices that reported that laboratory and/or imaging orders were not sent "electronically," but reported in follow-up questions that these orders were indeed (1) printed and faxed, (2) faxed electronically, or (3) transmitted electronically directly from the system. That is, for these practices, their responses to the year 1 validation survey were updated to reflect that laboratory and/or imaging orders were submitted electronically. Hence, these practices were not penalized

[16] After consulting with CMS, this definition was subsequently provided in the year 2 OSS validation survey.

[17] The online validation survey was programmed such that, if a practice responded that none of the three patients who had laboratory and/or imaging results reported to the practice (that is, the response to questions 2a or 3a was zero), they were skipped out of questions 2b and 2c, and questions 3b and 3c, respectively.

for likely not understanding that printing and faxing or scanning results/orders were considered electronic functions in the validation survey.

Unfortunately, as with the analysis of the year 1 validation survey, in year 2, even though a clear definition of electronic receipt of lab and imaging results and orders was provided (in validation survey questions 2a and 3a), practices continued to be confused by the term "electronic." As a result of this confusion, it was concluded that practices completed the OSS and validation survey under the assumption that printing and faxing or receiving and scanning results/orders do not constitute electronic functions. For this reason, the same sets of questions excluded in year 1 also were excluded from the analysis in year 2, and the same procedure was used in year 2 (as in year 1) for correcting practices' responses to validation questions related to laboratory and imaging orders.

2. Converting Response Categories to a Common Scale

After selecting the relevant OSS questions, different response categories were converted to a common scale for all questions used in the analysis; this common scale was used for the analysis in both years 1 and 2. Because the OSS and validation survey employed different response categories, all comparisons required some transformation and regrouping of the data. OSS survey questions included five response categories: (1) none; (2) some, but less than 1/4; (3) 1/4 or more, but less than 1/2; (4) 1/2 or more, but less than 3/4; and (5) 3/4 or more. In contrast, the validation questionnaire allowed for only four response categories: (1) no patients; (2) one out of the three patients; (3) two out of the three patients; and (4) all three patients.

To create a common scale, each OSS response (which, as noted above, was in the form of a range, such as "1/4 or more, but less than 1/2") was converted into a single percentage by using the midpoint of the range. For example, a response of "1/4 or more, but less than 1/2" was converted to 37.5 percent. The five OSS responses thus became (1) zero, (2) 12.5 percent, (3) 37.5 percent, (4) 62.5 percent, and (5) 87.5 percent. For each topic area, these transformed responses were then summed to yield the percentage of patients for whom the practice performed any of the functions in the topic area.[18] This percentage was then multiplied by three (since the validation survey asked about sets of three patients), and rounded to a whole number to calculate the number of patients out of three (0, 1, 2, or 3 patients) for whom the practice would have performed the function described in validation survey. Once the OSS responses were converted to the same patient scale as the responses in the validation survey, the results of the two instruments could be compared. These transformed OSS responses are referred to as the "OSS comparable responses."

For the comparisons of each set of questions described in Section B above, only those discrepancies in which the OSS comparable response was higher (that is, indicated greater use of

[18] After excluding all OSS questions about printing and faxing or scanning orders/results, some topic areas contained only one relevant OSS question. For example, for the laboratory results topic area, practices estimated the proportion of patients for whom they: (1) receive laboratory results by fax or mail and scan paper versions into the electronic system and (2) review laboratory test results electronically. Because the first question was excluded, practices' responses to the second OSS question were simply compared to their validation survey response.

a functionality) than the validation survey response were flagged. Discrepancies in which the opposite was true (that is, those in which the validation survey indicated a higher level of use of a function than did the OSS survey) were not flagged. The rationale for this is that, due to the secular trend in use of EHRs, using functions for more patients than reported in the OSS is a desired outcome and should not be penalized as a discrepancy.

3. Assigning Discrepancy Scores

A simple scoring system was devised to quantify the degree of discrepancy between OSS and validation survey responses; this scoring system was used in the analysis in both years 1 and 2. Specifically, if a practice reported in the validation survey that a function was used for zero patients, an OSS comparable response of up to one patient (or between 17 percent and 49 percent of patients) was considered a small discrepancy and given a score of one. An OSS comparable response of up to two patients (or between 50 percent and 82 percent of patients) was considered a medium discrepancy and given a score of two. Finally, an OSS comparable response of three patients (83 percent or more of patients) was a large discrepancy and assigned a score of three. Table F.2 shows all combinations of discrepancies between responses to the OSS and the validation survey, and illustrates how scores are distributed among these discrepancies.

Table F.2. Scoring System for Agreement of Responses Between the OSS and the Validation Questionnaire

| | | Response to Validation Questionnaire | | | |
|---|---|---|---|---|---|
| Comparable Response to OSS Survey Items | No patients | No patients Match | 1 patient Validation response higher | 2 patients Validation response higher | 3 patients Validation response higher |
| | Roughly 1 patient (between 17% and 49%) | Flag for discrepancy (1 point)* | Match | Validation response higher | Validation response higher |
| | Roughly 2 patients (between 50% and 82%) | Flag for discrepancy (2 points)* | Flag for discrepancy (1 point)* | Match | Validation response higher |
| | Roughly 3 patients (83% or more) | Flag for discrepancy (3 points)* | Flag for discrepancy (2 points)* | Flag for discrepancy (1 point)* | Match |

Source: Mathematica Evaluation Team.

Note: Entries with an asterisk (*) correspond to the discrepancy sizes used for scoring the discrepancy between the OSS and validation responses.

Discrepancy scores of between 0 and 3 for each of the six topic areas were then added to produce a total discrepancy score of between 0 and 18. Prior to the analysis, a total discrepancy score of 8 or above was defined as indicative of substantial discrepancies between a practice's OSS responses and validation survey responses, which were based on practices' reports of actual use of the EHR functionalities. The cutoff of 8 was somewhat arbitrarily selected between (1) a low level of discrepancy among the OSS and the validation surveys, and (2) a high level of discrepancy among the surveys. A low level of discrepancy (scores below or equal to 8) signaled minor differences in OSS and validation responses that were probably due to small

samples of patients selected for the validation survey and somewhat different time frames used in the OSS versus the validation survey. A high level of discrepancy (scores above 8) signaled major differences in the OSS and validation responses that were probably due to systematic differences between practices' reported functions in the OSS and their performed functions recorded in the validation survey.

## E. Results

Under the scoring system, of those practices that completed the validation survey in years 1 and 2, only one treatment practice in year 2 had a discrepancy score of 8 or higher (Table F.3). A careful review of the practice's responses to both the year 2 OSS and the year 2 validation survey found that this practice did not understand what constituted electronic receipt of lab and imaging orders and results when completing the OSS and the validation survey. As a result, when completing the OSS, this practice overstated its ability to electronically receive lab and imaging results and send lab and imaging orders.[19]

Table F.3. Distribution of Total Discrepancy Scores, by Study Group

| | Treatment Group Practices | | | Treatment Group Practices | | | Control Group Practices | | |
|---|---|---|---|---|---|---|---|---|---|
| Total Discrepancy Score | Year 1 Number of Practices | Percent-age | Cumulative Percentage | Year 2 Number of Practices | Percent-age | Cumulative Percentage | Year 2 Number of Practices | Percent-age | Cumulative Percentage |
| 0 | 37 | 60 | 60 | 39 | 67 | 67 | 38 | 55 | 55 |
| 1 | 9 | 15 | 74 | 9 | 16 | 83 | 16 | 23 | 78 |
| 2 | 5 | 8 | 82 | 4 | 7 | 90 | 4 | 6 | 84 |
| 3 | 6 | 10 | 92 | 5 | 9 | 98 | 5 | 7 | 91 |
| 4 | 1 | 2 | 94 | 0 | 0 | 98 | 4 | 6 | 97 |
| 5 | 1 | 2 | 95 | 0 | 0 | 98 | 1 | 1 | 99 |
| 6 | 3 | 5 | 100 | 0 | 0 | 98 | 0 | 0 | 99 |
| 7 | 0 | 0 | 100 | 0 | 0 | 98 | 1 | 1 | 100 |
| 8 | 0 | 0 | 100 | 0 | 0 | 98 | 0 | 0 | 100 |
| 9 | 0 | 0 | 100 | 1 | 2 | 100 | 0 | 0 | 100 |

Source: Year 1 OSS and Year 1 OSS Validation Survey, 2010; Year 2 OSS and Year 2 OSS Validation Survey, 2011.

There were few reported discrepancies among both treatment and control practices, which leads to the conclusion that the treatment group practices (in both year 1 and year 2) reported their usage of EHR functions in the OSS as accurately as control group practices (in year 2), which did not have the incentive to inflate their responses. More than half of the treatment group practices in both years (60 percent in year 1 and 67 percent in year 2) and more than half of the control group practices (55 percent in year 2) had total discrepancy scores of 0. About one-third (31 percent in year 1 and 33 percent in year 2) of treatment group practices and 42 percent of control group practices (in year 2) had discrepancy scores between 1 and 4, and only a small

[19] While the practice reported in the OSS that lab and imaging results were received electronically for three-quarters or more of its patients, and that lab and imaging orders were sent electronically for the majority of its patients (via printing and faxing, e-faxing, and direct electronic transmission), in its responses to similar questions in the validation survey, the practice reported lesser use of these electronic functions (for printing and faxing of results and orders—for only one out of three patients).

number of practices (4 treatment practices in year 1, 1 treatment practice in year 2, and 2 control practices in year 2) had scores above 4. For additional verification of our findings, a Chi-squared test was conducted; it confirmed that there were no significant differences in the distribution of discrepancy scores between treatment and control group practices.

## F. Discussion

The analysis found that, in both year 1 and year 2, treatment practices' OSS responses were largely similar to their responses to the validation survey (Table F.2). Because practices were required to attest to the validity of their answers and provide screen shots for each question of the validation survey, it is probable that practices' responses to the OSS and validation survey reflect their actual uses of the EHR, with the exception of the one treatment practice in year 1 that had a discrepancy score of 9. This finding suggests that, with this one exception, treatment practices did not inflate their responses to the OSS, and therefore received the appropriate incentive payment for their participation in the demonstration and use of EHRs. In addition, the findings show that control practices were also unlikely to inflate their answers to the OSS; this is to be expected, as they did not have the option to receive an incentive payment, and therefore had no interest in inflating their OSS responses in order to receive more payment.

In conclusion, it appears that a relatively simple accountability and transparency instrument can address CMS's needs efficiently and reliably. However, there remains confusion among practices as to the precise definition of the term "electronic." The results of both the OSS and the validation survey suggest that there are challenges in collecting reliable data on the use of EHRs, particularly when that use determines the level of incentive payments a practice would be eligible to receive.

This page is intentionally left blank.

# APPENDIX G

# DETAILS ON SAMPLE, MEASURES, DATA SOURCES, AND METHODS

This page is intentionally left blank.

# I. SAMPLE

This section describes the sample, including the rules for assignment of beneficiaries to treatment and comparison practices, as well as the number of practices enrolled in the demonstration and included in implementation and impact analyses.

## A. Rules for Assignment of Beneficiaries to Practices

The main sample consists of fee-for service (FFS) Medicare beneficiaries who were assigned by the implementation support contractor to demonstration or control practices at the end of the baseline period (2008), at the end of Year 1 (June 1, 2009, through May 31, 2010), and/or at the end of Year 2 (June 1, 2010, through May 31, 2011) based on claims submitted by those beneficiaries during each period.

Specifically, each beneficiary with the target chronic conditions (coronary artery disease [CAD], chronic heart failure [CHF], and diabetes) or other specified chronic conditions (Alzheimer's disease or other mental, psychiatric, or neurological disorders; any chronic cardiac/circulatory disease, such as arteriosclerosis, myocardial infarction, or angina pectoris/stroke; any cancer; arthritis and osteoporosis; kidney disease; and lung disease) was identified through the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis codes available in Medicare claims data (Wilkin et al. 2007). Each beneficiary who had one or more of these chronic conditions was assigned to the practice at which he or she received the plurality of evaluation and management (E&M) services.[20] In cases in which two or more practices provided an equal number of visits, the beneficiary was assigned to the practice with the most recent E&M visit. Finally, because many outcome measures were defined over a year-long period, beneficiaries were excluded (that is, not assigned to a practice) if they were not eligible for at least six months of the year. A beneficiary was not eligible during a particular month if he or she:

- Died
- Relocated out of the demonstration state
- Lacked either Part A or Part B coverage
- Had Medicare as the secondary payer due to work status
- Elected hospice coverage
- Enrolled in a Medicare Advantage or a Medicare coordinated care plan for more than six months of the demonstration year

[20] E&M services were identified using the following Current Procedural Terminology Codes in Medicare claims data: 99201–99215 (office or other outpatient service), 99301–99316 (nursing facility service), 99321–99333 (domiciliary, rest home, boarding home, or custodial care service), 99341–99350 (home service), 99381–99397 (preventive medicine service), 99401–99429 (counseling and/or risk factor reduction intervention) (Wilkin et al. 2007).

Table G.1 reports total sample numbers and beneficiary characteristics by state for treatment and comparison groups.

## B. Practice Enrollment and Analysis Sample

At the start of the demonstration in June 2009, 412 treatment group practices and 413 control group practices were enrolled. By the end of the EHRD's first year of implementation, 363 treatment and 405 control practices were participating in the demonstration, and 338 of these participating treatment practices completed the 2010 OSS. By the end of the EHRD's second year of implementation, 346 treatment and 389 control practices were participating in the demonstration, and 311 of these participating treatment practices completed the 2011 Office Systems Survey (OSS). The incentive payment analysis includes only the subset of participating treatment practices that submitted data in both years (see Chapter IV).

The claims-based impact analysis for the final evaluation of the EHRD demonstration includes 412 treatment and 410 control practices (Table G.2). These are all the practices for which the evaluation team has Medicare claims for at least one year. (Only 3 control practices randomized under the demonstration did not have any Medicare claims for any demonstration year, and were thus excluded from the analysis.)

Table G.1. Beneficiary Characteristics, Service Use, Medicare Expenditures, and Quality of Care for Beneficiaries Assigned to Practices at Baseline, by Site (Percentages, Unless Otherwise Noted)

| | Louisiana | | Maryland | | Pennsylvania | | South Dakota | |
|---|---|---|---|---|---|---|---|---|
| | Treatment Group | Control Group | Treatment Group | Control Group | Treatment Group | Control Group | Treatment Group | Control Group |
| Characteristic | | | | | | | | |
| Age in Years | | | | | | | | |
| 64 or younger | 21.20 | 20.28 | 13.46 | 13.29 | 13.62 | 13.24 | 11.37 | 10.73 |
| 65 to 79 | 55.50 | 54.18 | 58.82 | 59.09 | 45.95 | 45.44 | 54.31 | 54.53 |
| 80 or older | 23.30 | 25.54 | 27.72 | 27.62 | 40.43 | 41.32 | 34.32 | 34.74 |
| Male | 38.16 | 37.89 | 38.84 | 36.63 | 37.86 | 36.34 | 39.15 | 37.85 |
| Enrolled in an HMO | 2.65 | 2.91 | 2.86 | 2.84 | 6.75 | 7.22 | 1.12 | 0.86 |
| Reason for Medicare Eligibility | | | | | | | | |
| Age | 73.44 | 74.67 | 84.47 | 85.12 | 82.79 | 83.09 | 85.10 | 85.99 |
| Disability (Including ESRD) | 26.56 | 25.33 | 15.53 | 14.88 | 17.21 | 16.91 | 14.90 | 14.01 |
| Presence of Chronic Conditions[a] | | | | | | | | |
| CAD | 25.54 | 24.16 | 21.08 | 20.64 | 27.16 | 27.75 | 18.83 | 18.48 |
| CHF | 12.69 | 12.44 | 9.46 | 8.61 | 12.47 | 11.91 | 10.59 | 10.43 |
| Diabetes | 37.21 | 36.28 | 34.65 | 34.46 | 32.47 | 32.08 | 30.59 | 29.08 |
| Other chronic conditions | 95.57 | 95.52 | 96.32 | 96.40 | 95.49 | 96.04 | 92.31 | 93.08 |
| Beneficiary Race/Ethnicity | | | | | | | | |
| African American | 26.05 | 27.58 | 20.75 | 23.58 | 4.73 | 6.48 | 0.15 | 0.17 |
| Hispanic | 72.50 | 71.26 | 76.70 | 74.33 | 94.41 | 92.70 | 98.31 | 98.28 |
| White | 0.32 | 0.36 | 0.35 | 0.42 | 0.06 | 0.03 | 0.14 | 0.08 |
| Other | 1.04 | 0.72 | 2.10 | 1.63 | 0.71 | 0.71 | 1.30 | 1.40 |

| | Louisiana | | Maryland | | Pennsylvania | | South Dakota | |
|---|---|---|---|---|---|---|---|---|
| | Treatment Group | Control Group | Treatment Group | Control Group | Treatment Group | Control Group | Treatment Group | Control Group |
| Medicare Expenditures and Service Use | | | | | | | | |
| Total Medicare Expenditures (Dollars) | 12,277.01 | 11,656.25 | 11,774.88 | 11,202.88 | 12,312.99 | 11,530.45 | 8,583.80 | 8,637.25 |
| Number of Hospitalizations | 0.51 | 0.48 | 0.45 | 0.42 | 0.58 | 0.53 | 0.42 | 0.44 |
| Number of ER Visits | 0.61 | 0.55 | 0.40 | 0.41 | 0.49 | 0.48 | 0.48 | 0.46 |
| Number of Physician Visits | 19.64 | 19.07 | 20.68 | 20.57 | 20.98 | 20.83 | 17.63 | 18.50 |
| Quality-of-Care Measures | | | | | | | | |
| Among Beneficiaries with CAD (N=48,229): | | | | | | | | |
| Any blood test for cholesterol or lipids | 82.71 | 80.67 | 86.85 | 87.34 | 74.09 | 78.72 | 76.39 | 77.57 |
| Among Beneficiaries with Diabetes (N=46,908): | | | | | | | | |
| Any blood test for hemoglobin A1c | 85.99 | 85.05 | 89.89 | 89.26 | 85.07 | 87.25 | 92.27 | 92.96 |
| Any blood test for cholesterol or lipids | 84.79 | 81.10 | 89.71 | 89.87 | 83.00 | 85.72 | 86.69 | 84.57 |
| Any urine test for protein (microalbuminuria) | 76.89 | 77.97 | 82.72 | 83.35 | 80.76 | 80.13 | 82.81 | 83.91 |
| Among Female Beneficiaries Between Ages 40 and 69 (N=27,137): | | | | | | | | |
| Any screening for breast cancer | 60.06 | 63.44 | 66.86 | 68.02 | 59.79 | 61.21 | 69.72 | 68.81 |
| Number of Beneficiaries[b] | 36,963 | 31,603 | 54,243 | 50,201 | 31,457 | 29,668 | 23,217 | 24,462 |
| Number of Practices[c] | 103 | 97 | 125 | 127 | 138 | 141 | 43 | 44 |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries who were assigned to treatment (demonstration) or control practices at the end of the baseline period (CY 2008) of the demonstration.

[a]The sum of the chronic condition percentages is greater than 100 because beneficiaries can have more than one condition.

[b]The number of beneficiaries assigned to practices at baseline (CY 2008) are those beneficiaries with nonmissing information for the regression control variables (see Table G.2).

[c]The number of practices with baseline data is 818, or 7 less than the total number of 825 randomized practices. This is the full number of randomized practices with assigned beneficiaries during the baseline period.

CAD = coronary artery disease; CHF = congestive heart failure; COPD = chronic obstructive pulmonary disease; CY = calendar year; ER = emergency room; ESRD = end stage renal disease; EHRD = Electronic Health Records Demonstration; HMO = health maintenance organization.

Table G.2. Number of Practices and Assigned Beneficiaries per Study Group at Baseline, Year 1 and Year 2 in the Evaluation of the EHRD

| | Treatment | Control | Total |
|---|---|---|---|
| Number of Practices Randomized | 412 | 413 | 825 |
| Final Number of Practices Included in Claims-Based Analysis | 412 | 410 | 822 |
| Number of Practices Included in OSS[a] Analysis[b] | | | |
| Year 2 | 314 | 267 | 581 |
| Number of Practices Included in Incentive Payment Analysis | | | |
| Year 1 | 338 | NA | 338 |
| Year 2 | 311 | NA | 311 |
| Number of Beneficiaries Included in Claims-Based Analysis | | | |
| Baseline | 145,880 | 135,934 | 281,814 |
| Year 1 | 139,812 | 126,681 | 266,493 |
| Year 2 | 132,038 | 120,179 | 252,217 |

Source: Office Systems Survey (OSS), conducted in spring and summer 2011. Medicare claims data for the baseline period (CY 2008), and Year 1 and Year 2 of the demonstration; payment data for Year 1 and Year 2, provided by EHRD's implementation support contractor.

[a]Excludes the three control practices that had no Medicare claims for any demonstration year.

[b]Consists of all practices that completed the 2011 OSS, regardless of their participation at the time of the surveys.

[c]Consists of those eligible and participating treatment practices that completed the 2010 and 2011 OSS.

EHRD = Electronic Health Records Demonstration; NA = not applicable.

# II. MEASURES AND DATA SOURCES

This section describes the types of outcome measures and the accompanying data sources used in this report. These measures are related to electronic health record (EHR) and health information technology (IT) use; quality of care; potentially preventable hospitalizations; and Medicare expenditures and service use, incentive payments, and practice characteristics.

## A. EHR and Health IT Use

The EHR demonstration aimed to improve quality of care by increasing adoption and use of EHRs. Mathematica constructed OSS system scores, by domain, to summarize the use of select EHR functions. Some key EHR functions include recording patient data, ordering prescriptions, and ordering laboratory and other diagnostic tests and recording their results. The OSS also measures the practices' systems capability to send automated alerts, plan patient care, send patient reminders, and facilitate patient education. The OSS scores measure the degree to which a practice has used these functions. The evaluation assessed the impact of the demonstration payments on these scores for Year 2 of the demonstration, the year for which OSS data were collected for both treatment and control group practices.

Fifty-three EHR functions were scored based on practices' responses to questions on the 2010 and 2011 OSS. Most functions were scored on a 0 to 4 (five-point) scale. The response choices for most questions translated directly into practices' score for the function, with 0 being less desirable and representing no use of a function, and 4 indicating the function is used for "3/4 or more" of patients. In addition, the following five domains—or groups of several individual functions—were defined to represent the objectives of the functions queried in the OSS:

1. Completeness of Information
2. Communication About Care Outside the Practice
3. Clinical Decision Support
4. Increasing Patient Engagement/Adherence
5. Medication Safety

Each of these objectives is intuitively tied to care improvements. Some functions could contribute to more than one of the five objectives represented by the domains. However, in order to keep the scoring approach simple, each question was classified in only one domain representing its predominant objective. The predominant objective for each function was based on consensus among CMS and ASPE staff and CMS contractors. The steps used to score each domain were to: (1) sum the points for each question within the domain, and (2) calculate the percentage of possible points achieved in each domain. Note that the method for scoring each domain gives each function within the domain equal weight because all functions are scored on a 0 to 4 scale. The number of scored functions per domain ranges from 9 to 14.

In addition, the OSS summary score was calculated by multiplying each domain score by its weight, and summing the products. Domain weights were decided by CMS through consensus among involved CMS and ASPE staff, after considering input from Mathematica and ARC. The domain weighting scheme gives three domains slightly higher weights; this is based upon CMS's

understanding from a literature review conducted by ARC that at present, in general, evidence suggests the potential for savings from use of EHR functions related to electronic laboratory and radiology ordering, clinical decision support, and medication safety checks (domains 2, 3, and 5). See the OSS scoring plan (Appendix E) for more information regarding how OSS overall and domain scores were calculated.

## B. Quality-of-Care Measures

The demonstration was expected to improve physicians' adherence to recommended care guidelines for 26 quality-of-care measures that would be rewarded by incentive payments (Table G.3). Seven of these measures could be observed in the Medicare claims data. Under the demonstration's original design, one of the demonstration's quality data collection contractors would construct these seven measures. However, the demonstration was terminated before the demonstration quality data collection contractor constructed the measures. Mathematica implemented the code used by the data collection contractor to construct the same claims-based measures for Medicare Care Management Performance demonstration (MCMP) (Trisolini et al. 2007), and was able to reliably create five of the seven claims-based measures.[21] Mathematica also constructed several proxies for quality of care related to preventable hospitalizations. Table G.4 describes quality-of-care measures, quality of care proxies (such as preventable hospitalizations), and expenditure measures that were drawn from Medicare claims data for the analyses included in this report.

The demonstration's data collection support contractor retroactively assigned beneficiaries to treatment or control practices at the end of the baseline period, Year 1, or Year 2. Using the specifications in Trisolini et al. (2007), Mathematica determined whether assigned beneficiaries were eligible for each measure (for example, whether they had the qualifying diagnosis and were within the specified age range) and should therefore be included in the denominator for that measure. For both the treatment and the control group, all Medicare beneficiaries eligible for a particular measure according to their Medicare claims data were included in the denominator of that quality measure for this analysis.

The numerators of the claims-based quality measures were defined according to whether the beneficiary received appropriate tests. For example, one claims-based quality measure was defined as the percentage of beneficiaries with CAD who received a blood test for cholesterol or lipids (Table G.4).

[21] The diabetes dilated exam measure for diabetes patients could not be reliably constructed because the Current Procedural Technology Category II (CPT II) codes used to identify the "Patient identified as low risk for retinopathy (no evidence of retinopathy in the prior year)" portion of the Eye Exam criteria was not provided for individuals in our data file. Similarly, the left ventricular function quality measure for CHF patients could not be coded because there was a substantial change in the codes relating to cardiac ejection fractions and catheterization since the MCMP quality measures specifications were written.

Table G.3. The 26 Quality Measures Incentivized by the MCMP Demonstration, Measure Owners, Data Sources, and Availability for Comparison Group Practices

| Quality Measure | Measure Owner | Data Source: Medical Records | Data Source: Medicare Claims | Data Available for Control Group Practices |
|---|---|---|---|---|
| Whether Patients with CAD: | | | | |
| Were prescribed antiplatelet therapy | AMA | X | | No |
| Were prescribed a lipid-lowering therapy | AMA | X | | No |
| Were prescribed beta-blocker therapy, among those with prior myocardial infarction | AMA | X | | No |
| Received a lipid profile | AMA | X | X | Yes |
| Had most recent LDL cholesterol <130 mg/dl | CMS | X | | No |
| Were prescribed ACE inhibitor therapy, among those who also have diabetes and/or LVSD | AMA | X | | No |
| Whether Patients with CHF: | | | | |
| Had left ventricular function results recorded | AMA | X | | No |
| Had left ventricular ejection fraction tested (among those hospitalized with heart failure) | CMS | X | X | No |
| Had weight measurement recorded | AMA | X | | No |
| Had patient education class on disease management and health behavior change during one or more visits within a six-month period | AMA | X | | No |
| Were prescribed beta-blocker therapy, among those who also have LVSD | AMA | X | | No |
| Were prescribed ACE inhibitor therapy, among those who also have LVSD | AMA | X | | No |
| Were prescribed warfarin therapy, among those with paroxysmal or chronic atrial fibrillation | AMA | X | | No |
| Whether Patients with Diabetes: | | | | |
| Had blood test for HbA1c | NCQA | X | X | Yes |
| Had most recent A1c level >9 percent[e] | NCQA | X | | |
| Had blood pressure below 140/99 mm Hg | NCQA | X | | |
| Had LDL cholesterol test | NCQA | X | X | Yes |
| Had most recent LDL cholesterol <130 mg/dl | NCQA | X | | |
| Had test for microalbumin | NCQA | X | X | Yes |
| Had dilated retinal exam | NCQA | X | X | No |
| Had foot exam | NCQA | X | | |
| Whether Patients with Specified Chronic Conditions[a] Received Preventive Care Measures, Including: | | | | |
| Blood pressure measurement during last office visit | AMA | X | | No |

| Quality Measure | Measure Owner | Data Source: Medical Records | Data Source: Medicare Claims | Data Available for Control Group Practices |
|---|---|---|---|---|
| Breast cancer screening during current or previous year, among those younger than 69 | NCQA | X | X | Yes |
| Colorectal cancer screening during recommended period | AMA | X | | No |
| Influenza vaccination during September through February of year prior to measurement year, among those older than 50 | AMA | X | | No |
| Pneumonia vaccination, among those with a chronic condition older than 65 | NCQA | X | | No |

Source: Moreno et al. (2007).

[a]In addition to three primary target chronic conditions—CHF, CAD, and diabetes mellitus—the other specified eligible conditions are Alzheimer's disease or other mental, psychiatric, or neurological disorders; any chronic cardiac/circulatory disease (such as arteriosclerosis, myocardial infarction, or angina pectoris/stroke); any cancer; arthritis and osteoporosis; kidney disease; and lung disease. These conditions were identified through International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis codes available in Medicare claims data (Wilkin et al. 2007).

ACE = angiotensin-converting enzyme inhibitor; AMA = American Medical Association; CAD = coronary artery disease; CHF = congestive heart failure; CMS = Centers for Medicare & Medicaid Services; LDL = low-density lipoprotein; LVSD = left ventricular systolic dysfunction; NCQA = National Committee for Quality Assurance.

Table G.4. Claims-Based Outcome Measures Used in the Analysis of the EHRD Demonstration

| Measure | Data Collection Method |
|---|---|
| **Quality-of-Care Measures[a]** | |
| Among beneficiaries with CAD, received:<br>A lipid profile<br><br>Among beneficiaries with diabetes, had:<br>Blood test for hemoglobin A1c<br>LDL cholesterol test<br>Urine test for protein (microalbuminuria)<br><br>Among femaiale beneficiaries ages 40 to 69, had:<br>Breast cancer screening | Medicare claims data provided by EHRD's implementation support contractor for beneficiaries assigned to treatment (demonstration) or control practices at the end of either the baseline period (2008), Year 1 (June 1, 2009, through May 31, 2010),or Year 2 (June 1, 2010, through May 31, 2011) |
| **Potentially Preventable Hospitalizations[b]** | |
| Whether a beneficiary had any preventable hospitalizations related to:<br>Cardiac problems (for those with CAD)<br>CHF complications or fluid/electrolyte problems (for those with CHF)<br><br>Cardiac problems, diabetes complications, microvascular complications, or peripheral vascular or extremity complications (among those with diabetes)<br><br>Number of the following types of hospitalizations:<br>Cardiac hospitalizations (for those with CAD or diabetes)<br>CHF hospitalizations (for those with CHF)<br>Diabetes hospitalizations (for those with diabetes)<br>Hospitalizations for peripheral vascular or extremity complication (for those with diabetes)<br><br>Hospitalizations for microvascular complications (for those with diabetes) | Medicare claims data provided by EHRD's implementation support contractor for beneficiaries assigned to treatment (demonstration) or control practices at the end of either the baseline period (2008), Year 1 (June 1, 2009, through May 31, 2010), or Year 2 (June 1, 2010, through May 31, 2011) |
| **Medicare Expenditure and Service-Use Measures** | |
| Total Part A and Part B Medicare expenditures<br><br>Expenditures by type of service:<br>Inpatient hospitalization<br>Skilled nursing facility<br>Hospice<br>Physician office visit<br>Outpatient visit<br>Home health service<br><br>Service-use measures for:<br>Inpatient hospitalization<br>Emergency room visit<br>Physician office visit<br><br>Outpatient visit | Medicare claims data provided by EHRD's implementation support contractor for beneficiaries assigned to treatment (demonstration) or control practices at the end of either the baseline period (2008), Year 1 (June 1, 2009, through May 31, 2010), or at Year 2 (June 1, 2010, through May 31, 2011) |

[a]While 7 of the quality-of-care measures could be constructed from the Medicare claims data, only 5 of the 26 quality-of-care measures could be constructed from the available programming specifications. Detailed information on the source and construction of each of the 26 measures is available in Trisolini et al. (2007).

[b]Measures defined based on Medicare claims data and previously used in the evaluation of other CMS demonstrations, including the evaluation of the LifeMasters Supported Selfcare Demonstration (Esposito et al. 2008).

CAD = coronary artery disease; CHF = congestive heart failure; EHRD = Electronic Health Records Demonstration; LDL = low-density lipoprotein.

## C. Potentially Preventable Hospitalizations

Other outcome measures related to care quality included preventable hospitalizations, which were constructed from Medicare claims data. A beneficiary was considered to have had a potentially preventable hospitalization if he or she had any of the types of hospitalizations shown in Table G.4 during the demonstration period. A ll beneficiaries identified by the demonstration's implementation support contractor, ARC, as having CAD, CHF, or diabetes were included in the denominator for this measure.

## D. Medicare Expenditures and Service Use

The implementation support contractor provided Medicare claims data for beneficiaries assigned to demonstration and control practices at the end of the baseline period, Year 1, or Year 2. These data were used to construct annualized outcome measures for the relevant year related to expenditures and service use for beneficiaries assigned to treatment or control practices. These measures include total Medicare expenditures; Medicare expenditures by certain types of service; and use of selected Medicare services, such as emergency room visits and hospitalizations (Table G.4). Because Medicare home health services were provided under both Part A and Part B, each Medicare service type was classified into one of three broad categories: Part A (other than home health), Part B (other than home health), or home health services.

## E. Payment Measures

The demonstration's implementation support contractor provided data on i ncentive payments received by practices during each demonstration year. Data on baseline incentive payments for reporting EHR use were available for 338 treatment practices (of the 412 enrolled in the demonstration as of July 1, 2009) that submitted performance data for Year 1 b y completing the OSS. In addition, incentive payment data were available for 311 t reatment practices that completed the OSS in Year 2.

## F. Measures of Practice Characteristics Drawn from Application Data

Demonstration and control practices completed the EHRD application (provided in Appendix H). T he application asked about practice characteristics as well as practices' use of particular health IT functions. A pplication data were used to construct regression control variables and describe practice characteristics in the implementation analysis.

# III. METHODS

This section discusses the methods used to randomly assign practices to treatment and control groups, identify physicians and beneficiaries within practices, and estimate changes in key outcomes during the demonstration.

## A. Research Design

A key method for ensuring that valid estimates of impacts are obtained is the comparison group strategy–that is, identifying a sample of practices that will yield reliable estimates of what would have occurred to treatment group practices, and their physicians and beneficiaries, without the demonstration's systems, reporting, and performance payments. The evaluation used a stratified random assignment of practices to a treatment group or a control group within each site as the comparison group strategy. The evaluation estimated impacts of the demonstration at the site level. This was achieved by comparing average outcomes of treatment group practices with those of control group practices and accounting for the nesting of data. Specifically, because beneficiaries and physicians are nested within practices, which is the unit of intervention, the analysis accounts for this nesting. It is important to highlight that the analysis does not report practice-level impact estimates, but accounts for the hierarchical nature of the data. This section describes the randomization of practices to treatment and control groups, discusses how physicians and beneficiaries were allocated to practices, and provides a brief overview of how the evaluation team estimated impacts on practice, physician, and beneficiary outcomes.

### 1. Random Assignment of Practices to Treatment and Control Groups

The evaluation's stratified, randomized design randomly allocated each practice to the treatment or control group within specified strata. Stratification ensured that treatment group practices were similar to control group practices on key factors likely to be associated with outcomes of interest. Adopting a stratified design maximizes the precision of impact estimates and avoids compromising the credibility of the evaluation with chance imbalances across stratifying factors. Three stratifying variables were used to capture practice characteristics measured at baseline that were likely to be associated with EHR adoption, quality of care, and Medicare expenditures: (1) site, (2) practice size, and (3) whether the practice was located in a medically underserved area. Random assignment of practices to treatment and control groups also ensured that physicians practicing in and beneficiaries assigned to treatment and control group practices were, on average, similar with regard to baseline and other characteristics. More information on stratifying variables and steps to implementing the randomized design are available in Mathematica's randomization plan, submitted to CMS in 2008 (Appendix H).

### 2. Identification of Physicians Within Practices

To be eligible for the demonstration, practices completed an EHRD application form (found at the end of Appendix H) identifying physicians and other providers who agreed to participate in the demonstration. Although practices identified participating physicians just prior to random assignment, the composition of physicians within a practice changed over time as some retired or left and others joined. The implementation support contractor tracked physicians practicing in treatment group practices every year of the demonstration. At the end of Year 2, the contractor

also updated physicians practicing in control group practices. As a result, the beneficiary assignments used in the impact analysis reflected updated physician information in both treatment and control groups.

3. Identification and Assignment of Treatment and Control Group Beneficiaries to Practices

Measurement of care quality, Medicare service use and expenditures requires beneficiary-level data. The cross-section of beneficiaries served by treatment and control group practice physicians were identified each year to evaluate the cost neutrality of the demonstration and to estimate impacts on outcomes measured at the beneficiary level. The implementation support contractor was responsible for linking beneficiaries to treatment and control group practices. The contractor used an algorithm for allocating beneficiaries to one practice based on provider identification numbers available in claims data and reported in practices' demonstration application forms. The implementation support contractor then provided lists of beneficiaries associated with practices each year of the demonstration to Mathematica.

For both treatment and control group practices, the algorithm assigned each beneficiary represented in the claims files to the practice that provided the plurality of evaluation and management services during the year. As a tiebreaker for beneficiaries seen by more than one such practice, the algorithm assigned beneficiaries to practices based on whether the practice provided the beneficiary's most recent E&M visit, had the highest Medicare expenditures for that beneficiary in the year, and was located in a demonstration site.[22] This procedure avoided assigning beneficiaries to more than one practice.

## B. Estimation Methods for Changes in Outcomes

A difference-in-differences estimation method was used to estimate changes in outcome measures such as quality-of-care measures, annualized expenditures, and use of Medicare-covered services during the baseline year and each of the two demonstration years. With this approach, outcomes among beneficiaries assigned to either demonstration or control practices were compared before (the pre-period) and after (the post-period) the start of the demonstration in each demonstration year. The analysis controlled for (1) beneficiary demographic characteristics (age, sex, and race stated in Medicare enrollment files); (2) Medicare eligibility variables (including reason for enrollment [age or disability], dual-eligibility status, partial-year enrollment in a health maintenance organization [HMO], and new Medicare eligibility during the baseline period or any of the demonstration years); and (3) diagnoses (indicator variables for CAD, CHF, diabetes, any mental or behavioral health conditions, chronic kidney disease, chronic lung disease, any form of cancer, osteoporosis or arthritis, and other cardiac or circulatory diseases). The control variables are listed in Table G.5.

[22] The procedure codes for identifying E&M services are: 99201 through 99215 (office or other outpatient services); 99301 through 99316 (nursing facility services); 99321 through 99333 (domiciliary, rest home, boarding home, or custodial care services); 99341 through 99350 (home services); 99381 through 99397 (preventive medicine services); and 99401 through 99429 (counseling and/or risk factor reduction intervention) (Wilkin et al. 2007).

Table G.5. Control Variables Used in EHRD Analysis of Changes in Outcomes

| Categories | Variables |
|---|---|
| Beneficiary Demographic Characteristics | Age<br>Less than 50 years (omitted category)<br>50–64 years<br>65–69 years<br>70–74 years<br>75–79 years<br>80 years or more<br>Whether beneficiary is male<br>Race<br>African American<br>Hispanic<br>White<br>Other (omitted category) |
| Medicare Eligibility | Whether eligible for Medicare due to age<br>Whether eligible for Medicare due to disability<br>Whether dually eligible<br>Whether enrolled in an HMO for part of the year<br>Whether newly eligible for Medicare during baseline year<br>Whether newly eligible for Medicare during Jan. – June 2008<br>Whether newly eligible for Medicare during Year 1<br>Whether newly eligible for Medicare during Year 2 |
| Beneficiary Diagnoses (Chronic Conditions) | CAD<br>CHF<br>Diabetes<br>Any mental or behavioral health condition<br>Chronic kidney disease<br>Chronic lung disease<br>Any cancer<br>Osteoporosis or arthritis<br>Other cardiac or circulatory disease |
| Practice-Level Characteristics | Practice fixed effects (indicator for each practice) |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor.

CAD = coronary artery disease; CHF = congestive heart failure; EHRD = Electronic Health Records Demonstration; HMO = health maintenance organization.

Specifically, the following equation was estimated:

$$Y_{ipt} = \beta X_{ipt} + \gamma_1 T_1 + \gamma_2 T_2 + \theta_1 D_p T_1 + \theta_2 D_p T_2 + \mu_p + \varepsilon_{ipt}$$

where i, p, and t are suffixes for individual (beneficiary), practice, and time period, respectively; Y is any outcome variable; X is a vector of beneficiary characteristics; $T_1$ and $T_2$ are indicator variables for demonstration Years 1 and 2, respectively, that account for any change in an outcome between the pre-period and the post-periods across all practices; $D_p$ is an indicator variable for demonstration (that is, treatment) practices; $\mu_p$ are the practice fixed effects that account for all practice characteristics that do not change over time and might affect a beneficiary's service use and expenditures;[23] and $\varepsilon_{ipt}$ denotes the error term. In this estimation framework, $\theta_1$ and $\theta_2$ capture the differential changes in an outcome for beneficiaries in treatment practices versus those in comparison practices in Year 1 and Year 2, respectively, and are therefore the difference-in-differences estimates for changes in the outcome variable.

The regression analysis did not control explicitly for Hierarchical Condition Category (HCC) scores, since the calculation of such scores takes into consideration certain acute conditions the demonstration sought to prevent, such as heart attacks and acute diabetes. Hence, it would be inappropriate to control for these measures during the three demonstration periods. However, as mentioned above, there are separate control variables for several components in the calculation of HCC scores, such as beneficiary demographics, dual-eligibility status, and presence of several chronic conditions.

The demonstration pre-period includes the 2008 Medicare claims for beneficiaries who were assigned to practices during the baseline year. The post-period includes Medicare claims for demonstration Year 1 (June 1, 2009, through May 31, 2010) and Year 2 (June 1, 2010, through May 31, 2011) for beneficiaries who were assigned to demonstration practices during these periods.

Individual-level enrollment weights were constructed so that those beneficiaries who were observed for more months in the claims data received more weight. Specifically, weights were based on the number of months that a person was alive; enrolled in both Medicare Part A and Part B; and not in an HMO during a specific observation period (the baseline and demonstration years). Months that a beneficiary was in an HMO were excluded because no Medicare claims were available during these months. All observations in the expenditures analyses were weighted by these enrollment weights. In the analyses involving quality-of-care measures and potentially preventable hospitalizations, all observations that were truncated (because an individual died, enrolled in managed care, or became ineligible) were weighted. In Appendix J,

[23] The fixed effects are captured by including a vector of dummy variables in the model—one dummy variable for each practice. The use of practice fixed effects in the analysis helps reduce bias arising from unobserved practice-level characteristics that could be correlated with both treatment status and outcomes, such as the practice staff's level of motivation and skills for using health IT. The inclusion of practice fixed effects helps to remove any bias arising from such unobserved practice characteristics.

detailed difference-in-differences estimates are reported for the quality-of-care measures, potentially preventable hospitalizations, Medicare expenditures, and Medicare service use. The p-values reported in the tables (showing the difference-in-differences estimates) were drawn from the coefficients on treatment status ($\theta_1$ and $\theta_2$). Note that standard errors and the associated p-values were adjusted to account for the correlation in outcomes among beneficiaries in the same practice; that is, observations were clustered at the practice level, and the standard errors accounted for the clustering.

## C. Sensitivity Tests

For the analysis of claims-based quality-of-care measures and expenditure and service-use measures, a wide range of sensitivity tests were run to confirm that the findings in difference-in-differences estimates were not sensitive to the model specifications.

Findings on health IT impacts were verified by estimating impacts with and without non-response weights, and with and without baseline regressions controls. Also, nonresponse bias was assessed by comparing respondents to the 2011 OSS to the full sample of randomized practices that were eligible to respond to the OSS.[24]

As shown in Table G.6, compared to the full eligible population, OSS respondents were more likely to be large practices (with six or more physicians) and to have an EHR at baseline, but less likely to be located in a medically underserved area. However, the respondent population was similar to the eligible population across each of the other characteristics after the nonresponse weights were applied. For example, 43 percent of the full treatment group population (shown in column 1) and 49.7 percent of treatment group OSS respondents (column 2) had an EHR at baseline; however, using nonresponse weights, the percentage of treatment respondents with an EHR at baseline (43.4 percent, as shown in column 3) is nearly identical to the percentage in the full sample for the treatment group (43.0 percent).

Thus, this analysis suggests that while there are some differences between OSS respondents and the full sample, the weights adjust for any bias related to observed practice characteristics. However, the weights could not adjust for nonresponse bias related to unobserved practice characteristics. For example, it is possible that practices that had staff with health IT skills were disproportionately likely to respond to the 2011 OSS, since the OSS asked questions related to technology. If that were the case, then practices' use of health IT would likely be overstated on the OSS (because respondents would be more likely than nonrespondents to be sophisticated health IT users), particularly for the control group relative to the treatment group (since the response rate for the control group was lower than that of the treatment group).

[24] Seven of the 412 randomized treatment practices and one of the 413 randomized control practices were not eligible to receive the OSS because they did not meet the terms and conditions of the demonstration.

Table G.6 Comparison of 2011 OSS Respondents to Full Sample of Eligible Practices (Percentages)

| Practice Characteristics | Treatment Group Full Eligible Population (1) | Treatment Group Respondents Unweighted (2) | Treatment Group Respondents Weighted (3) | Control Group Full Eligible Population (4) | Control Group Respondents Unweighted (5) | Control Group Respondents Weighted (6) |
|---|---|---|---|---|---|---|
| Practice Size | | | | | | |
| 1 to 2 physicians | 30.1 | 24.8 | 29.5 | 34.2 | 29.2 | 35.6 |
| 3 to 5 physicians | 22.2 | 22.3 | 22.6 | 18.4 | 16.5 | 17.2 |
| 6 + physicians | 47.6 | 52.9 | 47.9 | 47.3 | 54.4 | 47.1 |
| Located in Medically Underserved Area | 29.1 | 27.7 | 29.5 | 29.6 | 28.5 | 29.2 |
| Had EHR at Baseline | 43.0 | 49.7 | 43.4 | 43.7 | 48.3 | 44.4 |
| Sample Size | 405 | 314 | 405 | 412 | 267 | 412 |

Source: Office Systems Survey (OSS), conducted in spring and summer 2011, and data drawn from applications practices submitted by practices to EHRD in 2008.

Notes: Seven of the 412 randomized treatment practices and one of the 413 randomized control practices were not eligible to receive the 2011 OSS because they did not meet the terms and conditions of the demonstration. Weighted sample sizes are reported for columns 3 and 6.

EHR = Electronic health record; EHRD = Electronic Health Records Demonstration.

## D. Statistical Power

Minimum detectable differences (MDDs) for the EHRD claims-based analysis were calculated based on the full intent-to-treat sample (pooled across states) of 405 treatment and 412 control practices (excluding only the 7 treatment group practices and 1 control group practice that did not meet the terms and conditions of the demonstration) and 290 patients per practice. Previous calculations (Dale et al. [2011]) suggested that for a two-sided test at the 5 percent level, the evaluation would have 80 percent power to detect a 5.7 percent change in Medicare expenditures and a 1.7 percentage point change in binary measures. These calculations were based on the following assumptions drawn from data for the MCMP demonstration: (1) an intra-cluster correlation coefficient (ICC) of .029; (2) a coefficient of variation of 1.7 (for continuous measures); (3) a mean of .5 (for binary variables); and (4) a regression $R^2$ of .1. When EHRD data became available, the evaluation team found that the ICC for EHRD was .037 and the coefficient of variation averaged about 2.2 for both inpatient hospitalizations and total expenditures. MDD calculations that assume the values drawn for the ICC and coefficient of variation (CV) are .037 and 2.2 respectively (the values drawn from actual EHRD data rather than from MCMP data) indicate that the evaluation has 80 percent power for a two-sided test at the 5 percent level to detect an 8.4 percent change in continuous variables (which translates to about $840 in total Medicare expenditures per beneficiary per year), and a 2 percentage point change in binary variables; for a two-tailed test at the 10 percent level (rather than the 5 percent level), minimum detectable effects are 7.5 percent for continuous variables and 1.7 percentage points for binary variables.[25]

For outcomes drawn from the 2011 OSS for the sample of 267 control group and 314 treatment group respondents, the evaluation has 80 percent power at the 5 percent level to detect a 10 percentage point effect for a binary variable with a mean of 0.5, assuming a regression $R^2$ of 0.3. (Previous calculations in Dale et al. [2011] reported a 9 percentage point MDD for outcomes drawn from the OSS; MDDs are slightly higher here because actual OSS response rates were slightly lower than expected.) The impacts of the demonstration on practices' use of health IT (as reported in Chapter V) were generally much larger than the MDDs, suggesting the evaluation was well powered for these outcome measures.

[25] MDDs for a single state are about twice as high as those for the analyses that are pooled across states.

This page is intentionally left blank.

# APPENDIX H

# RANDOMIZATION PLAN

This page is intentionally left blank.

MATHEMATICA
Policy Research, Inc.

P.O. Box 2393
Princeton, NJ 08543-2393
Telephone (609) 799-3535
Fax (609) 799-0005
www.mathematica-mpr.com

MEMORANDUM

TO: Lorraine Johnson

FROM: Lorenzo Moreno and Audra Wenzlow

DATE: 8/18/2008
EHRD-017

SUBJECT: Randomization Plan - REVISED

---

This memorandum summarizes our randomization plan of eligible physician practices into a treatment or a control group in each of 12 demonstration sites for the evaluation of the EHR demonstration. The document consists of four sections: (1) the rationale for choosing a stratified randomization design, (2) the stratifying variables, (3) our approach to implementing randomization, and (4) steps for implementing it.[1]

## A. RATIONALE FOR CHOOSING A STRATIFIED RANDOMIZATION DESIGN

Implementation of the demonstration will occur in two phases, one year apart, beginning in June 2009. The treatment group will receive financial incentives to (1) adopt and use electronic health records (EHRs); (2) report clinical measures; and (3) meet certain quality performance targets. The control group will receive none of these payments nor will be required to adopt and use an EHR.

CMS expects about 200 primary care physician practices from each of the 12 sites to enroll in the demonstration. Stratification of the sample in each site (that is, separate randomization of practices within each stratum) will ensure that the treatment and control groups are well balanced in particularly important factors (that is, practice characteristics). Although unrestricted randomization should generate an approximately equal number of treatment and control practices for the overall demonstration, formal stratification will ensure a balanced allocation over these

[1] The schedule of data flows between Mathematica Policy Research, Inc. (MPR) and Actuarial Research Corporation (ARC) needed to implement the randomization plan is detailed in "EHRD: Schedule for Data File Transfers for the Evaluation" (EHRD-021, dated August 18, 2008).

factors.[2] Indeed, stratification could be seen as a low-cost insurance against the chance of large imbalances, which could compromise the credibility of the evaluation.

CMS has a selected site coordinator in each of 12 sites.[3] Each site coordinator will assist CMS in efforts to recruit at least 200 eligible practices in its respective site. Recruitment for the four Phase I sites will begin fall 2008 and should be completed by the end of the year. It is anticipated that randomization of Phase I practices in each of the first four sites will be completed during the first quarter of calendar year 2009. Recruitment and randomization for Phase II practices are expected to occur one year later.

## B. STRATIFICATION VARIABLES

Ideally, the stratification variables would be predictors of success in implementing EHRs and, ultimately, meeting the demonstration's targets. However, if these variables are not important predictors of outcomes, it may still be important to achieve balance on practice characteristics to enhance the credibility of the design and reduce the need for model-based adjustments during the analysis. We will consider four key stratifying variables: (1) site, (2) practice size (that is, the total number of physicians in the practice), (3) urban or rural location of the practice, and (4) whether the practice already had an EHR at the time of application. We also discuss other variables that may be considered as potential stratifying variables. As noted, the data source for the stratifying variables will be the EHRD application form, which is attached:

1. Site. There will be considerable variation across the 12 selected sites on (1) physician regulations and practice patterns, (2) ongoing efforts by private and public organizations to implement EHRs and improve quality of care, (3) secular trends due to the phased implementation of the demonstration, and (4) general economic conditions in the practices' market or service area. Because of this variability, and because separate impact estimates will be produced for each site, it is important for each site to be a separate stratum. Thus, we will randomize the eligible practices for each site separately. Data for constructing this variable will come from the

[2] Reducing chance imbalances on important factors attenuates the loss of precision of the treatment estimate, particularly for small sample sizes (that is, under 100 per study group) (McEntegart 2003). It also improves the power of hypothesis tests (that is, statistical efficiency) (Greevy et al. 2004).

[3] The Phase I sites are located in the District of Columbia, Louisiana, Maryland, six counties in the Pittsburgh, Pennsylvania area, and South Dakota. (The South Dakota site also includes border counties in Iowa, Minnesota, and North Dakota.) The Phase II sites are located in Alabama, Delaware, selected counties in the Jacksonville, Florida area, Georgia, Maine, Oklahoma, selected counties in Wisconsin, and Virginia.

applications of all practices that define a site (and Question 3 for identifying geographic location).

2. Practice Size. Although the EHR demonstration will target small- to medium-size practices (that is, practices with 20 or fewer physicians),[4] practices with 5 or fewer physicians constitute the vast majority of all practices (Government Accountability Office 2008) and therefore are likely to comprise the majority of practices enrolled in the demonstration. Among this group, solo practices and those with two physicians, which represent about 83 percent of practices in the United States, are likely to be different than larger practices in their ability to implement EHRs aggressively (Blumenthal et al. 2006; Simon et al. 2007a). Other key outcomes, such as quality of care, patients' Medicare expenditures, and physician practice patterns may vary by practice size. Thus, we recommend balancing the sample along this likely predictor of success in implementing EHRs and, ultimately, meeting the demonstration's quality targets. Furthermore, we suggest limiting the number of categories for the practice-size variable to three (1 or 2 physicians, 3 to 5, and 6 to 20) to avoid the problem of small cells in each stratum.[5] However, we can revise this decision if the size distribution of enrolled practices differs markedly from our assumption. While it is not strictly necessary, we expect to use the same stratification categories in each site, unless that approach creates too much risk of an unbalanced assignment of practices. Data for constructing this variable will come from Question 1 in the revised application form.

3. Urban versus rural location. Another important factor that has been associated with success in implementing EHRs is the geographic location of the practice, specified as urban or rural. This factor measures the higher availability of resources for acquiring and using EHRs, better training and experience with health information technology, and less isolation from colleagues, among other considerations, for providers in urban

---

[4] However, practices with a few more physicians may be allowed to participate in the demonstration, if space permits, and they are otherwise eligible.

[5] As a stratifying variable for the random assignment process, we will define practice size as the total number of number of physicians in the practice. The demonstration will consider as eligible for participating in the EHR demonstration physicians and other providers (such as physician assistants and nurse practitioners) as reported on page 4 in the application form. However, the decision to define the size of the practice using only physicians, whether eligible for participating in the demonstration or not, stems from the need to use the same definition of practice size that is reported in the literature as a predictor of EHR adoption and use. Although the total number of physicians in the practice will be different from the number of providers eligible for the demonstration, these two counts will be highly correlated (by construction).

practices relative to those in rural practices (Grossman and Reed 2006; Simon et al. 2007b). Thus, we recommend balancing the sample along this important practice characteristic. Data for this variable will come from Question 3 in the practice application form, which refers to the practice's primary location. We will use the practice's primary location to define this stratifying factor because the primary location is the basis for identifying a practice with satellite offices and will be the unit of intervention (that is, the entity that will receive the incentive payments, if appropriate.) A variant of using urban versus rural location is to assess whether the practice is located in a medically underserved or health-professional shortage area, as defined by the Health Resources and Services Administration (HRSA). We have explored the feasibility of geocoding the practices' address for matching to HRSA's database and construct a dichotomous indicator of whether a practice is in a medically-underserved area or not. Construction of this stratifying variable is entirely feasible. Based on these considerations, CMS may want to recommend which of the two variables to use as the stratifying variable.

4. Whether the practice has an electronic health record system in the office. Practices that have acquired an EHR at the time of the application for the demonstration will have considerably more experience addressing the financial and technical challenges to adopt this technology than practices that have not acquired a system yet (Simon et al. 2007a and 2007b). Furthermore, practices that have acquired an EHR system may have received technical assistance for acquiring it and using it. Thus, because the effects of the EHRD might well differ between these two types of practices, we strongly recommend also balancing the sample along this factor.[6] Data for this variable will come from Question 9 in the application form.

Other practice characteristics available in the application form are: (1) practice setting (such as, hospital-based versus non-hospital based) and (2) the number of Medicare fee-for-service beneficiaries that use the practice as primary source. As noted below, because stratified randomization works best when the number of stratifying variables is restricted to four or five variables, most likely we will restrict these variables to those four discussed above.

[6] This factor could be defined in different ways to capture the practice's degree of sophistication with health information technology and use of EHRs. For example, we could use responses from Questions 10 and 11 for defining alternative measures. However, we will only use Question 9 so that the factor has only two categories to keep the stratification manageable.

## C. APPROACH TO PRACTICE ALLOCATION

As noted, the random allocation of practices into the treatment or the control group will happen at a point in time for each phase.[7] This greatly simplifies the implementation of the stratified randomization design because it will be feasible to know the sample size in each stratum before randomization. In this section, we discuss our approach for allocating the sample of eligible practices using stratified randomization. We also discuss an alternative approach—optimal matching—that we may implement if stratified randomization is not feasible.[8]

Stratified randomization (that is, random assignment of practices in each stratum using random permuted blocks of varying size) is the most common approach for implementing this design (Cochran and Cox 1957; McEntegart 2003; Woodward 2005). In addition, it is easy to implement and well-known statistical packages handle the allocation of practices across strata.[9] Furthermore, stratified randomization, by minimizing imbalances as part of the study design, reduces the need to adjust for any imbalances by including covariates in the statistical analysis model (that is, post stratification). However, stratified randomization is only manageable when the number of stratifying variables is less than five, so that the likelihood of cells with just one practice is negligible given the sample sizes for the demonstration (Woodward 2005). For instance, using the categories proposed for the four variables listed above, we would have 144 cells (= 12 x 3 x 2 x 2), or 12 cells per site. Thus, the sample of 2,400 practices, or 200 practices per site, will have an average of 17 practices per cell and may result in empty or sparse cells for strata defined by categories with low frequency, such as practices with, say, 6 to 20 physicians in a rural location and that have adopted an EHR system. For this reason, stratified randomization is recommendable only when the samples are relatively large (that is, at least 100 units per study group) and the number of strata does not exceed 15-20 cells per site, as would be the case if we restrict the number of stratifying factors to less than five.

An alternative to stratified randomization is to use optimal matching (Greevy et al. 2004; Rosenbaum 1989). This approach is suitable for balancing on a large number of covariates (that

---

[7] In contrast, randomization on a rolling basis occurs when the enrolled practices only become known over an extended period and must be randomized within a short interval of time after enrolling in the demonstration.

[8] We do not discuss in this memo other randomization methods, such as dynamic allocation, because they are only appropriate for randomization on a rolling basis.

[9] We plan to use ralloc in STATA®, a routine that provides a list of cases randomly assigned to a treatment or a control group in each stratum using a well-known algorithm for generating blocks of varying size within each stratum (Ryan 2008). The package is general enough to accommodate any blocked randomized design and to control the balance of all stratifying variables. Furthermore, it is straightforward to implement.

is, 10 or more). The approach consists of dividing the sample into pairs of practices, where the pairs are formed by an algorithm that identifies the optimal grouping based on a measure of the dissimilarity (or distance) between practices on all covariates.[10] Assignment to the treatment or the control group is done randomly for each optimally formed pair. The method is more suitable for continuous factors than for categorical factors, as those described in the previous section. Finally, it is difficult to explain to non-statisticians, it is computationally intensive, and there is limited experience with its use.[11]

We recommend using the straightforward stratified randomization to allocate physicians practices into the treatment or the control groups while balancing on the four factors discussed above. Should insufficiently populated cells (for example, strata with one or three practices) be numerous, we will consider combining categories to ensure we have at least four practices per stratum.[12] This grouping can only be done for factors with more than one category, so we will examine the size distribution of enrolled practices to decide whether our proposed categorization into three groups needs to be modified. For example, we may change the original categories proposed (1 or 2 physicians, 3 to 5, and 6 to 20) to two categories if the number of practices with 6 to 20 physicians would result in strata with one or three practices. Alternatively, we could drop one of the stratifying factors. Likewise, should the number of factors increase substantially, we would consider using optimal matching after we review the distribution of practices across cells defined by the factors in each site and confirm that the use of this approach has substantial technical advantages relative to stratified randomization.

## D. STEPS FOR IMPLEMENTING THE STRATIFIED RANDOMIZATION DESIGN

We will take the following steps for implementing the stratified randomized design, which involves data flows between MPR and ARC:

1. ARC will process the application forms and transmit the data to MPR with the identity of the practices being concealed (only an anonymous practice ID would be supplied with the original data)

---

[10] This grouping also is known as pairwise matching. SAS® includes a macro (%match) that implements matching using the optimal algorithm (Mandrekar and Mandrekar 2004).

[11] Furthermore, under certain conditions, pairwise matching may result in less power than an unmatched design (Martin et al. 1993).

[12] The STATA ralloc macro generates sufficient assignment sequences to complete the final block in the stratum, even if the number of units in the stratum is an odd number (Ryan 2008).

2. MPR will tabulate the data for the stratifying variables for each site
3. MPR will review the distribution of practices across strata in each site and decide whether and how to group the stratifying variables, if appropriate, as noted in the previous section
4. MPR will run the allocation program and would transmit to ARC the files with the allocation of practices to a treatment or a control group, by stratum, for each site
5. ARC will link the files with the practices' allocations to their list of practices
6. ARC will inform the practices of their assignment

These steps are consistent with the description of data flows between ARC and MPR, described in the final version of the document "EHRD: Schedule for Data File Transfers for the Evaluation" (memo EHRD-021, dated August 18, 2008). If needed, we will refine these steps once we begin reviewing application data for the practices in the first four sites.

Attachments: Electronic Health Records (EHR) Demonstration Application to Participate

cc: J. Schore, S. Felt-Lisk, M. Kovac, File

REFERENCES

Blumenthal, David, Catherine DesRoches, Karen Donelan, Timothy Ferris, Ashish Jha, Rainu Kaushal, Sowmya Rao, and Sara Rosenbaum. Health Information Technology in the United States: The Information Base for Progress. Washington, DC: The George Washington University, Massachusetts General Hospital, and Robert Wood Johnson Foundation, 2006.

Cochran, William G., and Gertrude M. Cox. Experimental Designs. New York: John Wiley, 1957.

Government Accountability Office, "Medicare Physician Payment. Care Coordination Programs Used in Demonstration Show Promise, but Wider Use of Payment Approach May be Limited." Washington, DC, February 2008.

Greevy, Robert, Bo Lu, Jeffrey H. Silber, and Paul Rosenbaum. "Optimal Multivariate Matching Before Randomization." Biostatistics, vol. 5, no. 2, 2004, pp. 263-275.

Grossman, Joy M., and Marie C. Reed. "Clinical Information Technology Gaps Persist Among Physicians." Center for Studying Health System Change, Issue Brief No. 106, November 2006.

Mandrekar Jayawant N., and Sumithra J. Mandrekar. "An introduction to matching and its application using the SAS® system. Paper 208-29." Proceedings of the 29th SAS Users Group International Conference (SUGI) 2004.

Martin, Donald C., Paula Diehr, Edward B. Perrin, and Thomas D. Koepsl. "The Effect of Matching on the Power of Randomized Community Intervention Studies." Statistics in Medicine, vol. 12, 1993, pp. 329-338.

McEntegart, Damian J. "The Pursuit of Balance Using Stratified and Dynamic Randomization Techniques: An Overview." Drug Information Journal, vol. 37, 2003, pp. 293-308.

Rosenbaum, Paul R. "Optimal Matching for Observation Studies." Journal of the American Statistical Association, vol. 84, no. 408, 1989, pp. 1024-1032.

Ryan, Philip. "ralloc – Random allocation of treatment in controlled trials." Available online [www.stata.com/statalist/archive/2007-10/msg00226.html]. Accessed on April 4, 2008.

Simon, Steven, Rainu Kaushal, Paul D. Cleary, Chelsea A. Jenter, Lynn A. Volk, Eric G. Poon, John Orav, Helen G. Lo, Deborah H. Williams, and David W. Bates. "Correlates of Electronic Health Record Adoption in Office Practices: A Statewide Survey." Journal of the American Medical Informatics Association, vol. 14, no. 1, Jan/Feb 2007a, pp. 110-117.

Simon, Steven, Rainu Kaushal, Paul D. Cleary, Chelsea A. Jenter, Lynn A. Volk, Eric G. Poon, John Orav, Elisabeth Burdick, Eric G. Poon, and David W. Bates. "Physicians and Electronic Health Records. A Statewide Survey." Archives of Internal Medicine, vol. 167, March 12, 2007b, pp. 507-512.

Woodward, Mark. Epidemiology. Study Design and Data Analysis, Second Edition. Boca Raton: Chapman & Hall/CRC, 2005.

# ELECTRONIC HEALTH RECORDS (EHR) DEMONSTRATION APPLICATION TO PARTICIPATE

The goal of the Electronic Health Records Demonstration (EHR) is to establish a 5-year pay-for-performance demonstration project with small and medium sized primary care physician practices to promote the adoption and use of certified EHRs to improve the quality of patient care for chronically ill Medicare patients. Doctors who meet or exceed performance standards established by CMS will receive incentive payments for managing the care of eligible Medicare beneficiaries. Practices incorporating greater use of health information technology into their office practices will be eligible to earn additional incentives.

Each practice applying to participate must have a designated staff person authorized to speak for the group, provide requested information, and to whom all correspondence will be directed. All physicians who are members of the practice and who wish to participate in the demonstration must sign the enclosed data sharing consent form agreeing to share data submitted to CMS and/or its contractors assisting in the implementation or evaluation of the demonstration.

**Those who wish to participate should fill out this form completely. Completing this form does not guarantee participation in the demonstration. CMS reserves the right to limit the number of practices that may participate.**

## Physician Office Information

For office use only

Name of Practice

**1.** How many physicians are part of this practice?_____
Of these how many primarily provide primary care *(general practice, family practice, gerontology, internal medicine)*? _____

**2.** Briefly describe your practice in terms of how it is organized, locations, services offered, affiliation with larger networks, etc. _____

_____

_____

_____

**3.** Address of primary practice location

| Street Address | | | Office Number |
|---|---|---|---|
| City | State | Zip | Country |

**4.** List all other locations that are part of this practice and participating in the demonstration

| Location #2 Name of Practice at this location | | | Office Number |
|---|---|---|---|
| Street Address | | | |
| City | State | Zip | Country |
| Location #3 Name of Practice at this location | | | Office Number |
| Street Address | | | |
| City | State | Zip | Country |

❑ Check here if additional locations. Attach information on additional pages

**5a.** Designated Contact Person

| Name of Designated Contact Person | | Title | |
|---|---|---|---|
| Street Mailing Address *(if different from primary practice location)* | | | |
| City | State | Zip | Country |
| Telephone | E-mail | | |

**5b.** Staff Person Responsible for Completion of this Application (If different from the "Designated Contact Person" in 5a above).

| Name of Designated Contact Person | | Title | |
|---|---|---|---|
| Street Mailing Address *(if different from primary practice location)* | | | |
| City | State | Zip | Country |
| Telephone | E-mail | | |

**6.** Secondary Contact Person (if applicable, for mailing purposes)

| Name of Secondary Contact Person | | Title | |
|---|---|---|---|
| Street Mailing Address *(if different from primary practice location)* | | | |
| City | State | Zip | Country |
| Telephone | E-mail | | |

**7.** Estimated number of Medicare Fee-For-Service patients that use your practice as primary source of care _____

**8.** All incentive payments associated with the demonstration will be made to the practice and not to individual physicians. Please provide information regarding the legal entity to which payments should be made, as specified below.

| Name of entity to which payments should be made | | | |
|---|---|---|---|
| Street Mailing Address *(if different from primary practice location)* | | | Practice Tax Identification Number |
| City | State | Zip | Country |

**9.** Do you have an electronic health record (EHR) in your office?
Yes ❑ (Please respond to questions that follow, **and then proceed to Question #11**)

If yes, what is the vendor and product? _____

Is this system certified by the Certification Commission for Health Information Technology (CCHIT)?
Yes ❑ No ❑ Unknown ❑

What is the date of certification?
2006 ❑ 2007 ❑ 2008 ❑ Unknown ❑ Other ❑ _____

No ❑ (Please go to Question #10)

**10.** If you do not currently have an EHR, when do you plan to implement an EHR?
0–6 months? ❑ 7–12 months? ❑ 13–24 months? ❑ Other? ❑ _____

Has an EHR product been selected?
Yes ❑ No ❑

If yes, what is the vendor and product? _____

Is this system certified by the Certification Commission for Health Information Technology (CCHIT)?
Yes ❑ No ❑ Unknown ❑

What is the date of certification?
2006 ❑ 2007 ❑ 2008 ❑ Unknown ❑ Other ❑ _____

**11.** If you have an electronic system in your office, please describe the type of health information technology **currently used** in your practice, either as part of an EHR or independently as a stand-alone product *(check all that apply)*:

❑ Electronic patient visit notes
❑ Electronic patient-specific problem lists
❑ Automated patient-specific alerts and reminders
❑ Electronic disease-specific patient registries
❑ Clinical decision support/automated references to best practices
❑ Patient e-mail
❑ Patient-specific educational materials
❑ On-line referrals to other providers
❑ Clinical messaging with other physicians
❑ Transmission of records to hospitals or other facilities

Laboratory tests:
❑ On-line order entry
❑ On-line results viewing

Radiology tests:
❑ On-line order entry
❑ On-line results (reports and/or digital films)

E-Prescribing:
❑ Printing and/or faxing Rx
❑ On-line Rx transmission to pharmacy

Other: _____ _____

# PHYSICIANS PARTICIPATING IN THE EHR DEMONSTRATION IN THIS PRACTICE

Practice Name

| Practice Group PIN number *(if applicable)* | Group NPI *(if applicable)* |
|---|---|
| | |

Please provide information listed in the chart below for all providers (includes nurse practitioners and physician assistants that provide primary care and bill Medicare independently) in this practice applying to participate in this demonstration.

| Physician Name (PRINT) | Specialty | Tax Identification Number* | Medicare Provider Identification Number (PIN) at this Location | Individual NPI–National Provider Identification number | Consent Form Attached (Y/N) |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |

* Provide the Tax Identification Number used by each physician when billing for Medicare services **as a member of this practice**.

# CONSENT TO SHARE DATA

As an applicant to the Electronic Health Records Demonstration project, I agree to comply with the requirements of this demonstration, including sharing all data submitted to CMS and/or its contractors assisting in the implementation or evaluation of the demonstration.*

Provider Name *(print)*

Provider Signature | Date

Medicare Provider Identification Number | Individual National Provider Identifier (NPI)

Provider Name *(print)*

Provider Signature | Date

Medicare Provider Identification Number | Individual National Provider Identifier (NPI)

Provider Name *(print)*

Provider Signature | Date

Medicare Provider Identification Number | Individual National Provider Identifier (NPI)

Provider Name *(print)*

Provider Signature | Date

Medicare Provider Identification Number | Individual National Provider Identifier (NPI)

* **This form must be signed by each participating physician in the practice. If additional signatures are necessary, please copy and submit additional signature sheets.**

MATHEMATICA
Policy Research, Inc.

P.O. Box 2393
Princeton, NJ 08543-2393
Telephone (609) 799-3535
Fax (609) 799-0005
www.mathematica-mpr.com

MEMORANDUM

TO: Lorraine Johnson

FROM: Audra Wenzlow, Lorenzo Moreno, Licia Gaber-Baylis

DATE: 2/24/2009
EHRD-053

SUBJECT: EHRD: Randomization of Phase I Practices

---

This memorandum describes the randomization of the 825 eligible physician practices that volunteered for Phase I of the EHR demonstration (EHRD) into treatment and control groups.[1] Beginning in June 2009, Phase I practices in the treatment group will receive financial incentives to (1) adopt and use CCHIT-certified electronic health records (EHRs); (2) report clinical measures; and (3) meet certain quality performance targets. Practices in the control group will receive none of these payments nor will be required to adopt and use an EHR system. The treatment and control allocation of each practice will be maintained throughout the course of the demonstration and its evaluation. A stratified randomization design was used to allocate each of the 825 eligible practices to a treatment or a control group for assessing the impact of the intervention. This document summarizes: (1) the stratifying variables, (2) the method used to randomize practices, and (3) the resulting distribution of treatment and control group practices across strata and other outcomes of interest.

Overall, randomization of Phase I EHR eligible practices was successfully implemented according to the proposed plan (see memorandum "Randomization Plan - REVISED" [EHRD-017], dated August 18, 2008). The allocation resulted in a balanced design of treatment and control practices across strata and across other variables measured at the time of application.

---

[1] Recruitment and randomization for Phase II practices are expected to occur one year later. The Phase I sites are located in Louisiana, Maryland and the District of Columbia, six counties in the Pittsburgh, Pennsylvania area, and South Dakota. The South Dakota site also includes border counties in Iowa, Minnesota, and North Dakota. The Phase II sites are located in Alabama, Delaware, selected counties in the Jacksonville, Florida area, Georgia, Maine, Oklahoma, selected counties in Wisconsin, and Virginia.

## A. STRATIFYING VARIABLES

As described in the randomization plan, stratification of the sample ensures that the treatment and control groups are well balanced in particularly important practice characteristics. Although unrestricted randomization within each site should generate an approximately equal number of treatment and control practices for the overall demonstration, formal stratification will ensure a balanced allocation over key practice characteristics.[2] Stratification could be seen as a low-cost insurance against the chance of large imbalances, which could compromise the credibility of the evaluation. Furthermore, by minimizing imbalances as part of the study design, stratified randomization reduces the need to rely on regression analysis or weighting to adjust for any imbalances (that is, post stratification control).

Ideally, the stratification variables are predictors of success in implementing EHRs and other outcomes of interest in the evaluation (for example, quality of care and beneficiary health expenditures). We stratified the practices according to (1) site, (2) practice size, and (3) whether the practice was located in a medically underserved area. The data source for the stratifying variables is the EHRD application form, which is attached, and a secondary dataset publicly available from the Health Resources and Services Administration as described below:

1. Site. There is considerable variation across states and regions in (1) physician regulations and practice patterns, (2) efforts to implement EHRs and improve quality of care, and (3) general economic conditions. Because of this variability, and because separate impact estimates will be produced for each site, each of the four Phase I sites was identified as a separate stratum and eligible practices in each site were randomized separately. Data for constructing this variable came from the EHR application file practice identifier, the second and third character of which identify the site as Louisiana (N=204), Maryland (N=255), Pennsylvania (N=279), or South Dakota (N=87).

2. Practice Size. The number of physicians in Phase I practices ranges from 1 to 27, reflecting the fact that the EHR targets small- to medium-size practices. Solo practices and those with two physicians represent about 83 percent of practices in the United States (Government Accountability Office 2008) but just over half (52.3 percent) of practices eligible for Phase I of the EHR demonstration. Research and common sense suggests that small practices are likely to be less able than larger practices to implement EHRs aggressively (Blumenthal et al. 2006; Simon et al. 2007a) and the characteristics of Phase

[2] Reducing chance imbalances on important factors attenuates the loss of precision of the treatment estimate, particularly for small sample sizes (that is, under 100 per study group) (McEntegart 2003). This greater precision also improves the power of hypothesis tests (that is, statistical efficiency) (Greevy et al. 2004).

I practices confirms this link. There was a strong association between practice size and whether the practice already had an EHR at the time of application among Phase I practices. The percentage with EHR systems increased from 35.9 for solo or 2-physician practices to 46.3 for practices with 3 to 5 physicians, and 58.2 for practices with 6 or more physicians. Although expenditure and quality outcomes data were not available for eligible Phase I practices at the time of this writing, these outcomes also may vary by practice size. Thus, we balanced the sample along this likely predictor of success in implementing EHRs and meeting the demonstration's evaluation outcomes. To avoid problems of small cell sizes in each stratum, we limited the number of categories for the practice size variable to three (1 or 2 physicians, 3 to 5, and 6 to 20). Data for constructing this variable came from Question 1 of the application form.[3]

3. Medically Underserved Area/Population (MUA/P). Another important factor that has been associated with the study's outcomes is the geographic location of the practice as measured by whether the practice is in a medically underserved area (MUA) or has a medically underserved population (MUP). MUA/Ps are areas or populations designated by HRSA as having "too few primary care providers, high infant mortality, high poverty and/or high elderly population."[4] MUA was found to be correlated with implementing EHRs as well as baseline quality outcomes in the Medicare Care Management Performance (MCMP) demonstration. To construct this variable, each practice's primary location was geocoded and merged with HRSA data by census tract. Addresses for which tracts were not available and those for which HRSA reported only metropolitan area information were manually entered into the HRSA website to determine their MUA/P status. Because MUA and MUP are not differentiated in the manual web-based data tool, we were unable to determine the MUA status alone of all practices. Therefore, we used the combined MUA/P measure identifying practices that are either in an MUA or an MUP. About 29 percent of Phase I practices met these criteria. Practices' primary

---

[3] As a stratifying variable for the random assignment process, we defined practice size as the total number of physicians in the practice. The demonstration will consider as eligible for participating in the EHR demonstration physicians and other providers (such as physician assistants and nurse practitioners) as reported on page 4 in the application form. However, the decision to define the size of the practice using only physicians, whether eligible for participating in the demonstration or not, stems from the need to use the same definition of practice size that is reported in the literature as a predictor of EHR adoption and use. Although the total number of physicians in the practice will be different from the number of providers eligible for the demonstration, these two counts will be highly correlated (by construction).

[4] Further information on HRSA MUA and MUP designations are described at http://bhpr.hrsa.gov/shortage/muaguide.htm.

location came from Question 3 of the application form and MUA/P data were obtained from the HRSA website.[5]

We also considered using urban/rural practice location and whether the practice already had an EHR at the time of application as stratifying variables. Neither the urban/rural variable nor the MUA/P variable was significantly associated with having an EHR once we controlled for site and practice size. However, these measures are likely to be strongly associated with other outcomes of interest in the evaluation. Because our urban/rural measure resulted in small strata sizes, and because MUA was found to be associated with key outcomes of interest in the MCMP project, we used MUA rather than urban/rural location as a stratifying variable in all sites.

To limit the number of strata, we also excluded having an EHR at the time of application from the stratification. We chose to exclude pre-demonstration EHR adoption from stratification for several reasons. First, both practice size and MUA have been shown to be important correlates of both EHR use and other demonstration outcomes, whereas it is unknown whether initial EHR use will be a good predictor of future EHR development and various other demonstration outcomes. Second, information on having an EHR system collected in the application was not verifiable. If practices viewed the information as potentially influencing their participation in the demonstration, the reported EHR system information may not be accurate. Excluding EHR use from the randomization process enabled us to test the balance of the randomized practices across this measure, as described below. Our evaluation will examine and control for the effect on outcomes of having an EHR at the time of application.

We used the same strata specifications across sites because practice size was consistently associated with EHR use in all sites and we expect that MUA/P will be associated with enrollee outcomes across sites. Using the same strata specifications across sites also made the randomization process more efficient and will simplify the post-randomization analyses.

## B. RANDOMIZATION METHOD

We implemented a straightforward stratified randomization to allocate physicians practices into the treatment or the control groups while balancing on the three factors discussed above. Stratified randomization is the most common approach for ensuring that random assignment results in a balanced design across strata (Cochran and Cox 1957; McEntegart 2003; Woodward

[5] MUA and MUP status by county and census track can are available at http://muafind.hrsa.gov/. MUA and MUP status for a specified address were obtained at http://datawarehouse.hrsa.gov/GeoAdvisor/ShortageDesignationAdvisor.aspx.

2005). The site, practice size, and MUA strata resulted in 24 cells (= 4 x 3 x 2), or 6 cells per site. We used the ralloc program version 3.5.2 in STATA®, a routine that provides a list of cases randomly assigned to a treatment or a control group in each stratum using a well-known algorithm for generating blocks of varying size within each stratum (Ryan 2008; Ryan 2009). The algorithm randomly assigns practices in each stratum using random permuted blocks of varying size. Our implementation utilized block of sizes varying between 2 and 10. The ralloc program generates sufficient assignment sequences to complete the final block in the stratum, even if the number of units in the stratum is an odd number. Any extra assignments that exceeded the number of practices in a stratum were discarded. As a result, the number of treatment and control group practices may not be equal in each stratum, even if there were an even number of practices in the stratum.

## C. DISTRIBUTION OF TREATMENT AND CONTROL GROUP PRACTICES ACROSS STRATA AND OTHER OUTCOMES OF INTEREST

The implementation of the randomization approach described above using the three stratification variables resulted in a nearly balanced allocation of treatment and control group practices within each stratum (Table 1). A successful allocation should result in treatment and control groups that are different in pre-implementation characteristics and study outcomes only by chance. We tested whether this was the case by comparing the percentage of treatment and control group practices with EHR systems at the time of application, overall and by site (Table 2). Although some chance differences were evident, they were small and none were statistically significant from zero. Similarly, there were no significant differences between the percentage of treatment group and the percentage of control group practices, overall or by site, that were in urban areas (data not shown). We conclude that the practices were successfully randomized in a way that will minimize chance differences between treatment and control groups on key outcomes during the evaluation. We anticipate using the same stratifying variables and methodology to randomize Phase II practices in early 2010.

A file identifying which Phase I practices have been allocated to the treatment group and the control group has been sent to Actuarial Research Corporation (ARC). The file contains the EHRD practice ID and a variable called treatment that is coded as "Treatment" for treatment group practice or "Control" for control group practices.

Attachment: Electronic Health Records (EHR) Demonstration Application to Participate

cc: J. Schore, S. Felt-Lisk, M. Kovac, File

TABLE 1

DISTRIBUTION OF PHASE I TREATMENT AND CONTROL GROUP PRACTICES ACROSS STRATA

| Site | Number of Physicians | Medically Underserved Area | Number of Practices | Number of Treatment Group Practices | Number of Control Group Practices |
|---|---|---|---|---|---|
| Louisiana | 1-2 physicians | No | 57 | 28 | 29 |
| | 1-2 physicians | Yes | 66 | 33 | 33 |
| | 3-5 physicians | No | 28 | 15 | 13 |
| | 3-5 physicians | Yes | 23 | 12 | 11 |
| | 6+ physicians | No | 18 | 10 | 8 |
| | 6+ physicians | Yes | 12 | 6 | 6 |
| Total | | | 204 | 104 | 100 |
| Maryland | 1-2 physicians | No | 111 | 55 | 56 |
| | 1-2 physicians | Yes | 20 | 10 | 10 |
| | 3-5 physicians | No | 58 | 29 | 29 |
| | 3-5 physicians | Yes | 11 | 6 | 5 |
| | 6+ physicians | No | 43 | 22 | 21 |
| | 6+ physicians | Yes | 12 | 5 | 7 |
| Total | | | 255 | 127 | 128 |
| Pennsylvania | 1-2 physicians | No | 108 | 54 | 54 |
| | 1-2 physicians | Yes | 45 | 23 | 22 |
| | 3-5 physicians | No | 65 | 33 | 32 |
| | 3-5 physicians | Yes | 25 | 12 | 13 |
| | 6+ physicians | No | 23 | 10 | 13 |
| | 6+ physicians | Yes | 13 | 6 | 7 |
| Total | | | 279 | 138 | 141 |
| South Dakota | 1-2 physicians | No | 17 | 8 | 9 |
| | 1-2 physicians | Yes | 8 | 4 | 4 |
| | 3-5 physicians | No | 28 | 14 | 14 |
| | 3-5 physicians | Yes | 2 | 1 | 1 |
| | 6+ physicians | No | 27 | 14 | 13 |
| | 6+ physicians | Yes | 5 | 2 | 3 |
| Total | | | 87 | 43 | 44 |
| TOTAL | | | 825 | 412 | 413 |

Source: Mathematica Policy Research, Inc. calculations using data from the EHRD application and other constructed variables, as described in the text.

TABLE 2

PERCENTAGE OF TREATMENT AND CONTROL GROUP PRACTICES USING EHR SYSTEM AT TIME OF APPLICATION, OVERALL AND BY SITE

| Site | Percentage Using EHR System at Time of Application | | | p-value ($H_o$: Difference is equal to zero) |
|---|---|---|---|---|
| | Treatment Group | Control Group | Difference | |
| All four sites | 42.5% | 43.6% | -1.1% | 0.75 |
| Louisiana | 39.4% | 43.0% | -3.6% | 0.60 |
| Maryland | 48.0% | 50.8% | -2.8% | 0.66 |
| Pennsylvania | 41.3% | 39.7% | 1.6% | 0.79 |
| South Dakota | 37.2% | 36.4% | 0.9% | 0.93 |

Source: Mathematica Policy Research, Inc. calculations using the practice allocation summarized in Table 1.

REFERENCES

Blumenthal, David, Catherine DesRoches, Karen Donelan, Timothy Ferris, Ashish Jha, Rainu Kaushal, Sowmya Rao, and Sara Rosenbaum. Health Information Technology in the United States: The Information Base for Progress. Washington, DC: The George Washington University, Massachusetts General Hospital, and Robert Wood Johnson Foundation, 2006.

Cochran, William G., and Gertrude M. Cox. Experimental Designs. New York: John Wiley, 1957.

Government Accountability Office, "Medicare Physician Payment. Care Coordination Programs Used in Demonstration Show Promise, but Wider Use of Payment Approach May be Limited." Washington, DC, February 2008.

Greevy, Robert, Bo Lu, Jeffrey H. Silber, and Paul Rosenbaum. "Optimal Multivariate Matching Before Randomization." Biostatistics, vol. 5, no. 2, 2004, pp. 263-275.

McEntegart, Damian J. "The Pursuit of Balance Using Stratified and Dynamic Randomization Techniques: An Overview." Drug Information Journal, vol. 37, 2003, pp. 293-308.

Ryan, Philip. "ralloc – Random allocation of treatment in controlled trials." Available online [www.stata.com/statalist/archive/2007-10/msg00226.html]. Accessed on April 4, 2008.

Ryan, Philip. "ralloc – an update." Available online [http://www.stata.com/statalist/archive/2008-06/msg00403.html]. Accessed on February 9, 2009.

Simon, Steven, Rainu Kaushal, Paul D. Cleary, Chelsea A. Jenter, Lynn A. Volk, Eric G. Poon, John Orav, Helen G. Lo, Deborah H. Williams, and David W. Bates. "Correlates of Electronic Health Record Adoption in Office Practices: A Statewide Survey." Journal of the American Medical Informatics Association, vol. 14, no. 1, Jan/Feb 2007a, pp. 110-117.

Simon, Steven, Rainu Kaushal, Paul D. Cleary, Chelsea A. Jenter, Lynn A. Volk, Eric G. Poon, John Orav, Elisabeth Burdick, Eric G. Poon, and David W. Bates. "Physicians and Electronic Health Records. A Statewide Survey." Archives of Internal Medicine, vol. 167, March 12, 2007b, pp. 507-512.

Woodward, Mark. Epidemiology. Study Design and Data Analysis, Second Edition. Boca Raton: Chapman & Hall/CRC, 2005.

# APPENDIX I

# SUPPLEMENTAL TABLES TO CHAPTER IV

This page is intentionally left blank.

Table I.1. Average Incentive Payments During the First and Second Years of the Demonstration, by Practice Characteristics (for All Practices That Responded to the OSS)

| Practice Characteristic | Average Payment (Dollars) in Year 1 | Average Payment (Dollars) in Year 2 |
|---|---|---|
| Site | | |
| Louisiana | $5,244 | $6,702 |
| Maryland | $7,692 | $10,726 |
| Pennsylvania | $4,263 | $5,941 |
| South Dakota | $8,285 | $11,923 |
| p-value | 0.002*** | 0.000*** |
| Practice Size | | |
| 1-2 physicians | $2,684 | $3,380 |
| 3-5 physicians | $5,499 | $8,187 |
| 6 or more physicians | $13,473 | $16,116 |
| p-value | 0.000*** | 0.000*** |
| Practice Affiliation | | |
| Unaffiliated | $5,934 | $7,949 |
| Affiliated[a] | $6,118 | $8,427 |
| p-value | 0.832 | 0.621 |
| Located in a Rural Area | | |
| Yes | $6,569 | $9,957 |
| No | $5,908 | $7,885 |
| p-value | 0.572 | 0.111 |
| Located in a Medically Underserved Area | | |
| Yes | $5,528 | $7,438 |
| No | $6,198 | $8,516 |
| p-value | 0.485 | 0.314 |
| Participating in Another EHR, Quality Improvement, or Quality Reporting Program | | |
| Yes | $7,056 | $8,534 |
| No | $4,215 | $6,349 |
| p-value | 0.001*** | 0.108 |
| Number of Practices[b] | 338 | 311 |

Source: Office Systems Survey (OSS), conducted in spring and summer 2010 and 20 11; payment data provided by EHRD's implementation support contractor for treatment group practices that submitted OSS data in 2010 and/or 2011; baseline characteristics from practice applications and HRSA's Area Resources File (ARF).

Notes: The p-values from testing the equality of means across binary variables are from t-tests. The p-values from testing the equality of means across practice characteristics for a variable with multiple (>2) categories (practice size) are from the F-test of an analysis of variance.

[a]Owned by a hospital, hospital system, or larger medical group, or affiliated with a larger medical group, independent practice association, physician hospital organization, or other entity.

[b]Includes eligible and participating treatment practices that completed the OSS.

***Statistically significant at 1 p ercent level; **statistically significant at 5 percent level; *statistically significant at 10 percent level.

EHR = Electronic health record; EHRD = Electronic Health Records Demonstration.

This page is intentionally left blank.

# APPENDIX J

# SUPPLEMENTAL TABLES TO CHAPTER V

This page is intentionally left blank.

# TABLES

This page is intentionally left blank.

Table J.1. Impacts of EHRD on Specific Health IT Functions, by Domain (Percentages)

| Functions | Predicted Treatment Group Mean | Predicted Control Group Mean | Impact | p-value |
|---|---|---|---|---|
| **Domain 1: Completeness of Information** | | | | |
| 4.1a. Paper records that have been transitioned to the EHR system. | 81.90 | 66.50 | 15.40 | 0.00 |
| 4.1b. Paper charts that were pulled for scheduled patient visits over the past month. | 52.99 | 46.13 | 6.86 | 0.10 |
| 4.1d. Clinical notes for individual patients | 83.79 | 68.59 | 15.21 | 0.00 |
| 4.1e. Allergy lists for individual patients | 86.25 | 71.99 | 14.26 | 0.00 |
| 4.1f. Problem or diagnosis lists for individual patients | 84.53 | 70.42 | 14.11 | 0.00 |
| 4.1g. Patient demographics (for example, age or sex) | 86.19 | 72.62 | 13.57 | 0.00 |
| 4.1h. Patient medical histories | 83.93 | 68.55 | 15.38 | 0.00 |
| 4.1i. Recording (or entering) laboratory orders into electronic system | 81.01 | 67.08 | 13.93 | 0.00 |
| 4.1j. Receiving laboratory results by fax or mail and scanning paper versions into electronic system | 79.91 | 62.60 | 17.31 | 0.00 |
| 4.1k. Reviewing laboratory test results electronically | 78.57 | 65.35 | 13.21 | 0.00 |
| 4.1l. Recording (or entering) imaging orders into electronic system | 77.40 | 63.43 | 13.97 | 0.00 |
| 4.1m. Receiving imaging results by fax or mail and scanning paper versions into electronic system | 77.96 | 63.86 | 14.10 | 0.00 |
| 4.1n. Reviewing imaging results electronically | 73.93 | 60.25 | 13.68 | 0.00 |
| 4.1o1. Recording that instructions or educational information were given to diabetes patients | 63.06 | 43.60 | 19.46 | 0.00 |
| 4.1o2. Recording that instructions or educational information were given to coronary artery disease patients | 48.09 | 28.00 | 20.08 | 0.00 |
| 4.1o3. Recording that instructions or educational information were given to congestive heart failure patients | 44.94 | 26.50 | 18.44 | 0.00 |
| 4.1o4. Recording that instructions or educational information were given to preventive care patients | 59.33 | 42.80 | 16.53 | 0.00 |
| 4.1p. Recording (or entering) prescription medications (new prescriptions and refills) into electronic system | 87.79 | 75.63 | 12.16 | 0.00 |
| **Domain 2: Communication of Care Outside the Practice** | | | | |
| Laboratory Orders | | | | |
| 4.2a. Print and fax laboratory orders to facilities outside the practice | 61.52 | 48.76 | 12.76 | 0.00 |
| 4.2b. Fax laboratory orders electronically from system, or order electronically through a portal maintained by facilities outside the practice | 36.76 | 30.94 | 5.82 | 0.14 |
| 4.2c. Transmit laboratory orders electronically directly from system to facilities outside the practice that have the capability to receive such transmissions MU-C | 35.73 | 35.12 | 0.61 | 0.87 |
| Imaging Orders | | | | |
| 4.2d. Print and fax imaging orders to facilities outside the practice | 68.82 | 56.59 | 12.22 | 0.00 |
| 4.2e. Fax imaging orders electronically from system, or order electronically through a portal maintained by facilities outside the practice | 27.64 | 25.60 | 2.05 | 0.58 |

Table J.1 (continued)

| Functions | Predicted Treatment Group Mean | Predicted Control Group Mean | Impact | p-value |
|---|---|---|---|---|
| 4.2f. Transmit imaging orders electronically directly from system to facilities outside the practice that have the capability to receive such transmissions MU-C | 19.48 | 22.42 | -2.94 | 0.39 |
| Laboratory Results | | | | |
| 4.2g. Transfer electronic laboratory results (received in non-machine readable form, such as an e-fax) directly into system | 25.75 | 21.27 | 4.48 | 0.21 |
| 4.2h. Enter laboratory results manually into electronic system in a searchable field (whether received by fax, mail, or phone) | 63.03 | 50.84 | 12.19 | 0.00 |
| 4.2i. Receive electronically transmitted laboratory results directly into system from facilities that have the capability to send such transmissions | 68.27 | 58.52 | 9.75 | 0.01 |
| Imaging Results | | | | |
| 4.2j. Transfer electronic imaging results (received in non-machine readable form, such as an e-fax) directly into system | 25.14 | 19.12 | 6.03 | 0.09 |
| 4.2k. Enter imaging results manually into electronic system in a searchable field (whether received by fax, mail, or phone) | 51.93 | 43.03 | 8.90 | 0.02 |
| 4.2l. Receive electronically transmitted imaging results directly into system from facilities that have the capability to send such transmissions | 46.52 | 40.47 | 6.05 | 0.12 |
| Referral and Consultation Requests | | | | |
| 4.2m. Enter requests for referrals to or consultation with other providers (for example, specialists, sub-specialists, physical therapy, speech therapy, nutritionists) | 70.15 | 57.70 | 12.46 | 0.00 |
| Sharing Information with Other Providers | | | | |
| 4.2n. Transmit medication lists or other medical information to other providers (for example, hospitals, home health agencies, or other physicians) MU-C | 58.54 | 43.82 | 14.72 | 0.00 |
| 4.2o. Transmit laboratory results to other providers (for example, hospitals, home health agencies, or other physicians) Results are sent as machine-readable data. | 42.62 | 36.53 | 6.08 | 0.12 |
| 4.2p. Transmit imaging results to other providers (for example, hospitals, home health agencies, or other physicians) Results are sent as machine-readable data. | 38.58 | 33.78 | 4.81 | 0.22 |
| 4.2q. Receive electronically transmitted reports directly into system, such as discharge summaries, from hospitals or other facilities that have the capability to send such transmissions | 47.56 | 40.07 | 7.49 | 0.06 |

Table J.1 (continued)

| Functions | Predicted Treatment Group Mean | Predicted Control Group Mean | Impact | p-value |
|---|---|---|---|---|
| Prescription Orders | | | | |
| 4.2r. Print prescriptions (new prescriptions and refills) on a computer printer and fax to pharmacy or hand to patient | 74.51 | 59.50 | 15.02 | 0.00 |
| 4.2s. Fax prescription orders (new prescriptions and refills) electronically from electronic system | 66.29 | 54.26 | 12.03 | 0.00 |
| 4.2t. Transmit prescription orders (new prescriptions and refills) electronically directly from system to pharmacies that have the capability to receive such transmissions MU-C | 86.77 | 71.83 | 14.94 | 0.00 |
| Domain 3: Clinical Decision Support | | | | |
| 4.3a. Enter information from clinical notes into documentation templates | 79.64 | 64.60 | 15.03 | 0.00 |
| 4.3b. View graphs of patient height or weight data over time MU-C | 74.40 | 59.19 | 15.21 | 0.00 |
| 4.3c. View graphs of patient vital signs data over time (such as blood pressure or heart rate) MU-C | 74.20 | 60.00 | 14.20 | 0.00 |
| 4.3d. Flag incomplete or overdue test results MU-CDS | 63.68 | 53.30 | 10.38 | 0.01 |
| 4.3e. Highlight out-of-range test levels MU-CDS | 67.62 | 58.12 | 9.50 | 0.01 |
| 4.3f. View graphs of laboratory or other test results over time for individual patients | 62.52 | 53.76 | 8.76 | 0.03 |
| 4.3g. Prompt clinicians to order necessary tests, studies, or other services MU-CDS | 63.60 | 52.97 | 10.63 | 0.01 |
| 4.3h1. Review and act on reminders at the time of a patient encounter regarding interventions, screening, or follow-up office visits recommended by evidence-based practice guidelines for diabetes patients MU-CDS | 59.48 | 42.61 | 16.87 | 0.00 |
| 4.3h2. Review and act on reminders at the time of a patient encounter regarding interventions, screening, or follow-up office visits recommended by evidence-based practice guidelines for coronary artery disease patients MU-CDS | 46.44 | 26.80 | 19.64 | 0.00 |
| 4.3h3. Review and act on reminders at the time of a patient encounter regarding interventions, screening, or follow-up office visits recommended by evidence-based practice guidelines for congestive heart failure patients MU-CDS | 44.42 | 26.30 | 18.11 | 0.00 |
| 4.3h4. Review and act on reminders at the time of a patient encounter regarding interventions, screening, or follow-up office visits recommended by evidence-based practice guidelines for preventive care patients MU-CDS | 59.50 | 42.63 | 16.86 | 0.00 |
| 4.3i. Reference information on medications being prescribed | 76.32 | 63.60 | 12.72 | 0.00 |
| 4.3j. Reference guidelines and evidence-based recommendations when prescribing medication for a patient | 58.85 | 52.70 | 6.14 | 0.12 |
| 4.3k. Search for or generate a list of patients requiring a specific intervention (such as an immunization) MU-M | 53.42 | 41.35 | 12.07 | 0.00 |
| 4.3l. Search for or generate a list of patients on a specific medication (or on a specific dose of medication) MU-M | 54.06 | 44.27 | 9.79 | 0.01 |

Table J.1 (continued)

| Functions | Predicted Treatment Group Mean | Predicted Control Group Mean | Impact | p-value |
|---|---|---|---|---|
| 4.3m. Search for or generate a list of patients who are due for a lab or other test in a specific time interval MU-M | 47.28 | 37.88 | 9.40 | 0.01 |
| 4.3n. Search for or generate a list of patients who fit a set of criteria, such as age, diagnosis, and clinical indicator value. | 55.33 | 42.70 | 12.63 | 0.00 |
| Domain 4: Use of System to Increase Patient Engagement/Adherence | | | | |
| 4.4a. Manage telephone calls | 80.52 | 63.74 | 16.78 | 0.00 |
| 4.4b. Exchange secure messages with patients | 30.78 | 29.81 | 0.97 | 0.80 |
| 4.4c. Allow patients to view their medical records online MU-M | 29.37 | 25.84 | 3.53 | 0.27 |
| 4.4d. Allow patients to provide information online to update their records | 28.98 | 27.27 | 1.70 | 0.62 |
| 4.4e. Allow patients to request appointments online | 30.93 | 30.90 | 0.03 | 0.99 |
| 4.4f. Allow patients to request referrals online | 22.99 | 24.16 | -1.17 | 0.73 |
| 4.4g1. Produce hard copy or electronic reminders for diabetes patients about needed tests, studies, or other services (for example, immunizations) MU-M | 50.57 | 39.29 | 11.28 | 0.00 |
| 4.4g2. Produce hard copy or electronic reminders for coronary artery disease patients about needed tests, studies, or other services (for example, immunizations) MU-M | 36.77 | 24.67 | 12.09 | 0.00 |
| 4.4g3. Produce hard copy or electronic reminders for congestive heart failure patients about needed tests, studies, or other services (for example, immunizations) MU-M | 35.69 | 23.80 | 11.90 | 0.00 |
| 4.4g4. Produce hard copy or electronic reminders for preventive care patients about needed tests, studies, or other services (for example, immunizations) MU-M | 47.42 | 37.59 | 9.83 | 0.01 |
| 4.4h1. Generate written or electronic educational information to help diabetes patients understand their condition or medication MU-M | 55.71 | 39.92 | 15.79 | 0.00 |
| 4.4h2. Generate written or electronic educational information to help coronary artery disease patients understand their condition or medication MU-M | 39.94 | 24.93 | 15.01 | 0.00 |
| 4.4h3. Generate written or electronic educational information to help congestive heart failure patients understand their condition or medication MU-M | 38.39 | 23.78 | 14.61 | 0.00 |
| 4.4h4. Generate written or electronic educational information to help preventive care patients understand their condition or medication MU-M | 52.12 | 38.82 | 13.30 | 0.00 |
| 4.4i1. Create written care plans (personalized to patient's condition or age/gender for preventive care) to help guide diabetes patients in self-management | 43.50 | 29.15 | 14.35 | 0.00 |
| 4.4i2. Create written care plans (personalized to patient's condition or age/gender for preventive care) to help guide coronary artery disease patients in self-management | 33.94 | 19.40 | 14.54 | 0.00 |
| 4.4i3. Create written care plans (personalized to patient's condition or age/gender for preventive care) to help guide congestive heart failure patients in self-management | 30.66 | 17.32 | 13.35 | 0.00 |

Table J.1 (continued)

| Functions | Predicted Treatment Group Mean | Predicted Control Group Mean | Impact | p-value |
|---|---|---|---|---|
| 4.4i4. Create written care plans (personalized to patient's condition or age/gender for preventive care) to help guide preventive care patients in self-management | 39.01 | 27.77 | 11.24 | 0.00 |
| 4.4j1. Prompt provider to review patient self-management plan (or patient-specific preventive care plan) with the diabetes patient during a visit | 41.79 | 28.23 | 13.56 | 0.00 |
| 4.4j2. Prompt provider to review patient self-management plan (or patient-specific preventive care plan) with the coronary artery disease patient during a visit | 34.00 | 19.30 | 14.70 | 0.00 |
| 4.4j3. Prompt provider to review patient self-management plan (or patient-specific preventive care plan) with the congestive heart failure patient during a visit | 31.38 | 18.04 | 13.34 | 0.00 |
| 4.4j4. Prompt provider to review patient self-management plan (or patient-specific preventive care plan) with the preventive care patient during a visit | 38.03 | 28.61 | 9.43 | 0.01 |
| 4.4k1. Modify self-management plan (or patient-specific preventive care plan) as needed following a diabetes patient visit | 40.06 | 28.98 | 11.07 | 0.00 |
| 4.4k2. Modify self-management plan (or patient-specific preventive care plan) as needed following a coronary artery disease patient visit | 32.69 | 19.12 | 13.57 | 0.00 |
| 4.4k3. Modify self-management plan (or patient-specific preventive care plan) as needed following a congestive heart failure patient visit | 30.52 | 18.43 | 12.10 | 0.00 |
| 4.4k4. Modify self-management plan (or patient-specific preventive care plan) as needed following a preventive care patient visit | 37.99 | 28.77 | 9.22 | 0.01 |
| 4.4l. Identify generic or less expensive brand alternatives at the time of prescription entry | 78.72 | 61.76 | 16.96 | 0.00 |
| 4.4m. Reference drug formularies of the patient's health plans/pharmacy benefit manager to recommend preferred drugs at time of prescribing MU-M | 75.60 | 58.03 | 17.58 | 0.00 |
| Domain 5: Medication Safety | | | | |
| 4.5a. Maintain medication list for individual patients | 87.69 | 72.68 | 15.02 | 0.00 |
| 4.5b. Generate new prescriptions (that is, system prompts for common prescription details including medication type and name, strength, dosage, and quantity) MU-C | 87.91 | 71.67 | 16.23 | 0.00 |
| 4.5c. Generate prescription refills (that is, system allows provider to reorder a prior prescription by revising original details associated with it, rather than requiring re-entry) MU-C | 87.38 | 72.42 | 14.96 | 0.00 |
| 4.5d. Select individual medication for prescription (for example, from a drop-down list in the electronic system) MU-C | 87.59 | 71.05 | 16.53 | s0.00 |
| 4.5e. Calculate appropriate dose and frequency, or suggest administration route based on patient parameters such as age, weight, or functional limitations | 61.55 | 48.35 | 13.20 | 0.00 |
| 4.5f. Screen prescriptions for drug allergies against the patient's allergy information MU-C | 86.23 | 67.99 | 18.24 | 0.00 |

Table J.1 (continued)

| Functions | Predicted Treatment Group Mean | Predicted Control Group Mean | Impact | p-value |
|---|---|---|---|---|
| 4.5g. Screen new prescriptions for drug-drug interactions against the patient's list of current medications MU-C | 86.22 | 69.64 | 16.59 | 0.00 |
| 4.5h. Check for drug-laboratory interaction. | 37.68 | 32.10 | 5.58 | 0.16 |
| 4.5i. Check for drug-disease interaction | 48.84 | 38.49 | 10.35 | 0.01 |
| Number of Practices (Weighted) | 405 | 412 | | |
| Number of Practices (Unweighted) | 324 | 268 | | |

Sources: Office Systems Survey (OSS) conducted in spring and summer of 2011, and data drawn from the applications practices submitted to EHRD in 2008.

Notes: Reported means are regression-adjusted. Regressions control for state, whether the practice was in a medically underserved area , practice size, and health IT-related variables practices reported on the application to the demonstration. Observations for treatment and control group practices are adjusted for nonresponse to the 2011 OSS and for demonstration attrition. The weighted sample reflects all randomized practices, except for seven treatment practices and one control practices that were determined by CMS to be ineligible prior to the demonstration. Eighty percent (324 of 405) of eligible randomized treatment practices and 65 percent (268 of 412) of eligible randomized control group practices responded to the Year 2 OSS. MU-C indicates a function related to a Stage 1 meaningful use core set item; MU-M indicates a function related to a Stage 1 meaningful use menu set item; MU-CDS indicates a function likely related to implementation of one clinical decision support rule, a core set item. See Chapter III for additional information on meaningful use items.

Table J.2. Impacts of EHRD on the Percentage of Beneficiaries Receiving Appropriate Quality of Care During Year 1 and Year 2 of the EHRD, by Site (Percentages Unless Otherwise Indicated)

| | All Sites | | | Louisiana | | Maryland | | Pennsylvania | | South Dakota | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Quality-of-Care Measures | Control Group Mean at Baseline | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 |
| Among Beneficiaries with CAD (N=182,365): | | | | | | | | | | | |
| Any blood test for cholesterol or lipids | 81.6 | -0.6 | -0.2 | -1.1 | 0.5 | -0.3 | -0.3 | -0.6 | -1.1 | -0.3 | 0.2 |
| p-value | | 0.214 | 0.738 | 0.305 | 0.678 | 0.594 | 0.788 | 0.457 | 0.227 | 0.791 | 0.829 |
| Among Beneficiaries with Diabetes (N=276,881): | | | | | | | | | | | |
| Any blood test for HbA1c | 88.3 | -0.2 | 0.0 | 0.3 | 0.1 | -0.7 | -0.7 | -0.3 | -0.1 | 0.3 | 1.2* |
| p-value | | 0.460 | 0.917 | 0.671 | 0.940 | 0.137 | 0.362 | 0.640 | 0.927 | 0.528 | 0.053 |
| Any blood test for cholesterol or lipids | 85.8 | -0.5 | 0.2 | -1.6 | -0.6 | -0.3 | 0.5 | -0.5 | 0.1 | 0.9 | 1.0 |
| p-value | | 0.245 | 0.633 | 0.175 | 0.615 | 0.386 | 0.350 | 0.467 | 0.924 | 0.326 | 0.499 |
| Any urine test for protein (microalbuminuria) | 81.5 | 0.6 | 0.9 | 0.9 | 0.6 | 0.8 | 2.2** | 0.4 | -0.6 | -0.1 | 0.0 |
| p-value | | 0.147 | 0.139 | 0.343 | 0.650 | 0.189 | 0.027 | 0.577 | 0.516 | 0.901 | 0.994 |
| Among Female Beneficiaries Between Ages 40 and 69 (N=130,072): | | | | | | | | | | | |
| Any screening for breast cancer | 65.6 | -0.6 | -0.1 | 0.9 | 1.0 | -1.7** | -1.1 | -0.6 | -1.2 | -0.4 | 2.1 |
| p-value | | 0.243 | 0.874 | 0.400 | 0.497 | 0.034 | 0.287 | 0.611 | 0.356 | 0.711 | 0.245 |
| Number of Observations[a] | | 800,524 | | 188,905 | | 306,007 | | 174,331 | | 131,281 | |
| Number of Practices | | 822 | | 201 | | 255 | | 279 | | 87 | |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment and control group practices at the end of the baseline period (2008); at the end of Year 1 (June 1, 2009, through May 31, 2010), and/or at the end of Year 2 (June 1, 2010, through May 31, 2011).

Note: Reported coefficients and p-values are from difference-in-differences regression models that control for beneficiary demographic characteristics, diagnoses, reason for Medicare eligibility, practice-fixed effects, and indicators for whether the observation is in the Year 1 or the Year 2 period. Standard errors are robust to the clustering of patients within practices, and reported p-values correspond to the respective standard errors and coefficients on treatment status. Sample size (N) reflects sample that is pooled across four states. Each beneficiary has up to three observations (baseline, Year 1, and Year 2). A negative coefficient estimate suggests that the treatment group experienced a net reduction (due to smaller improvement or greater decline) in a quality measure relative to the control group in Year 1 or Year 2, holding all other covariates constant

Reported means are regression-adjusted. Regressions control for state, MUA, practice size, and health IT-related variables practices reported on the application to the demonstration. Observations for treatment and control group practices are adjusted for non-response to the 2011 OSS and for demonstration attrition. The weighted sample reflects all randomized practices, except for seven treatment practices and one control practices that were determined by CMS to be ineligible prior to the demonstration. Eighty percent (324 of 405) of eligible treatment practices and 65 percent (268 of 412) of eligible control group practices responded to the Year 2 OSS.

*** Statistically significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

CAD = coronary artery disease; CHF = congestive heart failure; EHRD=Electronic Health Records Demonstration; HbA1c = hemoglobin A1c.

Table J.3. Change in the Percentage of Beneficiaries Receiving Appropriate Quality of Care during the EHRD, by Level of Health IT Improvement

| | No Condition-Specific Health IT Improvement[a] | Moderate Condition-Specific Health IT Improvement[a] | Large Condition-Specific Health IT Improvement[a] |
|---|---|---|---|
| Any Blood Test for Cholesterol or Lipids (among Beneficiaries with CAD) | | | |
| Treatment | -1.4 | -0.9 | 0.6 |
| Control | -1.1 | -2.4 | -1.6 |
| Any Blood Test for HbA1c (among Beneficiaries with Diabetes) | | | |
| Treatment | 3.1 | 2.9 | 5.1 |
| Control | 3.3 | 3.7 | 1.2 |
| Any Blood Test for Cholesterol or Lipids (among Beneficiaries with Diabetes) | | | |
| Treatment | -2.0 | 0.2 | 0.1 |
| Control | -1.1 | -1.4 | -2.1 |
| Any Urine Test for Protein (microalbuminuria) (among Beneficiaries with Diabetes) | | | |
| Treatment | 4.1 | 4.7 | 1.7 |
| Control | 3.1 | 0.8 | -0.7 |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment and control group practices at the end of the baseline period (2008); at the end of Year 1 (June 1, 2009 through May 31, 2010), and/or at the end of Year 2 (June 1, 2010 through May 31, 2011).

Notes: Numbers are based on unadjusted means, as opposed to regression-adjusted means. Change in percentage of beneficiaries with hospitalizations is calculated as the change in practices' average hospitalization rates for beneficiaries with these conditions from baseline to Year 2. A chi-squared significance tests showed no statistical differences between low-, moderate-, and high-adopting practices in either treatment or control.

Total number of observations included in the analysis is 401,574 beneficiaries with at least one of the following conditions: CAD, CHF, and diabetes. Sample sizes for treatment practices with no improvement, moderate improvement, and large improvement in health IT use related to CAD were 220, 46, and 45, respectively. Sample sizes for treatment practices with no improvement, moderate improvement, and large improvement in health IT use related to CHF were 226, 42, and 43, respectively. Sample sizes for treatment practices with no improvement, moderate improvement, and large improvement in health IT use related to diabetes were 186, 70, and 55, respectively.

Sample sizes for control practices with no improvement, moderate improvement, and large improvement in health IT use related to CAD were 215, 19, and 21, respectively. Sample sizes for control practices with no improvement, moderate improvement, and large improvement in health IT use related to CHF were 217, 17, and 21, respectively. Sample sizes for control practices with no improvement, moderate improvement, and large improvement in health IT use related to diabetes were 192, 39, and 24, respectively.

[a]Condition-specific health IT improvement refers to a change from limited or no health IT use at baseline to some health IT use tailored to beneficiaries with CAD, CHF, and diabetes during the demonstration period. Practices classified as having no improvement did not adopt any new health IT functions related to these conditions, practices classified as having moderate improvement adopted between 2 and 4 health IT functions, and practices classified as having large improvement adopted between 5 and 7 heal th IT functions. Functions included using health IT to record that educational materials were distributed to patients regarding their condition, review and act on reminders at the time of the patient encounter, produce reminders for tests and other services, generate electronic educational information, create written care plans, review self-management plans, and modify self-management plans

Table J.4. Impacts of EHRD on the Percentage of Beneficiaries with Preventable Hospitalizations During Year 1 and Year 2 of the EHRD, by Site (Percentages)

| | All Sites | | | Louisiana | | Maryland | | Pennsylvania | | South Dakota | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Beneficiary Subgroup | Control Group Mean at Baseline | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 |
| Among Beneficiaries with CAD (N=182,365): | | | | | | | | | | | |
| Any hospitalization related to CAD | 8.0 | -0.1 | -0.4 | 0.2 | -1.3** | 0.0 | 0.4 | -0.2 | -0.8 | -0.8 | -0.9 |
| p-value | | 0.819 | 0.186 | 0.751 | 0.027 | 1.000 | 0.407 | 0.752 | 0.188 | 0.432 | 0.427 |
| Among Beneficiaries with CHF(N=86,694): | | | | | | | | | | | |
| Any hospitalization related to CHF | 17.2 | 0.2 | -0.5 | 0.4 | -1.0 | -0.1 | -0.4 | 2.5* | 0.6 | -2.9** | -1.6 |
| p-value | | 0.767 | 0.497 | 0.719 | 0.406 | 0.958 | 0.787 | 0.053 | 0.669 | 0.035 | 0.397 |
| Among Beneficiaries with Diabetes (N=276,881): | | | | | | | | | | | |
| Any hospitalization related to diabetes | 8.3 | -0.1 | -0.4 | 0.4 | -0.8* | 0.0 | -0.1 | -0.2 | 0.0 | -0.9* | -0.8 |
| p-value | | 0.772 | 0.155 | 0.377 | 0.095 | 0.992 | 0.841 | 0.696 | 0.984 | 0.073 | 0.231 |
| Among Beneficiaries with CAD, CHF, or diabetes (N=401,574): | | | | | | | | | | | |
| Any hospitalization related to these conditions | 11.8 | 0.0 | -0.5** | 0.2 | -1.2*** | 0.3 | 0.0 | -0.2 | -0.2 | -1.1** | -0.9 |
| p-value | | 0.902 | 0.050 | 0.601 | 0.005 | 0.404 | 0.968 | 0.692 | 0.748 | 0.045 | 0.178 |
| Number of Observations[a] | | 800,524 | | 188,905 | | 306,007 | | 174,331 | | 131,281 | |
| Number of Practices | | 822 | | 201 | | 255 | | 279 | | 87 | |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment and control group practices at the end of the baseline period (2008); at the end of Year 1 (June 1, 2009, through May 31, 2010), and/or at the end of Year 2 (June 1, 2010, through May 31, 2011).

Note: Reported coefficients and p-values are from difference-in-differences regression models that control for beneficiary demographic characteristics, diagnoses, reason for Medicare eligibility, practice-fixed effects, and indicators for whether the observation is in the Year 1 or the Year 2 period. Standard errors are robust to the clustering of patients within practices, and reported p-values correspond to the respective standard errors and coefficients on treatment status. Sample size (N) reflects sample that is pooled across four states. Each beneficiary has up to three observations (baseline, Year 1, and Year 2). Less than 1 percent of the sample was dropped from the analysis due to missing control variables. A negative coefficient estimate suggests that the treatment group experienced a net reduction, which is a favorable outcome, in the number of preventable hospitalizations in Year 1 or Year 2 relative to the control group, holding all other covariates constant.

*** Statistically significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

CAD = coronary artery disease; CHF = congestive heart failure; EHRD=Electronic Health Records Demonstration.

Table J.5. Impacts of EHRD on the Number of Preventable Hospitalizations During Year 1 and Year 2 of the EHRD, by Site

| Number of Preventable Hospitalizations | Control Group Mean at Baseline | All Sites | | Louisiana | | Maryland | | Pennsylvania | | South Dakota | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 |
| Among Beneficiaries with CAD (N=182,365): | | | | | | | | | | | |
| Cardiac hospitalizations | 0.088 | -0.002 | -0.005 | 0.005 | -0.013** | -0.003 | 0.003 | -0.005 | -0.011* | -0.007 | -0.006 |
| p-value | | 0.650 | 0.144 | 0.434 | 0.037 | 0.652 | 0.575 | 0.477 | 0.067 | 0.483 | 0.614 |
| Among Beneficiaries with CHF(N=86,694): | | | | | | | | | | | |
| Hospitalizations for CHF | 0.222 | 0.000 | -0.005 | 0.007 | 0.003 | -0.009 | -0.019 | 0.025 | 0.010 | -0.029 | -0.011 |
| p-value | | 0.979 | 0.670 | 0.726 | 0.890 | 0.642 | 0.384 | 0.177 | 0.656 | 0.102 | 0.593 |
| Hospitalizations for fluid/ electrolyte problems | 0.012 | 0.000 | -0.001 | 0.003 | 0.002 | 0.001 | -0.004 | 0.001 | -0.002 | -0.005 | -0.001 |
| p-value | | 0.874 | 0.568 | 0.392 | 0.652 | 0.863 | 0.378 | 0.841 | 0.641 | 0.361 | 0.847 |
| Among Beneficiaries with Diabetes (N=276,881): | | | | | | | | | | | |
| Cardiac hospitalizations | 0.037 | 0.000 | -0.002 | 0.003 | -0.004 | 0.000 | 0.000 | 0.001 | -0.002 | -0.004 | -0.003 |
| p-value | | 0.853 | 0.313 | 0.379 | 0.189 | 0.898 | 0.970 | 0.851 | 0.510 | 0.408 | 0.524 |
| Diabetes hospitalizations | 0.019 | -0.002 | -0.003* | 0.001 | -0.003 | 0.000 | -0.001 | -0.009** | -0.008* | -0.002 | -0.001 |
| p-value | | 0.170 | 0.086 | 0.888 | 0.418 | 0.901 | 0.785 | 0.018 | 0.064 | 0.492 | 0.842 |
| Hospitalizations for peripheral vascular or extremity complication | 0.082 | -0.003 | -0.002 | 0.002 | -0.001 | 0.001 | -0.002 | -0.008 | -0.001 | -0.013* | -0.007 |
| p-value | | 0.385 | 0.556 | 0.708 | 0.899 | 0.851 | 0.752 | 0.267 | 0.929 | 0.059 | 0.402 |
| Hospitalizations for microvascular complication | 0.003 | 0.000 | 0.000 | -0.001 | -0.001 | 0.000 | 0.001 | -0.001 | -0.001 | 0.002 | 0.001 |
| p-value | | 0.953 | 0.866 | 0.489 | 0.504 | 0.914 | 0.295 | 0.687 | 0.542 | 0.117 | 0.642 |
| Number of Observations[a] | | 800,524 | | 188,905 | | 306,007 | | 174,331 | | 131,281 | |
| Number of Practices | | 822 | | 201 | | 255 | | 279 | | 87 | |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment and control group practices at the end of the baseline period (2008); at the end of Year 1 (June 1, 2009, through May 31, 2010), and/or at the end of Year 2 (June 1, 2010, through May 31, 2011).

Table J.5 (continued)

Note: Reported coefficients and p-values are from difference-in-differences regression models that control for beneficiary demographic characteristics, diagnoses, reason for Medicare eligibility, practice-fixed effects, and indicators for whether the observation is in the Year 1 or the Year 2 period. Standard errors are robust to the clustering of patients within practices, and reported p-values correspond to the respective standard errors and coefficients on treatment status. Sample size (N) reflects sample that is pooled across four states. Each beneficiary has up to three observations (baseline, Year 1, and Year 2). Less than 1 percent of the sample was dropped from the analysis due to missing control variables. A negative coefficient estimate suggests that the treatment group experienced a net reduction, which is a favorable outcome, in the number of preventable hospitalizations in Year 1 or Year 2 relative to the control group, holding all other covariates constant.

*** Statistically significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

CAD = coronary artery disease; CHF = congestive heart failure; EHRD=Electronic Health Records Demonstration.

Table J.6. Impacts of EHRD on the Percentage of Beneficiaries with Preventable Hospitalizations Related to Diabetes, CAD, or CHF During Year 1 and Year 2 of the EHRD, by Practice Subgroup (Percentages)

| | All Sites | | | | |
|---|---|---|---|---|---|
| Practice Subgroups | Control Group Mean at Baseline | Year 1 | Year 2 | Number of Observations[a] | Number of Practices[b] |
| Practice Size | | | | | |
| 1 or 2 physicians | 12.0 | -0.5 | -1.0** | 109,661 | 429 |
| p-value | | 0.264 | 0.039 | | |
| 3 or more physicians | 11.7 | 0.1 | -0.3 | 291,913 | 393 |
| p-value | | 0.577 | 0.279 | | |
| Caseload | | | | | |
| < 100 beneficiaries per physician | 12.6 | 0.3 | 0.0 | 148,500 | 434 |
| p-value | | 0.470 | 0.943 | | |
| ≥ 100 beneficiaries per physician | 11.3 | -0.2 | -0.8** | 252,747 | 384 |
| p-value | | 0.467 | 0.012 | | |
| Used EHR/Health IT Before EHRD | | | | | |
| Yes | 11.4 | -0.1 | -0.8** | 194,907 | 355 |
| p-value | | 0.666 | 0.030 | | |
| No | 12.1 | 0.0 | -0.2 | 206,667 | 467 |
| p-value | | 0.877 | 0.519 | | |
| Original Reason for Medicare Eligibility | | | | | |
| Aged | 11.2 | 0.0 | -0.3 | 320,439 | 822 |
| p-value | | 0.882 | 0.186 | | |
| Disabled | 14.2 | 0.0 | -0.9* | 81,135 | 817 |
| p-value | | 0.973 | 0.077 | | |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment and control group practices at the end of the baseline period (2008); at the end of Year 1 (June 1, 2009, through May 31, 2010), and/or at the end of Year 2 (June 1, 2010, through May 31, 20p11).

Note: Reported coefficients and p-values are from difference-in-differences regression models that control for beneficiary demographic characteristics, diagnoses, reason for Medicare eligibility, practice-fixed effects, and indicators for whether the observation is in the Year 1 or the Year 2 period. Standard errors are robust to the clustering of patients within practices, and reported p-values correspond to the respective standard errors and coefficients on treatment status. Sample size (N) reflects sample that is pooled across four states. Each beneficiary has up to three observations (baseline, Year 1, and Year 2). Less than 1 percent of the sample was dropped from the analysis due to missing control variables. A negative coefficient estimate suggests that the treatment group experienced a net reduction, which is a favorable outcome, in the number of preventable hospitalizations in Year 1 or Year 2 relative to the control group, holding all other covariates constant.

[a]The number of observations included in impact estimates is restricted to only beneficiaries with diabetes, CAD, or CHF.

[b]Caseload information was unavailable for a small number of practices that had no beneficiaries assigned during baseline. As a result, these practices (and their corresponding observations) were excluded from impact estimates related to caseload. In addition, five practices had no beneficiaries (with diabetes, CAD, or CHF) who qualified for Medicare due to a disability. As such, the number of practices included in impact estimates for beneficiaries with disabilities is 817 (as opposed to 822).

*** Statistically significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

CAD = coronary artery disease; CHF = congestive heart failure; EHRD=Electronic Health Records Demonstration.

Table J.7. Change in Percentage of Beneficiaries with Preventable Hospitalizations During the EHRD, by Level of Health IT Improvement

| | No Condition-Specific Health IT Improvement[a] | Moderate Condition-Specific Health IT Improvement[a] | Large Condition-Specific Health IT Improvement[a] |
|---|---|---|---|
| Change in CAD-Related Hospitalizations | | | |
| Treatment | -0.4 | -0.2 | -2.0 |
| Control | 0.1 | 0.1 | 0.3 |
| Change in CHF-Related Hospitalizations | | | |
| Treatment | 0.9 | 0.7 | -0.2 |
| Control | 0.8 | 0.7 | 0.5 |
| Change in Diabetes-Related Hospitalizations | | | |
| Treatment | 0.0 | -0.2 | -2.2 |
| Control | 0.4 | 0.0 | 0.4 |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment and control group practices at the end of the baseline period (2008); at the end of Year 1 (June 1, 2009, through May 31, 2010), and/or at the end of Year 2 (June 1, 2010, through May 31, 2011).

Notes: Numbers are based on unadjusted means, as opposed to regression-adjusted means. Change in percentage of beneficiaries with hospitalizations is calculated as the change in practices' average hospitalization rates for beneficiaries with these conditions from baseline to Year 2. A chi-squared significance test showed no statistical differences between low-, moderate-, and high-adopting practices in either treatment or control groups.

Total number of observations included in the analysis is 401,574 beneficiaries with at least one of the following conditions: CAD, CHF, and diabetes. Sample sizes for treatment practices with no improvement, moderate improvement, and large improvement in health IT use related to CAD were 220, 46, and 45, respectively. Sample sizes for treatment practices with no improvement, moderate improvement, and large improvement in health IT use related to CHF were 226, 42, and 43, respectively. Sample sizes for treatment practices with no improvement, moderate improvement, and large improvement in health IT use related to diabetes were 186, 70, and 55, respectively.

Sample sizes for control practices with no improvement, moderate improvement, and large improvement in health IT use related to CAD were 215, 19, and 21, respectively. Sample sizes for control practices with no improvement, moderate improvement, and large improvement in health IT use related to CHF were 217, 17, and 21, respectively. Sample sizes for control practices with no improvement, moderate improvement, and large improvement in health IT use related to diabetes were 192, 39, and 24, respectively.

[a]Condition-specific health IT improvement refers to a change from limited or no health IT use at baseline to some health IT use tailored to beneficiaries with CAD, CHF, and diabetes during the demonstration period. Practices classified as having no improvement did not adopt any new health IT functions related to these conditions, practices classified as having moderate improvement adopted between two and four health IT functions, and practices classified as having large improvement adopted between five and seven health IT functions. Functions included using health IT to record that educational materials were distributed to patients regarding their condition, review and act on reminders at the time of the patient encounter, produce reminders for tests and other services, generate electronic educational information, create written care plans, review self-management plans, and modify self-management plans.

CAD = coronary artery disease; CHF = congestive heart failure; EHRD = Electronic Health Records Demonstration; IT = information technology.

Table J.8. Impacts of EHRD on Total Medicare Expenditures per Beneficiary During Year 1 and Year 2 of the EHRD, by Practice Subgroup (in U.S. Dollars)

| | All Sites | | | | |
|---|---|---|---|---|---|
| Practice Subgroup | Control Group Mean at Baseline | Year 1 | Year 2 | Number of Observations[a] | Number of Practices[b] |
| Practice Size | | | | | |
| 1 or 2 physicians | 11,836.61 | 308.30 | 336.72 | 211,573 | 429 |
| p-value | | 0.164 | 0.225 | | |
| 3 or more physicians | 10,968.08 | -8.26 | 48.55 | 588,951 | 393 |
| p-value | | 0.956 | 0.763 | | |
| Caseload | | | | | |
| < 100 beneficiaries per physician | 11,468.61 | -129.42 | 19.14 | 299,810 | 434 |
| p-value | | 0.523 | 0.926 | | |
| ≥ 100 beneficiaries per physician | 11,034.91 | 205.78 | 186.73 | 500,140 | 384 |
| p-value | | 0.195 | 0.315 | | |
| Used EHR/Health IT Before EHRD | | | | | |
| Yes | 10,845.66 | -43.29 | 155.08 | 398,569 | 355 |
| p-value | | 0.811 | 0.439 | | |
| No | 11,559.11 | 191.26 | 86.09 | 401,955 | 467 |
| p-value | | 0.276 | 0.655 | | |
| Original Reason for Medicare Eligibility | | | | | |
| Aged | 10,293.68 | 50.77 | 161.16 | 651,356 | 822 |
| p-value | | 0.676 | 0.250 | | |
| Disabled | 15,410.60 | 133.30 | -63.18 | 149,168 | 820 |
| p-value | | 0.679 | 0.863 | | |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment and control group practices at the end of the baseline period (2008); at the end of Year 1 (June 1, 2009, through May 31, 2010), and/or at the end of Year 2 (June 1, 2010, through May 31, 2011).

Note: Reported coefficients and p-values are from difference-in-difference regression models that control for beneficiary demographic characteristics, diagnoses, reason for Medicare eligibility, practice-fixed effects, and indicators for whether the observation is in the Year 1 or the Year 2 period. Service use observed only during months that beneficiaries were alive and not in managed care. Standard errors are robust to the clustering of patients within practices, and reported p-values correspond to the respective standard errors and coefficients on treatment status. A negative coefficient estimate suggests that the treatment group experienced a net reduction in Medicare service use in Year 1 or Year 2 relative to the control group, holding all other covariates constant.

[a]Each beneficiary has up to three observations (baseline, Year 1, and Year 2). Less than 1 percent of the sample was dropped due to missing control variables.

[b]Caseload information was unavailable for a small number of practices that had no beneficiaries assigned during baseline. As a result, these practices (and their corresponding observations) were excluded from impact estimates related to caseload. In addition, two practices had no beneficiaries who qualified for Medicare due to a disability. As such, the number of practices included in impact estimates for beneficiaries with disabilities is 820 (as opposed to 822).

EHRD = Electronic Health Records Demonstration; IT = information technology.*** Statistically significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table J.9. Impacts of EHRD on Medicare Expenditures per Beneficiary During Year 1 and Year 2 of the EHRD, by Site (in U.S. Dollars)

| | | | | Louisiana | | Maryland | | Pennsylvania | | South Dakota | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Selected Expenditure Outcome | Control Group Mean at Baseline | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 |
| Total Expenditures | 11,196 | 79.64 | 128.61 | 514.9** | 421.5 | 78.66 | 71.97 | -190.47 | 226.58 | -222.41 | -151.56 |
| p-value | | 0.532 | 0.361 | 0.035 | 0.131 | 0.715 | 0.764 | 0.523 | 0.434 | 0.291 | 0.588 |
| All Part A (Excluding Home Health) | 5,960 | 87.98 | 64.45 | 219.37 | 149.72 | 185.51 | 116.6 | -22.92 | 221.12 | -197.39 | -303.71 |
| p-value | | 0.37 | 0.568 | 0.245 | 0.504 | 0.278 | 0.543 | 0.914 | 0.310 | 0.267 | 0.174 |
| Selected Part A Services | | | | | | | | | | | |
| Inpatient hospital | 4,910 | 103.55 | 92.19 | 291.57* | 230.55 | 205.26 | 155.12 | -41.68 | 139.98 | -212.14 | -220.73 |
| p-value | | 0.230 | 0.349 | 0.084 | 0.255 | 0.178 | 0.381 | 0.821 | 0.387 | 0.132 | 0.257 |
| Skilled nursing facility | 966 | -18.72 | -30.63 | -65.73 | -72.66 | -28.49 | -52.01 | -5.15 | 84.74 | 39.02 | -84.76 |
| p-value | | 0.494 | 0.372 | 0.253 | 0.330 | 0.485 | 0.270 | 0.934 | 0.323 | 0.581 | 0.278 |
| Hospice | 84 | 3.15 | 2.89 | -6.47 | -8.17 | 8.74 | 13.48* | 23.91 | -3.61 | -24.27** | 1.78 |
| p-value | | 0.571 | 0.661 | 0.602 | 0.660 | 0.204 | 0.066 | 0.103 | 0.810 | 0.025 | 0.906 |
| All Part B Expenditures (Excluding Home Health) | 4554 | -30.12 | 64.83 | 177.37** | 242.82*** | -89.37 | -23.05 | -168.75 | -5.37 | -24.23 | 149.02 |
| p-value | | 0.504 | 0.240 | 0.012 | 0.003 | 0.159 | 0.751 | 0.253 | 0.970 | 0.662 | 0.393 |
| Selected Part B Services | | | | | | | | | | | |
| Physician services | 2690 | -12.83 | -13.53 | 122.09*** | 90.41* | -21.19 | -27.12 | -145.34 | -74.09 | -14.09 | -28.18 |
| p-value | | 0.704 | 0.698 | 0.002 | 0.051 | 0.493 | 0.446 | 0.276 | 0.572 | 0.675 | 0.520 |
| Outpatient services | 1534 | -23.78 | 68.55* | 32.14 | 128.89** | -65.25 | 11.47 | -40.63 | 30.35 | -8.03 | 177.36 |
| p-value | | 0.288 | 0.077 | 0.448 | 0.024 | 0.139 | 0.833 | 0.269 | 0.454 | 0.832 | 0.277 |
| Durable medical equipment | 330 | 6.49 | 9.8 | 23.14 | 23.52 | -2.94 | -7.4 | 17.22 | 38.38** | -2.12 | -0.16 |
| p-value | | 0.358 | 0.244 | 0.248 | 0.321 | 0.716 | 0.429 | 0.203 | 0.032 | 0.860 | 0.992 |
| Part A and Part B Services | | | | | | | | | | | |
| Home health services | 682 | 21.78 | -0.67 | 118.17*** | 28.96 | -17.48 | -21.57 | 1.2 | 10.82 | -0.79 | 3.13 |
| p-value | | 0.139 | 0.971 | 0.008 | 0.639 | 0.258 | 0.254 | 0.956 | 0.671 | 0.952 | 0.828 |
| Number of Observations[a] | | 800,524 | | 188,905 | | 306,007 | | 174,331 | | 131,281 | |
| Number of Practices | | 822 | | 201 | | 255 | | 279 | | 87 | |

Table J.9 (continued)

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment and control group practices at the end of the baseline period (2008); at the end of Year 1 (June 1, 2009, through May 31, 2010), and/or at the end of Year 2 (June 1, 2010, through May 31, 2011).

Note: Reported coefficients and p-values are from difference-in-difference regression models that control for beneficiary demographic characteristics, diagnoses, reason for Medicare eligibility, practice-fixed effects, and indicators for whether the observation is in the Year 1 or the Year 2 period. Service use observed only during months that beneficiaries were alive and not in managed care. Standard errors are robust to the clustering of patients within practices, and reported p-values correspond to the respective standard errors and coefficients on treatment status. A negative coefficient estimate suggests that the treatment group experienced a net reduction in Medicare service use in Year 1 or Year 2 relative to the control group, holding all other covariates constant.

[a]Each beneficiary has up to three observations (baseline, Year 1, and Year 2). Less than 1 percent of the sample was dropped due to missing control variables.

*** Statistically significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table J.10. Impacts of EHRD on Selected Medicare Service Use per Beneficiary During Year 1 and Year 2 of the EHRD, by Site

| | All Sites | | | Louisiana | | Maryland | | Pennsylvania | | South Dakota | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Selected Service Use Outcomes | Control Group Mean at Baseline | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 |
| Number of Inpatient Hospital Days | 3.361 | 0.034 | 0.003 | 0.121 | -0.015 | 0.075 | 0.040 | -0.050 | 0.042 | -0.052 | -0.046 |
| p-value | | 0.541 | 0.963 | 0.379 | 0.918 | 0.324 | 0.671 | 0.722 | 0.716 | 0.489 | 0.664 |
| Number of Inpatient Hospital Stays | 0.474 | 0.000 | -0.002 | 0.010 | 0.010 | -0.002 | -0.011 | -0.008 | 0.002 | -0.002 | 0.001 |
| p-value | | 0.975 | 0.792 | 0.481 | 0.442 | 0.828 | 0.316 | 0.489 | 0.859 | 0.904 | 0.951 |
| Number of Emergency Room Visits | 0.469 | 0.006 | 0.011 | 0.019 | 0.012 | 0.008 | 0.003 | 0.000 | 0.009 | -0.007 | 0.034* |
| p-value | | 0.364 | 0.145 | 0.304 | 0.543 | 0.413 | 0.798 | 0.991 | 0.552 | 0.625 | 0.082 |
| Number of Physician Visits | 20.051 | 0.038 | 0.016 | 0.321* | 0.388* | -0.175 | -0.279 | 0.054 | 0.227 | 0.099 | -0.007 |
| p-value | | 0.650 | 0.879 | 0.053 | 0.077 | 0.210 | 0.110 | 0.760 | 0.211 | 0.611 | 0.979 |
| Number of Outpatient Visits | 4.554 | 0.021 | 0.015 | 0.052 | 0.099 | 0.002 | -0.028 | -0.080 | -0.191 | 0.150 | 0.291 |
| p-value | | 0.644 | 0.866 | 0.535 | 0.433 | 0.974 | 0.802 | 0.371 | 0.130 | 0.309 | 0.446 |
| Number of Observations[a] | | 800,524 | | 188,905 | | 306,007 | | 174,331 | | 131,281 | |
| Number of Practices | | 822 | | 201 | | 255 | | 279 | | 87 | |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment and control group practices at the end of the baseline period (2008); at the end of Year 1 (June 1, 2009, through May 31, 2010), and/or at the end of Year 2 (June 1, 2010, through May 31, 2011).

Note: Reported coefficients and p-values are from difference-in-difference regression models that control for beneficiary demographic characteristics, diagnoses, reason for Medicare eligibility, practice-fixed effects, and indicators for whether the observation is in the Year 1 or the Year 2 period. Service use observed only during months that beneficiaries were alive and not in managed care. Standard errors are robust to the clustering of patients within practices, and reported p-values correspond to the respective standard errors and coefficients on treatment status. A negative coefficient estimate suggests that the treatment group experienced a net reduction in Medicare service use in Year 1 or Year 2 relative to the control group, holding all other covariates constant.

[a]Each beneficiary has up to three observations (baseline, Year 1, and Year 2). Less than 1 percent of the sample was dropped due to missing control variables.

*** Statistically significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table J.11. Impacts of EHRD on Number of In-Patient Hospital Stays During Year 1 and Year 2 of the EHRD, by Practice Subgroup

| | All Sites | | | | |
|---|---|---|---|---|---|
| Practice Subgroup | Control Group Mean at Baseline | Year 1 | Year 2 | Number of Observations[a] | Number of Practices[b] |
| Practice Size | | | | | |
| 1 or 2 physicians | 0.497 | -0.008 | -0.002 | 211,573 | 429 |
| p-value | | 0.464 | 0.865 | | |
| 3 or more physicians | 0.466 | 0.002 | -0.002 | 588,951 | 393 |
| p-value | | 0.744 | 0.809 | | |
| Caseload | | | | | |
| < 100 beneficiaries per physician | 0.494 | -0.007 | 0.001 | 299,810 | 434 |
| p-value | | 0.439 | 0.895 | | |
| ≥ 100 beneficiaries per physician | 0.462 | 0.004 | -0.004 | 500,140 | 384 |
| p-value | | 0.587 | 0.651 | | |
| Used EHR/Health IT Before EHRD | | | | | |
| Yes | 0.450 | -0.001 | 0.001 | 398,569 | 355 |
| p-value | | 0.861 | 0.893 | | |
| No | 0.498 | 0.001 | -0.005 | 401,955 | 467 |
| p-value | | 0.938 | 0.586 | | |
| Original Reason for Medicare Eligibility | | | | | |
| Aged | 0.430 | -0.002 | -0.001 | 651,356 | 822 |
| p-value | | 0.708 | 0.913 | | |
| Disabled | 0.677 | 0.006 | -0.004 | 149,168 | 820 |
| p-value | | 0.727 | 0.781 | | |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment and control group practices at the end of the baseline period (2008); at the end of Year 1 (June 1, 2009, through May 31, 2010), and/or at the end of Year 2 (June 1, 2010, through May 31, 2011).

Note: Reported coefficients and p-values are from difference-in-difference regression models that control for beneficiary demographic characteristics, diagnoses, reason for Medicare eligibility, practice-fixed effects, and indicators for whether the observation is in the Year 1 or the Year 2 period. Service use observed only during months that beneficiaries were alive and not in managed care. Standard errors are robust to the clustering of patients within practices, and reported p-values correspond to the respective standard errors and coefficients on treatment status. A negative coefficient estimate suggests that the treatment group experienced a net reduction in Medicare service use in Year 1 or Year 2 relative to the control group, holding all other covariates constant.

[a]Each beneficiary has up to three observations (baseline, Year 1, and Year 2). Less than 1 percent of the sample was dropped due to missing control variables.

[b]Caseload information was unavailable for a small number of practices that had no beneficiaries assigned during baseline. As a result, these practices (and their corresponding observations) were excluded from impact estimates related to caseload. In addition, two practices had no beneficiaries who qualified for Medicare due to a disability. As such, the number of practices included in impact estimates for beneficiaries with disabilities is 820 (as opposed to 822).

*** Statistically significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table J.12. Impacts of EHRD on Selected Medicare Service Use Rates During Year 1 and Year 2 of the EHRD, by Site (Percentages)

| Selected Service Use Outcomes | All Sites | | | Louisiana | | Maryland | | Pennsylvania | | South Dakota | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Control Group Mean at Baseline | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 | Year 1 | Year 2 |
| Utilization of Inpatient Hospital Stays | 26.6 | -0.1 | -0.3 | -0.1 | -0.4 | -0.1 | -0.6 | 0.0 | -0.2 | -0.1 | 0.5 |
| p-value | | 0.847 | 0.350 | 0.818 | 0.488 | 0.808 | 0.213 | 0.986 | 0.665 | 0.933 | 0.552 |
| Utilization of Emergency Room Visits | 26.4 | -0.2 | 0.2 | 0.6 | 0.3 | -0.2 | -0.1 | -1.0** | 0.0 | -0.5 | 1.1 |
| p-value | | 0.343 | 0.410 | 0.289 | 0.646 | 0.524 | 0.851 | 0.040 | 0.956 | 0.388 | 0.200 |
| Utilization of Physician Visits | 100 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| p-value | | 0.845 | 0.237 | 0.959 | 0.660 | 0.958 | 0.273 | 0.865 | 0.611 | 0.794 | 0.473 |
| Utilization of Outpatient Visits | 77.7 | -0.1 | -0.3 | 0.4 | -0.5 | -0.4 | 0.2 | -1.0*** | -1.4** | 1.2 | 0.5 |
| p-value | | 0.729 | 0.499 | 0.444 | 0.514 | 0.397 | 0.817 | 0.009 | 0.047 | 0.257 | 0.790 |
| Utilization of Hospice Visits | 1.4 | 0.0 | 0.0 | -0.1 | -0.1 | 0.1* | 0.1 | 0.2 | 0.2 | -0.5*** | -0.2 |
| p-value | | 0.971 | 0.752 | 0.603 | 0.547 | 0.063 | 0.281 | 0.308 | 0.363 | 0.005 | 0.403 |
| Utilization of Skilled Nursing Facilities | 7.1 | -0.1 | -0.1 | -0.6* | -0.4 | -0.2 | -0.5* | 0.3 | 0.3 | 0.4 | 0.5 |
| p-value | | 0.764 | 0.474 | 0.069 | 0.339 | 0.404 | 0.074 | 0.397 | 0.454 | 0.322 | 0.285 |
| Utilization of Home Health Services | 12.5 | 0.2 | -0.2 | 1.06** | 0.1 | -0.2 | -0.4 | 0.2 | -0.3 | 0.1 | 0.0 |
| p-value | | 0.229 | 0.289 | 0.023 | 0.895 | 0.423 | 0.166 | 0.592 | 0.523 | 0.787 | 0.940 |
| Utilization of Durable Medical Equipment | 35.8 | 0.2 | -0.1 | 0.5 | 0.0 | 0.0 | -0.2 | 0.2 | 0.1 | 0.0 | 0.2 |
| p-value | | 0.514 | 0.775 | 0.321 | 0.974 | 0.962 | 0.595 | 0.660 | 0.880 | 0.945 | 0.812 |
| Number of Observations[a] | | 800,524 | | 188,905 | | 306,007 | | 174,331 | | 131,281 | |
| Number of Practices | | 822 | | 201 | | 255 | | 279 | | 87 | |

Source: Medicare enrollment and claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to treatment and control group practices at the end of the baseline period (2008); at the end of Year 1 (June 1, 2009, through May 31, 2010), and/or at the end of Year 2 (June 1, 2010, through May 31, 2011).

Table J.12 (continued)

Note: Reported coefficients and p-values are from difference-in-difference regression models that control for beneficiary demographic characteristics, diagnoses, reason for Medicare eligibility, practice-fixed effects, and indicators for whether the observation is in the Year 1 or the Year 2 period. Service use observed only during months that beneficiaries were alive and not in managed care. Standard errors are robust to the clustering of patients within practices, and reported p-values correspond to the respective standard errors and coefficients on treatment status. A negative coefficient estimate suggests that the treatment group experienced a net reduction in Medicare service use in Year 1 or Year 2 relative to the control group, holding all other covariates constant.

[a]Each beneficiary has up to three observations (baseline, Year 1, and Year 2). Less than 1 percent of the sample was dropped due to missing control variables.

*** Statistically significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table J.13. Rationale and Results of Sensitivity Tests for Medicare Expenditures and Service Use Measures for Years 1 and 2

| Sensitivity Test | Rationale for the Test | Results |
|---|---|---|
| Estimate model without practice-fixed effects | Were results sensitive to the model specification? | Results were similar to main findings. |
| Trim extreme values of Medicare expenditures (those above the 98th percentile) to the value of the 98th percentile | Were results sensitive to outliers (very high expenditures) in the distribution of Medicare expenditures? | Outpatient expenditures for Year 2 were no longer statistically significant. |
| Estimate models that control for individual fixed effects (indicator variables for each beneficiary) by restricting the sample to beneficiaries appearing in at least two waves of the data including baseline | Were results sensitive to controlling for both observed and unobserved beneficiary characteristics that are time invariant? | Outpatient expenditures for Year 2 were no longer statistically significant; durable medical equipment expenditures became statistically significant (and positive) for Year 2.<br><br>The probability of a physician visit became statistically significant and negative for Year 1. |
| Use a logarithmic transformation to estimate Part A, Part B, and total expenditures | Were results sensitive to large outliers in expenditure measures? | Part B expenditures became statistically significant (and positive) for Year 2. |

Source: Mathematica's calculations based on Medicare claims data provided by EHRD's implementation support contractor for all beneficiaries with any of the specified chronic conditions who were assigned to demonstration and control practices at the end of the baseline period (2008); at the end of Year 1 (June 1, 2009, through May 31, 2010), and/or at the end of Year 2 (June 1, 2010, through May 31, 2011).

Note: "Estimates" refer to the estimates obtained from difference-in-differences models described in Appendix G. These estimates reflect the difference between the treatment and control groups in the change in outcomes between baseline, Year 1, and Year 2.

MATHEMATICA
Policy Research

www.mathematica-mpr.com

Improving public well-being by conducting high-quality, objective research and surveys

Princeton, NJ ■ Ann Arbor, MI ■ Cambridge, MA ■ Chicago, IL ■ Oakland, CA ■ Washington, DC