Contract No.: HHSM-500-2005-00025I
MPR Reference No.: 6514-224

MATHEMATICA
Policy Research, Inc.

Program Evaluation of the 9th Scope of Work QIO Program: Evaluation Methodology, Conceptual Framework, and State Specific Provider Environment Task

Final Report

April 29, 2010

Arnold Chen
Andrew Clarkwest
Sarah Croake
Sue Felt-Lisk
Kate Stewart
Christianna Willams*

****Abt Associates***

Submitted to:

Centers for Medicare and Medicaid Services
Room S3-10-04
7500 Security Blvd,
Baltimore, Maryland 21244

Project Officer:
Cynthia Pamon

Government Task Leader:
Bob Kambic

Submitted by:

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

Project Director:
Myles Maxfield

CONTENTS

Chapter		Page
	EXECUTIVE SUMMARY	xi
I	INTRODUCTION	1
	A. BACKGROUND AND POLICY CONTEXT	1
	B. BRIEF DESCRIPTION OF THE 9TH SCOPE OF WORK	3
	1. National Themes	3
	2. Subnational Themes	10
	3. Summary of 9th SOW Themes	15
	C. OVERVIEW OF RESEARCH QUESTIONS AND GOALS OF THE EVALUATION	20
	D. CHALLENGES TO THE EVALUATION	20
	E. GUIDE TO THE REST OF THIS REPORT	21
II	CONCEPTUAL FRAMEWORK FOR THE 9TH SOW	23
	A. OVERVIEW	23
	B. ASSESSING AND DESCRIBING THE FRAMEWORK	28
	1. Inputs to QIO Activities	28
	2. QIO Activities	32
	3. Environment	32
	4. Outcomes	39
III	DESIGNS FOR IMPACT AND COST-BENEFIT/ COST-EFFECTIVENESS ANALYSES	41
	A. COMMON IMPACT ESTIMATION APPROACHES ACROSS THEMES AND SUBTHEME COMPONENTS	42
	1. Regression Discontinuity	42
	2. Matching and Comparison	45
	3. Trend Analyses	47
	B. COMMON DESCRIPTIVE ANALYSES ACROSS THEMES AND SUBTHEME COMPONENTS	47

CONTENTS *(continued)*

Chapter **Page**

C. THEME-BY-THEME ANALYSIS PLAN....48
1. Beneficiary Protection Theme--Assisting Hospitals with RHQDAPU....48
2. Patient Safety Theme....49
3. Prevention Disparities....62
4. Prevention....63
5. Care Transitions....65
6. Chronic Kidney Disease....69

D. COMMON COST-EFFECTIVENESS AND COST-BENEFIT METHODS....70

IV MECHANISMS ANALYSES....75

A. GATHERING DATA ON QIOS' ACTIVITIES....75
1. QIO Survey....75
2. Discussions with QIO Partner Organizations....76
3. Review of SDPS Documents....76
4. Focus Groups of Medicare Beneficiaries....77

B. GATHERING DATA ON STATE PROVIDER ENVIRONMENTS....78
1. Selection of Case Studies....78
2. Selection of Providers and Community Health Leaders Within Case Studies....79
3. Discussion Topics....80
4. Descriptions of Provider Environment....82

C. MECHANISMS AND IMPACTS....84
1. Which Types of QIO Program Were Most Effective?....84
2. Linking Provider Environments to Impacts....87
3. Which Interventions Work for Whom, and in What Circumstances?....87

V CONCLUSIONS....89

A. SELECTED EVALUATION CHALLENGES OF THE 9TH SOW AND IMPLICATIONS FOR THE 10TH SOW AND FUTURE SOWS....89

B. REPORTING RESULTS....91
1. Synthesizing Results....91
2. The Current Evaluation in the Context of Previous Studies of the QIO Program....93
3. Forthcoming Reports....93

C. PROJECT TIMELINE....94

REFERENCES....99

CONTENTS *(continued)*

Chapter		**Page**
APPENDIX A:	LOGIC MODELS FOR INDIVIDUAL 9TH SOW THEMES	
APPENDIX B:	TECHNICAL DETAILS OF ESTIMATION IN REGRESSION DISCONTINUITY DESIGNS	
APPENDIX C:	DEVELOPMENT OF ANALYTIC FILE FOR PATIENT SAFETY MEASURES RELATED TO NURSING HOME QUALITY	
APPENDIX D:	MINIMUM DETECTABLE IMPACT CALCULATIONS FOR CARE TRANSITIONS ANALYSES	
APPENDIX E:	HOW EVALUATION ACTIVITIES ADDRESS IOM AND NORC RECOMMENDATIONS	

TABLES

Table **Page**

I.1 NUMBERS OF PROVIDERS RECRUITED FOR SELECTED PATIENT SAFETY THEME COMPONENTS AS OF SEPTEMBER 2009 6

I.2 NUMBERS OF HOSPITALS RECRUITED FOR THE METHICILLIN RESISTANT STAPH AUREUS (MRSA) PATIENT SAFETY COMPONENT AS OF SEPTEMBER, 2009 8

I.3 NUMBERS OF PARTICIPATING PRACTICES RECRUITED FOR THE QIO 9TH SOW PREVENTION THEME AS OF FEBRUARY 2009 11

I.4 NUMBER OF PARTICIPATING PRACTICES RECRUITED FOR QIO 9TH SOW PREVENTION DISPARITIES THEME AS OF FEBRUARY 2009 12

I.5 SUMMARY OF 9TH SOW THEMES AND SUBTHEME COMPONENTS 16

II.1 OVERVIEW OF THE QIO PROGRAM AND EXPECTED BENEFITS FOR BENEFICIARIES 24

II.2 PROVIDER ENVIRONMENT TOPICS COVERED BY DATA COLLECTION EFFORTS 35

III.1 PERCEIVED VALUE OF RHQDAPU MEETINGS AMONG SURVEYED HOSPITALS WITH AT LEAST ONE SUCH MEETING (PERCENTAGE OF HOSPITALS) 49

III.2 PATIENT SAFETY THEME OUTCOME MEASURES FOR DESCRIPTIVE AND IMPACT ANALYSES 52

III.3 SELECTION MEASURES FOR PATIENT SAFETY IMPACT ANALYSES USING REGRESSION DISCONTINUITY DESIGN 54

III.4 POTENTIAL COVARIATES FOR IMPACT ANALYSES OF PATIENT SAFETY IN NURSING HOMES (PHYSICAL RESTRAINTS, PRESSURE ULCERS) 55

III.5 POTENTIAL COVARIATES FOR SCIP/HF HOSPITAL 56

III.6 MINIMUM DETECTABLE EFFECTS FOR REGRESSION DISCONTINUITY ANALYSES IN THE PATIENT SAFETY THEME 60

III.7 OUTCOME MEASURES TO BE USED IN THE IMPACT ANALYSES OF THE PREVENTION DISPARITIES THEME 62

TABLES *(continued)*

Table **Page**

III.8 POTENTIAL COVARIATES FOR IMPACT ANALYSES OF PREVENTION DISPARITIES THEME 64

III.9 POTENTIAL COVARIATES FOR MATCHING INTERVENTION COMMUNITIES WITH CONTROL COMMUNITIES FOR THE CARE TRANSITIONS THEME 67

III.10 POTENTIAL COVARIATES FOR IMPACT ANALYSES OF CARE TRANSITIONS THEME 68

III.11 RANGE OF MINIMUM DETECTABLE IMPACTS FOR CARE TRANSITIONS ANALYSES 69

III.12 POTENTIAL COMMUNITY CHARACTERISTICS FOR MATCHING INTERVENTION COMMUNITIES WITH COMPARISON COMMUNITIES FOR THE CHRONIC KIDNEY DISEASE THEME 71

III.13 ANALYTIC APPROACHES TO 9TH SOW THEMES AND SUBTHEME COMPONENTS 73

IV.1 OVERVIEW OF PARTNERSHIPS AND STATE-LEVEL EXPERIENCE REPORTED BY PARTNERS FOR THE CT AND CKD THEMES 77

IV.2 DISCUSSION TOPICS FOR SITE VISITS, BY TYPE OF ORGANIZATION OR PROVIDER 81

IV.3 MEAN QIO THEME LEADER AGREEMENT WITH EACH STATEMENT ABOUT THE PROVIDER ENVIRONMENT (3 = HIGHEST POSSIBLE AGREEMENT 84

IV.4 IMPACT ESTIMATES FOR HOSPITAL OUTCOMES BY QIO PROGRAM TYPOLOGY, PROVIDER ENVIRONMENT, AND PROVIDER CHARACTERISTIC 88

V.1 UPCOMING DELIVERABLE REPORTS FOR 9TH SOW EVALUATION 95

FIGURES

Figure **Page**

II.1 DRAFT CONCEPTUAL MODEL OF THE QIO PROGRAM 27

II.2 THE QIO ENVIRONMENT 29

III.1 REGRESSION DISCONTINUITY ILLUSTRATION 43

IV.1 PATRIOT SYSTEM SCREEN FOR ENTRY OF QIO ACTIVITIES FOR PATIENT SAFETY THEME 86

V.1 OVERVIEW OF THE UPCOMING SCHEDULE FOR THE EVALUATION 97

EXECUTIVE SUMMARY

The Quality Improvement Organization Program of the Centers for Medicare & Medicaid Services (CMS) is a key component of CMS's agenda for ensuring and improving quality of care for Medicare beneficiaries. As required by the Social Security Act, CMS contracts with a nationwide network of independent quality improvement organizations (QIOs) to help health care providers deliver high quality care to Medicare beneficiaries. The most recent contract, the 9th Statement of Work (SOW), runs from August 2008 through July 2011. Forty-three QIOs are carrying out the 9th SOW in the 50 states, plus the District of Columbia, Puerto Rico, and the U.S. Virgin Islands.

The importance of the QIO Program's functions and the magnitude of its budget (over one billion dollars for the 9th SOW) make its evaluation essential. Several recent assessments of the QIO program, including a Congressionally mandated overview by the Institute of Medicine (IOM) and a study sponsored by the Assistant Secretary for Planning and Evaluation (ASPE), have found inconclusive evidence of the program's effectiveness and have recommended further research. In response, CMS contracted with Mathematica Policy Research in September 2008 to independently design and conduct an evaluation of the 9th SOW. This report contains our design and approach to the evaluation.

One of the major challenges facing previous studies of the QIO program has been the stringent statutory and regulatory restrictions on QIOs' releasing information on the identities of providers who work with the QIOs; even CMS is not permitted access to this information. Previous government and academic studies have all pointed out the difficulties in evaluating the program's effectiveness when the identities of participating health care providers must remain secret. For the current evaluation, Mathematica has recently executed subcontracts with each of the QIOs under which Mathematica will assist the QIOs in determining the effectiveness of their services. Federal regulations require that QIOs disclose to subcontractors all information necessary for the subcontractors' work. Other challenges for the evaluation have included completing arrangements for the specially configured computers and network connections through which data access must occur, and obtaining access to certain specialized datasets necessary for the evaluation.

Another challenge to the evaluation is the sheer diversity of interventions and goals of the 9th SOW. The evaluation must essentially conduct multiple smaller evaluations, one for each type of intervention. Each of these smaller evaluations has its own design, data sources, and analytic approach. The SOW is organized into six themes—three covering the entire nation (national themes), and three involving selected states (sub-national themes)—spanning a wide range of topics and care settings, from acute hospital care, to physician office outpatient care, to long-term care. Several of the themes are further subdivided into many "subtheme components" that are only loosely related. The specific requirements for QIO recruitment of providers to work with also varies greatly by theme and subtheme component. For some components, QIOs were to recruit from lists of providers with relatively poor baseline performance. In other components, the QIOs were free to recruit any providers that met certain criteria, and in yet others QIOs were

to work with entire communities or directly with beneficiaries. The table below provides a highly abridged overview of the 9th SOW.

OVERVIEW OF THEMES AND SUBTHEMES IN THE QIO 9TH SOW

Theme/Component	Provider Recruitment	Goal of QIO Assistance
Beneficiary Protection		
Various review activities mandated by law and regulation	Not applicable	Address beneficiary complaints, address quality concerns, meet program requirements
Assisting with RHQDAPU	Recruit hospitals	Increased reporting to RHQDAPU
Patient Safety Theme[a]		
Hospital SCIP/HF	Recruit hospitals with baseline performance below cutoff	Improve process quality measures for surgical care and heart failure
Hospital methicillin-resistant staph aureus (MRSA) infections	Recruit hospitals reporting to specialized CDC data system	Reduce rates of MRSA infections
Nursing home pressure ulcers	Recruit nursing homes with baseline rates above cutoff	Reduce rates of pressure ulcers
NH physical restraints	Recruit nursing homes with baseline rates above cutoff	Reduce rates of physical restraints
Nursing homes in need	Recruit small handful of nursing homes with especially serious quality deficiencies	Reduce rates of pressure ulcers and physical restraints
Drug safety	Recruit wide variety of drug plans and healthcare providers	Reduce rates of drug-drug interactions and potentially inappropriate medications
Prevention Theme		
Cancer screenings/vaccinations	Recruit primary care physician practices using electronic health records that meet certain functionality requirements	Increase rates of cancer screening and vaccinations
Prevention—Disparities Theme		
Diabetes monitoring	Recruit primary care physician practices serving underserved Medicare beneficiaries with diabetes	Increase rates of recommended tests for diabetes care
Beneficiary diabetes self-management education	Recruit underserved Medicare beneficiaries to participate in a special several week long group diabetes self-care program	Improve self-care
Care Transitions Theme		
Working with intervention communities	Defined geographic community and all healthcare providers willing to participate	Reduce rates of hospital readmissions
Prevention—Chronic Kidney Disease Theme		
Urinary microalbumin testing	Recruit primary care physician practices	Increase statewide rates of recommended urine tests in diabetes
Treatment with ACE-I/ARB drugs	Recruit primary care physician practices	Increase statewide rates of prescription of recommended drugs for diabetes that lower risk of CKD
Arteriovenous fistula	Recruit kidney specialist practices	Increase statewide rates at of use of arteriovenous fistulae at initiation of hemodialysis

[a]Hospital pressure ulcers was originally also a component of the patient safety theme but was discontinued by CMS in February 2010.

The major research questions for each of these smaller theme and component evaluations, and then for the evaluation as a whole are:

1. What is the impact of the program on the quality of care for Medicare beneficiaries?

- What is the cost-effectiveness of the program? What factors mediate costs and benefits, and cost-effectiveness?
- Has the program narrowed health care disparities for underserved beneficiaries?

2. Which interventions work? Which interventions work for whom (which providers and which patients), and in what circumstances?

How might the program be improved to provide greater value?

IMPACT ANALYSES WILL RELY ON REGRESSION DISCONTINUITY AND MATCHED COMPARISON GROUP METHODS

We will draw on multiple secondary data sources as well as conduct a national survey of hospitals and nursing homes with 1,250 completed surveys each. Simple comparisons of the outcomes of providers that participated with the QIOs to those of providers who did not are likely to lead to a misleading picture of QIO impacts. QIOs may have sought out providers with greater motivation and resources for quality improvement to participate, or ones with previous success in implementing such projects. Providers willing to work with QIOs might likewise have stronger desire and better means to improve quality. Any observed improvements in quality between participating and nonparticipating providers might then all be due to these underlying differences, rather than from any effects of the QIO program.

We will apply (1) *regression discontinuity*, and (2) *matched comparison group* approaches, two statistical and econometric techniques developed to attribute program effects when simple participant/non-participant comparisons might not yield accurate results. Differences in how providers were recruited for the various subtheme components will determine which approach is appropriate. For some subtheme components we will not be able to estimate program impacts at all, because there is no valid comparison group. In these cases we will present descriptive statistics on time trends in outcomes for the relevant providers.

Regression discontinuity (RD) approaches compare the outcomes of the treatment and comparison groups when assignment to the treatment is determined by a cutoff value on a score, so that individuals with scores on one side of the cutoff receive the intervention while those on the other side do not. Matched comparison group (MCG) approaches compare the treatment group to a comparison group that has been constructed to be as similar as possible to the treatment group on all observed characteristics. In addition to estimating main impacts between intervention and comparison groups, we will also look for evidence that QIO activities may have caused greater or lesser quality improvements among Medicare beneficiaries belonging to racial and ethnic subgroups. The following table shows the *primary* analytic approach for each theme and subtheme component (each theme and subtheme component will also have additional secondary approaches, such as descriptive and qualitative analyses, but these are not shown).

ANALYTIC APPROACHES TO 9TH SOW THEMES AND SUBTHEME COMPONENTS

Theme/Subtheme	Primary Analysis
Patient Safety Theme	
Hospital SCIP/HF	RD
Hospital MRSA	Descriptive
Nursing Home Pressure Ulcer	RD
Nursing Home Physical Restraint	RD
Nursing Homes in Need	Descriptive
Drug Safety	Descriptive
Prevention Theme	
Working with PPs on cancer screening and vaccinations	Descriptive
Prevention—Disparities Theme	
Working with PPs	RD
Beneficiary DSME	Qualitative
Care Transitions Theme	
Working with intervention communities	MCG
Prevention—CKD Theme	
Urinary microalbumin testing	MCG
Treatment with ACE-I/ARB drugs	Qualitative
AV fistula	MCG

RD=Regression discontinuity
MCG=Matched comparison group
Descriptive=Descriptive statistics from survey data or descriptive time trends of quality measures
Qualitative=Findings from focus groups or semi-structured interviews of QIO staff, beneficiaries, and providers

We are still exploring the details of how the prevention disparities QIOs implemented the recruitment of providers and it is still possible that statistical power may be too low for the proposed regression discontinuity analyses of this subtheme. Between the RD and MCG approaches, RD is considered the stronger one that is more likely to yield true impact estimates. Unfortunately, it cannot be applied to all themes and components. It can also have problems with low statistical power. MCG is more prone to bias; one can never be certain that the matched comparison group truly reflects what would have happened to the intervention group in the absence of the program.

THE EVALUATION WILL CALCULATE MEASURES OF COST-EFFECTIVENESS AND COST-BENEFIT

To examine the cost-effectiveness of the program, we will search the cost-effectiveness literature for data to convert the various estimated intervention effects into (1) effects on Medicare health care expenditures (for example, reduced costs from averted hospitalizations, or increased costs from increased longevity) and (2) effects on health ("life years" [LYs] gained, or "quality-adjusted life years" [QALYs] gained). There will likely not be published studies available for all of the outcomes. We will combine the projected program effects on Medicare expenditures with the corresponding Medicare spending on the QIO program. We can then combine effects on health with effects on spending to calculate cost-effectiveness ratios—the number of Medicare dollars expended to achieve a QALY. In a related series of analyses we will also calculate cost-benefit ratios or net cost-benefit differences, in which both benefits and costs are expressed as dollars.

DETERMINING WHICH INTERVENTIONS WORK AND UNDER WHAT CIRCUMSTANCES

Our overall strategy is to (1) compute QIO-specific impact estimates, (2) classify QIOs into a "typology," and (3) correlate typologies and impacts. Using state-specific samples, we will calculate impact estimates for each QIO for each subtheme component and outcome using the original underlying methodology (that is, RD or matched comparison groups). We will not consider statistical significance in this preliminary step.

In order to classify QIOs we must first describe the QIOs' interventions. We will survey all QIOs nationwide (QIO directors and theme leaders) through a self-administered, web-based instrument. The survey will gather detailed information on QIOs' major activities to accomplish theme goals, their experience with the contract and CMS-sponsored supports for their work, their processes for recruiting providers (for applicable themes), and their input as to how the program could be improved. Through telephone discussions we will also learn about the experiences of organizations partnering with QIOs in the care transitions and CKD subnational themes, and we will conduct focus groups of beneficiaries who received diabetes self-management education from the QIOs in the prevention disparities theme.

We envision a two-step approach to classifying QIOs: (1) an initial exploration of possible quantitative or statistical approaches (for example, a principal components or classification and regression tree analysis of the QIO survey and other data—sample sizes may preclude such approaches, however), and (2) independent implicit reviews by members of the research team of all the descriptive data on QIOs. Researchers' implicit reviews may also reveal important commonalities between QIOs. We will assess the reproducibility of researchers' reviews through inter-rater reliability statistics and consolidate results from the two approaches.

To correlate QIO typologies to impacts we will first construct simple matrices consisting of the QIOs in rows, rank ordered by size of impacts, and their typologies in the columns and look for patterns of certain typologies. We will next divide the impacts into quantiles (quartiles, quintiles, and so on) and calculate the percentages of QIOs of a particular type in each quantile. Finally, we will restrict the matrices and descriptive percentages to those QIOs with statistically significant impacts and assess the feasibility of using regression models to further explore the relationship between QIO "types" and impacts. These steps will generate hypotheses for which we can then search our qualitative data for corroborating evidence.

We will also study associations between provider environments and QIO impacts. The survey of QIOs includes items designed to capture their understanding of their provider environments. To further understand the range of state-specific provider environments and how these environments affect QIOs' work, we will conduct 10 "case studies" of QIO programs and the stakeholders in their states from late 2010 through the spring of 2011. Case studies will include week-long site visits during which we will meet with QIO staff, providers (representatives of hospitals, nursing homes, and physician practices), and key spokespersons for a state's hospital, nursing home and physician communities. We will parallel the typology-impact analysis by creating matrices in which QIOs are again in the rows and rank ordered by impact size, but provider environment summary indexes or classifications based on the survey data are now in the columns. As we did with the QIO typologies analysis, we will then move on

to calculations of the percentages of provider environment types in each quantile of the impact distribution, restriction to statistically significant impacts or impacts of a minimum size, and consideration of regression models that correlate impact size with provider environment. We will then combine these analyses with our qualitative data from the case studies.

The final research question (which interventions work for whom, and in what circumstances?) asks about the conjunction of (1) QIO program type, (2) providers, and (3) provider environment. We will combine qualitative and quantitative approaches as there are simply insufficient data to attempt three-way interactions on different QIO program types across different provider environments acting on different provider types. For example, we will construct and visually inspect matrices that display QIO typologies down the rows, provider environment categories in the columns, and provider type-specific estimates in each cell. We will look for patterns of larger or smaller impacts among the cells. Obviously, the number of combinations of provider environment features, and provider characteristics that we will be able to examine is limited, and our survey and interview findings will help guide us in the factors to be assessed. The qualitative data will prove key in bolstering any hypotheses that arise from our tabular analyses.

OPTIONS TO INCREASE THE EVALUABILITY OF THE 10th SOW AND FUTURE SOWS

We have noted the limitations of the impact analyses in the current evaluation and how there will be persistent uncertainties about the accuracy and validity of many of the impact estimates. We discuss two alternative options for selecting the participating providers that QIOs work with. These options would strengthen program and contract evaluation in future SOWs. In the first option, from among a pool of providers that would benefit from the QIO program (such as those below a certain performance threshold), CMS would randomly pick providers for QIOs to work with. QIOs would attempt to recruit *all* providers in this group, and would also be evaluated on the outcomes of all providers in the group, even if some refused to participate. This would ensure the comparability of the participating and nonparticipating groups. The potential drawback of this approach is that it fails to take advantage of useful information that QIOs may possess on which providers might be most helped by their intervention, and which might be most willing to cooperate. In the second option, CMS would again create a pool of providers suitable for QIO intervention, but then randomly divide it into two pools, one of participating provider *candidates* and the other of providers not eligible for QIO services. QIOs would select a set of providers to work with from the candidate pool. QIOs' performance would be evaluated by comparing the outcomes for the *entire* pool of participating provider candidates (not just those selected as participating providers) to the entire ineligible pool. Using both pools in their entirety leads to an unbiased comparison, unlike a comparison of only the participating providers to the pool of ineligibles. This approach allows QIOs discretion in picking participating providers but has the disadvantages of increased data collection costs and diminished statistical power.

REPORTING RESULTS

Reports on the 9th SOW evaluation will require a challenging synthesis of results from the multiple studies of themes and subtheme components. As noted, each theme and component

targets different providers and care settings. The strength of evidence for each result will vary. For each theme, we will first assess the proportion of outcomes subsumed by the theme that exhibit favorable impacts, the size and statistical significance of those impacts, and the susceptibility of the estimators to bias. We will review estimate of cost-effectiveness and cost-benefit for specific subtheme components. We will then assess the extent to which the implementation of the theme followed the steps and logic models originally planned. We will then enter summaries of all of these component-specific assessments into a series of matrices in which the rows are the subtheme components and the columns are summaries of the individual assessments listed above, namely--estimated impacts on different outcomes; size, statistical significance, robustness and likely unbiasedness of these impacts; measures of cost-effectiveness and cost-benefit; faithfulness to the logic models and to implementation as planned; and mechanisms/environment/provider findings. Since impact analyses, cost-effectiveness/cost-benefit analyses, and mechanisms analyses may not be feasible for all of the components, some of the cells may remain blank. Inspection and analysis of these matrices will help us to answer each research question for each of the subtheme components. Finally, we will consider whether we can build these individual subtheme component assessments into an overall assessment.

Obviously, an overall assessment is straightforward if all component evaluations are either uniformly positive or uniformly negative. Such a scenario is highly unlikely, however. It will be tempting to boil the wealth of findings from the matrices described earlier into a single, simple message (such as the 9th SOW "worked" or "did not work"). However, such a single message risks discarding an enormous amount of information; it might mask, for example, that a few things worked extremely well, while others looked promising but evidence was weak. On the other hand, a complex list of findings qualified by numerous caveats is also not helpful to decisionmakers. Although the nature of specific tradeoffs must await the findings of our analyses, we will work with CMS to produce concise, policy relevant reports that fairly represent the complexity of results while providing clear guidance and recommendations.

We also briefly consider how the current evaluation relates to the Institute of Medicine's recommendations on the QIO program and to the recent NORC study sponsored by ASPE. The Mathematica evaluation, by its existence and scope, meets the IOM recommendation for an external evaluation. The evaluation design meets several of the specifics the IOM recommended as well, including analyses to attribute quality improvements to the QIO's intervention, "mechanisms" analyses to examine the relative effectiveness of various types of interventions, cost-effectiveness analyses, and assessment of the QIOs' role relative to other quality improvement organizations. The IOM report also made several recommendations on the management of the QIO program. The current evaluation will assess the success of CMS efforts to follow some of these recommendations, through QIO directors' perceptions of the core contract and the criteria for evaluation of contract performance, communications from CMS, the contract timeframe, and contract modifications. However, the evaluation will not assess other topics raised by the IOM report, such as the QIO selection process or the incentives contained within QIO and QIOSC contracts. The evaluation will also not address broader recommendations from the IOM and ASPE on the functioning of the QIO data system and the regulation of data sharing.

Specific upcoming reports include a summary report of QIOs' attainment of the mid-course milestones in their contracts, and a report on findings from the evaluation's surveys of hospitals,

nursing homes and QIO staff. In late September of 2010 we will submit a detailed draft outline (including chapter headings and table shells or dummies) for the interim report that is due in early February of 2011. The February 2011 interim report will contain results of quantitative descriptive and impact analyses. The final evaluation report, due in October 2011, will update the quantitative analyses of the February report with more recent data; present results of all of the qualitative components of the study, the mechanisms analysis, and the cost-effectiveness and cost-benefit analyses; and conclude with a synthesis of all analyses of the evaluation and future implications and recommendations. This schedule assumes that all of the QIO- and CMS-furnished data necessary for the evaluation are accurate and available in time for report analysis and preparation.

Report	**Due Date**
Status of QIOs' achievement of their milestones	10 weeks after receipt of access to 18-month scores determined by CMS
Survey report on partner's experience of service by the QIOs and report on the survey of QIOs	24 weeks after OMB Clearance (Anticipated due date of December 21, 2010)
Preliminary draft outline (including chapter headings) and set of dummy tables	September 27, 2010
Final outline and set of dummy tables following receipt of CMS comments	October 25, 2010
Draft interim impact report containing quantitative descriptive and impact analyses	February 1, 2011
Final interim impact report	February 15, 2011
Draft final report with updated quantitative results using more recent data, qualitative findings, mechanisms analysis, cost-effectiveness/cost-benefit analyses, synthesis, and conclusions.	September 19, 2011
Final report	October 3, 2011

I. INTRODUCTION

The Quality Improvement Organization Program of the Centers for Medicare & Medicaid Services (CMS) is a key component of CMS's agenda for ensuring and improving quality of care for Medicare beneficiaries. As required by Sections 1152 through 1154 of the Social Security Act, CMS contracts with a nationwide network of independent quality improvement organizations (QIOs) to help health care providers deliver high quality care to Medicare beneficiaries.[1] The contracts last for three years, with each contract cycle called a scope of work, or SOW. The 9th SOW began on August 1, 2008, and will end August 31, 2011. With budgets of roughly $1.1 to $1.2 billion dollars for the current and preceding SOWs, the QIO program is the single largest investment in quality improvement infrastructure—public or private—in the nation.

CMS has contracted with Mathematica Policy Research to independently design and conduct an evaluation of the 9th SOW. This report contains our design and approach to the evaluation.

A. BACKGROUND AND POLICY CONTEXT

The importance of the QIO Program's functions and the magnitude of its budget make evaluation of its effectiveness essential. Understanding the program's overall effectiveness and identifying its most successful components or activities are prerequisites to improving the program as a whole. Moreover, given the influence of the Medicare program on the American health care system, the QIO Program can lead to better care not only for Medicare beneficiaries but for all Americans.

In the Medicare Prescription Drug, Improvement, and Modernization Act of 2003 (P.L. 108-173), Congress mandated the Institute of Medicine (IOM) to conduct an overview of the QIO Program, including a review of "the extent to which quality improvement organizations improve the quality of care for Medicare beneficiaries" (Institute of Medicine 2006). Following an extensive review of scientific literature published between 1995 and 2005, the IOM concluded that "although the quality of care received by Medicare beneficiaries has improved somewhat, researchers have been unable to attribute these changes to the QIO program." The IOM could not determine whether this lack of evidence for QIO impacts was due to the methodological limitations of many of the studies reviewed, and to the difficulty of disentangling the effects of QIO activities from the many other concurrent quality improvement efforts, or to a true lack of program effectiveness (IOM 2006). The IOM report also recommended that CMS periodically commission independent, external evaluations of the program. In his 2006 Report to Congress responding to the IOM's recommendations, the Secretary of Health and Human Services agreed on the need for strengthened methods of program evaluation (Leavitt 2006), and CMS

[1] The current report focuses on the impacts of the QIO Program on quality improvement. Other missions of the QIO Program include protecting beneficiaries' rights by reviewing and investigating complaints and appeals, and protecting the Medicare Trust Funds by ensuring that Medicare pays only for services and goods that are reasonable, necessary, and provided in the most appropriate setting.

commissioned the current evaluation. Chapter V of this report discusses how the current evaluation relates to IOM's recommendations and to recommendations from other studies.

At around the same time that IOM was preparing its report, the Assistant Secretary for Planning and Evaluation (ASPE) was studying options for evaluating the effectiveness of the QIO Program. ASPE contracted with the National Opinion Research Center (NORC) to develop a richer inventory and description than previously available of QIOs' activities and strategies, and to assess alternative designs for potential future evaluations of the QIO Program. NORC's literature review for this project on the impacts of the QIO program reached the same conclusions as IOM's, namely, that the literature is ambiguous on the effectiveness of the program and that previous studies have suffered from a variety of methodological problems. NORC's report concluded with several design options and recommendations for further research on the QIO Program (Sutton et al. 2007).

One of the major challenges facing previous studies of the QIO program has been the stringent statutory and regulatory restrictions on QIOs' releasing information on the identities of providers who work with the QIOs (Social Security Act, 42 CFR Part 480). In its QIO Manual, CMS has distilled these restrictions into the following instructions to QIOs—"you cannot disclose information that explicitly identifies institutions [or] practitioners [with whom you are working]…without their consent" (CMS 2009); even CMS is not permitted access to this information.[2]

Historically, these restrictions date from a time when the primary job of QIOs (that is, their predecessor organizations, the Professional Standards Review Organizations and the Peer Review Organizations) was to conduct punitive provider reviews of utilization and practice patterns. However, it is extremely difficult to evaluate the program's effectiveness when the identities of health care providers who participate in the program must remain a secret. Both the NORC report (Sutton 2007) and a General Accountability Office (GAO) report on QIOs' efforts to improve nursing home quality (GAO 2007) specifically highlighted the problems to program evaluation that the QIO confidentiality restrictions pose. The NORC report pointed out that publicly reported, detailed data on individual providers' quality performance are increasingly common (such as on CMS's Hospital Compare website), and the GAO report called for CMS to revise the confidentiality regulations to facilitate better evaluation of the QIO program.

For the current evaluation, Mathematica has executed subcontracts with each of the QIOs; federal regulations do require a QIO to disclose to a subcontractor information that is necessary for the subcontractor to provide specified services to the QIO (42 CFR Part 480 Section 135). In the 9th SOW contract, CMS has specified that each QIO must seek Mathematica's assistance in demonstrating that improvements in outcome measures are attributable to its (the QIO's) interventions. In order to do so, the QIO must subcontract with Mathematica. Mathematica will provide additional assistance to QIOs by determining and making recommendations on which QIO interventions appeared to be most effective.

[2] The major exceptions are that QIOs must disclose information containing provider identifiers to licensing, accreditation, or certification agencies as necessary for them to carry out their functions as outlined under state law, and to the Office of the Inspector General and the General Accountability Office as necessary for them to fulfill their statutory responsibilities; these disclosures must occur "onsite" at the QIO (42 CFR 480.140).

B. BRIEF DESCRIPTION OF THE 9TH SCOPE OF WORK

We provide here only a short summary of the work and activities required of the QIOs in their 9th SOW contracts. Chapter II provides a more detailed overview of the goals and objectives of the QIO Program and the 9th SOW, the context in which the program operates, and the mechanisms or pathways through which desired outcomes are to be achieved.

There are 43 QIO contractors carrying out the 9th SOW under 53 contracts (one for each of the 50 states, plus the District of Columbia, Puerto Rico, and the U.S. Virgin Islands).[3] The 9th SOW contracts were first awarded, and the 9th SOW officially began, in August 2008. However, an extensive modification of the 9th SOW was executed in July 2009.

The 9th SOW is organized into six themes—three covering the entire nation (national themes), and three involving selected states (sub-national themes)—spanning a wide range of topics and care settings. There are also six "QIO Support Contractors" (QIOSCs) that are providing specialized theme-specific support to QIOs and to CMS; all but one of these QIOSCs are also QIOs.

QIOs were to recruit sets of providers with whom to work for each theme or theme component. For each QIO, the recruitment targets for the different provider types and themes were negotiated with CMS and specified in the QIO's contract.

1. National Themes

The three national themes are: (1) beneficiary protection, (2) patient safety, and (3) prevention.

a. Beneficiary Protection

Under this theme, QIOs conduct certain activities required of the QIO program by statute and regulation. These include utilization reviews, quality-of-care reviews, reviews of beneficiary appeals of provider notices, and reviews of potential anti-dumping cases. As appropriate, QIOs also mediate disputes between beneficiaries and providers, apply provider sanctions, and cooperate with state agencies that inspect and certify providers and with other CMS contractors that monitor the appropriateness of Medicare payments. The MITRE Corporation recently assessed much of this beneficiary protection work under a separate contract with CMS, and Mathematica will not evaluate these previously studied activities.[4]

[3] Throughout the remainder of this report we will use the term "state" broadly to include the 50 states and the three non-state jurisdictions of the District of Columbia, Puerto Rico, and the U.S. Virgin Islands. Furthermore, we will use "QIO" interchangeably with "state," although six QIOs hold contracts for two jurisdictions and two QIOs for three jurisdictions.

[4] One of the subtasks of the beneficiary protection theme was apparently not studied by MITRE. In this subtask, QIOs are to encourage hospitals to submit quality data for the Reporting Hospital Quality of Annual Payment

b. Patient Safety

The patient safety theme encompasses seven components:[5]

1. Improving hospital care (rates of recommended processes of care) for surgical safety and heart failure (known as the Surgical Care Improvement/Heart Failure [SCIP/HF] component)[6]
2. Reducing hospital rates of health-care-associated methicillin-resistant Staphylococcus aureus (MRSA) infections
3. Reducing rates of pressure ulcers in nursing homes
4. Reducing rates of physical restraint use in nursing homes
5. Assisting a very small set of nursing homes (nursing homes in need, or NHIN) with severe quality deficits (roughly one facility per year for each QIO)
6. Improving drug safety (rates of drug-drug interactions [DDIs] and potentially inappropriate medications [PIMs] for the elderly) in a wide variety of settings
7. Improving rural providers' rates of pressure ulcers (in rural hospitals and nursing homes) and physical restraints (in rural nursing homes).

The 9th SOW refers to the subtasks above as "components" of the patient safety theme. For another theme, however, the 9th SOW uses the term "clinical focus areas" to describe subtasks within the theme. To avoid using multiple terms (such as components, clinical focus areas, subtasks, and so on) all meaning the same thing, in this report we will use the terms *subtheme components* or *components* to refer to subtasks within a theme.

Hospital SCIP/HF, nursing home pressure ulcers, and nursing home physical restraints components (the first, third, and fourth components above). Recruitment of providers (scheduled to be completed by September 30, 2008) for these three components was highly structured. CMS rank ordered providers in each state on baseline values of the relevant quality indicators, established minimum threshold scores for each indicator, and created lists of all providers falling short of these thresholds. Each QIO had provider recruitment targets for these three components negotiated with CMS. The QIOs were to recruit at least 85 percent of their targets from the lists; the remaining 15 or fewer percent could be providers not on the lists

(continued)
Update (RHQDAPU), and to provide to interested hospitals technical assistance and training in the use of the CMS Abstraction and Reporting Tool (CART) and its associated electronic QualityNet Exchange reporting system. As described later, we will assess hospitals' perceptions of this technical assistance.

[5] An eighth component specified in the original 9th SOW was reducing rates of pressure ulcers in hospitals. However CMS discontinued this component in early February 2010.

[6] SCIP is the acronym for the Surgical Care Improvement Project. HF is short for heart failure.

(providers above the thresholds). We will refer to hospitals and nursing homes that agree to work with the QIOs as participating providers (PPs).[7] Table I.1 shows the numbers of providers recruited for each of these four subtheme component as of September 2009.

QIOs' interventions for these components consist primarily of training and education of the staff of participating hospitals and nursing homes. The QIOs also collect quality indicator data from the PPs and provide quarterly feedback on provider performance. CMS first trained two or three QIO staff members (called national QI leaders) in effective "action generating" meeting techniques; these national QI leaders then returned to their home QIOs to train additional QIO staff. QIO staff are sponsoring trainings and meetings for both individual and multiple PPs in approaches to improving quality in these components. Finally, the QIOs are coordinating quality improvement communities (QI communities) consisting of providers, private and public organizations, state agencies, patients, and other quality and patient safety stakeholders to advance patient safety statewide and foster a culture of safety in health care facilities.

MRSA Component. For this component, QIOs were to recruit hospitals participating in the Centers for Disease Control and Prevention's (CDC's) National Health Safety Network (NHSN) program (specifically, an aspect of the NHSN called the Multidrug Resistant Organism, or MDRO, module).[8] Since hospital participation in the NHSN is confidential, QIOs had to publicize to all hospitals statewide the opportunity to work with the QIO on the MRSA component in order to have NHSN participating hospitals self-identify to QIOs (with the exception states that mandate hospital reporting through the NHSN). A hospital entity participating in the 9th SOW MRSA component need not be the entire facility, but can be an individual unit or location within the hospital (for example, a medical critical care or cardiac surgery intensive care unit), although participation is limited to one unit per hospital. Because of the confidential nature of NHSN data, participating hospitals and the QIOs had to execute signed agreements allowing the QIOs and the patient safety QIOSC access to view and analyze the hospitals' NHSN-MDRO data. CMS' RFP for QIOs contained information on the numbers of hospitals in each state reporting to the NHSN, and these numbers are reproduced in Table I.2.

QIOs were to assist hospitals in the MRSA component by training the hospitals' staff in a special program called TeamSTEPPS, sponsored by the Agency for Healthcare Research and Quality (AHRQ) and the Department of Defense (DoD). TeamSTEPPS aims to improve patient safety within health care facilities by teaching health care professionals special communication and teamwork skills (AHRQ 2009). The QIOs were to send two staff members to undergo

[7] Although recruited providers for the patient safety theme are technically referred to within the QIOs' own internal data system as identified participants (IPs), we will call them PPs to be consistent with the terminology of several of the later themes and to adopt a single term across all of the themes.

[8] The NHSN is a nationwide, confidential, web-based standardized reporting system sponsored by the CDC. It allows national estimation and monitoring of health-care-associated adverse events (including health-care-associated infections, HAI), and provides feedback to participating health care facilities for quality improvement and benchmarking purposes. Most health care facilities participate voluntarily, although some states have mandated that all hospitals statewide perform public reporting of HAIs and have required state hospitals to report through the NHSN. In addition, NHSN participating facilities may choose to report to one or more "modules," which focus on different adverse events (such as device-associated infections, procedure-associated infections, and so on). As noted, the 9th SOW focused on hospitals participating in the MDRO module.

TABLE I.1

NUMBERS OF PROVIDERS RECRUITED FOR SELECTED PATIENT SAFETY THEME COMPONENTS AS OF SEPTEMBER 2009

State	Nursing Home-Pressure Ulcers	Nursing Home-Physical Restraints	Hospitals-SCIP/HF
AK	3	2	1
AL	19	6	19
AR	23	83	13
AZ	8	11	15
CA	36	41	22
CO	6	24	7
CT	13	7	3
DC	4	1	2
DE	2	1	2
FL	33	45	16
GA	41	37	28
HI	1	2	3
IA	12	8	6
ID	2	6	5
IL	27	24	22
IN	19	27	18
KS	29	9	12
KY	8	10	20
LA	33	45	22
MA	19	24	4
MD	26	18	8
ME	8	6	1
MI	30	42	10
MN	8	11	9
MO	25	29	7
MS	15	21	8
MT	10	1	4
NC	27	76	15
ND	5	2	2
NE	1	4	3
NH	7	9	N/A
NJ	58	18	4
NM	8	21	9
NV	6	12	11
NY	70	17	28
OH	45	48	20
OK	72	104	12
OR	13	14	11
PA	35	22	29
PR	3	N/A	13
RI	11	4	1
SC	13	30	10
SD	11	4	3
TN	9	27	22
TX	28	90	80
UT	4	24	8
VA	43	8	16
VI	1	N/A	2
VT	14	N/A	1
WA	13	6	8

TABLE I.1 (*continued*)

State	Nursing Home-Pressure Ulcers	Nursing Home-Physical Restraints	Hospitals-SCIP/HF
WI	23	10	5
WV	16	8	4
WY	3	1	3
Total	**999**	**1,100**	**607**

Source: SDPS/QIONet Program Progress Reports report generated on September 9, 2009.

Note: QIOs were to focus on recruiting from lists of nursing homes and hospitals whose performance at the start of the 9th SOW on specific measures did not meet certain cutoffs. Each QIO had a target number of providers to recruit and was to recruit 85 percent of its participating providers from the lists; the remaining 15 percent could be providers not on the lists. The cutoffs were as follows:

- Nursing home pressure ulcers—facilities whose rates of pressure ulcers among high-risk long-stay residents during 2 out of the 3 quarters from 2006 Q4 through 2007 Q2 were 20 percent or higher (that is, exceeded by 14 or more percentage points the goal of no more than 6 percent).

- Nursing home physical restraints—facilities whose rates of physical restraints among long-stay residents during 2 out of the 3 quarters from 2006 Q4 through 2007 Q2 were 11 percent or higher (that is, exceeded by 8 or more percentage points the goal of no more than 3 percent).

- Hospital SCIP/HF—hospitals whose appropriate care measure (ACM) score for the SCIP-Infection 1 and SCIP-Infection 3 measures in 2006 Q4 was 62.5 percent or lower, *and* whose ACM score in 2007 Q1 was 64 percent or lower (that is, both ACM scores fell short by 30 or more percentage points of the achievable benchmarks of care [ABC] rates for these two quarters of 92.5 and 94 percent, respectively).

As noted in the text, the QIOs originally also recruited hospitals to work on reducing pressure ulcers in hospitalized patients, but CMS discontinued this component in early February 2010.

TABLE I.2

NUMBERS OF HOSPITALS RECRUITED FOR THE METHICILLIN RESISTANT STAPH AUREUS (MRSA) PATIENT SAFETY COMPONENT AS OF SEPTEMBER, 2009

State			
AK	1	NC	15
AL	4	ND	2
AR	2	NE	2
AZ	5	NH	4
CA	10	NJ	11
CO	22	NM	2
CT	5	NV	2
DC	1	NY	60
DE	5	OH	10
FL	8	OK	31
GA	8	OR	5
HI	2	PA	31
IA	3	PR	2
ID	2	RI	2
IL	8	SC	40
IN	4	SD	2
KS	2	TN	29
KY	7	TX	3
LA	5	UT	2
MA	5	VA	10
MD	10	VI	1
ME	4	VT	4
MI	22	WA	10
MN	2	WI	12
MO	6	WV	4
MS	6	WY	2
MT	2		
		Total	**459**

Source: SDPS/QIONet Program Progress Report report generated on September 9, 2009.

Note: For the MRSA component of the patient theme of the QIO 9th SOW, QIOs were to recruit hospitals reporting on the multidrug resistant organism (MDRO) module of the Centers' for Disease Control and Prevention's National Healthcare Safety Network program.

TeamSTEPPS Master Training (free training was offered by AHRQ and DoD until August 2009); these master trainers would then train other QIO staff and PP hospital staff.[9] In coordinating the previously mentioned QI communities, QIOs should also include MRSA reduction efforts.

Drug Safety. The drug safety component intervention allows QIOs considerable flexibility in selecting providers to work with and interventions to pursue. QIOs were to seek partnerships with Medicare providers and practitioners, Medicare Advantage (Medicare Part C) plans, and Part D prescription drug plans (PDPs) in order to decrease rates of DDIs and PIMs as measured in Part D claims data. The nature of these partnerships was not specified. QIOs could offer staff time, data, lists of public websites and resources, and general quality improvement expertise and tools.

Nursing Homes in Need (NHIN). In the final component of the patient safety theme, the QIOs were to work intensively with a small, highly selected group of nursing homes in particular need of quality improvement to reduce rates of (1) pressure ulcers and (2) use of physical restraints. The 9th SOW anticipated each QIO would work with roughly one NHIN every 12 months for a total of three NHINs over the three-year SOW contract. QIOs were to select NHINs from CMS's list of special focus facility (SFF) nursing homes. CMS designates as SFFs nursing homes with a longstanding history (at least three years) of many serious quality issues. These facilities are then surveyed by the state survey agencies twice as frequently as other nursing homes; those failing to correct deficiencies and exhibit improvement are subject to monetary fines and, ultimately, to termination from the Medicare and Medicaid programs. QIOs were to start recruitment among facilities designated as SFFs for at least six months, but the 9th SOW also provided a series of contingency steps—in case the QIO's initial choice refused to participate, in case no facilities designated as SFFs for at least six months agreed to participate, and so on.

The QIOs were to conduct site visits and prepare root cause assessments (RCA) of the nursing homes' quality problems. The QIOs were to then develop action plans for the facilities to reduce the two targeted quality indicators (pressure ulcers and physical restraints). The RCAs and action plans may address a wide range of issues, include nursing homes' management, financial status, staffing, staff communication, care processes, and so on.

Rural-Focused Patient Safety Project. This project is a new component in the 9th SOW modification executed in mid-July 2009. A number of selected QIOs awarded this project are to assist rural nursing homes in improving rates of pressure ulcers and physical restraints. We continue to work with CMS on receiving the full documentation for this project.

[9] This "train the trainer" model for the MRSA component was similar to that for the hospital SCIP/HF, nursing home pressure ulcer, and nursing home physical restraints components. However, for the hospital SCIP/HF, nursing home pressure ulcer, and nursing home physical restraints components, the training was in "action generating effective meeting management techniques," with training provided by CMS. For the MRSA component, the training was in the TeamSTEPPS program, with training provided by AHRQ and DoD.

c. Prevention

This theme aims to improve rates of mammography and colorectal cancer screening, and of pneumococcal and influenza vaccination among primary care practices. The QIOs were to recruit primary care physician practices (called participating practices or PPs) that had implemented an electronic health record (EHR) certified by the Certification Commission for Healthcare Information Technology (CCHIT). Furthermore, the EHRs had to have certain care management capabilities (such as the ability to create problem or diagnosis lists or to identify patients fitting specific age or clinical characteristics), and these capabilities had to have been implemented for at least one of a set of conditions (such as hypertension or diabetes). PPs had to sign a consent form agreeing to implement the EHR care management capabilities for the cancer screenings and vaccinations for most of their patients, and to report their EHR data on these preventive care measures. The QIOs were also to identify a set of nonparticipating practices (NPs) that met all criteria for PPs, but did not agree to the activities required of the PPs. The number of NPs had to be between 50 and 125 percent of the PP target. Table I.3 shows the number of PPs recruited as of February 2009.

As with many of the other themes, the QIOs are helping the PP practices through education and technical assistance. Possible QIO activities include completion of on-site assessments, consultation on redesign of practice workflows, provision of educational tools and resources, and training in teambuilding and quality improvement techniques.

2. Subnational Themes

The three subnational themes are (1) prevention—disparities, (2) care transitions, and (3) prevention—chronic kidney disease (CKD).

a. Prevention—Disparities

The goal of this theme is to improve diabetes care among underserved Medicare beneficiaries. CMS directed six states to undertake this theme—the District of Columbia (DC), Georgia (GA), Louisiana (LA), Maryland (MD), New York (NY), and the Virgin Islands (VI).[10] QIOs' activities' were to both recruit and then assist PPs, and to provide diabetes self-management education (DSME) to beneficiaries.[11]

PPs had to meet the following criteria: (1) at least 25 percent of their Medicare patients with diabetes belonged to underserved groups, and (2) the average of their performance on "diabetes measures" had to be below the "median performance" for the state. The SOW did not specify whether it was the median performance of all practices statewide or only of practices meeting the

[10] Underserved included persons of the following racial and ethnic minorities: African American and Hispanic/Latino, Asian/Pacific Islander, or American Indian/Alaska Native. In practice, most of the beneficiaries served in this theme are African American and Hispanic/Latino.

[11] As noted, QIOs have traditionally focused on working with providers; the direct provision of DSME to beneficiaries is a new role for QIOs.

TABLE I.3

NUMBERS OF PARTICIPATING PRACTICES RECRUITED FOR THE QIO 9TH SOW PREVENTION THEME AS OF FEBRUARY 2009

State	Number of Practices	State	Number of Practices
AL	22	MS	11
AK	4	MT	10
AZ	25	NE	10
AR	37	NV	16
AZ	25	NH	16
CA	24	NJ	40
CO	21	NM	17
CT	21	NY	115
DC	9	NC	0
DE	12	ND	10
FL	67	OH	62
GA	30	OK	40
HI	11	OR	27
IA	13	PA	60
ID	16	PR	7
IL	54	RI	15
IN	30	SC	25
KS	10	SD	8
KY	19	TN	49
LA	18	TX	102
ME	14	UT	28
MD	21	VI	6
MA	92	VT	9
MD	21	VA	25
ME	14	WA	30
MI	41	WV	14
MN	10	WI	14
MS	11	WY	8
MO	37		
		Total	**1,432**

Source: SDPS/QIONet Program Progress Reports report generated February 11, 2009.

first criterion, and also did not specify the diabetes measures to include in the average performance.

Each QIO had to recruit a target number of PPs that all together served a specified range of minority Medicare beneficiaries with diabetes. The specified range varied by state, depending on the state's population of minority Medicare beneficiaries with diabetes, but was set so that the combined number of underserved Medicare beneficiaries belonging to all PPs would not exceed around 2,500 to 3,000 beneficiaries, although there could be fewer.[12] Table I.4 shows the number of PPs recruited for this theme, by state, as of February 2009 (we learned that the total number as of August 31, 2009 was 551).

The QIOs are to help all PPs increase rates of hemoglobin A1c testing, diabetic eye examination, and lipid testing, and to help those PPs reporting to the Physician Quality Reporting Initiative (PQRI) to improve rates of blood pressure control. The QIOs are to submit weekly reports to CMS on which PPs are reporting to PQRI (but are not expected to encourage PPs' participation in PQRI). In addition, CMS has retained a disparities data contractor, Masspro, to collect *clinical* data (laboratory results for hemoglobin A1c and lipids, blood pressure and weight readings, presence of diabetic retinopathy, and documentation of communication between the ophthalmologist and the primary care physician) through abstraction of PPs' medical charts. The SOW only asks QIOs to cooperate with this contractor; it does not require QIOs to help PPs with the new reporting process or to assist PPs in improving the clinical measures.

TABLE I.4

NUMBER OF PARTICIPATING PRACTICES RECRUITED FOR QIO 9TH SOW PREVENTION DISPARITIES THEME AS OF FEBRUARY 2009

State	Number of Recruited Practices
DC	113
GA	166
LA	35
MD	128
NY	82
VI	5
Total	**544**

Source: King, Terris. "Health Disparities Program." Presentation at American Health Quality Association annual conference, February 2009, Tampa, FL. [http://www.ahqa.org/pub/uploads/KingAHQAFeb2009Disparities.ppt] accessed August 30, 2009.

Note: The total recruitment as of August 31, 2009 was 551 practices.

[12] QIOs had to recruit enough PPs so that the number of underserved Medicare beneficiaries with diabetes belonging to the PPs equaled a variable percentage of all underserved Medicare beneficiaries with diabetes in the state. The percentages varied inversely with the population of underserved Medicare beneficiaries with diabetes in the state. States with relatively small numbers of underserved Medicare beneficiaries with diabetes (less than 15,000 beneficiaries) had to recruit enough PPs that together served at least 15 percent of underserved Medicare beneficiaries with diabetes in the state (thus between 0 and 2,250 beneficiaries). In contrast, states with a relatively large number of underserved Medicare beneficiaries with diabetes (between 25,000 and 59,999 for example) had to recruit enough PPs that together served at least 5 percent of underserved Medicare beneficiaries with diabetes in the state (thus, between 1,250 and 3,000).

The QIOs are also to recruit minority Medicare beneficiaries with diabetes to receive DSME. The QIOs can provide one of two CMS-approved DSME programs—either Project Dulce, developed by the Scripps Institute, or the Diabetes Education Empowerment Program (DEEP), developed by the University of Illinois at Chicago. No Medicare claims will be submitted for these DSME services, since CMS is already funding them through the QIO program.

The majority of the Medicare beneficiaries undergoing the DSME are not patients of the PPs. Although PPs were encouraged to refer their underserved Medicare patients with diabetes to the QIOs' DSME programs, the referral rates among busy PPs was quite low. QIOs thus began recruiting Medicare beneficiaries from non-PP sources, such as community organizations or local agencies; these beneficiaries did not necessarily belong to PPs. However, in some cases, beneficiaries' participation in the DSME program apparently made the physicians of these beneficiaries aware of the opportunity to work with the QIO on the prevention disparities theme; some of these physicians who met the eligibility criteria for the theme became PPs.

b. Prevention—Chronic Kidney Disease

This theme's broad objective is to improve selected aspects of prevention and treatment for chronic kidney disease (CKD). This theme was awarded competitively to 10 states on the basis of their proposals: Florida (FL), GA, Missouri (MO), Montana (MT), Nevada (NV), NY, Rhode Island (RI), Tennessee (TN), Utah (UT), and VI. Although the CKD theme is formally described as consisting two tasks in the QIOs' SOW: (1) clinical quality improvement, and (2) community collaboration, the community collaboration activities are not really a separate task but underlie and reinforce the clinical quality improvement activities. The community collaboration activities consist of QIOs assembling and/or sustaining state and local coalitions to work towards systematic quality improvement for CKD prevention and care across the state. The QIOs are to build new partnerships and strengthen existing ones with a wide range of organizations, foster increased involvement by coalition members, and leverage members' resources.

The clinical quality improvement work in turn consists of three subtasks or "clinical focus areas," in which the QIOs are to encourage physicians to: (1) perform annual urinary microalbumin testing for beneficiaries with diabetes; (2) treat beneficiaries with diabetes, early CKD (stages 1-4), and hypertension with angiotensin converting enzyme inhibitor (ACE-I) or angiotensin II receptor blocker (ARB) drugs; and (3) refer beneficiaries (with or without diabetes) nearing hemodialysis for arteriovenous (AV) fistula placement. For the first two clinical focus areas (urinary microalbumin testing and ACE-I/ARB treatment of early CKD and hypertension), QIOs are to work with primary care physicians and other physicians (such as endocrinologists) who care for beneficiaries with diabetes. For the third focus area (increased use of AV fistulas), QIOs are to target primary care physicians, nephrologists, and general and vascular surgeons for recruitment. Again, the QIOs' interventions for these clinical focus areas consist of education, consultation, and technical assistance. Through their work with both individual providers and practices, and with the state and regional coalitions, QIOs are expected to effect changes in outcome measures (urinary microalbumin testing, ACE-I/ARB prescription, and AV fistula use) for *all* Medicare beneficiaries in the state who are eligible for the measures.

c. Care Transitions

The last theme focuses on reducing hospital readmissions among beneficiaries discharged from an acute hospital stay. The care transitions theme was also awarded competitively to 14 states—Alabama (AL), Colorado (CO), FL, GA, Indiana (IN), LA, Michigan (MI), Nebraska (NE), New Jersey (NJ), NY, Pennsylvania (PA), RI, Texas (TX), and Washington (WA). Each QIO selected a "geographic area" or "community" with which to work; the SOW anticipated that most QIOs would define their target community by a list of zip codes, although QIOs could also include geopolitical boundaries, hospital service areas (HSAs), or hospital referral regions (HRRs).[13] The SOW also provided extensive guidelines on baseline area characteristics to consider and on power calculations to ensure that the selected communities would be able to detect certain minimum effect sizes on rehospitalization rates. The intervention communities are:

- Alabama: Tuscaloosa
- Colorado: Northwest Denver
- Florida: Miami
- Georgia: Metro Atlanta East
- Indiana: Evansville
- Louisiana: Baton Rouge
- Michigan: Greater Lansing Area
- Nebraska: Omaha
- New Jersey: Southwestern New Jersey (Burlington, Camden and Gloucester counties)
- New York: the Upper Capitol Region (Warren, Washington, Rensselaer, Schenectady and Saratoga counties)
- Pennsylvania: southwest Pittsburgh
- Rhode Island: Providence
- Texas: lower Rio Grande Valley (Brownsville, Harlingen, and Weslaco)
- Washington: Whatcom County

The 14 intervention communities are served by about 70 hospitals.

To help provide context and a rough benchmark for any changes in hospital admissions among the intervention communities, the Care Transitions QIOSC, Colorado Foundation for

[13] HSAs and HRRs were defined by the Dartmouth Atlas Project. An HSA is a local hospital care market. An HRR is a regional tertiary care health care market.

Medical Care (CFMC), identified three to four comparison communities around the country for each intervention community. These comparison communities were not selected through any rigorous or formal quantitative matching procedure.

The QIOs submitted a strategic plan of their interventions for this theme at the end of September 2008. The QIOs are to consider enlisting a wide array of stakeholders (such as state and local agencies, health care purchasers and payers, advocacy organizations, hospitals, nursing homes, physician practices, home health agencies, and so on). The QIOs are then to select a variety of interventions from among a broad list in the 9th SOW of care transition interventions with some evidence of effectiveness. This list includes hospital discharge, post-discharge follow-up, and enhanced inter-provider communication interventions aimed at both patients and clinicians. The QIOs are to lead the community collaboratives in the implementation of these interventions.

Although the QIOs were originally to encourage collaborating health care providers to use a Web based tool called the Continuity Assessment Record and Evaluation (CARE) instrument, in June of 2009 CMS announced that it could no longer support the Web-based CARE instrument. CMS instead encouraged the care transitions communities to consider using a paper-based version. The CARE instrument is a standardized patient assessment tool with which clinicians in different care settings can share patients' recent medical history, and health and functional status over the Internet. It was originally developed by RTI under contract to CMS for use in the ongoing Medicare Post Acute Care Reform Demonstration as a means of uniformly recording Medicare beneficiaries' clinical status and needs in different acute and post-acute care settings in order to assess Medicare's various acute and post-acute payment systems. CMS, RTI, and the care transition QIOs used provider feedback to develop a new Handover Management section for the CARE instrument especially for the care transitions theme (CMS 2009; CIMRO of Nebraska 2009). (CMS 2009).

3. Summary of 9th SOW Themes

The 9th SOW is clearly a complex program. It consists of five broad themes, but each theme includes multiple distinct subtheme components. Table I.5 summarizes these subtheme components by the different providers, recruitment procedures, interventions, and outcome measures involved.

C. OVERVIEW OF RESEARCH QUESTIONS AND GOALS OF THE EVALUATION

The evaluation of the 9th SOW encompasses three general research questions:

1. What is the impact of the program on the quality of care for Medicare beneficiaries (either nationally or subnationally)?
   - How do program costs and benefits compare, and what is the cost-effectiveness of the program? What factors mediate costs and benefits, and cost-effectiveness?

TABLE I.5

SUMMARY OF 9TH SOW THEMES AND SUBTHEME COMPONENTS

Theme/Component	Targeted Participants[b]	Method of Recruitment[b]	Other Groups[c]	QIO Interventions	Targeted Outcomes/Goals
Beneficiary Protection					
Multiple utilization, quality of care, beneficiary appeal reviews[a]	No targeting or recruitment involved	No targeting or recruitment involved	--	Case reviews of quality of care, utilization, and potential anti-dumping cases; handling of appeals; quality improvement activities; alternative dispute resolution; sanction activities; other related activities	Beneficiary satisfaction, timeliness of case reviews
Assisting hospitals with RHQDAPU	Hospitals	RHQDAPU volunteer hospitals	--	Technical assistance	Increased reporting to RHQDAPU
Patient Safety Theme[d]					
Hospital SCIP/HF	Hospitals	SCIP/HF state pool/cutoff[e]	--	National QI leaders "train the trainers" model[f] Provider education QI communities[g]	SCIP HF
Hospital methicillin-resistant staph aureus (MRSA) infections	Hospitals	MRSA volunteer hospitals[h]	--	TeamSTEPPS "train the trainers" model Provider education QI communities	Hospital MRSA
NH PrU[i]	Nursing Homes	NH PrU state pool/cutoff[e]	--	National QI leaders "train the trainers" model Provider education QI communities	NH PrU
NH physical restraint (PhyR)	Nursing Homes	NH PhyR state pool/cutoff[e]	--	Training (national QI leaders) Provider education QI communities	NH PhR
Nursing Homes in Need	Nursing Homes	CMS's special focus facility list[j]	--	Intensive assistance Root cause analyses Action plans	NH PrU NH physical restraints
Drug Safety	• Medicare providers and practitioners • Medicare Advantage (Medicare Part C) plans • Part D prescription drug plans	Drug safety volunteer entities[k]	--	Wide range of possible assistance--staff time, data, lists of public websites and resources, QIOs' general quality improvement expertise and tools	Drug-drug interactions Potentially inappropriate medications

TABLE I.5 (*continued*)

Theme/Component	Targeted Participants[b]	Method of Recruitment[b]	Other Groups[c]	QIO Interventions	Targeted Outcomes/Goals
Prevention Theme					
Cancer screenings/vaccinations	PCP practices with EHRs	Prevention volunteer practices[l]	Prevention NPs[m]	Provision to practices of: • Education • Consultation • Technical assistance	Mammography Colorectal cancer screening Influenza vaccinations Pneumococcal vaccinations
Prevention—Disparities Theme					
Diabetes monitoring	PCP practices serving underserved	Disparities pool/cutoff[n]	--	Provision to practices of: • Education • Consultation • Technical assistance	Hemoglobin A1c testing Diabetic eye examination Lipid testing (among PQRI practices) Improve rates of blood pressure control
Beneficiary DSME	Underserved beneficiaries	Volunteer beneficiaries[o]	--	DSME: • Project Dulce • Diabetes Education Empowerment Program (DEEP)	Number of beneficiaries trained
Care Transitions Theme					
Working with intervention communities	Communities	QIOs defined their intervention communities (lists of zip codes and /or geopolitical units, hospital service areas, or hospital referral regions)	QIOSC-selected comparison communities[p]	Build community coalitions to implement one or more care transitions interventions involving: • "Coaching" beneficiaries at hospital discharge • Post-discharge follow-up and education of beneficiaries • Increasing communication between hospital and post-acute providers	Hospital readmissions
Prevention—CKD Theme[q]					
Urinary microalbumin testing	PCP practices	Urinary microalbumin volunteer practices	--	Provision to practices of: • Education • Consultation • Technical assistance	Urinary microalbumin testing
Treatment with ACE-I/ARB drugs	PCP practices	ACE-I/ARB volunteer practices	--	Provision to practices of: • Education • Consultation • Technical assistance	Treatment with ACE-I/ARB drugs

TABLE I.5 (*continued*)

Theme/Component	Targeted Participants[b]	Method of Recruitment[b]	Other Groups[c]	QIO Interventions	Targeted Outcomes/Goals
AV Fistula	Nephrology practices/other physician practices	AV fistula volunteer practices	--	Provision to practices of: • Education • Consultation • Technical assistance	ESRD patients starting hemodialysis via AV fistula, or ESRD patients starting hemodialysis with AV fistula in place, even if not mature
Community Collaboration	Wide range of organizations to form statewide or regional coalitions and partnerships	CKD volunteer organizations[q]	--	Build and/or sustain state or local coalitions and partnerships with a wide range of organizations to: • Advance one or more of the Task 1 clinical focus areas • Work towards systematic quality improvement in CKD prevention and care	System-level change

Source: QIOs' 9th SOW contracts: original dated August 1, 2008, and contract modification dated July 9, 2009.

[a]Not part of this evaluation.

[b]"Targeted Participants" and "Method of Recruitment" vary widely from theme to theme. Most themes and subtheme components target health care providers (such as hospitals, nursing homes, and physician practices) but one component targets Medicare beneficiaries and other themes target organizations ranging from advocacy groups and professional physician societies to Medicare Part D prescription drug plans. Some themes and subtheme components required the QIOs to clearly identify "participating providers" that had to formally agree to work with the QIO; other components only required QIOs to organize willing providers and organizations into coalitions to work on topics without a formal commitment to participate or enroll with the QIO.

[c]"Other Groups" refers to comparison groups that the 9th SOW specifically describes will be constructed by CMS or its contractors for CMS to use in evaluating QIOs' contract performance.

[d]The QIOs' contract modification of July 2009 added a new patient safety theme component, "Rural-Focused Patient Safety Projects." We are still working with CMS on gathering information on this component and have not included it in this table.

[e]CMS created lists of hospitals and nursing homes whose performance on certain quality indicators fell below pre-specified cutoffs. QIOs were to recruit at least 85 percent of their providers from these lists; the remaining 15 percent or less of providers could come from providers not on the lists.

[f]CMS provided training to two or three staff members from each QIO (national quality improvement leaders) in effective meeting management techniques. These staff members were to return to their home QIOs to train additional staff, and QIO staff would then train provider staff.

[g]QIOs were also to create and foster "Communities of Practice"—state and regional collaborations of providers and stakeholders dedicated to improving quality and to learning from each others' experiences.

TABLE I.5 (*continued*)

[h]Hospitals participating in CDC's National Healthcare Safety Network-Multidrug Resistant Organisms (NHSN-MDRO) reporting module that were willing to share their NHSN-MDRO data with QIOs and to work with them on reducing MRSA infections.

[i]The original QIO 9th SOW also included a hospital pressure ulcers component as well, but this was discontinued by CMS in February 2010.

[j]SFF List maintained by CMS includes nursing homes with persistent, severe quality deficiencies.

[k]Various health care providers who agree to work with the QIO in improving the two drug safety measures. Unlike many of the other themes, there are no formal distinctions between participating and non-participating providers.

[l]Primary care physician practices that possess and use electronic health records (EHRs) and that are willing to commit to improving performance on the prevention measures.

[m]Primary care physician practices that meet all the same eligibility criteria of PPs for participating in the prevention theme but do not wish to commit to improving performance on the prevention measures. However, the Prevention NPs will still receive technical assistance from the QIOs on using their EHRs more effectively.

[n]Practices with the following characteristics: (1) underserved Medicare beneficiaries with diabetes must be $\geq$ 25 percent of Medicare beneficiaries with diabetes in practice, and (2) practice's "average of diabetes measures" must be "within the lower 50th percentile for the state."

[o]In many cases QIOs were recruiting beneficiaries directly from various community settings to participate in diabetes self-management education (DSME), rather than through practices recruited to participate in the Disparities Prevention theme..

[p]The QIO Support Contractor (QIOSC) identified a group of generally similar comparison communities through a heuristic process in order to provide a rough context or benchmark for the intervention communities. These comparison communities were not selected through a formal quantitative matching process.

[q]The Prevention—CKD Theme includes community collaboration activites which are described as a separate subtask, but in practice these activites apply to the entire theme. QIOs are to build and foster state or local coalitions and partnerships with a wide range of organizations interested in CKD; these coalitions would then work towards systematic quality improvement in CKD prevention and care and system-level changes.

- Do impacts differ for underserved beneficiaries and non-underserved beneficiaries (has the program narrowed health care disparities)?

2. Assuming there are impacts, which interventions work (what are the *mechanisms* of impacts)? Which interventions work for whom (which providers and which patients), and in what circumstances?

3. How might the program be improved to provide greater value?

   - Can key activities be more standardized across QIOs in a way that would improve the impact?

These three questions form a hierarchy in terms of increasing generality and level of assessment. The first question naturally leads into a series of detailed analyses of whether each of the various themes and subtheme components have resulted in impacts, although the rigor of the impact analyses that can be achieved across subtheme components varies greatly, given the extremely wide variety of activities, interventions, and participants. The second question leads to a higher level of analysis in the consideration of impacts both within and across themes and subtheme components, and across QIOs and providers, to identify whether certain interventions or types of intervention might be more successful than others. The third question draws on findings from the second question—if specific interventions or activities are indeed found to be more effective for specific topics or providers, then broader dissemination of these lessons might lead to improvement of the QIO program. However, the third question may also lead to the highest level analyses on whether underlying structural features of the program, such as methods of contracting with QIOs, performance incentives for QIOs, organization of CMS to supervise QIOs, and the basic missions and goals of the program, might also be improved.

D. CHALLENGES TO THE EVALUATION

The evaluation faces multiple challenges. The first is the challenge of evaluating an extremely broad, heterogeneous set of activities. Although referred to as "the" QIO Program, the variety of topics, interventions, and participants in the 9th SOW makes it more of a collection of multiple programs, and the overall evaluation thus actually comprises several separate, though interconnected, smaller program evaluations. As mentioned, the rigor of the impact analyses for some of these "smaller programs" will vary widely.

A second challenge lies in a few residual gaps in our knowledge about how providers were recruited for (1) the prevention disparities theme and (2) the rural patient safety theme. As discussed in Chapter III, the prevention disparities QIOs were to recruit practice sites that both served a high percentage of underserved beneficiaries with diabetes and that fell below the state median in performance in measures of diabetes test utilization (hemoglobin A1c tests, lipid tests, and eye examinations). However the QIOs had considerable leeway in how they implemented these criteria, and we are still in the process of learning what each QIO did. Chapter III also describes our uncertainty over whether rural providers for the patient safety theme were recruited based on a rank ordering of performance or on some other basis. Mathematica has been working with CMS and relevant QIOs to clarify these issues. Our analytic approach to the evaluation is likely to evolve as we gain further understanding of these issues.

A third challenge faces the research question on mechanisms of impacts. As described later, the basic approach is to correlate variations in impacts across states or provider types on the one hand, with different types of QIO activities and interventions on the other. However, even if there appears to be variation in impacts for some of the themes or subtheme components, our ability to distinguish whether one state's impacts is statistically significantly different from those of other states is likely to be limited, especially for states with few providers. Sorting the wide variety of QIOs' activities, described by narrative text and survey responses, into clear categories will also prove difficult. Disentangling whether certain activities may have led to larger impacts, certain types of providers may have responded better than others, or certain contextual factors may have contributed to greater effects will require a combination of quantitative and qualitative approaches.

A fourth major challenge lies in the overall synthesis of findings. As noted above, in many ways the evaluation consists of several smaller program evaluations. It may turn out that one theme or subtheme component appears highly successful, while another appears less so. As discussed further on, we will have to decide how to weigh various considerations in the synthesis of results from each of these smaller evaluations—the strength of evidence, the size of effects, and the potential importance for Medicare beneficiaries and the Medicare program.

E. GUIDE TO THE REST OF THIS REPORT

There are five chapters to this report. Chapter II outlines a conceptual framework and logic model for the 9th SOW and explains how we will assess and describe the framework—for example, whether and how anticipated pathways in fact took place, and whether and how QIOs' environment and context may have affected their activities. Chapter III describes the impact analyses for each of the themes and subtheme components. Chapter IV discusses our approach to determining whether specific strategies or mechanisms undertaken by the QIOs may have been more effective, and whether certain types of providers or settings may have responded more strongly than others. Chapter V comments on improving the evaluability of future SOWs in light of the challenges facing this evaluation, and explains how we will synthesize the various findings from different themes and methodologies to yield overall evaluation findings. It also discusses the current evaluation in light of previous recommendations by the IOM and NORC studies, and how the challenges facing the current evaluation might inform the design of the upcoming 10th SOW. Finally, Chapter V outlines the forthcoming reports and deliverables and the project timeline. A complete description of data collection plans and copies of instruments are in the Paperwork Reduction Act (PRA) supporting statement for the evaluation (Kovac et al. 2010).

II. CONCEPTUAL FRAMEWORK FOR THE 9TH SOW

This chapter presents a conceptual framework for the 9th SOW. We first review the goals and objectives of the SOW. We then describe the resources and inputs for QIO activities, the QIOs' expected activities, the context and environment in which QIOs operate, and the pathways and mechanisms by which the ultimate desired outcomes are to be achieved.[1]

A. OVERVIEW

To design an evaluation of the QIO Program 9th SOW, we need to first understand the program conceptually. CMS identifies the core functions of the QIO Program as: (1) improving quality of care for beneficiaries; (2) protecting the integrity of the Medicare Trust fund by ensuring that Medicare pays only for services and goods that are reasonable and necessary and that are provided in the most appropriate setting; and (3) protecting beneficiaries by expeditiously addressing individual complaints, such as beneficiaries' complaints, provider-based notice appeals, violations of the Emergency Medical Treatment and Labor Act (EMTALA), and other related responsibilities as articulated in QIO-related law.

At the highest level, we can summarize the QIO Program's primary quality improvement aim in a single sentence: With CMS direction and support, contracted Quality Improvement Organizations in 53 states/jurisdictions provide resources and consultation to health care organizations to catalyze improvements in quality of care and patient safety, improving beneficiaries' health.

In practice, the program is complex and ambitious. The program's quality improvement goals encompass six distinct focus areas related to the national theme of patient safety in hospitals and nursing homes, as well as the national theme of preventive services in physician practices. In addition, all QIOs must implement beneficiary protection activities, including review of potential quality-of-care problems and supporting hospital public reporting for the Reporting Hospital Quality Data for Annual Payment Update (RHQDAPU). Some QIOs are also contracted under the program to undertake additional activities to reduce disparities, improve care transitions, and prevent and better treat chronic kidney disease.

Table II.1 summarizes the many required QIO activities under each theme and patient safety subtheme, and lists the potential benefits beneficiaries may ultimately experience from those activities. Although the focus of the program is Medicare beneficiaries, in fact, the general public is expected to benefit as well, because when providers improve their practice, they tend to do so practice-wide rather than for a segment of their patient population.

[1] Section A of this chapter was previously provided to CMS in a memorandum dated June 12, 2009, but Section B is new.

TABLE II.1

OVERVIEW OF THE QIO PROGRAM AND EXPECTED BENEFITS FOR BENEFICIARIES

QIO PROGRAM THEME AND SCOPE	**QIO ACTIONS**	**POTENTIAL BENEFITS FOR BENEFICIARIES**
1. Patient Safety: All states		
Pressure Ulcers Physical Restraints Surgical Care Improvement Project • Lower performers (<25% of hospitals and nursing homes in each state)	Works with a set of hospitals (pressure ulcers and surgical care improvement) and nursing homes (pressure ulcers and physical restraints) in each state whose performance is substantially below target levels, to assist them to improve. Uses trainings/meetings to facilitate change; works with executive leadership to initiate additional commitments to QI in their facilities; measures patient safety and quality culture in hospitals and nursing homes and helps them use the survey results to improve; provides improvement tools and guidance on using them; provides feedback to providers on their quality measure data Also works with provider associations and other health care organizations who can help advance patient safety goals, adding value to their efforts	Fewer long-stay nursing home residents should be getting pressure sores and/or be physically restrained. Fewer patients should be getting pressure sores while in the hospital. Hospitals should improve processes of care related to surgical infection prevention, and appropriate medication for heart failure patients and patients on beta blockers; this should lead to fewer patients with infections after surgery and better outcomes for heart failure patients and those on beta blockers.
Methicillin-resistant Staphylococcus aureus (MRSA) • Number of providers who work with the QIO varies widely by state	Works with hospitals that voluntarily report MRSA to the CDC, training provider staff in TeamSTEPPS, a method for effecting change in health provider organizations, and supporting their improvement efforts with tools and resources. Recruits more hospitals to report MRSA to CDC.	Lower chance of MRSA infection, as more hospitals measure and improve their rates of infection and transmission.
Drug Safety • Number and intensity of projects QIO is involved in varies by state	Works in partnership with a set of providers, Medicare Advantage Health Plans, and Prescription Drug-Sponsor Plans (PDPs) who share desire to reduce drug-drug interactions and prescribing of inappropriate medications. Provides information, tools, guidance, staff time and/or data to further the shared objective.	Less chance beneficiaries will be prescribed inappropriate medications and/or will experience a drug-drug interaction that could lead to an adverse event.
Nursing Homes in Need • One nursing home in need selected by CMS in each state each year	Works in-depth with one poor-performing nursing home in each state in each year. Assists each nursing home in identifying the root causes of its problems and developing and implementing an action plan to address them.	Residents of the poorly-performing nursing homes the QIOs work with will experience improved care.

TABLE II.1 *(continued)*

QIO PROGRAM THEME AND SCOPE	QIO ACTIONS	POTENTIAL BENEFITS FOR BENEFICIARIES
2. Prevention • All states: participating practices include 4 to 125 practices with EHRs in each state	Assist physician practices in use of their electronic health records system to improve delivery of preventive services.	Beneficiaries more likely to get timely breast cancer screening, colorectal cancer screening, flu immunization, pneumococcal immunization.
3. Prevention: Disparities • 6 states: In each, practices must serve a minimum (1-15%) of the state's Medicare underserved diabetes population	Works with participating physician practices and other organizations to increase availability and use of diabetes self-management education (DSME)	Patients' knowledge and skills improve with respect to self-management of diabetes among underserved beneficiaries with the disease, resulting in healthier lives with fewer medical problems.
4. Care Transitions • 14 states, one community per state	Works with health care providers, advocacy and service organizations, major purchasers and payers, regional health initiatives, etc. in a selected community, to reduce the rate of hospital readmissions. Assists the health providers in using a specific instrument (CARE) to share critical information during transitions from the hospital.	Beneficiaries receive better care after discharge from the hospital and are less likely to have to be readmitted to the hospital within 30 days.
5. Prevention: Chronic Kidney Disease • 11 states	In a community, develops a strategic plan and works with a broad range of community leaders and providers to prevent and treat CKD accompanied by diabetes and hypertension. Works with providers to incorporate relevant clinical standards into their health information systems.	More patients with diabetes are tested for CKD annually (in accordance with guidelines), allowing for earlier identification and treatment and preventing and delaying ESRD. Patients with diabetes and hypertension and early stage CKD are more likely to be taking medications in accordance with guidelines. When dialysis is required, a higher proportion of patients will receive an AV fistula (best) as the first dialysis treatment
6. Beneficiary Protection • Nationwide	When notified of a potential problem, performs case reviews to identify quality problems; when quality of care concerns are confirmed at the highest level, follows up to ensure a plan to improve the concern is adopted by the relevant provider Supports public reporting of quality data by hospitals	Beneficiaries who had bad healthcare experiences may gain better peace of mind by having their cases reviewed by a neutral third party. If the beneficiary issue is confirmed as a quality concern, the beneficiary may be satisfied that the provider will be required to plan follow-up action to improve the situation. As hospital quality becomes more transparent, it improves.

Figure II.1 provides a conceptual model of the QIO program through a different lens. In this model, the focus is less on the specific benefits the beneficiaries may experience, and more on *how* the results are expected to be achieved. First, looking at the column level and reading across from left to right, we can see that (I) inputs to QIO activities will shape QIO activities (II); QIO activities will be implemented in an environment (III), which will mediate the extent to which they cause the intended reactions within the health system (IV), and thereby improve outcomes including improved quality, improved beneficiary health, and potential savings for the Medicare program (V).

Noteworthy observations from the figure include:

- A well-specified CMS contract, information and tools to support appropriate interventions, and QIO organizational factors (such as qualified staff and management) are the three critical inputs to QIO activities.
- While QIO activities are heavily focused on health care providers, QIOs are also required to work with other health care organizations, such as health plans, and provider or professional associations, and beneficiaries.
- In addition to their main mission, QIOs are required to report on their activities and outcomes to CMS, often through the PATRIOT system, for purposes of CMS oversight, evaluation, and program refinement.
- The environment within which the QIOs must operate is complex. Each of the boxes shown—provider environment (culture, infrastructure, and data), payment environment, legal/regulatory environment, public reporting environment, and non-QIO quality activity and resources—may have interaction effects (either synergies or dampening effects) with the QIO activities that influence their impact.
- In other words, provider, community, and beneficiary reactions to the QIO activities may depend on the activities themselves and on other influences from the environment.
- Ultimately, quality and patient safety measures shown by the program should improve, beneficiaries' health should improve, and the better health may save money for the Medicare program through reduced health care needs.

The evaluation will collect information about the entire program framework shown, so as to understand not only whether beneficiary outcomes improve as expected due to the program, but also which factors within the framework contributed to and hampered success. The two types of influencing factors we will be examining most closely will be (1) those factors within CMS control—that is, the contract-related features, the QIOSC structure and activities providing information and tools to support the QIO interventions, and the required reports and reporting mechanisms, and (2) the environment—particularly the provider environment and non-QIO quality activity and resources, since those factors were mentioned as important in our case studies pertaining during the 8th SOW evaluation.

FIGURE II.1: DRAFT CONCEPTUAL MODEL OF THE QIO PROGRAM

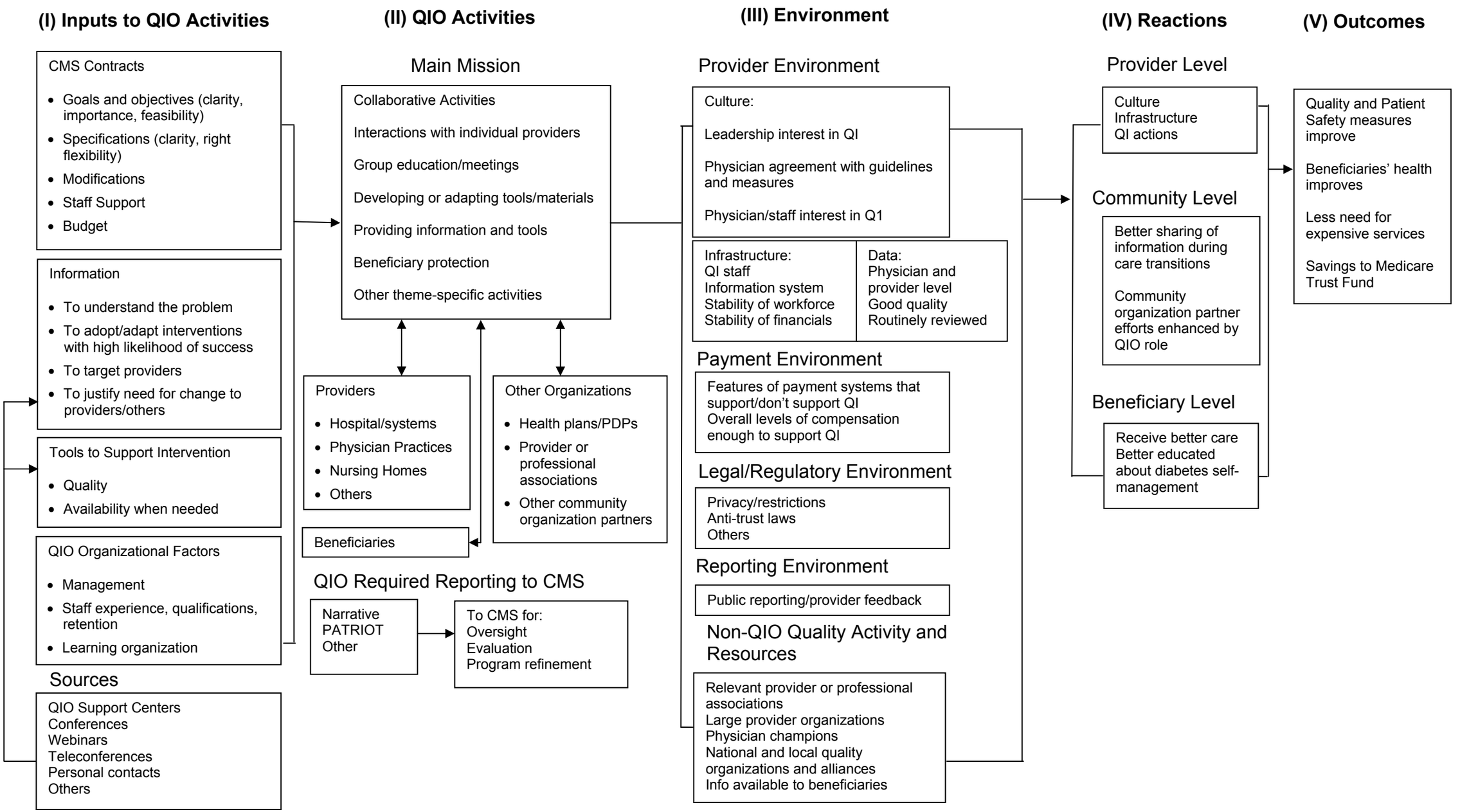


Figure II.2 focuses on the QIO's operating environment. In this figure, the QIO is in the center, and the figure shows the CMS-funded organizations it works with (on the left), the organizations within its state/local environment (the center box), and the organizations and major factors in the national environment that also may affect its work and impact (around the outside of the state/local box). Relative to Figure II.1, Figure II.2 expands the detail shown regarding the CMS-supported infrastructure for the program, and separates and details the national versus state/local environment. Since QIOs are state-specific, studying the relationships between them and the other entities in their state/local environments, and how these affect the QIOs' impacts, is an important part of the evaluation. The figure also recognizes that QIOs may often be working with subcontractors. Our QIO survey will include a request for a list of subcontractors and their main purposes, to understand the full set of entities whose activities are funded under the program.

As described in Chapter I, the 9th SOW is further divided into separate themes and theme components. Appendix C provides logic models for each theme of the QIO program. These serve as schematic, summary representations of the material in the QIO contract pertaining to each theme. They do not include the environment or inputs to the QIO activities, because they are meant to represent only contract-required activities and expected outcomes. They were useful references for us as we designed the evaluation, taking into account the contract requirements related to each theme.

B. ASSESSING AND DESCRIBING THE FRAMEWORK

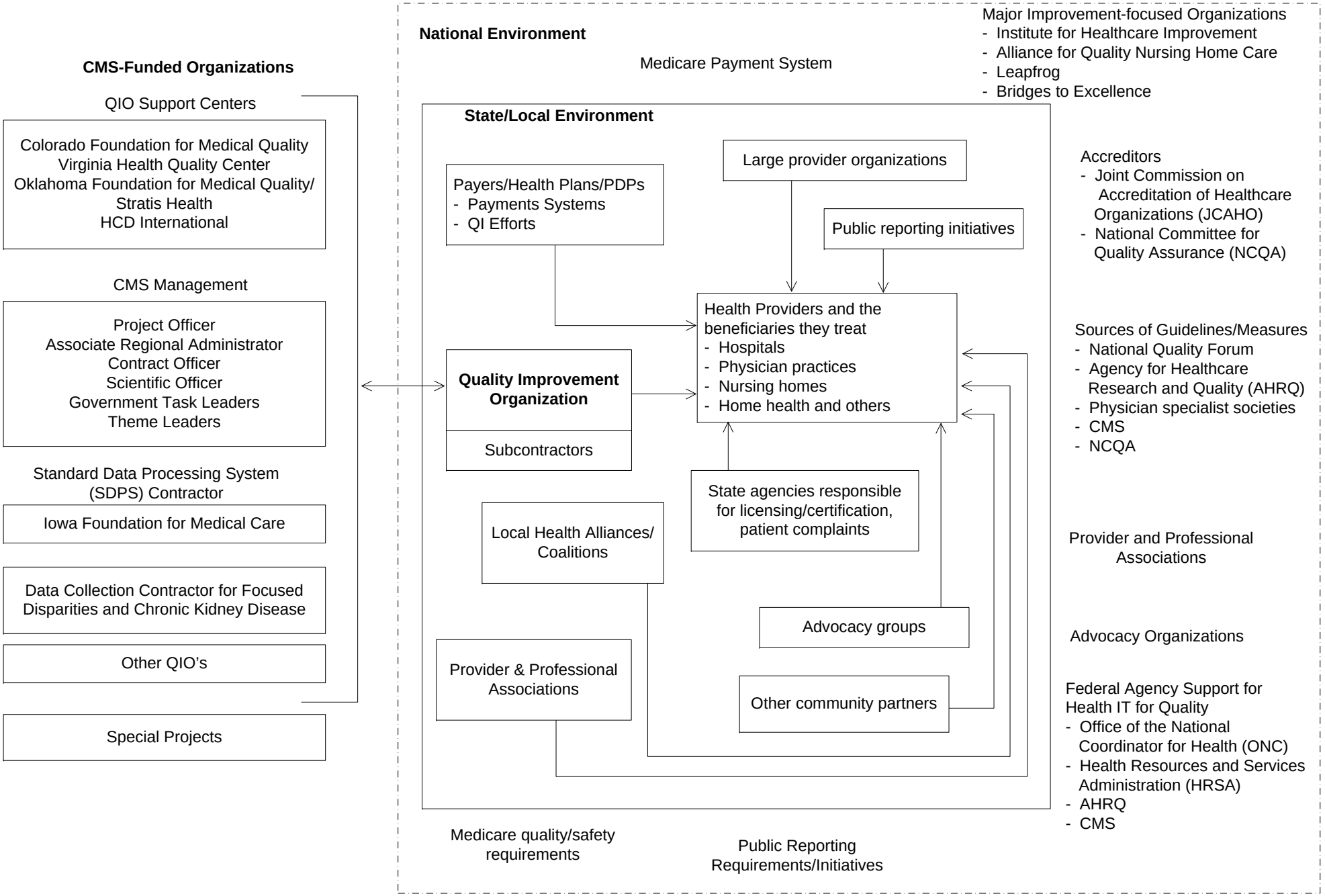
Below we explain how we will assess and describe each of the columns of Figure II.1 and the relationships between them—were activities implemented and pathways followed as anticipated and diagrammed? Chapter III explains how we will quantify whether the program produced the desired outcomes (impact analyses), and Chapter IV discusses our approach to determining whether specific elements in some of the columns may have had led to greater effects than others (for example, within "Group education/meetings" in Column II, QIO Activities, were there particular types of education that seemed more effective, or particular providers who appeared more responsive?)

1. Inputs to QIO Activities

Effective QIO activities depend upon a set of inputs that include clear and well-specified CMS contracts, information and tools to support design and implementation of their interventions, and the healthy functioning of the QIO organization itself. The main data source for this information is the QIO survey.

CMS Contracts. The experience of QIOs with CMS contracts will be captured through the QIO web survey from QIO theme leaders. Specific survey topics covered are listed in Box II.1.

FIGURE II.2: THE QIO ENVIRONMENT



Box II.1: QIO Survey Topics Related to QIOs' Experience with Their CMS Contract

QIO Theme Leaders:
Clarity of the contract and other official documents
Sufficiency of resources in relation to goals
Attainability of improvement targets
Meaningfulness of improvement targets
Reasonableness of time frame
Clarity of method for evaluating the QIO
Importance of focus areas of the contract
Contract well-focused on providers whose improvements will impact quality in the state
Knowledge base of CMS oversight personnel relative to their responsibilities
Supportiveness and helpfulness of the CMS Project Officer
CMS Project Officer understands the QIO's interventions
Clarity of communication by CMS personnel
Consistency of communication among CMS personnel
Effort required to implement contract modification(s)
Value of contract modifications in improving the contract
Does the QIO recommend any changes to:
    Focus of QIO contract
    How QIOs are evaluated

For the 10 case study states, regarding any items that are negative toward the experience with the contract, we will ask the respondent to tell us more about the problem and whether it significantly lessened the results they were able to achieve (and if so, why). We will ask the QIO director and all theme leaders in the case study states to identify any barriers to the QIO's effectiveness that stem from the contract or CMS procedures. In addition, we will ask the respondents to elaborate on any negative responses and any "excellent" responses regarding the knowledge base and communications among CMS oversight staff. To encourage frank responses, we will not associate individual or state names with specific comments.

Information and Tools. Information critical to the effectiveness of the QIOs comes from CMS-sponsored sources such as QIOSCs and annual meetings held for some themes, and from non-CMS sources. The QIO survey is the data source for the evaluation to understand the extent to which information and tools from CMS and other sources supported the program well. Theme leaders are the primary respondents since information and tools are theme-specific, as Box II.2 shows.

For the case study states, we will probe on any negative responses on the QIO theme leader survey questions on information and tool supports, to explore the types of data they felt they needed but did not have and whether they believe this significantly lessened the results they were able to achieve. We will ask what factors led them to rate some information sources as having high value and others as having low value. Regarding support from the QIOSC, we will ask all the theme leaders what the QIOSC contributed to their ability to work effectively on their theme. For the relevant themes we will ask how useful they found the "change package" that was

developed centrally by CMS,[2] and whether they benefited from the annual in-person meetings held by CMS.

If the QIO directors suggested one or more improvements to CMS-funded tools or resources, we will ask them to elaborate about their ideas and how improvements might help QIOs better facilitate quality and patient safety in the health system.

Healthy Functioning of the QIO Organization. Another key input to effective QIO activities is a healthy QIO organization, including sound management, staff with strong experience and qualifications for their positions, and organizational learning processes so that mistakes are not repeated and the level of effectiveness improves over time. Management is difficult to measure; we plan to assess it qualitatively through the case studies. For the 12 case study states, we hope to have enough information to be able to assess whether any shortcomings (such as many missed project milestones, need for extensive CMS or QIOSC assistance, or failure to achieve process and outcome goals) stem in part from management issues. In our past experience with case studies, any serious management problems tend to become obvious through the interviews.

In order to be able to identify staffing factors that might be associated with QIO success, the QIO survey will ask, for each theme, about the qualifications and experience level of the staff who work most directly with provider organizations. In addition, the theme leaders will indicate their level of agreement with three staff-related items: (1) that QIO staff assigned to this theme have the right substantive expertise and experience; (2) that an adequate number of QIO staff have been available to perform work on this theme; and (3) that the QIO has been able to retain key staff working on this theme (that is, turnover has not been a problem).

Box II.2: QIO Survey Topics Related to Tools and Information Supporting QIO Activities

QIO Directors:
Improvements they would suggest to CMS-sponsored tools or resources
Any change needed in the program's emphasis on QIOSCs

Theme Leaders:
Sufficiency of data to:
- Understand the problem the intervention is addressing
- Support intervention design
- Identify disparities related to the theme
- Identify which interventions are working elsewhere
- Adequately justify the intervention to providers and others

Value of information received from a list of sources (QIOSCs, QualityNet conferences, etc.)
Quality of tools and other resources to support interventions
Timeliness of availability of tools and other resources supporting the interventions
Functionality of measurement tools (how well they work)
Need for adaptations to existing tools or resources
Need to create new tools or resources

[2] Early discussion with the patient safety theme leader at CMS indicated CMS is developing a set of resource tools called a change package, to assist QIOs in working effectively with providers on the patient safety theme.

Finally, whether the QIOs are learning organizations can be most effectively ascertained for the case study QIOs by first reviewing their quarterly reports submitted to CMS on this topic, and then discussing what we learned during the site visits. On each site visit interview with a QIO director, we will summarize what we learned from their quarterly reports. We will then ask them to confirm our summary and to elaborate on anything puzzling or particularly interesting from the review.

2. QIO Activities

Because the QIO contracts offer considerable flexibility in how the QIOs achieve their goals, our preliminary information indicates substantial variation in the emphasis and specific activities QIOs and their subcontractors undertake to achieve their goals. The QIO survey aims to capture the variation in the mechanisms and emphasis in the field, in order to exploit it in our analysis of what worked for whom and under what circumstances. Box II.3 lists the types of activities related to QIOs' main mission that we will capture on the survey for each theme. It also includes items that capture the theme leaders' perceptions about what motivates providers in their state to improve. Differences in perceptions about provider motivation may help explain differences in activities, which would be important to producing appropriately nuanced findings on our research questions. In the case studies, we will follow up on the QIO survey responses to discuss why they rated various activities as high- and low-value.

In addition to the activities related to their main mission, QIOs also are required to report to CMS on their activities. The required reporting structure and frequency varies by theme. It is beyond the scope of this evaluation to fully assess the structure of reporting and its value to CMS and the program. However, because reporting activities represent a significant QIO responsibility under the 9th SOW, we have included items in the QIO survey to assess the QIOs' experience with the system they use to report data to CMS (PATRIOT), and to assess the level of effort they are devoting to reporting requirements each month (Box II.4).

3. Environment

The evaluation's plans for measuring (and assessing) the role of the state-specific provider environment were described in a memorandum to CMS dated August 11, 2009 and are repeated here in Section a below for completeness of the evaluation design within this document. Since column III of the conceptual framework of the 9th SOW program includes much more than the provider environment (shown in more detail in Figure II.2 above), we have added text here to explain how the evaluation will take into account the other relevant parts of the environment: payment environment; legal/regulatory environment; reporting environment; and non-QIO quality activity and resources.

Box II.3: QIO Survey Topics Related to QIO Activities (Main Mission)

QIO Directors
Any change needed in how QIOs are expected to work with other providers?
Any change needed in how QIOs are expected to work with other health care organizations?

Theme Leaders
For each of the following, indicate if it is a major or minor component (or not applicable), and how important it is (very, somewhat, or not important) to improving quality or patient safety for the theme:

Collaborative Activities
Forming new provider collaborations
Forming new collaborations including organizations other than providers
Contributing to existing collaborations
Supporting a large organization (such as a health delivery organization or health plan) in its efforts to improve

Interactions with Individual Providers
Problem-solving or strategizing with individual providers at their request
Problem-solving or strategizing with individual providers during meetings the QIO initiated
Making presentations on-site at individual providers
Interacting with top leadership of provider organizations
Helping integrate clinical guidelines into health information systems
Helping providers better use their health information systems to better support QI
Discussing providers' own performance with them
Training staff within provider organizations

Group Education/Meeting Activities
Providing educational or shared learning sessions via telephone
Large regional or statewide in-person meetings
Routinely providing provider-specific data to providers with benchmarks
Notifying providers of quality improvement-related opportunities sponsored by others
Summarizing quality improvement tips or information in a QIO or provider association newsletter, in paper or electronic format

Business Case Focus
Developing or incorporating information into materials, talks, consultations, etc. regarding the business case for quality improvement relevant to this theme

Care Transitions Theme Only:
Encouraging and training on use of the CARE instrument
Use of a transitions coach

Prevention – Disparities Theme Only:
Obtaining clinical EHR-based data from practices
Recruiting and training community health workers
Implementing diabetes self-management education for beneficiaries with diabetes

For Prevention – Disparities theme only:
- Mechanism used to recruit beneficiaries for DSME
- Urban/rural nature of the geographic area targeted under this theme

Agreement with statements regarding key motivators for quality improvement:
- Business case for quality, when clear, is a key motivator
- Pay-for-performance efforts are a key motivator
- Motivational speakers are effective motivators
- Public reporting is a key motivator for improvement

Box II.4: QIO Survey Topics Related to QIO Reporting

QIO Theme Leaders
Smoothness of functioning of the PATRIOT system
- In the first six months of the contract
- After the first six months of the contract

Number of hours spent fulfilling CMS reporting requirements in an average month
- Senior staff
- Mid-level staff
- Junior staff

a. Provider Environment

As shown in Figure II.1, the provider environment encompasses three domains: (1) professional culture regarding quality improvement, (2) infrastructure to support QI, and (3) the availability and use by providers of timely, relevant data to monitor improvement. These three sets of factors in the provider environment are expected to affect providers' receptivity to QIO information or advice, their interest in quality improvement, and their ability to make desired improvements. Information on the provider environment in all 53 QIO jurisdictions will be collected through the QIO survey (since 50 of the jurisdictions are states, they will be referenced as states hereafter in this memo for simplicity). More in-depth information will be collected for 12 states through case studies, and nationally and for large states through the hospital and nursing home surveys.

QIO Survey.[3] For the portion of the QIO survey relevant to the provider environment, theme leaders are the relevant respondents.[4] Theme leaders in each state will be asked to take a statewide perspective regarding the types of providers relevant to their theme—to think beyond the smaller set of providers they have worked with to achieve the goals for their theme.. Often theme leaders have broad-based experience working with a large cross-section of providers in their state over many years, making them a useful resource on the statewide provider environment. Although we are still pre-testing the survey, theme leaders interviewed for the pretests thus far have felt competent to comment with a statewide perspective. Most of the relevant survey questions ask whether they strongly agree, agree, disagree, or strongly disagree with the statements corresponding to the topics shown in Table II.2, as they relate to the respondent's specific theme. Questions regarding the prevalence and role of large provider organizations in driving quality in the state do not follow an agree/disagree format because responses are tailored to each question. Although the information will be useful to the evaluation, it will represent opinions of these individuals rather than objective data on the topic (which do not exist).

[3] See the PRA supporting statement (Kovac et al. 2010).

[4] Theme leaders are individuals designated by the QIO to be responsible for leading QIO quality improvement work with providers and others relevant to the following themes or patient safety subthemes in the 9th SOW: Patient Safety – Pressure Ulcers; Patient Safety – Surgical Care Improvement Project; Patient Safety – Methicillin-resistant Staphylococcus aureus (MRSA); Drug Safety; Nursing Homes in Need; Prevention; Prevention – Disparities; Prevention – Chronic Kidney Disease; and Care Transitions.

TABLE II.2

PROVIDER ENVIRONMENT TOPICS COVERED BY DATA COLLECTION EFFORTS

Provider Environment Topic	QIO Survey	Hospital and Nursing Home Surveys	Case Study Respondents				
			QIOs	Hospitals	Nursing Homes	Physician Practices	Community Health Leaders[a]
Motivation/Culture							
Provider organizations' interest in quality, and impact of this	X	X	X	X	X	X	
Perception among providers of a strong business case for quality	X			X	X	X	
Factors motivating providers to improve quality			X	X	X	X	
Willingness among providers to share information on QI (and impact and factors underlying that)			X	X			
Role of large provider organizations in the state in driving quality	X	X					
Adequacy of number of physician champions willing to help facilitate improvement	X						
Data							
How commonly providers regularly review data on their performance	X	X				X	
Infrastructure							
Extent to which information system issues remain a barrier to improvement	X	X					
Extent to which providers have staff who are educated and qualified to support improvement efforts	X	X					
Workforce instability (turnover) is a barrier to improvement	X	X					
Provider Culture-Related Reasons for Poor Performance (where it exists)							
Physician disagreement with relevant guidelines/measures	X	X					
Physician disagreement with establishing care routines based on guidelines	X						
Corporate chain managers who do not believe in establishing care routines based on guidelines	X						

Table II.2 (continued)

Provider Environment Topic	QIO Survey	Hospital and Nursing Home Surveys	Case Study Respondents: QIOs	Provider Environment Topic	QIO Survey	Hospital and Nursing Home Surveys	QIOs
Characteristics Affecting QIO Impact							
Characteristics of provider environment that make providers particularly receptive to QIO initiatives			X				X
Characteristics of provider environment that make it particularly challenging for QIO to assist providers			X				X

[a]A community health leader may be, for example, the leader of a regional quality coalition within the state or the leader of a provider association that has been active in quality improvement efforts.

Hospital and Nursing Home Surveys. The evaluation team will conduct surveys of hospitals and nursing home quality improvement directors, including some from facilities that work with the QIO on quality improvement and others from facilities that do not. The surveys, to be conducted during May through August 2010, will be an important source of information on the provider environment. While our proposed sample sizes were driven by minimum detectable differences for national impact estimates, we will also explore state-specific estimates for some of the larger states. For the 1,250 hospitals and 1,250 nursing homes expected to complete the survey, we will have information about their interactions with the QIO, their own characteristics and culture, their infrastructure for QI, their use of data, and their outcomes on QIO-targeted measures. These data will allow for powerful analysis of the relationships between provider characteristics and QIO impacts, as described below and considered further in the forthcoming evaluation design report.

Case Studies of QIO Programs. The 9th SOW evaluation team plans to conduct site visits to 12 state QIO programs during November 2010 through May 2011 (the selection of QIO programs for site visit is described in Chapter IV). The site interview guides will stimulate discussion of the state provider environment with QIOs, hospitals, nursing homes, physician practices, and community health leaders (topics summarized in Table II.2). We will first screen provider respondents as to whether they sometimes talk with other peer providers in the state about these topics; we will only further probe about the state's provider environment with those that do.

b. Payment Environment

Widespread recognition that the provider payment environment does not support high quality care has led to CMS, private payer, and state-based efforts to better support quality through value-based purchasing, pay-for-performance, and most recently, bundled payments

(Massachusetts). In addition, the overall level of payments may affect providers' ability or willingness to engage in quality improvement activities.

The evaluation plans to identify whether the payment environment is playing a role in supporting or detracting from quality improvements and whether it plays a role in QIOs' performance. Box II.5 lists the relevant primary data collection topics for the evaluation, by data source. In addition, we plan to use secondary data on provider income or operating margin to the extent feasible for hospitals, nursing homes, and physicians, to represent the net effect of the payment environment.

c. Legal/Regulatory Environment

The evaluation will explore the role of the legal/regulatory environment through the case studies. Not enough is known about how the legal/regulatory environment may be affecting quality improvement and QIOs' ability to influence it to include it in our more structured data collection efforts (especially given the need to limit the length of the instruments to encourage response). For example, the IOM and NORC reports identified lack of data-sharing as a problem inhibiting quality improvement due to legal and regulatory restrictions on data. HIPAA and anti-trust laws are other relevant legal domains. As illustrated for the payment environment in Box II.5, the case studies will include discussion of factors that motivated quality improvement and factors that inhibit it, and we will look for whether the legal and regulatory environment (and the specific legal domains mentioned above) appears in these discussions.

Box II.5: Primary Data Collection Topics Related to Payment Environment

Theme Leader Survey
Extent to which ongoing pay-for-performance efforts are a key motivation for QI in this state
Level of agreement that poor performers often have financial and management problems

Hospital and Nursing Home Surveys
Extent to which resource constraints are a barrier to improvement

Case Studies
QIO Directors and Team Leaders
Reasons for sufficient or insufficient motivation to improve among the provider community

Hospital (H) and Nursing Home (NH) QI Directors and Physicians (MD) – Identify if payment environment is part of the provider's story about:

Their motivation to make improvements on QIO program-relevant measures (H, NH, MD)
Remaining barriers to achieving optimal performance on the QIO program-relevant measures (H, NH, MD)
How quality fits into the provider's overall business strategy (H, NH)
The main factors that led it to improve its performance over past three years (if improved) (H, NH)
The main reason it does not at present give itself a high score for overall quality and safety (H, NH)
One or two changes that could most improve performance (H, NH)

Community Health Leaders - Identify if payment environment is part of the health leader's story about:

Remaining key barriers to improvement
Characteristics of the provider environment that make providers particularly receptive to QIOs
Characteristics of the provider environment that make it challenging to assist providers

d. Reporting Environment

Public reporting of quality data can be an important influence in providers' quality improvement (Paez et al. 2009), and could also interact with QIOs' ability to foster improvement with providers. Therefore, our QIO Theme Leaders survey will ask if state-level public reporting exists relevant to each theme in the state. Also, we know what data are nationally publicly reported. Ideally, public reporting would enhance the desire by providers to improve, and the QIO would assist them in accomplishing improvement. If so, it should amplify the QIO impact for measures that are publicly reported. Similar to other aspects of the environment, we will identify whether public reporting is mentioned as a factor motivating improvements, particularly if it is cited as a factor that enhances QIOs' ability to work with providers on improvements.

e. Non-QIO Quality Activity and Resources

QIOs are only one of many players attempting to positively influence quality of care, as shown in Figures II.1 and II.2. The efforts of national-level players such as the Alliance for Quality Nursing Home Care, and the Institute for Healthcare Improvement, are well known to the evaluation staff and their efforts will be recognized in the analysis plan.[5] However, the state-level players who may be important will vary. Therefore, our Theme Leader Survey includes a significant component to understand the other important players in the state and their roles (Box II.6).

In the case studies, we will follow up on the survey information to learn more about the types of activities of these other organizations and the relative role of the QIO. In addition, we will discuss with hospitals, nursing homes, and drug safety organizations their interactions with external organizations around quality or patient safety improvement, to determine which of these interactions had an important influence on the provider's quality or safety-related efforts and how care changed as a result.

Box II.6: Quality Improvement Actors Other than the QIO, and Their Roles

Theme Leader Survey

Role of state agency most relevant to each theme (regulatory oversight, actively engaged in fostering quality improvement, or both)

For up to two provider or professional associations most relevant to each theme:

- Presence (or not) of at least one staff member with major responsibility and time devoted to QI
- Association sponsors (or not) a quality-focused entity like a Quality Council or Quality Institute
- QIO and association work jointly on one or more QI efforts substantial in scope
- QIO staff speak at association-sponsored meetings at least annually
- Association and QIO staff talk at least quarterly to avoid duplication of effort
- Association works on entirely different QI projects
- Association works with a different set of providers than the QIO
- Association primarily focuses on quality reporting rather than QI

For up to two large provider health delivery organizations in the state most relevant to this theme:

- Extent to which headquarters of the organization drives quality in owned or affiliated organizations

Adequate number (or not) of physician champions willing to help facilitate improvement on key measures for each theme

List of up to three other external organizations whose efforts are proving important to achieving improvements on each theme

[5] For example, we will look for patterns in QIO effectiveness on program-relevant measures also targeted by these groups vs. those not also targeted by these groups.

f. Reactions

Figure II.1 shows that we expect reactions to occur to QIO activities at the beneficiary, community, and provider levels in order to produce changes in outcomes. Our primary data collection efforts will be key to identifying such reactions (Box II.7).

4. Outcomes

Finally, column V of Figure II.1 contains the anticipated end results of the 9th SOW. The impact analyses, discussed in the next chapter, will analyze whether or not the 9th SOW in fact led to these desired outcomes.

Box II.7: Primary Data Collection Topics Identifying Reactions to QIO Activities

Hospital and Nursing Home Surveys
Did any meetings with the QIO lead to any changes at the hospital that ultimately improved patient care?
If yes, for which measures (if specific to measures)
Did educational materials or tools from the QIO lead to changes at the hospital that ultimately improved care?
If yes, for which measures (if specific to measures)
Extent to which data feedback from the QIO are shared with hospital/nursing home physicians and staff
Has feedback from QIO identified a quality issue not known, heightened attention to issues already known, or otherwise been important to the hospital's quality efforts?
If the hospital participated in the Hospital Leadership Quality Assessment Tool (HLQAT), were any changes made as a result that strengthened quality at the hospital; were they important or not very important changes?

Case Studies
Hospitals, Nursing Homes, Physicians, and Drug Safety Organizations
Any changes made as a result of interactions with the QIO (H, NH, MD, DSO)
Any observation of improvements in the condition of your patients who attended the DSME training, that you believe were attributable to the class (MD)
Any changes made that ultimately improved care as a result of HLQAT (H)

Partner Organizations
Partner organization operational changes resulting from collaborative participation
Changes in care resulting from the work of the collaborative

Focus Groups
Changes in health behaviors and knowledge related to diabetes

III. DESIGNS FOR IMPACT AND COST-BENEFIT/ COST-EFFECTIVENESS ANALYSES

This chapter focuses on the evaluation's approach to studying the impacts of the 9th SOW. There are, however, themes and subtheme components for which impact analyses cannot be done; this chapter also describes these situations and the descriptive analyses that we plan to do.

By program impacts, we mean outcomes that were caused by the program. The ideal situation for inferring causation is one in which we can compare outcomes of Medicare providers participating in the 9th SOW and receiving assistance from the QIOs (called the "intervention" or "treatment" state) to an otherwise similar group of providers not exposed to the 9th SOW (called the "counterfactual" or "control" state); any differences in outcomes must then be due to the program (Rubin 1974).[1] Such a situation holds true in the setting of an experiment in which providers are randomly assigned to either receive or not receive the program; because of the random assignment, participating providers in each group must be otherwise the same.

When experiments are not possible, as in the 9th SOW, inferring program impacts from comparisons between program participants and nonparticipants becomes less straightforward. Providers recruited to receive assistance from QIOs, and those not so recruited may differ in important ways that can affect their outcomes and thus confound the interpretation of observed differences in outcomes. QIOs may have sought out providers with greater motivation and resources for quality improvement, or ones with previous success in implementing such projects. Providers willing to work with QIOs may likewise have stronger desire and better means to improve quality. It may be these underlying and unmeasured characteristics that actually cause any observed outcomes of increased care quality, and not the QIO program. Attributing simple differences in outcomes between participants and nonparticipants to the program thus risks so-called "biased" estimates of program impacts, that is, a systematic overestimation of QIO program impacts.[2] A wide variety of statistical and econometric techniques have thus been developed that go beyond simple participant/nonparticipant comparisons in attempts to avoid or minimize bias in estimating program impacts from nonexperimental situations.

Two such approaches—(1) "regression discontinuity" and (2) "matching"—are relevant for several themes and subtheme components. Section A provides general descriptions of these approaches and their strengths and weaknesses, and Section C outlines the details of their application to specific themes and analyses.

[1] We focus our discussion on providers because, as explained in Chapter II (Figure II.1), the QIO program's primary efforts are in assisting Medicare providers; resultant improvements in providers' care delivery then lead to improvements in beneficiary outcomes.

[2] We have discussed here the example of overestimation of program impacts, because that bias seems more likely given how QIOs and providers agree to work with each other, but there are also programs and program evaluation analyses in which the bias may be towards *under*estimation of program impacts.

As mentioned, there are also themes and subtheme components for which impact analyses cannot be done. In most cases, this is because there is no separate group of nonparticipant providers that can serve as a reasonable control or counterfactual condition, but there are also instances in which there are no data available on nonparticipants, the interventions are highly variable, or the numbers of participants are very small. Where impact analyses cannot be done, we will describe time trends of outcomes among providers for which we have data. We cannot infer program impacts from such descriptive trends; we could only do so if we knew for certain beforehand what time trends would have been in the absence of the program, which of course we cannot know. Section B also describes where, why, and how we will do these descriptive analyses.

A. COMMON IMPACT ESTIMATION APPROACHES ACROSS THEMES AND SUBTHEME COMPONENTS

1. Regression Discontinuity

We plan on using regression discontinuity (RD) designs to estimate program impacts for three of the patient safety theme components—(1) SCIP/HF in hospitals, (2) pressure ulcers in nursing homes, and (3) physical restraints in nursing homes—and for the prevention disparities theme.[3] RD is considered possibly the strongest type of quasi-experimental design (Lee and Lemieux 2009) and has been found to perform well in reproducing results of randomized controlled trials (Cook and Wong 2008). We discuss general aspects of RD designs here and then in Section C below provide details of RD analyses that are specific to each particular 9th SOW theme or subtheme component.

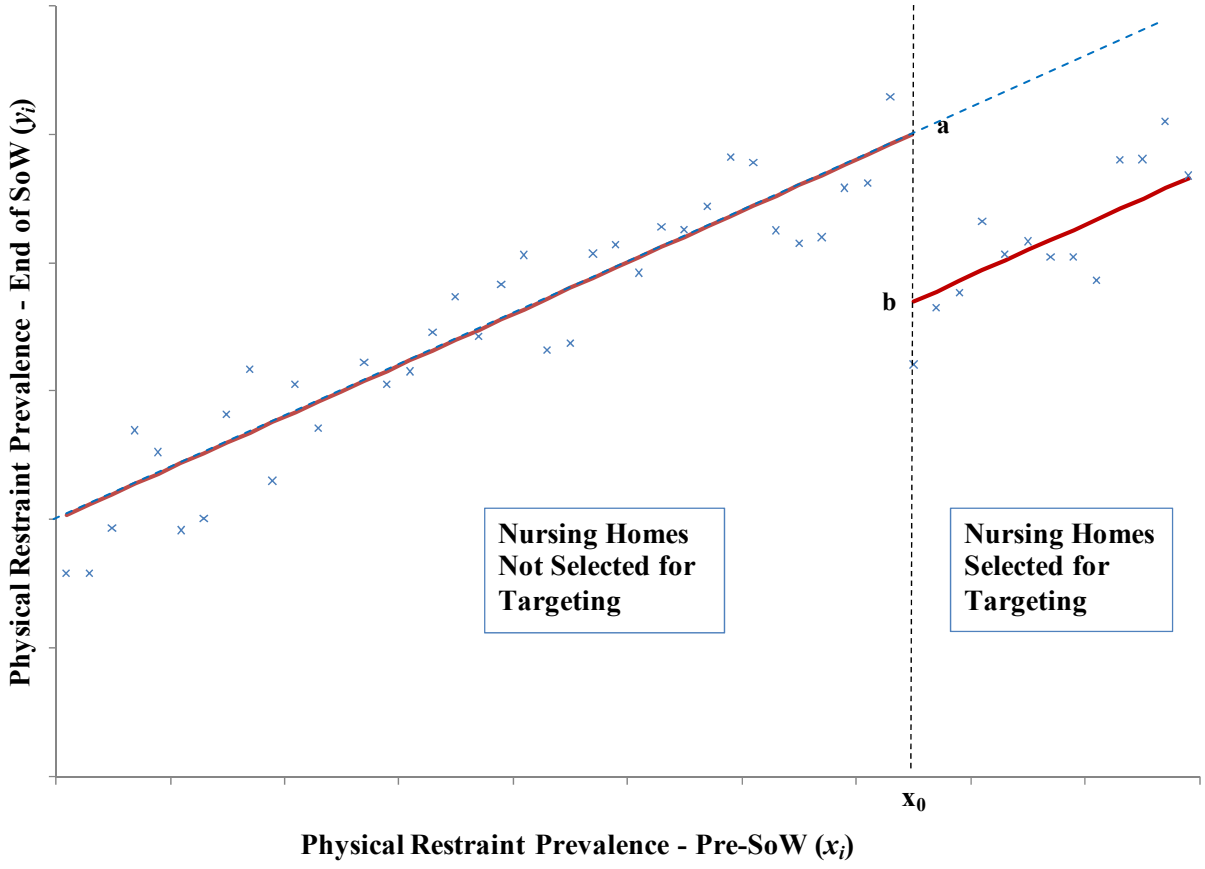
Regression discontinuity designs can be used in situations where assignment to a treatment is based on some selection measure, x_i, with those to one side of the cutoff value, x_0, being assigned to the treatment group ($D_i = 1$) and those to the other side to the control group ($D_i = 0$). For instance, in working with nursing homes to reduce the use of physical restraints under the patient safety theme, QIOs were instructed to recruit primarily from among nursing homes with physical restraint rates 8 or more percentage points above the goal of 3 percent (in other words, the cutoff was 11 percent). Nursing homes with baseline scores above the 11 percent cutoff had a much higher probability of becoming participating providers (PPs) than did those below the cutoff.

The intuition for the design is illustrated in Figure III.1. The figure contains hypothetical observations and fitted lines for the relationship between nursing homes' baseline rates of physical restraint use (x_i) and their rates at follow-up (the outcome variable, y_i). The small x's

[3] Because of small sample sizes or inconsistent methods of defining the cutoff, the RD design may prove infeasible for the prevention disparities theme, and we also discuss below the possibility of using matching methods to study the prevention disparities theme.

FIGURE III.1

REGRESSION DISCONTINUITY ILLUSTRATION



in the chart represent the average end-of-SOW physical restraint rates for providers with given pre-SOW rates. In the figure, providers with higher baseline levels of x_i tend to also have higher values of y_i, as we would expect. The line to the left of the selection cutoff, x_0, represents the actual association between x_i and y_i for observations not subject to the treatment. The dotted line to the right of x_0 is an extrapolation of expected outcomes in the absence of the program for nursing homes with values of x_i higher than the cutoff threshold. However, those observations are, in fact, subject to the treatment and their *actual* outcomes are far below that dotted line. The estimated regression for observations above the cutoff is the solid line to the right of x_0. Note that the regression line is smooth at all points other than the sudden downward shift at x_0, which reflects the results of the program impact on reducing physical restraint use.

The impact estimate in a regression discontinuity design is derived by comparing outcomes for observations just above and just below x_0. In the case of this example, nursing homes just above and below the cutoff are very similar in baseline characteristics that are correlated with or predict outcomes, so differences in outcomes would be attributable to the fact that they differ in their exposure to the treatment (QIO intervention). Impact estimates in an RD analysis are based on the vertical distance between the two trend lines (the incongruous/"discontinuous" jump) at the point x_0. In Figure III.1, that is the vertical distance between points *a* and *b*.

The strength of RD for inferring program impacts results is that, unlike other quasi-experimental techniques, it is not necessary to assume or simply hope that estimates are unconfounded by other potential factors that are unobserved and may be correlated with both treatment status and the outcome, because the identified variation in treatment status is fully observed and understood. As long as the agents being selected are unable to precisely impact their treatment status, assignment for those just above and below the cutoff is "as good as randomized" (Lee and Lemieux 2009). Consequently, causal attribution in RD analyses is strongest and estimation most straightforward when there are many observations near the selection cut-point. However, RD analyses often require use of observations farther from the cut-point in order to have a sample large enough to produce sufficient statistical power. This introduces the need to model the functional form between the selection variable and the outcome. If the relationship is nonlinear, failure to appropriately model it can lead to biased estimates of the size of the discontinuity at the cut-point, that is, of the program impact. Impact estimates must also take into account the fact that some providers below the cutoff may become PPs, while some of those above will not end up being PPs. Appendix B presents technical details of RD estimation related to both of these issues, and describes our approaches to dealing with them and to verifying the validity of the estimates.

A second limitation of RD analyses is of external validity. The impact estimates are generally considered to be relevant only to providers with baseline levels near the cut-point. To the extent that QIO impacts vary depending on the baseline performance of providers, such variation would not be detected using RD analyses.

a. Impacts for Subgroups

We will conduct subgroup analyses to address whether relevant themes and subtheme components have differential effects on certain subgroups of beneficiaries or providers. We will estimate subgroup impacts using interaction terms in the regression models. Those are variables where the treatment indicator is multiplied by the subgroup measure. For instance, if individuals are the unit of analyses and we are interested in whether impacts are greater for African Americans than for other racial/ethnic groups, we would include in the model the treatment indicator, an indicator with a value of 1 if the beneficiary is identified as African American, and a third measure that is the product of the other two variables. That third term will, consequently, take the value of 1 for African American beneficiaries served by provider on the J17 list and a value of 0 for all other beneficiaries. If program impacts on African American beneficiaries are no different than those on other beneficiaries, the regression coefficient for that term will have a value statistically indistinguishable from zero. A coefficient with a value statistically different from 0 would reflect differential impacts by race/ethnicity.

Analyses of beneficiaries will focus on racial and ethnic disparities, but will also investigate cross-region and urban-rural disparities. Our subgroup indicators for these three areas will be:

- ***Race/Ethnicity:***
  - Individual-level: Binary indicators for White (non-Hispanic), African American (non-Hispanic), Hispanic, and other non-Hispanic, respectively.
  - Provider-level: Binary indicators for high proportion of beneficiaries (where available) or residents in the same county who are White (non-Hispanic), African American (non-Hispanic), Hispanic, and other non-Hispanic, respectively.
- ***Region***: Binary indicators for the four major Census regions (East, West, South, Midwest) and non-state territories (as a group).
- ***Urbanicity***: Binary indicator for whether the provider/beneficiary is located/lives in a metropolitan area.

Analyses of providers will focus on provider characteristics that have been found to affect quality performance, such as for-profit or not-for-profit status, bed size, teaching status (for hospitals), and so on.

2. Matching and Comparison

For the chronic kidney disease (CKD) component and the care transitions theme, in which the QIO interventions are designed to affect care at the state and community-levels, respectively,

we propose using matching techniques to identify comparison communities[4] that are as similar as possible to intervention communities on all measured characteristics that might correlate with outcomes. These comparison groups serve as our estimate of the counterfactual conditions; the goal of the matching process is thus to identify comparison communities who differ from intervention communities only in their exposure to the QIO technical. Several general approaches for matching have been developed, and we describe below in Section C the specific matching procedures proposed for these themes.

We plan to evaluate changes in outcomes "pre and post" intervention, as measured by Medicare claims data on all patients attributed to intervention and comparison communities participating in the care transitions theme and CKD component. It is important to note that our "post" period coincides with the 9th SOW, as we are conducting this evaluation concurrently with QIO 9th SOW activities. As a result, a limitation of our analysis plan is that we are unable to measure impacts after the full three-year period of the 9th SOW and that we may thus underestimate the impact of the QIO program if additional time is required for changes in provider behavior and for such changes to then lead to changes in patient outcomes.

For both the CKD component and the care transitions theme, we will conduct descriptive analyses comparing outcomes between intervention and comparison communities both pre and post to evaluate whether intervention and comparison communities were at different levels of the outcome measures at baseline and whether they had different patterns of change between baseline and follow-up. We may also conduct descriptive pre and post analyses by specific sub-populations within these communities—for example, minority and underserved populations. These descriptive analyses are helpful to understand differences in the experiences of intervention and comparison providers over the 9th SOW, but should not be interpreted as impact estimates; for these, we require regression-adjusted analyses that further adjust for differences in community and patient characteristics between intervention and comparison communities.

We will estimate these regression models on patient-level data that include variables derived from Medicare claims and administrative data, including the outcomes of interest (described in more detail below), demographic characteristics (such as age, sex, and race), comorbidities (for example, ischemic heart disease, congestive heart failure, stroke, diabetes, cancer, and so on), and prior Medicare service use (hospitalizations, physician visits, Medicare nursing home care, and so on).

Each patient is from an intervention or comparison community; because the same communities are included both pre and post, many of the same patients are likely to be included both pre and post. We will adjust the standard errors on our estimates to account for the repeated measures on patients. We will evaluate whether patients treated by intervention communities had better outcomes relative to comparison groups post-intervention using difference-in-difference analyses. Specifically, we will estimate the following types of models:

[4] For the CKD theme, we plan to match all counties within a CKD state to comparison counties in other states, as our prior experience with matching suggests it is much easier to identify counties that are similar to each other than states.

(1) $\hat{Y}_{ij} = \beta_0 + \beta_1 * Int.Comm_{ij} + \beta_2 * post_i + \beta_3 * Int.Comm_{ij} * post_i + X'_{ij}\beta_x + \varepsilon_{ij}$

In equation (1) above, $\hat{Y}_{ij}$ represents the outcome for the i^{th} patient in the j^{th} community, $Int.Comm_{ij}$ is a dummy variable that signifies whether the i^{th} patient is from intervention community j, $post_i$ is a dummy variable for whether the outcome measured for the i^{th} patient is measured at the pre or post period, and $Int.Comm_{ij} * post_i$ is an interaction term whose coefficient β_3 captures the QIO program's impact on the outcome. Specifically, the β_3 coefficient represents the association between the QIO intervention and outcomes after adjusting for differences in outcomes at baseline as well as community characteristics and patients' demographic and health characteristics represented by the vector $\boldsymbol{X'_{ij}}$. The sections below highlight how we plan to construct the outcome variables for each theme and key issues related to the regression models for the care transitions and chronic kidney disease themes.

The impacts model described in equation (1) provides us with the national average impact of the QIO program on Medicare beneficiaries. The model includes all relevant beneficiaries attributed to all communities, with all beneficiaries weighted equally; as a result, the model implicitly weights QIOs that recruited larger communities more heavily than QIOs with smaller communities. This implicit weighting is appropriate for evaluating the nationwide impacts of the QIO program (we discuss alternative weighting schemes in the next chapter).

3. Trend Analyses

For the core prevention theme, we plan to compare quarterly estimates of rates of preventive care services provided to eligible Medicare beneficiaries by PPs and NPs. We view this comparison as a descriptive trend analysis rather than a full impact analysis. As described in Chapter I, the NPs for the prevention theme had to meet the same relatively restrictive eligibility criteria as the PPs (that is, having implemented a CCHIT-certified EHR and using the EHR to perform care management for at least one chronic condition), but were not required to commit to improvement on the prevention measures. NPs also receive technical assistance from QIOs on EHR use.

NPs thus do not reflect the counterfactual condition of practices identical to PPs but not receiving any QIO assistance. Rather, any observed differences between the two groups represent the combined effects of the underlying and unobservable motivation or ability of the PP practices that were willing to commit to improvement targets, and of any additional assistance that QIOs provided to PPs beyond what was provided to NPs.

B. COMMON DESCRIPTIVE ANALYSES ACROSS THEMES AND SUBTHEME COMPONENTS

We will present basic descriptive analyses for nearly all themes and subtheme components, whether or not we can do impact analyses, and we provide a brief overview of these here to avoid repeated explanations of common approaches. In the discussion below, we will then only

mention features of the descriptive analyses that are unique to a particular theme or subtheme component; otherwise the reader should assume that we will complete and present the general descriptive analyses described in this section.

In general, we will describe characteristics of the providers (such as hospitals, nursing homes, physician practices, and so on) and of the communities (for care transitions and CKD) who work with the QIOs and, where possible, the characteristics of the providers and communities who do not. Descriptive results may include baseline levels of outcomes, racial/ethnic composition of patients/local residents, provider size, and ownership type. We will examine results at state, regional, and national levels and provide summary findings.

We will also analyze changes in outcome measures from the baseline to the follow-up period. The baseline period is shortly before the start of the 9th SOW in August 2008; for example, depending on data availability and the specific outcome measures, this might be the calendar year from August 1, 2007 to July 31, 2008. The follow-up period will generally be the most recent year available. Again, we will review disaggregated results by region, urbanicity, and provider or regional characteristics (such as areas or providers with high and low proportions of racial/ethnic minority residents) in order to present summary findings.

C. THEME-BY-THEME ANALYSIS PLAN

1. Beneficiary Protection Theme--Assisting Hospitals with RHQDAPU

QIOs' technical assistance to hospitals for the Reporting Hospital Quality Data for Annual Payment Update (RHQDAPU) program, a subtask within the beneficiary protection theme, is one of the subtasks for which an impacts analysis is not possible because of the absence of a valid comparison group. All QIOs are charged with helping all the hospitals in their state with this task. The RHQDAPU is tied to hospitals' Medicare reimbursement; in federal fiscal year 2007 nearly 95 percent of all eligible hospitals successfully participated in the reporting program and received the full payment update for fiscal year 2008 (Centers for Medicare & Medicaid Services 2009). However, because the subtask represents a substantial part of one of the main themes in the 9th SOW, it is important that the evaluation document QIOs' assistance and hospitals' perceptions of that assistance.

Under the RHQDAPU initiative, originally initiated by the Medicare Modernization Act of 2003 and revised by the Deficit Reduction Act of 2005, hospitals that do *not* submit quality data to the Hospital Compare database experience reductions in their Medicare Annual Payment Updates. RHQDAPU is the major reason why the Hospital Compare database is well-populated, providing comparative quality data for hospitals nationally, available to the public online. In turn, public reporting has been linked to improved outcomes (Paez et al. 2009).

To describe the role of the QIOs, we plan to use data from the diaries of contacts between QIOs and RHQDAPU participating hospitals that QIOs are required to submit. These diaries document the technical assistance provided by each QIO to these hospitals with dates and summaries of each contact. To identify the perceived value of these interactions from the hospitals' perspectives, the hospital survey includes items that ask if RHQDAPU was a reason for any in-person or phone meetings with the QIO during the 9th SOW, and if so, how valuable

to the hospital this type of meeting was (highly valuable, of moderate value, not valuable). To identify the perceived value of this work from the QIO's perspective, we will ask the QIO directors on our site visits to describe how important they perceive this part of their work to be (and why), and to tell us what they think would have been the result over the past year if they had not assisted hospitals as they did. We will also speak with the CMS theme leader relevant to this work to tap his or her knowledge of what assistance has been provided and any issues, to ensure an accurate description of this work in the final evaluation report. Table III.1 provides an example of how the descriptive data from the survey may be displayed in the final report.

Contrasting the perceptions of QIOs' assistance held by hospitals that were reporting to RHQDAPU with those held by hospitals that were not might provide additional information on QIOs' efforts. However, as noted above, over 95 percent of eligible hospitals nationwide were already reporting in 2007, and the percentage may be even higher by 2008 or 2009, so there may not be enough non-reporting hospitals to conduct such a comparison. We will assess the percentages of hospitals that are not reporting both nationwide and at a state level to see if we can produce such tabulations.

2. Patient Safety Theme

As described in Chapters I and II, the patient safety theme consists of several discrete components that are not closely related. This section will present the details of quantitative analyses—both impact and descriptive—that are specific to each component of the patient safety theme. We first present the outcome measures that will be analyzed, and then the impact and descriptive analyses.

TABLE III.1

PERCEIVED VALUE OF RHQDAPU MEETINGS AMONG SURVEYED HOSPITALS WITH AT LEAST ONE SUCH MEETING (PERCENTAGE OF HOSPITALS)

	High Value (n=)	Moderate Value (n=)	No Value (n=)
All Hospitals with at Least One RHQDAPU Meeting			
Bed Size Category			
<50			
50-99			
100-249			
250 +			
Urban/Rural			
Urban			
Rural			
Hospital System Affiliation			
Affiliated			
Unaffiliated			

a. Outcome Measures

Most of the outcome measures for the patient safety theme analyses will come from secondary data sources. However, some outcome measures will come from a national survey of hospitals and nursing homes that we describe briefly here. A full description of the survey is in the PRA supporting statement for the evaluation (Kovac et al. 2010).

From May through August 2010, we will conduct a computer-assisted telephone interview (CATI) survey of 1,250 hospitals and 1,250 nursing homes, including facilities that work with QIOs and those that do not. The respondents will be the facilities' quality improvement directors. The key survey topics, which will be asked of all providers whether or not they formally worked with their local QIO,[5] include—level and types of contacts with the QIO, perceived value of QIO services, quality initiatives on the same topics as in the 9th SOW, non-QIO quality initiatives, sources for quality information, and barriers to further improvement. Our sampling design aims to (1) support national estimates of survey responses, and (2) support regression discontinuity (RD) impact estimates for specific survey items. For the hospital survey, we will stratify hospitals by whether their baseline Surgical Care Improvement Project (SCIP) appropriate care measure (ACM) scores are above or below the cutoff and allocate half of the sample to each stratum. We will then develop a second set of strata within each of these primary strata to allow oversampling of the hospitals "near" the cutoff, defined in terms of percentiles of each stratum ranked by the SCIP ACM scores. Our definition of "near" will depend on the distribution of hospital scores (which we just obtained on September 9, 2009) above and below the cutoff. In each explicit stratum, we will select an equal probability sample of hospitals. To improve the distributional characteristics of the samples, within the explicit strata we will also use implicit stratification of the following variables: CMS regions, location in an urban or rural area, for-profit status, and number of beds (quintiles of the distribution) (Chromy 1978).[6]

The target populations for the survey consist of hospitals or nursing homes that are certified to provide Medicare-Medicaid services and thus listed in the CMS Provider of Services (POS) file. The approximate population sizes are 4,500 hospitals and 16,000 nursing homes.[7] For each survey, the sampling frame will be constructed from the most recent version of the POS to include variables needed for sample selection and the computation of weights. We will exclude any providers that are no longer in service. Our goals are completed surveys from 1,250 hospitals and 1,250 nursing homes. Assuming a 70 percent response rate, we will draw samples of 1,785

[5] We will ask the first two questions of all facilities, even those not formally working with the QIOs. In the 8th SOW, some health care providers not officially participating with QIOs still had contact with QIOs (Clarkwest et al. 2009; Narayanan et al. 2008).

[6] Implicit stratification, also known as sequential random sampling or Chromy's method, is a sampling method in which the units to be sampled are first sorted by key covariates, and then randomly sampled in a way that evenly distributes covariate values throughout the sample and avoids excessive concentrations of any particular value.

[7] This number of hospitals includes roughly 950 Critical Access Hospitals (CAHs). Our understanding is that even QIOs that have not been awarded the special Rural Focused Patient Safety Projects are working with CAHs, and we thus plan to include CAHs in the sample universe. As noted earlier, we are still learning the details of the Rural Focused Patient Safety Projects. If our understanding is incorrect, we will exclude CAHs, and the sample universe of hospitals will number roughly 3,550.

hospitals and 1,785 nursing homes. Survey results will be appropriately weighted to yield national estimates.

Table III.2 presents the list of outcome measures for the evaluation, by theme component. The majority of the measures are also used by CMS in its assessment of QIOs' contract performance, with the exception of the nursing homes in need component. Although the NHIN component focuses on improving rates of pressure ulcers and use of physical restraints, it is also designed to broadly improve management and care in troubled facilities. We will thus construct additional outcome measures that are composites of all deficiency items in the areas of "Resident Behavior and Facility Practices" (5 items) and "Quality of Care" (25 items). We also note that we will not have access to data from the CDC's National Healthcare Safety Network Multi-drug Resistant Organism Module (NHSN-MDRO) for the MRSA subtheme component. Appendix C contains additional details on the specifications for the construction of some of these outcome measures and of the necessary data sources.

b. Impact Analyses

We will conduct impact analyses for the following three components of the patient safety theme:

1. Physical Restraints–Nursing Homes
2. Pressure Ulcers–Nursing Homes
3. SCIP/HF–Hospitals

For each of these components, QIOs were required to select at least 85 percent of their PPs from a designated list of providers. Providers were included on the list based on having baseline levels of quality of care that were worse than an explicit cutoff level. Because the probability of being selected as a PP varies substantially depending on which side of the cutoff a provider is on, we will estimate impacts for those components using a regression discontinuity design. Section A.1 of this chapter described the basics of the RD design and its fundamental reliance on knowledge of the selection measure and cutoff. The general design will be identical for all components, but the particular selection measure differs for each. Table III.3 presents a short description of how the list of targeted providers was created, the selection variable (x_i), and cutoff (x_0) level for each component.

As with a random assignment experiment, covariate adjustment is not necessary in a well-specified RD model. Conditional on the selection measure, treatment status is uncorrelated with any other baseline covariate. However, as with a randomized experiment, inclusion of covariates that are associated with the outcome can be useful for improving the precision of impact estimates. Tables III.4 and III.5 presents the covariates and control variables we will consider including in our analyses.

TABLE III.2

PATIENT SAFETY THEME OUTCOME MEASURES FOR DESCRIPTIVE AND IMPACT ANALYSES

Component and Outcome Measure	Data Source
Nursing Homes	
Pressure Ulcers	
Percentage of high-risk long-stay residents who have pressure sores	MDS
Physical Restraints	
Percentage of long stay residents who were physically restrained	MDS
General	
Frequency of contact with QIO	Nursing HomeSurvey
Presence of internal quality improvement efforts in specific measures[a]	Nursing HomeSurvey
Whether has ever analyzed performance data in specific measures to identify underlying causes ("root cause analysis")	Nursing HomeSurvey
Whether has undertaken various quality improvement strategies for specific measures[a]	Nursing HomeSurvey
Whether reports adequate leadership and resources for quality improvement in specific measures[a]	Nursing HomeSurvey
Whether faces barriers to quality improvement	Nursing HomeSurvey
Hospitals[b]	
SCIP/HF	
Surgery patients on a beta blocker prior to arrival who received a beta blocker during the perioperative period	Hospital Compare
Prophylactic antibiotic received on time (INF-1)	Hospital Compare
Percent who received prophylactic antibiotics recommended for their specific surgical procedure (INF-2)	Hospital Compare
Prophylactic antibiotics discontinued within 24 hours after surgery end time (INF-3)	Hospital Compare
Cardiac surgery patients with controlled 6 a.m. postoperative serum glucose (INF-4)	Hospital Compare
Surgery patients with appropriate hair removal (INF-6)	Hospital Compare
Surgery patients with recommended VTE prophylaxis ordered (VTE-1)	Hospital Compare
Surgery patients who received appropriate vte prophylaxis within 24 hours prior to surgery to 24 hours after surgery (VTE-2)	Hospital Compare
Heart failure patients with left ventricular systolic dysfunction without ACEI and ARB contraindications who are prescribed ACEI/ARB at discharge (HF 3)	Hospital Compare
Risk-adjusted 30-day heart failure mortality rate	PIHOEM[c]
Postoperative sepsis (PSI-13)	PIHOEM[c]
Postoperative wound dehiscence in abdominopelvic surgical patients (PSI-14)	PIHOEM[c]
SCIP/HF and MRSA	
Frequency of contact with QIO	Hospital Survey
Receipt of educational materials or tools from QIO	Hospital Survey
If received materials, perceived value	Hospital Survey
Presence of internal quality improvement efforts for specific measures[d]	Hospital Survey
Presence of internal quality improvement efforts in specific measures[d]	Hospital Survey
Whether has ever analyzed performance data in specific measures to identify underlying causes ("root cause analysis")	Hospital Survey

TABLE III.2 (*continued*)

Component and Outcome Measure	Data Source
Whether has undertaken various quality improvement strategies for specific measures[d]	Hospital Survey
Whether reports adequate leadership and resources for quality improvement in specific measures	Hospital Survey
Whether faces barriers to quality improvement	Hospital Survey
Prescription Drug Safety	
Drug-drug interactions[e]	PIM/DDI[f]
Potentially inappropriate medications[e]	PIM/DDI[f]
NHIN	
Percentage of high-risk long-stay residents who have pressure sores	MDS
Percentage of long stay residents who were physically restrained	MDS
Deficiencies in resident behavior and facility practices[g]	OSCAR
Deficiencies in quality of care[h]	OSCAR

Note: MDS=Minimum Data Set, contains data submitted by nursing homes on patients' clinical conditions
Nursing Home Survey and Hospital Survey are surveys to be fielded by Mathematica as part of this evaluation.
Hospital Compare=CMS' publicly reported data on hospitals' quality performance
VTE=venous thromboembolism
ACEI/ARB=angiotensin converting enzyme inhibitor and angiotensin II receptor blocker drugs
PIHOEM=Production and Implementation of Hospital Outcome and Efficiency Measures
MRSA=methicillin-resistant Staph aureus
PIM/DDI=potentially inappropriate medications/drug-drug interactions
OSCAR=Online Survey, Certification, and Reporting database
PSI=Patient Safety Indicator—a set of measures based on claims data developed by the Agency for Healthcare Quality and Research (AHRQ) for inpatient hospital safety

[a]Specific measures include—physical restraints, pressure ulcers, influenza vaccination, pneumococcal vaccination, urinary tract infections, urinary catheter use, depression or anxiety, moderate to severe pain, patient mobility, weight loss, and help with daily activities.

[b]INF-1 through INF-6, VTE-1 and VTE-2, and HF-3 are the abbreviated names for specific Hospital Compare measures (for example, measures of apppropriate selection and timing of perioperative prophylactic antibiotics, glycemic control in post-cardiac surgery patients, appropriate preoperative hair removal, appropriate ordering and perioperative receipt of VTE prophylaxis, and appropriate drug therapy in patients with systolic heart failure)

[c]PIHOEM is a project to produce an expanded set of outcome measures for the Hospital Compare dataset, performed by Mathematica under separate contract to CMS.

[d]Specific measures include the SCIP/HF measures listed above in footnote b, as well as MRSA infection and transmission rates.

[e]As determined by specific criteria for which drugs may interact with each other and which drugs are considered potentially inappropriate in elderly patients.

[f]The PIM/DDI indicators are created from Part D claims data for the QIO program by CMS data contractors; these data reside on the SDPS/QIONet data system.

[g]A composite of 5 items on the state survey form (F0221-F0225) and recorded in the OSCAR database that contains data from state survey agencies on nursing home deficiencies.

[h]A composite of 25 items on the state survey form (F0309-F0333) and recorded in the OSCAR database.

TABLE III.3

SELECTION MEASURES FOR PATIENT SAFETY IMPACT ANALYSES USING REGRESSION DISCONTINUITY DESIGN

Description of Target List Criteria	Selection Variable	Cutoff
Pressure Ulcers (PrU)–NH		
Nursing homes that during 2 out of the 3 quarters from 2006 Q4 through 2007 Q2 had results 14 or more percentage points away from the goal of no more than 6 percent of high-risk long-stay (HRLS) residents having pressure sores.	NH's HRLS PrU rate for 2nd highest of the 3 quarters	20 percent
Physical Restraints (PhyR)–NH		
Nursing homes that during 2 out of the 3 quarters from 2006 Q4 through 2007 Q2 had results 8 or more percentage points away from the goal of no more than 3 percent of long-stay (LS) residents being physically restrained.	NH's LS PhyR rate for 2nd highest of the 3 quarters	11 percent
SCIP/HF–Hospitals		
Hospitals that had an Appropriate Care Measure (ACM) score 30 points or more below the Achievable Benchmarks of Care rate for the two most recent quarters.	Hospital's best (highest) ACM score for the 2 quarters	62.5 percent (Q4 2006) 64.0 percent (Q1 2007)

Note: NH=nursing home.
SCIP/HF=Surgical Care Improvement Project/Heart Failure

TABLE III.4

POTENTIAL COVARIATES FOR IMPACT ANALYSES OF PATIENT SAFETY IN NURSING HOMES (PHYSICAL RESTRAINTS, PRESSURE ULCERS)

Variable	Data Source
County-Level Characteristics	
MDs per 1,000 Population	ARF; Census
RNs per 1,000 Population	ARF; Census
Per Capita Income	ARF; Census
Located in a Metropolitan Area	ARF; Census
Percentage of Population	
Age 0 to 19	ARF; Census
Age 65 and over	ARF; Census
With 4 years college	ARF; Census
Uninsured	ARF; Census
At or below poverty level	ARF; Census
Hispanic	ARF; Census
Black	ARF; Census
Provider-Level Characteristics	
Ownership Type	
For Profit, Individual, or Partnership	NH Compare
Government	NH Compare
Non-Profit, Corporation	NH Compare
Non-Profit, Church	NH Compare
Non-Profit, Other	NH Compare
Large Nursing Home	NH Compare
Located within a Hospital	NH Compare
Resident and Family Councils Present	NH Compare
Baseline Quality Measures	NH Compare
Physical restraint prevalence[a]	NH Compare
Pressure ulcer prevalence[b]	NH Compare
Improvement in ambulation	NH Compare
Improvement in pain interfering with activity	NH Compare
Improvement in transferring	NH Compare

[a]To be used as covariate in RD models estimating impacts on pressure ulcers.
[b]To be used as covariate in RD models estimating impacts on physical restraints.

TABLE III.5

POTENTIAL COVARIATES FOR SCIP/HF HOSPITAL
PATIENT SAFETY IMPACT ANALYSES

Variable	Data Source
County-level Characteristics	
MDs per 1,000 Population	ARF; Census
RNs per 1,000 Population	ARF; Census
Per Capita Income (logarithm)	ARF; Census
Located in a Metropolitan Area	ARF; Census
Percentage of Population	
Age 0 to 19	ARF; Census
Age 65 and Over	ARF; Census
With 4 Years College	ARF; Census
Uninsured	ARF; Census
At or Below Poverty Level	ARF; Census
Hispanic	ARF; Census
Black	ARF; Census
Provider-level Characteristics	
Large Hospital	Hospital Compare
Acute Care Hospital	Hospital Compare
Ownership Type	Hospital Compare
Non-Profit, Church	Hospital Compare
Non-Profit, Other	Hospital Compare
Non-Profit, Private	Hospital Compare
Government	Hospital Compare
Baseline Outcomes[a]	Hospital Compare
Inf Composite[b]	Hospital Compare
VTE Composite[c]	Hospital Compare
Heart Attack ACM Composite[d]	Hospital Compare
Pneumonia ACM Composite[e]	Hospital Compare

Note: SCIP/HF=Surgical Care Improvement Project/Heart Failure
ARF=Area Resource File
ACM=Appropriate Care Measure

[a]To be used as additional covariates in regressions where change in that measure is not the outcome of interest. In RD regressions where change in that measure is the outcome of interest, the baseline indicator is the selection variable and is a required component of the RD regression, not an additional regressor.

[b]Average of rates of infections measures—prophylactic antibiotic received within one hour prior to surgical incision, prophylactic antibiotic selection for surgical patients, prophylactic antibiotics discontinued within 24 hours after surgery end time (48 hours for cardiac patients), cardiac surgery patients with controlled 6 a.m. postoperative serum glucose, and surgery patients with appropriate hair removal.

TABLE III.5 (*continued*)

[c]Average of rates of venous thromboembolism (VTE) measures—surgery patients with recommended VTE prophylaxis ordered, and surgery patients with receipt of appropriate VTE prophylaxis within 24 hours prior to surgery to 24 hours after surgery.

[d]Average of rates of provision of aspirin at arrival, prescription of aspirin at discharge, appropriate prescription of ACE inhibitor or ARB at discharge, timely provision of beta blocker at arrival, and prescription of beta blocker at discharge for heart attack patients.

[e]Average of rates of oxygenation assessment, pneumococcal vaccination, and timely provision of initial antibiotic after admission for pneumonia patients.

Weighting. Our impact estimates will be produced using a pooled sample for observations across all QIOs. Our primary impact of interest is the average impact per beneficiary served by the program nationwide. Consequently, for our nationwide impact estimates, we will weight observations in ways that treat all beneficiaries equally. For outcomes with beneficiaries as the unit of observation—such as hospital pressure ulcer outcomes that are calculated from Medicare claims data—each beneficiary will have equal weight. In cases where the provider is the unit of observation, providers will be weighted by the number of beneficiaries they serve.

When all beneficiaries served are weighted equally, greater weight is implicitly given to larger states than to smaller ones. For the purposes of a nationwide impact estimate, this is appropriate. However, for other purposes it is more appropriate to weight each state equally. The QIO program is run independently in different states and practices vary across states. One important task of the evaluation is to identify which QIO practices are associated with more positive impacts. And each QIO is equally important for answering that question. Consequently, for those analyses of mechanisms, we will give proportionally more weight to observations from smaller states, such that each state QIO receives equal weight in the analyses.

Minimum Detectable Impacts (MDIs). The precision of the impact estimates will vary across outcomes depending on sample sizes and will also depend in part on factors—most importantly bandwidth choice—that will be determined after the follow-up data are collected. For a given sample size, MDIs are larger with an RD design than with a randomized controlled trial (RCT) because the variance of the impact estimate is greater due to the correlation between the treatment indicator and the selection variable, which must be controlled for in the regression estimates. MDIs are calculated based on the MDI values for an RCT study, inflated by the linear RD "design effect" (Schochet 2008):

$$(2)\ \textit{Design Effect} = \frac{(1-R_1^2)}{(1-R_0^2)} * \frac{1}{(1-R_{T|Score}^2)}$$

where R_1^2 is the regression *R*-squared value for the RD impact model, R_1^2 is the regression *R*-squared value under an experimental design with the same covariates, and $R_{T|Score}^2$ is the *R*-squared value when the treatment indicator, *T*, is regressed on the selection measure (and an intercept). The first ratio in the design effect is essentially 1, since the same explanatory variables would be used in either an RD or a random assignment experimental design. We treat it as 1 in our calculations. The second ratio in the equation is what drives the design effect.[8] The

[8] The term $1/(1-R_{T|Score}^2)$ appears because, by construction, treatment status and assignment scores are correlated in the RD regression model, but not in the random assignment experimental model. This correlation tends to be quite large in absolute value, which substantially increases the variance estimates under the RD design. Intuitively, the treatment effect is net of the score variable. Thus, the substantial collinearity between the treatment status and score variables reduces the information contained in the treatment status variable, which lowers the effective sample size for analysis.

MDI of an RD design is found by multiplying the MDI of an RCT (with the same sample size) by the square root of the RD design effect. As explained further in Appendix B, a key decision in an RD analysis is the "bandwidth" to use, which is essentially a decision on how far from the cutoff observations can be and still be included in the analysis. Wider bandwidths mean larger samples and greater power, but may also increase the risk of biased estimates.

In Table III.6 we present illustrative calculations for MDIs using four different bandwidths in order to provide a sense for the extent to which precision declines as bandwidth narrows. The first of the four is the "full" bandwidth—that is, retaining all observations. Moving from the "full" to the "wider" bandwidths removes the tails of the distributions. This tends to have fairly modest effects on MDIs, and in their simulations, Lee and Lemieux (2009) find that the greatest reduction in bias comes from removal of the tails. Once the bandwidth starts entering the thicker parts of the distribution, sample size and precision fall much more rapidly, while there is likely to be progressively less concomitant improvement in unbiasedness. In the table, the italicized "Narrower" rows contain the bandwidth that best reflects our expectations for the sample we will use. In each case the bandwidth is less (and generally much less) than half of the original range of the selection variable. For instance, nursing home pressure ulcer rates range from 0 to 100 percent, but our "narrower" bandwidth includes only observations with values between 14 and 26 percent. Roughly 35 percent of the sample falls within that range. The MDIs for the bandwidths we anticipate using reflect moderate effect sizes, between 0.35 and 0.45 standard deviations. As noted, those could be reduced nontrivially if we were to use somewhat wider (though still highly restricted) bandwidths.

d. Patient Safety Theme Components for Which Impact Analyses Are Not Possible

Impact analyses are not possible for the following patient safety theme components: (1) drug safety, (2) MRSA, and (3) NHIN. The drug safety component requires QIOs to collaborate in an unspecified manner with a wide and unspecified variety of providers. The types of providers who work with QIOs vary greatly from state to state, including physicians, pharmacies, prescription drug plans, hospitals, long term care facilities and community health centers. Thus there is no way to aggregate treatment groups for a national analysis. For MRSA, the outcome data infections will only be available for participating hospitals, who have agreed to give the QIO program access to the results. The signed agreements do *not* extend to Mathematica and we thus have no access to these data.

Due to the timing of the evaluation, for NHIN we will only have data on the one nursing home per state with which each QIO worked in the first year of the 9th SOW. The recruitment process for these nursing homes was highly idiosyncratic—recall from Chapter I that QIOs were to approach one of a small list of selected facilities (CMS' selected focus facilities or SFF list) with particularly severe quality problems; if that facility refused, the QIO followed guidelines to approach another facility, and so on. The sample sizes for NHIN are too small for formal statistical testing, and the nature of the recruiting process undermines the potential to identify a credible comparison group for a formal impacts analysis.

TABLE III.6

MINIMUM DETECTABLE EFFECTS FOR REGRESSION DISCONTINUITY ANALYSES IN THE PATIENT SAFETY THEME

Component Outcome Measure	Bandwidth (Min/Max Values)	Sample Size (PP/Non-PP)	Minimum Detectable Impact (Percentage Points)
Pressure Ulcers – NH			
Pressure Sores (High-Risk Long-Stay)	Full (0, 100)	999 / 14,708	1.3
	Wider (10, 30)	874 / 9,373	1.8
	Narrower (14,26)[a]	*665 / 4,884*	*2.4*
	Very Narrow (18, 22)	318 / 1,395	4.3
Physical Restraints – NH			
Physical Restraints (Long Stay Residents)	Full (0, 100)	1,100 / 14,607	1.0
	Wider (3, 19)	874 / 7,822	1.4
	Narrower (6,16) [a]	*665 / 4,172*	*1.9*
	Very Narrow (9, 13)	318 / 1,498	3.2
SCIP/HF – Hospitals			
Inf composite[b]	Full (-92.5, 7.5)	607 / 2,862	2.4
	Wider (-60, 0)	512 / 2,416	2.7
	Narrower (-50, -10) [a]	*395 / 1,563*	*3.5*
	Very Narrow (-40, -20)	228 / 712	4.7

Note: The minimum detectable impact (MDI) formula used to calculate the above is as follows:

$$2.80 \cdot \sigma_y \sqrt{(1 - R_1^2)\left(\frac{1}{N_T} + \frac{1}{N_U}\right)} \cdot \frac{1}{PP_T - PP_N} \cdot \sqrt{\frac{1-R_1^2}{1-R_0^2} \cdot \frac{1}{1-R^2_{T|Score}}}$$

where σ_y is the variance of the outcome variable, R_1^2 is the R^2 of the RD impact regression, R_0^2 is the R^2 value under an experimental design, $R^2_{T|Score}$ is the R^2 value when the cutoff indicator is regressed on the baseline selection variable, N_T and N_U are the respective sample sizes on and off the J-17 targeting lists (that is, to either side of the cutoff), PP_T and PP_U are the respective proportions of the J-17 and non-J-17 samples that are PPs, and 2.80 is the multiple of the standard error of the impact estimate for a two-tailed significance level of .05 80% power. We calculate predicted values for σ_y based on the values at the end of the 8th SOW, adjusted using the assumption that change in σ_y during the 9th SOW will continue at the same rate as during the 9th SOW. The values of R_0^2 and R_1^2 are assumed to be identical, both equaling the R^2 values obtained in impact estimate regressions for those outcomes in the 8th SOW. Sample sizes above and below the cutoff are calculated using baseline data for nursing hospitals and nursing homes, respectively, were provided by the Oklahoma Foundation for Medical Quality (OFMQ) and the Colorado Foundation for Medical Care (CFMC). The proportions of providers in each group are calculated assuming that 85% of PPs are selected from the J-17 targeting pools, and that within the non-targeting group QIOs do not recruit any PPs outside the "Wider" bandwidth (that is, those with very positive performance and little room to improve at baseline).

TABLE III.6 (*continued*)

---

[a]The italicized row represents our best guess as to the likely bandwidth used in the impact analyses. Other bandwidths are presented for the purpose of comparison.

[b]Inf composite is the hospital's average value on the SCIP Inf-1 (timely provision of prophylactic antibiotic before surgery) and SCIP Inf-3 (timely discontinuance of antibiotic after surgery) measures. Note that this differs from the baseline selection measure—the ACF measure of what proportion of patients received both measures—which requires patient-level data to calculate. The Inf composite outcome variable is an average that can be calculated from hospital-level data.

3. Prevention Disparities

As described earlier, under the prevention disparities theme six QIOs are working to improve diabetes management among underserved minority populations. The QIOs are helping PPs to improve rates of specific recommended diabetes care processes and providing specific types of diabetes self-management education (DSME) to Medicare beneficiaries. We plan for an impact analysis only of the QIOs' work with PPs in this theme. Identification information for the beneficiaries completing the DSME training is not available to us. Therefore, we cannot conduct any kind of beneficiary-level impact analysis. Under its contract as the disparities data contractor MassPro is conducting an analysis of surveys of knowledge of diabetes management that beneficiaries complete prior to and after training, and is also collecting clinical information from medical charts in primary care physicians' offices.

a. Outcome Measures

Table III.7 presents the outcome measures for the prevention disparities impacts study, all of which come from Medicare claims data.

b. Impact Analyses

We estimate the causal effects of QIOs' work with PPs on the three utilization outcomes listed in Table III.7. Those impacts, if any, are most likely to occur directly through QIOs' encouraging PPs to perform more regular testing of beneficiaries with diabetes. As noted in Chapter I, there is only modest overlap between the Medicare patients of PPs and the Medicare beneficiaries attending the DSME programs.[9]

TABLE III.7

OUTCOME MEASURES TO BE USED IN THE IMPACT ANALYSES OF THE PREVENTION DISPARITIES THEME

Type of Outcome	Data Source
Patient received HbA1c testing within the past 12 months	Quarterly Diabetes Analytic Files
Patient received a diabetic eye exam within the past 12 months	Quarterly Diabetes Analytic Files
Patient received lipid testing within the past 12 months	Quarterly Diabetes Analytic Files

Note: Quarterly Diabetes Analytic Files refer to files created from Medicare claims data by CMS data contractors for the QIO program; these files contain binary indicators of receipt of HbA1c testing, diabetic eye exam, and lipid testing among Medicare fee-for-service beneficiaries with diabetes. These files reside on the SPDS/QIONet system.

[9] As described in Chapter I, the QIOs are recruiting beneficiaries to participate in DSME both directly from community organizations and by going through PPs. The beneficiaries recruited directly from the community do not necessarily belong to PP practices. According to personal communications by CMS staff, as of late August 2009, only roughly 30 percent of beneficiaries recruited to DSME are tied to PPs; the remainder were recruited through senior centers, faith-based organizations, and so on, and are not tied to PPs.

Because PP eligibility is determined by meeting strict cutoffs, we also propose estimating impacts in the prevention disparities theme using a regression discontinuity design. Section A.1 describes the role of the selection variable (x_i) and cutoff (x_0). Because there are two selection criteria in the patient disparities theme, there are two distinct potential selection measures that could be used in the RD analyses. We propose conducting separate analyses using each of them, as there is no reason to expect one to produce less valid results than the other. One set of analyses will use the practice's percentile on the baseline diabetes measures as the selection variable (x_i), with the 50th percentile being the cutoff (x_0) at which the discontinuity is measured. The second set will use the proportion of a PP's beneficiaries who are from underserved racial/ethnic populations as the selection variable and 25 percent as the cut-point at which impacts are evaluated. Table III.8 lists the covariates we anticipate including in our models.

Minimum Detectable Impacts. Five-hundred ninety-one participating physician practices have been recruited for the prevention disparities theme, a moderately large number. However, we do not yet have information on the size of the pool of practices that each met the eligibility criteria of having underserved Medicare beneficiaries with diabetes represent over 25 percent of all Medicare beneficiaries with diabetes in the practice. The minimum detectable impacts depend crucially on the size of this pool. The larger the pool, the smaller the proportion of PPs within the pool, and in turn, the smaller the jump in that proportion at the selection threshold. If the pool is relatively small, then MDIs are likely to be reasonable. However, if the pool of eligibles is large, the MDIs could be too large, in which case we will conduct a descriptive trends analysis (Section A.3).

4. Prevention

As described in Section B.1. of Chapter I, providers participating in the prevention theme as either PPs or NPs were required to have implemented and be using a certified electronic health record (EHR) with specific minimum functionality requirements at the onset of the 9th SOW, and had to agree to report EHR-derived results on colorectal and breast cancer screening and flu and pneumonia vaccinations. PPs and NPs could be in either solo or group practices. Solo practitioners had to be full time primary care providers, while in participating group practices at least 40 percent of full time physicians had to be primary care physicians. Analyses for this theme will include descriptive trend analyses (Section A.3) of quarterly rates of preventive care services provided to PP and NP eligible patients from baseline through the most recent quarter of data available.

TABLE III.8

POTENTIAL COVARIATES FOR IMPACT ANALYSES OF
PREVENTION DISPARITIES THEME

Variable	Data Source
County-Level Characteristics	
MDs per 1,000 Population	Area Resource File; Census
RNs per 1,000 Population	Area Resource File; Census
Per Capita Income	Area Resource File; Census
Located in a Metropolitan Area	Area Resource File; Census
Percentage of Population	
Age 0 to 19	Area Resource File; Census
Age 65 and over	Area Resource File; Census
With 4 years college	Area Resource File; Census
Uninsured	Area Resource File; Census
At or below poverty level	Area Resource File; Census
Hispanic	Area Resource File; Census
Black	Area Resource File; Census
Provider Characteristics	
Group size	To be determined[a]
Percentage primary care physicians	To be determined[a]
Percentage minority: practice Medicare panel	Medicare enrollment database (EDB)
Percentage dual eligible: practice Medicare panel	EDB
Individual-Level Characteristics	
HbA1c testing within the past 12 months	Medicare Claims, Part B
Diabetic eye exam within the past 12 months	Medicare Claims, Part B
Lipid testing within the past 12 months	Medicare Claims, Part B
Age	Medicare Claims, Part B
Sex	Medicare Claims, Part B
Race/Ethnicity	Medicare Claims, Part B

Note: Data sources for constructing comparison group practices have not yet been determined. Sources under consideration include—the Community Tracking Survey (CTS) physician survey, non-participating practices practices who signed up for the DOQ-IT program, comparison practices in CMS's evaluation of the Medicare Care Management Performance (MCMP) demonstration, comparison practices in the Medicare Electronic Health Records demonstration (EHRD), physicians participating in the PQRI program.and the Medical Group Management Association (MGMA) survey.

[a]Characteristics of participating practices to come from QIOs' internal Program and Theme Reporting Information Online Tool (PATRIOT) data. Sources of data on characteristics of comparison practices will depend on the data sources used to construct the comparison group of practices. Some sources may have direct information on group characteristics. In other cases we may need to use an algorithm developed by researchers at the Center for Studying Health System Change (Pham et al. 2009) which involves pulling sample physicians' Medicare claims, extracting Tax Identification Numbers (TINs) from these claims, and using the TINs to perform a second Medicare claims pull. This last step identifies both the physicians belonging to the different practices and the beneficiaries treated by those practices. Depending on the data sources involved, an additional step may be necessary of cross-walking old Medicare Unique Provide Identifier Numbers (UPINs) to the new individual physician NPIs before the first Medicare claims pull.

The outcome variables for the prevention theme are indicator variables for whether each beneficiary received mammography, colorectal cancer screening, flu vaccination, and pneumonia vaccination within each measurement year, as identified using relevant Current Procedural Terminology (CPT) and Healthcare Common Procedure Coding System (HCPCS) procedure codes in Medicare claims data. The construction of the dependent variables and analyses on these outcomes will be limited to patient populations for which these measures apply.[10] For example, analyses of mammography will obviously be limited to women. We plan to use the analytic files and Program Progress Reports developed for the QIO contract evaluation for core prevention so that our measures of receipt of appropriate preventive care services are consistent with contract evaluation analyses. Since the 9th SOW is also focused on reduction of health care disparities, we will also conduct additional descriptive trend analyses to see if time trends in the prevention measures for underserved beneficiaries differ from those for non-underserved beneficiaries.

5. Care Transitions

As described in Chapter I, the units of intervention for the care transitions theme are communities, with the definition of community varying across the 14 participating QIOs.[11] Our general strategy is to identify comparison communities that match the treatment communities on key specific hospital-based and local health-system-related characteristics as well as population demographics that suggest that Medicare beneficiaries in the comparison communities experience patterns of health care and likelihood of hospitalization and readmission similar to those of beneficiaries living in intervention communities. It is clearly not possible to find comparison communities that are exactly like intervention communities in all respects except for the absence of the QIO intervention, but we do expect to find communities with very similar patterns of Medicare utilization.

To identify comparison communities, we plan to use the following process: we will first assign all acute care hospitals in the contiguous US to the county (or counties) in which their service area predominantly falls. We will identify all counties that fall within the care transitions intervention communities. Using readmission rates for hospitals in each county and the prevalence of AMI, CHF and pneumonia, we plan to identify potential matches for care transitions intervention counties using cluster analysis. Cluster analysis is a method used to identify groups with similar characteristics (Kaufman and Rousseeuw, 1990). Thus, the cluster analysis will identify potential comparison counties that are similar to care transitions counties in terms of prevalence of AMI, CHF, and pneumonia and readmission rates for these conditions. Once we have this list of potential comparison counties, we will match each intervention county to two comparison communities based on county-level characteristics such as county size,

[10] It is possible that rates of influenza vaccination for fall 2009 through summer 2010, our follow-up period, will be higher than in previous years due to the publicity surrounding the H1N1 flu. However, both PPs and NPs should be equally affected by these events.

[11] For example, Healthcare Quality Strategies, Inc., the New Jersey QIO, identified Virtua HealthSystem in southwestern New Jersey as its intervention community, while FMQAI in Florida identified its intervention community based on zip codes in the Miami area.

percent dual eligible, per capita income, rural or urban status, number of primary care physicians, and other local health characteristics (for example, rates of adult smoking and obesity and air pollution particulate matter days per year). The rationale for matching each intervention community to two comparison counties is to test the robustness of our regression analyses. Specifically, we plan to estimate regression models of outcomes pre- and post-QIO intervention (see section III.A.2 for description of our models) once using one of the comparison communities and a second time using the remaining comparison community. If the impact estimates from both models are similar, we can be relatively confident that our matching process worked well, and the estimates reflect the impact of QIO interventions; in contrast, if the impact estimates from the two models differ considerably, we know that our results are sensitive to the selection of comparison communities and we should be cautious in interpreting the impact estimates.

We plan to use several data sources for matching. To measure readmission rates, we will use data from Hospital Compare on readmissions for AMI, CHF and pneumonia (Mathematica is producing these rates for CMS under a separate contract). County-level measures of disease prevalence and other socio-demographic characteristics may be obtained from the Area Resource File (ARF), the County Health Rankings Project funded by the Robert Wood Johnson Foundation, and the Community Health Status Indicators report available on the US Department of Health and Human Services website.

Outcome and Control Variables. The outcome variables for the care transitions theme include all-cause readmission to a hospital within 30 days of discharge among patients admitted for (1) acute myocardial infarction (AMI), (2) congestive heart failure (CHF), and (3) pneumonia. We plan to measure outcomes separately for each condition as well as pooled across the three conditions. The data for these analyses will come from the Production and Implementation of the CMS Hospital Outcomes and Efficiency Measures (PIHOEM) II project that Mathematica is conducting for CMS under a separate contract. Our regressions will control for a variety of community characteristics (Tables III.9 and III.10).

Minimum Detectable Impacts. Assuming an ICC of 0.04, the projected MDIs for the various outcomes under the care transitions theme are relatively large (Table III.11).[12] For example, we will have power only to detect a 24 percentage point difference in readmission rates for all hospitalizations in unadjusted analyses (other assumptions underlying Table III.11 are shown in Appendix D). As noted above, the reason for the large MDIs is due to the small number of intervention communities for this theme. Previous analyses from the Dartmouth Atlas of health care suggest that small-area variations are important predictors of health care utilizations; as a result, we expect community-level factors to explain much of the variation in outcomes, and it may be more difficult to detect impacts for the care transitions theme.

[12] If the readmission rate is around 20 percent or 0.2, the overall variance in the outcome is 0.16 (from the formula *p*(1-p))*. If we assume that roughly 95 percent of the sites (that is, a range that extends two standard deviations from the mean) have readmission rates falling between 0.15 and 0.25, the between-site standard deviation is 0.025, and the between-site variance is 0.025 squared or 0.00625. The ICC, which is the between-site variance divided by the total variance, is 0.00625/0.16, or about 0.04.

TABLE III.9

POTENTIAL COVARIATES FOR MATCHING INTERVENTION COMMUNITIES WITH CONTROL COMMUNITIES FOR THE CARE TRANSITIONS THEME

Variable	Data Source
Community Characteristics	
Urban vs. rural	Area Resource File
Population Size (total population and population aged 65+)	Area Resource File
Uninsured rate (total population)	Area Resource File
Number of acute care beds per 1,000 persons	Area Resource File
Medicare/Medicaid inpatient discharges	Area Resource File
Total Medicare/Medicaid inpatient days	Area Resource File
Primary care physicians per 1,000 population	Area Resource File
RNs per 1,000 population	Area Resource File
Percentage Hispanic	Area Resource File
Percentage African American	Area Resource File
Percentage below poverty level poverty (total population and population aged 65 and over)	Area Resource File
Per capita income	Area Resource File
Percentage high school education (among 65+ pop)	Area Resource File
Percentage of adults who smoke	County Health Rankings[a]
Percentage of adults who are obese	County Health Rankings
Air pollution – particulate matter days per year	County Health Rankings

[a]Publicly available data from the County Health Rankings project (Robert Wood Johnson Foundation and University of1 Wisconsin Population Health Institute 2010).

TABLE III.10

POTENTIAL COVARIATES FOR IMPACT ANALYSES OF CARE TRANSITIONS THEME

Variable	Data Source
Dependent Variables	
Readmission within 30 days for acute myocardial infarction (AMI), congestive heart failure (CHF), and pneumonia	PIHOEM data
Independent Variables	
Key Independent variables	
Whether intervention or comparison community	QIO
Other Control Variables	
Age	Medicare enrollment database (EDB)
Sex	EDB
Race	EDB
Dual-eligible	EDB
Comorbidities (for example, cancer, dementia, COPD)	Medicare claims (inpatient, outpatient, part B, SNF, home health)
County level-data[a]:	
Percent poverty (among 65+ pop)	ARF
Percent high school education (among 65+ pop)	ARF

Note: PIHOEM=Production and Implementation of Hospital Outcome and Efficiency Measures, a project to produce an expanded set of outcome measures from Medicare claims data for the Hospital Compare dataset, performed by Mathematica under separate contract to CMS.

ARF=Area Resource File

[a]Because we match on these county-level characteristics, we may not need to adjust for them in our regression analyses. We will test whether including control variables for county-level characteristics affects our primary covariates of interest.

TABLE III.11

RANGE OF MINIMUM DETECTABLE IMPACTS FOR CARE TRANSITIONS ANALYSES

Outcome Measures[a]	Unadjusted Analyses	Regression-Adjusted Analyses
AMI-related hospitalizations	0.188	0.168
CHF-related hospitalizations	0.208	0.186
Pneumonia-related hospitalizations	0.188	0.169

[a]Assumes that 20 percent of hospitalized patients are readmitted for all conditions, except CHF, for which we assumed that 27 percent of hospitalized patients are readmitted.

6. Chronic Kidney Disease

The QIOs for Florida, Georgia, Missouri, Montana, Nevada, New York, Rhode Island, Tennessee, Utah, and the U.S. Virgin Islands (VI) are working on this theme. Given the likely extremely small sample sizes for the VI, and the unique features of the health care environment there, we will restrict our main impacts analyses to the 10 states, although we will include the VI in descriptive analyses.

As described in Chapter I, the CKD theme essentially consists of three "clinical focus areas" in which the QIOs are to encourage physicians to: (1) perform annual urinary microalbumin testing for beneficiaries with diabetes, (2) treat beneficiaries with diabetes, early CKD (stages 1-4), and hypertension with ACE-I or ARB drugs, and (3) refer beneficiaries nearing hemodialysis for AV fistula placement. At the same time, the QIOs are to pursue a variety of "community colloboration" activities—assembling and/or sustaining state or local coalitions to work towards systematic quality improvement for CKD prevention and care. The QIOs are to build new partnerships and strengthen existing ones with a wide range of organizations, foster increased involvement by coalition members, and leverage members' resources. The community collaboration activities are intended to lead to statewide changes in the three clinical focus areas.

We plan to conduct impact analyses for only the first and third clinical focus areas (urinary microalbumin testing and AV fistula placement). Our understanding from our discussions with CMS project officers and theme leads is that the Part D data used to measure the prescription of ACE-I/ARB drugs for the second CKD clinical focus area simply do not adequately capture beneficiaries' ACE-I/ARB usage. Rather than go through their Part D prescription drug plan, many beneficiaries get their ACE-I/ARB drugs filled at Walmart (or other such large retailers) through their discounted drug programs. Other beneficiaries may receive free supplies of drugs from pharmaceutical company assistance plans. The general strategy for evaluating these two clinical focus areas is again a comparison of intervention communities against matched comparison communities.

We plan to match all counties in CKD states to counties in non-CKD states. The reason for matching on counties rather than at the state-level is to increase the external validity of our analyses. For example, there are no states that look like Florida overall, but there may be counties in other states that look similar to counties in Florida based on health status, health care utilization patterns and other socio-demographic characteristics. The matching process will rely on a county-level database, such as the ARF, along with other databases with detailed county-level data on health care status and health care utilization (for example, detailed information the Dartmouth Atlas, the County Health Rankings Project funded by the Robert Wood Johnson

Foundation, and the Community Health Status Indicators report available from the US Department of Health and Human Services) merged onto it. In addition, we will aggregate baseline data from the CKD analytic files to measure county-level rates of urinary microalbumin testing and AV fistula placement, and use these as matching criteria.

We will estimate propensity score models for intervention counties to identify the matched comparison counties, as these models allow us to incorporate all of the potential matching covariates into a single index and have well-developed methods for assessing "nearness" and match quality (Dehejia and Wahba 2002). We will use standard techniques, such as nearest-neighbor, interval, or caliper matching, to identify at least one, if not multiple, comparison counties. We will assess the quality of the matching based on the distribution of observable characteristics used for the matching process between intervention and comparison counties. Table III.12 describes various county-level characteristics that we will use for the matching process.

D. COMMON COST-EFFECTIVENESS AND COST-BENEFIT METHODS

A final component of the impact analyses is the translation of results into common measures of clinical benefit and dollars. The findings of the preceding impact analyses will all be in terms of the dependent variables for each theme or subtheme component, for example, differences in the rates of pressure ulcers among nursing home residents or cholesterol testing among primary care patients with diabetes.

To compare diverse impacts of health and health care interventions, a large body of cost-effectiveness literature has developed methods to convert intervention effects into "life years" (LYs) gained, or "quality-adjusted life years" (QALYs) gained. The same literature also generally seeks to express in terms of dollars the net resources or efforts required to achieve these gains in LYs or QALYs (Gold et al. 1996). The various yields of very disparate health care interventions can then all be expressed as "dollars per QALY," allowing comparisons between, say, a program to increase bicycle helmet use among children and a program to reduce falls among nursing home residents, and allowing conclusions to be drawn on which program is more "cost-effective."

Cost-benefit studies are related to cost-effectiveness studies. Both try to convert health effects of different interventions into a common metric; however, where cost-effectiveness studies state program effects as LYs or QALYs and report study results as dollars per QALY, cost-benefit analyses attempt to express program *effects* in dollars and thus report study results as dollars per dollar (a cost-benefit ratio) or net differences in dollars (net costs or benefits). Obviously, attaching dollar figures to life years gained can be difficult and controversial. Alternatively, some studies only consider health intervention effects on payers' or insurers' health expenditures, which may also be controversial (for example, shortening peoples' life spans may in some instances actually save money for health care payers).

Assuming a theme yields favorable impacts on the main outcome variable, we will conduct literature searches to translate these impacts into effects on QALYs. Using clinical trial and epidemiologic data along with simulation methods, many studies have extrapolated intermediate clinical physiologic outcomes, such as blood pressure or cholesterol lowering, into effects on

TABLE III.12

POTENTIAL COMMUNITY CHARACTERISTICS FOR MATCHING INTERVENTION COMMUNITIES WITH COMPARISON COMMUNITIES FOR THE CHRONIC KIDNEY DISEASE THEME

Variable	Data Source
Baseline county-level rate of microalbumin testing	Chronic Kidney Disease-1 Analytic Dataset[a]
Baseline county-level rate of AV fistula placement	Chronic Kidney Disease-3 Analytic dataset[a]
Baseline county-level rate of diabetes among adults	Centers for Disease Control and Prevention
Urban vs. rural	Area Resource File
Population Size (total population and population aged 65+)	Area Resource File
Uninsured rate (total population)	Area Resource File
Number of acute care beds per 1,000 persons	Area Resource File
Medicare/Medicaid inpatient discharges	Area Resource File
Total Medicare/Medicaid inpatient days	Area Resource File
Primary care physicians per 1,000 population	Area Resource File
Nephrologists per 1,000 population	Area Resource File
RNs per 1,000 population	Area Resource File
Percentage Hispanic	Area Resource File
Percentage African American	Area Resource File
Percentage below poverty level poverty (total population and population aged 65 and over)	Area Resource File
Per capita income	Area Resource File
Percentage high school education (among 65+ pop)	Area Resource File
Percentage of adults who smoke	County Health Rankings[b]
Percentage of adults who are obese	County Health Rankings
Intensity of medical care services provided to Medicare beneficiaries by Hospital Referral Region (HRR) and/or Health Service Area(HSA)	Dartmouth Atlas

[a] Chronic Kidney Disease-1 and Chronic Kidney Disease-2 Analytic Datasets refer to files created from Medicare claims and CMS 2728 data by CMS data contractors for the QIO program; these files contain binary indicators of receipt of urine microalbumin testing, and dialysis through an AV fistula among Medicare fee-for-service beneficiaries with diabetes. These files reside on the SPDS/QIONet system.

[b]Publicly available data from the County Health Rankings project (Robert Wood Johnson Foundation and University of1 Wisconsin Population Health Institute 2010).

morbidity and life expectancy. For each theme and subtheme component that has significant impacts, we will conduct literature searches to assess conversion of these impacts into QALYs and dollars. A major challenge to this effort is that many 9th SOW outcomes are measures of the *processes* of care, such as whether or not a certain test was done, or a drug was prescribed when indicated. However, though these steps in care are necessary, they are not sufficient; just because cholesterol, hemoglobin A1c, or retinal backgrounds are tested does not mean that risk factors are better controlled or disease progression slowed. As we have been doing for a series of memos to assist CMS in its planning for the QIO Programs' 10th SOW (see for example, Wrobel and Maxfield 2009a and 2009b; Gimm et al. 2009; Schmitz et al. 2009), we will make a range of reasonable assumptions where possible on how completion of various process of care measures might lead to favorable effects on avoidance or delay of complications.

We will also assess the possibility of doing cost-benefit analyses that require conversion of QIO program impacts into dollar figures. We will conduct literature searches for studies that have developed and attached dollar figures to beneficial effects from health quality interventions such as those in the 9th SOW. The number of studies is likely to be low, limiting the number of cost-benefit analyses we can do. We will also consider restricting our attention to savings to the Medicare program, as was done for many of the 10th SOW memos, and thus comparing Medicare program savings due to the 9th SOW against Medicare expenditures on the 9th SOW.

Finally, we will combine the QALYs and dollar savings from each theme and subtheme with information on the costs devoted to the 9th SOW. Using CMS's Financial Information and Vouchering System, which tracks QIO contract budgets and expenditures, we will present results on the number of Medicare QIO dollars expended to achieve a QALY and the cost-benefit ratio for the QIO program.

E. CONCLUDING REMARKS

Table III.13 provides a summary overview of the analytic approaches for each theme and subtheme component. The variety of themes and subtheme components in the 9th SOW means that each essentially requires its own analytic approach. For some components, impact analyses are not possible and we will perform descriptive analyses. For the other components, the two main approaches will either be a regression discontinuity design or propensity score matching.

Unfortunately, uncertainty will remain over the validity and accuracy of several of the impact estimates, despite the use of sophisticated statistical and econometric techniques. Regression discontinuity is a stronger design than propensity score matching and is much more likely to yield accurate estimates of program impacts. However, regression discontinuity is only feasible in four subtheme components. Only the weaker matched comparison group approach is possible for the two themes and subtheme components. In fact, statistical power may turn out to be excessively low for the regression discontinuity analyses for the prevention disparities themes; if so, we may have to consider descriptive trend analyses as well. We await further information on sample sizes and pools of eligible providers for this component. Chapter V discusses how the challenges facing the evaluability of the 9th SOW may hold lessons for the design of the 10th and future SOWs.

TABLE III.13

SUMMARY OF ANALYTIC APPROACHES TO 9TH SOW THEMES AND SUBTHEME COMPONENTS

Theme/Subtheme	Quantitative Impact Analyses	Descriptive Statistics	Quantitative Mechanism Analyses	Qualitative Analyses
Patient Safety Theme				
Hospital SCIP/HF	RDD	x[a]	x	x
Hospital MRSA	--	x[b]	--	x
Nursing Home Pressure Ulcer	RDD	x[a]	x	x
Nursing Home Physical Restraint	RDD	x[a]	x	x
Nursing Homes in Need	--	x	--	x
Drug Safety	--	x	--	x
Prevention Theme				
Working with PPs on cancer screening and vaccinations	--	x	--	x
Prevention—Disparities Theme				
Working with PPs	RDD[c]	x[c]	x	x
Beneficiary DSME	--	--	--	x
Care Transitions Theme				
Working with intervention communities	MCG	x	x	x
Prevention—CKD Theme				
Urinary microalbumin testing	MCG	x	x	x
Treatment with ACE-I/ARB drugs	--	x	--	x
AV fistula	MCG	x	--	x

[a]From both the hospital and nursing surveys, and the other outcome data.
[b]From the hospital survey
[c]If the regression discontinuity design is underpowered, we will consider a descriptive analysis.

ACE-I/ARB= angiotensin converting enzyme inhibitor/angiotensin II receptor blocker drugs.
AV= arteriovenous
CKD= chronic kidney disease.
DSME= diabetes self-management education.
MCG=Matched comparison group design
MRSA= Methicillin-resistant Staphylococcus aureus.
PP= participating provider.
RDD=Regression discontinuity design.
SCIP/HF=Surgical Care Improvement Project/Heart Failure

IV. MECHANISMS ANALYSES

The evaluation also seeks to describe and understand the activities and interventions undertaken by the QIOs and the environment in which they operate, and to identify which types of QIO program are most effective, and for whom—what we have called the "mechanisms" of QIO action. Clearly, no evaluation of the QIO program would be complete without a full documentation of QIOs' interventions and activities, and previous studies of the QIO program have pointed out this need. The study by NORC of the 7th and 8th SOW of the QIO Program attempted to construct an inventory of QIO activities through interviews with CMS staff, a review of some QIO internal documents (NORC was granted only limited access to the internal SDPS/QIONet data system), and visits to QIOs' public websites (Sutton et al. 2007). The report concluded that "there were very few details on the technical assistance that was offered or specific interventions QIOs implemented," and that "…for the overwhelming majority of tasks [under the SOWs], large gaps exist in the data…efforts to locate details…often proved futile." A full description of the contexts and environments in which QIOs function is also essential. For example QIO interventions may have a better chance of effectiveness in states in which provider leaders are more interested in quality improvement and there are qualified/educated staff to support quality improvement.

Finally, synthesizing the above analyses to understand which types of QIO program appeared to work best, for which types of providers, and under what circumstances will provide important policy information for decision makers on how to improve and better target the national QIO program to achieve maximum effectiveness. This chapter describes our approaches to these efforts.[1]

A. GATHERING DATA ON QIOS' ACTIVITIES

To describe QIOs' activities in the 9th SOW, we will pursue several data collection efforts: (1) a nationwide survey of QIO staff, (2) discussions with partner organizations working with QIOs on the care transitions and CKD subnational themes, (3) a review of key QIO documents, and (4) focus groups of beneficiaries participating in the prevention disparities subnational theme. Volume II contains copies of all survey instruments and discussion guides as well detailed descriptions of the data collection efforts.

1. QIO Survey

We plan to field two self-administered web-based surveys in the late summer of 2010 to all 53 QIOs (since this is the universe of QIOs, there is no sampling involved): (1) the QIO director

[1] Some of the material in this chapter has been previously reported in the OMB submission report for the Paperwork Reduction Act (PRA) (Kovac et al. 2009) and in a memo to CMS (Felt-Lisk 2009). Details about survey procedures such as advance letters, follow-up of nonrespondents, toll-free numbers, help desk, and so on, are in the PRA report.

survey, and (2) the QIO theme leader survey (the QIO staff leaders responsible for each of the themes in the 9th SOW—patient safety, prevention, and so on). Based on the high salience of the 9th SOW evaluation to QIOs, we anticipate a 100 percent response rate from the 53 QIO directors. From the 402 theme leaders, we anticipate an 85 percent response rate to yield 342 completed surveys.[2] The QIO theme leader survey focuses on collecting data on QIO activities and interventions for each of the themes. The QIO director survey focuses on the state provider environment for quality improvement.

2. Discussions with QIO Partner Organizations

We will hold telephone discussions during November 2010 through February 2011, with providers and organizations ("partner organizations") that are working with the QIOs for two of the subnational themes, the 14 QIOs in the care transitions theme, and the 10 in the CKD theme. We will speak with up to 200 care transitions and CKD partners (around 100 per theme) about their perceptions of the value of their work with the QIOs. Although the discussions will provide a broad picture of partners' experiences, they will not constitute a scientific survey.

We will identify potential discussants through a two-step process. First, we will choose eight states randomly from among regional lists of the states participating in each of these themes, to ensure as much regional variation as possible. For the Prevention–CKD theme, we will then both obtain a list of all theme partners from the PATRIOT database as well as ask the selected QIOs to name all theme partners (and information for contact persons) and to briefly describe each partner's role in achieving theme objectives. Next, we will screen every partner listed by the QIO for their level of engagement with the QIO and whether the QIO had any influence on their activities. For those who indicated a significant level of engagement or QIO influence, we will complete the full discussion protocol. For the Care Transitions theme, we will also ask the QIO to identify all partnered organizations. Based on the national numbers of participating provider organizations for this theme, we anticipate needing to select up to 14 partner organizations from a longer list. Our goal is to achieve a diverse mix of partner health care organizations that encompass the bulk of care transitions for their respective communities. Table IV.1 shows the topics we plan on covering in the discussions. The discussions themselves will cover partners' perceptions of: the role of the QIO in the partnership and in any changes in care, strategies that were effective in improving care, lessons learned, and the durability of quality improvements.

3. Review of SDPS Documents

Many of the QIOs' deliverables in their 9th SOW contracts are narrative and descriptive documents that are then uploaded into the SDPS/QIONet internal QIO data system, where they reside in a Document Storage application. For example, for the patient safety theme, the QIOs are to provide quarterly reports on the effectiveness of various educational tools, requests from non-recruited providers for quality improvement assistance, and completed trainings and meetings. For the CKD theme, the QIOs are to report on activities that led to system change for

[2] As noted in Chapter I, some QIOs hold contracts for more than one state, but in all cases there are dedicated directors and staff for each state.

TABLE IV.1

OVERVIEW OF PARTNERSHIPS AND STATE-LEVEL EXPERIENCE REPORTED BY PARTNERS FOR THE CT AND CKD THEMES

	Care Transitions	**Prevention: CKD**
Partnership Mean Total Number of Partners Per State (of 8 selected states per theme) States (of 8 per theme) where number of provider partners is: <10 10-19 20-29 30 or more		
Decision-making: Number of states (of 8 per theme) where decision-making is: QIO-dominated Consensus-based		
Reported Value of QIO: Number of states (of 8 per theme) where at least three-fourths of the interviewed partner organizations reported the QIO's activities were valuable to furthering the goals of the initiative Examples of how QIO added value		
Breadth of Changes Reported: Number of states (of 8 per theme) where more than half the interviewed partner organizations made operational changes to improve care as a result of the initiative Most common types of changes reported Summary of evidence or anecdotes of improved care resulting from the initiative Strategies considered most successful Strategies considered least successful Most common challenges cited		

the specific urinary microalbumin and ACE-I/ARB subtasks, as well as on overall system level changes from the community collaboration task. We will review these documents (with a focus on summary or concluding reports) to gain an overall understanding of the various activities and changes described by the QIOs. In addition the document reviews will inform our partner organization discussions and our case study site visits.

4. Focus Groups of Medicare Beneficiaries

We will conduct four focus groups of beneficiaries who have participated in the diabetes self-management education (DSME) programs that are part of the prevention disparities theme. The provision of these DSME programs directly to beneficiaries represents a new role for QIOs, whose quality improvement work in previous SOWs has focused nearly entirely on providers. The DSME programs are best evaluated by listening to beneficiaries who received the services, just as we will seek feedback from providers who receive QIO assistance in the other themes. As described above, two of the states we choose for site visits will be ones participating in the

prevention disparities theme; we will hold the focus groups during the weeks of our site visits to these states, with two focus groups per state.

Each focus group will comprise 8 to 10 Medicare beneficiaries who have received DSME provided by the QIO. Upon approval from the Office of Management and Budget (OMB), we plan to contact the QIOs that are participating in the disparities work to make them aware of our plans to conduct beneficiary focus groups. We will ask the QIOs to begin to inform beneficiaries about this at the time of their training and to incorporate a statement giving them permission to share their contact information for purposes of the evaluation. We plan to obtain the list of beneficiaries who participated in the DSME along with their contact information in advance from the QIOs. We will begin recruiting participants roughly eight weeks before the scheduled dates of each focus group. The PRA report contains details on the recruiting process, scheduling, logistics, and incentive payments.

B. GATHERING DATA ON STATE PROVIDER ENVIRONMENTS

As mentioned above, the QIO director survey asks about the state provider environment for quality improvement. For example, there are questions on provider interest in quality improvement and on the availability of clinician leaders to champion quality efforts.

The other major effort to understand state provider environments is our set of "case studies" of QIO programs and of the stakeholders in their states. We will perform 10 case studies over a seven-month period, from November 2010 through May 2011 (months 28 through 34 of the 9th SOW). We do not anticipate that discussants' perceptions from the later site visits in spring 2011 will differ systematically from the earlier site visits in late 2010, as the QIOs' interventions should be mature by November 2010, and providers and stakeholders will have been exposed to them for over two years.

Case studies will include week-long site visits. During the site visits, in addition to meeting with QIO staff, we will speak with providers (representatives of hospitals, nursing homes, and physician practices) and what we are calling "community health leaders." Community health leaders are key individuals representing the hospital community (for example, a state hospital association representative), the nursing home community, and the physician community (for example, a representative of a primary care physicians' professional association such as the local American Academy of Family Physicians chapter).

1. Selection of Case Studies

Although we want to pick 10 states that provide a good representation certain characteristics, the goal is not to draw a scientific sample from which to estimate population parameters. The criteria for the 10 case studies are that they:

1. Include at least two states that are participating in the Prevention-Disparities theme
2. Include at least two states that are participating in the Prevention-CKD theme (which may or may not overlap with number 1 above)

3. Include at least three states participating in the Care Transitions theme (which may or may not overlap with criteria 1 and 2 above)

4. Represent equally the four U.S. regions of Northeast, Midwest, South, and West

5. Represent variation in state Medicare populations

We will divide the 49 continental U.S. states (48 states plus the District of Columbia) into 16 cells as defined by the four regions, state Medicare populations (above or below the median Medicare population across states), and participation status in any of the three subnational themes (21 states participate in at least one theme and 28 participate in none). A few cells have only one state (for example, the cell for the South region, Medicare population below the median, and participation in any subnational theme only has Louisiana), but most of these cells have three to five states. We will randomly select cells without replacement (meaning once a cell has been used, we will not use it again), and then draw states, one at a time, from within each selected cell. After selecting six states in this fashion, we will assess the mix of states for the desired characteristics (especially participation in the subnational themes). If it appears from our initial six selections that we may not fulfill the above criteria, we will revise the selection process (drawing the next three states from cells in which states are participating in CKD, for example) in order to meet the criteria. We reserve the option to select up to one state purposively in conjunction with CMS, while maintaining the above criteria.

2. Selection of Providers and Community Health Leaders Within Case Studies

Once we have selected the case studies, we will identify the providers and community health leaders within the states with whom to speak.

a. Providers

We will ask the selected QIOs to provide lists of the providers they worked with on each theme and subtheme, and the evaluation team will select and secure participation from organizations on the lists.[3] However, we will not talk to providers working on the care transitions and CKD themes because we will be speaking with them in the QIO partner organization discussions described in Section 4 below. The steps in the process are as follows:

1. Create one list for each provider type (hospitals, nursing homes, physician practices) of providers who worked with the QIO on any theme or subtheme, along with their city/state locations.

2. Examine city/state locations to identify the locations of participating providers that are feasible to visit on a single visit and include geographic diversity. Typically, this

[3] MPR will have legal access to these names because CMS and the QIOs are executing contract modifications to permit this to occur.

would include selecting two cities within a half-day drive of one another, with a rural area between them. One of these cities would be near the location of the QIO. Providers that are feasible to visit would include those within a 40-minute drive from either of the two cities plus those in the rural area between them.

3. For each type of provider, create a table showing the providers in geographically feasible locations (per Step 2), indicating the theme/subtheme(s) each worked on with the QIO.

4. Use the tables to select:

   - Three hospitals, including hospitals working on all the patient safety subthemes that involve hospitals.
   - Four nursing homes, including: one that worked with the QIO on pressure ulcers, one that worked with the QIO on physical restraints, one that worked on both pressure ulcers and physical restraints, and one that worked with the QIO on the nursing home in need subtheme. Because there will only be two nursing homes in need to select from, we may opt to talk with one of these organizations by telephone if their locations are too geographically dispersed to visit.
   - Two physician practices that worked with the QIO on the prevention theme. In the two states that include a prevention disparities theme, two physician practices that worked with the QIO on that theme.

b. Community Health Leaders

To get perspectives different from those of QIOs and their immediate partners, we will also meet with key people representing the hospital, nursing home, and physician communities (for example, a state chapter head of a primary care physicians' professional association such as the Academy of Family Physicians). We will identify these key contacts through the QIO during the scheduling process.

3. Discussion Topics

During our site visits, we will probe into the provider environment and factors that hinder or help quality improvement efforts and QIOs' work (Table IV.2). Before each site visit we will review SDPS documents and ask QIOs for updated survey responses and other information. We have planned for several hours of on-site interviews at each QIO to cover the core topics, to allow for a discussion of each theme and subtheme as well as a broad-based discussion with the QIO director. CMS officials responsible for overseeing the 9th SOW recommended we allot a substantial amount of time given the variety and breadth of work in the 9th SOW. In particular, they pointed to the range of topics in the patient safety theme, encompassing different provider settings (such as nursing homes and hospitals), different stages of understanding (Methicillin-resistant Staphylococcus aureus [MRSA] and drug safety are new), and different approaches.

TABLE IV.2

DISCUSSION TOPICS FOR SITE VISITS, BY TYPE OF ORGANIZATION OR PROVIDER

Topic	QIOs	Hospitals	Nursing Homes	Physician Practices	Community Health Leaders[a]
Motivation/Culture					
Provider organizations' interest in quality, and impact of this	X	X	X	X	
Perception among providers of a strong business case for quality		X	X	X	
Factors motivating providers to improve quality	X	X	X	X	
Willingness among providers to share information on QI (and impact and factors underlying that)	X	X			
Role of large provider organizations in the state in driving quality					
Adequacy of number of physician champions willing to help facilitate improvement					
Data					
How commonly providers regularly review data on their performance				X	
Infrastructure					
Extent to which information system issues remain a barrier to improvement					
Extent to which providers have staff who are educated and qualified to support improvement efforts					
Workforce instability (turnover) is a barrier to improvement					
Provider Culture-Related Reasons for Poor Performance (where it exists)					
Physician disagreement with relevant guidelines/measures					
Physician disagreement with establishing care routines based on guidelines					
Corporate chain managers who do not believe in establishing care routines based on guidelines					
Characteristics Affecting QIO Impact					
Characteristics of provider environment that make providers particularly receptive to QIO initiatives	X				X
Characteristics of provider environment that make it particularly challenging for QIO to assist providers	X				X

[a]Community health leaders are key representatives of the hospital, nursing home, and physician communities, for example, representatives of the state hospital and nursing home associations or of a relevant state medical organization (such as the state chapter of the American Academy of Family Physicians).

We will ask community health leaders about their views on the QIO's work, including its impact on health care, what activities by the QIO had greater and lesser value in fostering improvements, the state quality environment and how it affected the QIO's work and success, their advice to make the QIO Program more effective, and remaining barriers to further improvement in the state.

Providers that worked with the QIO will tell us how they got involved in the initiative and whether and how the experience may have affected their operations and quality of care. They will explain which QIO activities had greater and lesser value for them and what lessons were learned as a result of the initiative with the QIO. We will probe provider respondents to understand in some depth their way of thinking about quality improvement. For example, what factors do providers say are influencing their organization's motivation (or lack thereof) to improve quality? These factors—especially when combined with detailed information from QIO staff about their experience in working with the providers—may point to recommendations for addressing some of the remaining barriers through the QIO program. Then we will talk through the provider's quality improvement story, to review with representatives the provider's performance trend on the measure(s) of interest; we will ask for their views about which actions at which points led to performance improvements on the measure or failed to do so. These stories are expected to yield insights for the evaluation into the role of the QIO, other factors both within and outside the hospital, and how all of these factors played into the observed trends in performance. In addition, we will discuss the state's quality environment and remaining barriers to improvement and will obtain the provider's advice to CMS on how to improve the program. Details of our plans to enlist organizations' participation and to prepare for each visit, and of the site visit scheduling process are contained in the PRA supporting statement (Kovac et al. 2010).

4. Descriptions of Provider Environment

The QIO survey will provide data on provider environments for all 50 states. We will explore creation of summary indexes from the survey data, for example, an index of the supportiveness of the state's provider environment for quality improvement. The specifics will be determined after examining the distributions of the data received, but one likely approach is to:

1. Use the numbers 0 to 3 to correspond to the ordered responses Strongly Disagree to Strongly Agree, inverting the number order where necessary so that "3" always represents the response most supportive of QI
2. For each respondent, add the numbers for responses to seven items covering: provider interest in quality, perceived business case for quality, provider review of quality data, information systems barriers, qualified staff available, stable workforce, and adequate physician or other health professional champions

3. Rescale the index score to a 0-10 scale, taking into account any missing items[4]

Table IV.3 shows one possible format for summarizing results in the final report. Alternatively, we may also consider creating a series of dummy variables from the QIO survey data by setting responses that show strong agreement or agreement with relevant statements to 1 and others to 0.[5] For example, we expect the following state-level provider characteristics may be associated with greater QIO impact on a particular theme's measures:

- High versus low provider interest in quality
- High versus low adoption of information systems that facilitate quality improvement
- Stable versus high turnover provider workforces
- Adequate versus scarce supply of physician or other health professional quality improvement champions

In addition, for the 10 case study states we will have a wealth of qualitative data in the form of interview notes and summaries of documents. We will use Atlas.ti qualitative analysis software to facilitate retrieval and organized analysis of all of this information. We will code notes at the paragraph level for each topic listed in the site visit interview guide. We then can efficiently retrieve and review discussants' comments, coupling searches with other codes so as to look, for example, at similarities and differences in provider environments or partner experiences or their relations to different topical themes.

The research team will generate and discuss possible relationships and insights from the qualitative data through an iterative process. Using the interview data, the team will consider and explore alternative ideas and expansions upon original hypotheses. Text tables or matrices will be used for illustrating findings. For example, a finding that provider environments varied widely across site visit states but fell into three main categories might be illustrated by a table with a column for each type of environment, and rows for the characteristics of the environment (such as rural or urban location, degree of consolidation of provider organizations, leadership interest in quality, perception of a business case). An analysis of partner experiences might include a matrix in which the columns were states and the rows were responses to questions in the care transitions protocols, with additional rows for potentially important state characteristics (such as number of partners in total or baseline rate of hospital readmissions). In reporting our findings, we will support such adjectives as "some" or "many" with references to the numbers of sites or respondents that corroborate or refute particular contentions.

[4] For example, if all items are complete the index score runs from 0 to 21. A score of 7 would be converted to a rescaled score of 3.3. However, if two items are incomplete, the index score for that person runs from 0 to 15, in which case a score of 7 would be rescaled to a score of 4.7.

[5] Depending on the distribution of responses, we may also explore setting only the "strongly agree" response to 1, and set others to 0.

TABLE IV.3

MEAN QIO THEME LEADER AGREEMENT WITH EACH STATEMENT ABOUT THE PROVIDER ENVIRONMENT (3 = HIGHEST POSSIBLE AGREEMENT

	All Themes	Hospitals: SCIP/HF	Hospitals: MRSA	Nursing Homes: Physical Restraints	Nursing Homes: Pressure Ulcers	Physician Practices: Prevention
Senior leaders at providers care about their quality performance related to this theme						
Providers regularly review data on their performance related to this theme						
Providers perceive a strong business case for quality improvement on the measures important to this theme						
Many providers are motivated to improve Providers have staff who are qualified to support improvement efforts						
Limitations of provider information systems are not a large barrier to improvement						
Workforce turnover is not a large barrier to improvement						
Adequate physician champions						
Mean Summary Index (0-10) of Supportiveness of Provider Environment (10=most supportive)						

C. MECHANISMS AND IMPACTS

The most policy-relevant yet challenging analyses will be those that attempt to distinguish (1) which QIO programs were most effective, and (2) which interventions work for whom and in what circumstances?

1. Which Types of QIO Program Were Most Effective?

Our overall strategy for addressing this question is to take advantage of variations in impact size across QIOs by exploring whether QIO programs with larger impacts possess certain features that QIO programs with smaller impacts do not. Note that we are not attempting to answer the related but different question of whether certain QIO activities or interventions are more effective than others (for example, whether in person workshops are more effective than, say, web conferences or workflow assessments). Answering this second question would require the existence of mutually exclusive groups of participating providers that were exposed to only one of each type of activity of interest. First, it is highly unlikely that providers would receive only one type of intervention. Second, there is the selection problem described in Chapter III; that is, providers that receive more of a certain type of intervention may be systematically

different than those that do not, so that differences in outcomes may not be due to the intervention at all.

a. Calculating QIO-Specific Impacts

The first step is documenting the variation in the size of QIOs' impacts..[6] We will calculate QIO-specific impact estimates for each subtheme component and outcome using the underlying methodology (that is, RD or propensity score matching). Sample sizes for these QIO-specific estimates will obviously be smaller than for the national estimates, but we will not focus on statistical significance in this first step.

b. Developing a Typology of QIO Programs

The second step is characterizing QIO programs along various dimensions. We currently envision a two step process. In the first, we will explore possible quantitative or statistical approaches to data reduction (for example, a principal components or classification and regression tree analysis of the QIO survey and other data—sample sizes may preclude such approaches, however).

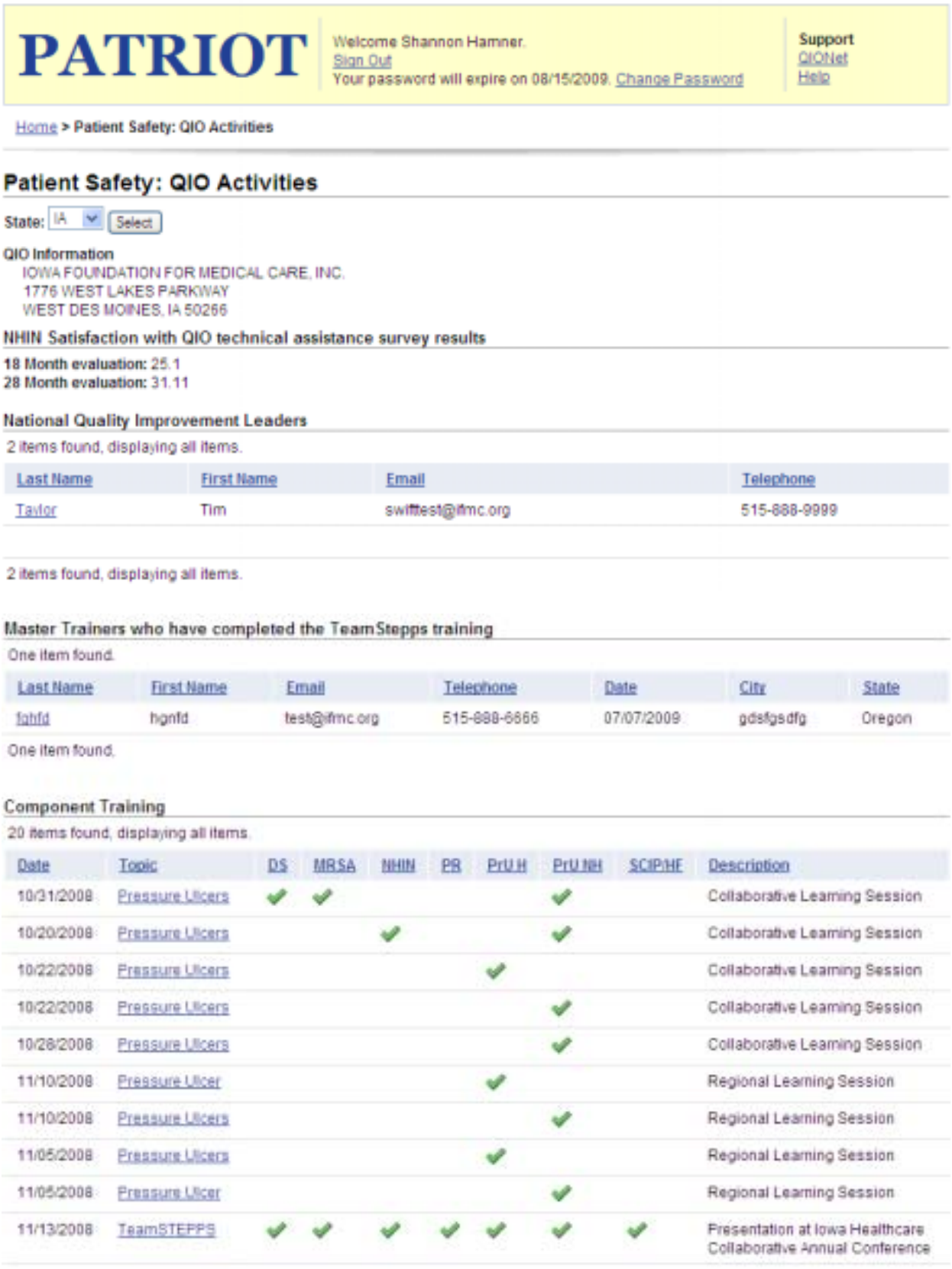
In the second, we will rely on four members of the research team independently reviewing all of the available data on QIO activities to implicitly develop a QIO classification scheme. The QIO Theme Lead survey asks several specific questions about the activities and interventions pursued by QIOs for each theme, for example, use of educational tools and resources, creation of collaborations with multiple organizations, working with individual providers, holding group educational and meeting events, and so on. The QIOs also report on their activities to the SDPS/QIONet intranet using a web browser application called the Program and Theme Reporting Information Online Tool (PATRIOT). For example, PATRIOT has a screen called "Patient Safety: QIO Activities" on which QIOs can record various trainings along with topics and descriptions (Figure IV.1). However, the level of detail is somewhat limited.[7] The hospital and nursing home surveys ask respondents about the nature and frequency of their contacts with their local QIOs; we will sort these survey data by state/QIO and link them with the other QIO-

[6] We may consider a quick exploration of alternate weighting in the main impact analyses as a way of assessing variation in impacts across QIOs. The main impact analyses described in Chapter III will implicitly weight each QIO by a measure of size, such as the numbers of patients treated by each QIO's PPs. We could also weight each QIO to contribute equally; that is, QIOs with fewer than average patients will receive weights greater than one, and vice-versa. Differences in the direction and/or magnitude of the impact estimates compared to the unweighted suggest differences in QIO impacts across states (or more specifically, that impacts differ between small and large QIOs).

[7] For example, for the patient safety "training information" data in PATRIOT, there are single text fields for TRAINING_TOPIC (with entries such as "Restraint Collaborative Learning Session 1," or "Pressure Ulcer Assessment") and for LESSONS_LEARNED_DURING_TRAINING (with entries such as "Participants are more engaged when they are provided real success stories from other peers," or "Nursing home culture is open to and embracing of TeamSTEPPS.")

FIGURE IV.1

PATRIOT SYSTEM SCREEN FOR ENTRY OF QIO ACTIVITIES FOR PATIENT SAFETY THEME



PATRIOT

Welcome Shannon Hamner.
Sign Out
Your password will expire on 08/15/2009. Change Password

Support
QIONet
Help

Home > Patient Safety: QIO Activities

Patient Safety: QIO Activities

State: IA Select

QIO Information

IOWA FOUNDATION FOR MEDICAL CARE, INC.
1776 WEST LAKES PARKWAY
WEST DES MOINES, IA 50266

NHIN Satisfaction with QIO technical assistance survey results

18 Month evaluation: 25.1
28 Month evaluation: 31.11

National Quality Improvement Leaders

2 items found, displaying all items.

Last Name	First Name	Email	Telephone
Taylor	Tim	swifttest@ifmc.org	515-888-9999

2 items found, displaying all items.

Master Trainers who have completed the TeamStepps training

One item found.

Last Name	First Name	Email	Telephone	Date	City	State
fghfd	hgnfd	test@ifmc.org	515-888-6666	07/07/2009	gdsfgsdfg	Oregon

One item found.

Component Training

20 items found, displaying all items.

Date	Topic	DS	MRSA	NHIN	PR	PrU H	PrU NH	SCIP/HF	Description
10/31/2008	Pressure Ulcers	✔	✔				✔		Collaborative Learning Session
10/20/2008	Pressure Ulcers			✔			✔		Collaborative Learning Session
10/22/2008	Pressure Ulcers					✔			Collaborative Learning Session
10/22/2008	Pressure Ulcers						✔		Collaborative Learning Session
10/26/2008	Pressure Ulcers						✔		Collaborative Learning Session
11/10/2008	Pressure Ulcer					✔			Regional Learning Session
11/10/2008	Pressure Ulcers						✔		Regional Learning Session
11/05/2008	Pressure Ulcers					✔			Regional Learning Session
11/05/2008	Pressure Ulcer						✔		Regional Learning Session
11/13/2008	TeamSTEPPS	✔	✔	✔	✔	✔	✔	✔	Presentation at Iowa Healthcare Collaborative Annual Conference

level data. Finally, we will have a great deal of qualitative and narrative data from QIO memos and our interviews and discussions on QIOs' activities. Using all of these data sources, we will develop a "typology" or categorization of different QIO activities. After synthesizing the categories independently developed by researchers, we will assess the reproducibility of their classifications through analyses of inter-rater reliability.

c. Linking Typologies to Impacts

Lastly, once we have state-level impacts on the one hand, and QIOs' grouped into a typology on the other, we will start with simple visual inspections of matrices consisting of the QIOs in rows, rank ordered by size of impacts, and their typologies in the columns. We will look for patterns of certain typologies appearing more or less frequently among the larger impacts. We can then divide the impacts into quantiles (quartiles, quintiles, and so on) and calculate the percentages of QIOs in each quantile that belong to certain types. To confirm these initial impressions, we will then restrict the matrices and descriptive percentages to those QIOs with statistically significant impacts or to those with impacts that exceed a certain threshold of the distribution. We will assess the feasibility of regression models that correlate impact size as a function of typology descriptors, although our ability to do so will be limited by sample sizes. We will have at the most around 50 observations, but once we begin examining specific themes or restricting to QIOs with statistically significant impacts or with impact estimates beyond a certain size, we will very likely have far fewer data points. Once these quantitative approaches have helped us develop some early ideas on which QIO interventions or activities may have been associated with larger impacts, we will search our qualitative data for corroborating evidence. For example, we will search our interviews for evidence of whether discussants perceived certain broad QIO strategies corresponding to specific typologies as being particularly effective or well received. Although a provider-level analysis comparing providers whose QIOs did pursue interventions of interest with those whose QIOs did not might appear to have larger sample sizes, it cannot overcome the limitation that the unit of intervention is still the state.

2. Linking Provider Environments to Impacts

We will extend the above methodology to examining associations between provider environments and QIO impacts, starting with visual inspection of matrices in which QIOs are again in the rows and rank ordered by impact size, but now with provider environment summary indexes or classifications in the columns (as in Table IV.3). As we did with the QIO typologies analysis, we will then move on to calculations of the percentages of provider environment types in each quantile of the impact distribution, restriction to statistically significant impacts or impacts of a minimum size, and consideration of regression models that correlate impact size with provider environment. We will then combine these analyses with our qualitative data.

3. Which Interventions Work for Whom, and in What Circumstances?

Finally we will move on to the full question above. The complexity of this question, which asks about the conjunction of three separate factors—(1) QIO program type, (2) providers, and (3) provider environment—also indicates the challenge of finding answers. Having developed a

typology of QIO programs, we now want to know whether QIO programs of a certain type work better with certain types of providers or in specific types of provider environments. We do not believe that a straightforward, purely quantitative approach to this question is feasible. For example, we would need data with sufficient sample sizes containing QIO program types A, B, and C, each operating in provider environments E, F, and G, and within each of these combinations, groups of provider types X, Y, and Z. The regression models would have to incorporate a variety of three-way interaction terms to reflect the interplay of the three factors of interest.

We will combine qualitative and quantitative approaches. For example, we will construct and visually inspect matrices that display QIO typologies down the rows, provider environment categories in the columns, and provider type-specific estimates in each cell. We will look for patterns of larger or smaller impacts among the cells. Table IV.4 shows how such a matrix might appear:

Obviously, the number of combinations of provider environment features, and provider characteristics that we will be able to examine is limited, and our survey and interview findings will help guide us in the factors to be assessed. The qualitative data will prove key in bolstering any hypotheses that arise from our tabular analyses.

TABLE IV.4

IMPACT ESTIMATES FOR HOSPITAL OUTCOMES BY QIO PROGRAM TYPOLOGY, PROVIDER ENVIRONMENT, AND PROVIDER CHARACTERISTIC

		High Adoption of Information Technology		Low Adoption of Information Technology	
QIO Program Type		Stable Provider Workforces	High Turnover Provider Workforces	Stable Provider Workforces	High Turnover Provider Workforces
Type A	Not-for-profit: For-profit:				
Type B	Not-for-profit: For-profit:				
Type C	Not-for-profit: For-profit:				

V. CONCLUSIONS

This chapter first reviews some lessons that the current design challenges for 9th SOW evaluation may hold for future SOWs. It then describes the reports that will be forthcoming from the current evaluation and the timeline and key milestones for the remainder of the project.

A. SELECTED EVALUATION CHALLENGES OF THE 9TH SOW AND IMPLICATIONS FOR THE 10TH SOW AND FUTURE SOWS

"Are the QIOs accomplishing what CMS wants them to accomplish—are their efforts improving care?" is a central question for the QIO program, whether for the purpose of program or contract evaluation. Planning documents for the 9th SOW have referred to this question as the one of "attribution" to the QIO program (Leavitt 2006). As explained in Chapter III, quantitative estimates of impacts for the program evaluation will not be feasible for several themes and subtheme components. Among the remaining themes and components, the rigor of the program evaluation impact estimates will vary widely despite the application of sophisticated statistical and econometric techniques. In particular, the impact analyses of the prevention CKD and care transitions component subthemes or themes rely on matched comparison group designs, which are of relatively lesser rigor because of the remaining uncertainty over whether the comparison communities may differ from the intervention communities on important but unobserved characteristics. In terms of contract evaluation, QIOs' performance on nearly all themes is evaluated by improvement in quality measures at remeasurement relative to baseline, but it is unknown whether these improvements might have occurred in the absence of QIOs' efforts. These program and contract evaluation challenges may hold lessons for the design of the 10th SOW and beyond.

The key to assessing QIO effects is knowing what the outcomes of the PPs and intervention communities would have been without QIOs' efforts. That knowledge can only come through observation of groups of physician practices and communities that are just like the PPs and intervention communities, except for the QIOs' assistance. Unfortunately, the discretion that the QIOs had in the 9th SOW in selecting the participating providers (PPs) and intervention communities makes identification of such comparison groups extremely difficult. Aware of the criteria by which CMS will be evaluating their contract performance, QIOs naturally and understandably had incentives to (1) pick PPs and communities that were most likely to improve anyway (not necessarily those who needed the most help) and (2) pick NPs and comparison communities that were least likely to improve (even if, under the 9th SOW, NPs and comparison communities are not used in contract evaluation). The QIOs' should not be blamed for such behavior, which simply reflects efforts to meet contract performance standards. A further specific difficulty in the core prevention theme is that the NP group receives QIO assistance with improving data collection and reporting. To the extent that such efforts also help improve quality of care, they will obscure or dilute any differences in outcomes between PPs and NPs. It will thus be difficult to say whether what the outcomes of the PPs and intervention communities would have been without the QIOs. Similarly, the program evaluation can never be certain that the outcomes of comparison practices and communities selected for impact estimation truly indicate the outcomes of the PPs and intervention communities, either.

In future SOWs, for themes or components that involve QIOs working with specified PPs, CMS could consider two options to avoid the problems above. In the first, CMS would first create pools of providers using selected criteria, for example, low baseline performance on quality measures. Geography could be considered as well, if there were concerns about the costs of QIOs having to work with distant or scattered providers. CMS would then randomly select sets of PPs from these pools that QIOs would be required to work with. The remaining providers would receive no QIO services and would serve as comparisons. The random selection process ensures the comparability of the PP and NP groups, and the exclusion of NPs from QIO services avoids the potential dilution of program effects present in the current prevention theme. The potential drawback of this approach is that it fails to take advantage of useful background information that QIOs may have on which providers might be most helped by their intervention, and which providers might be most willing to cooperate.

In the second option, CMS would again create a pool of providers suitable for QIO intervention, but then randomly divide it into two pools, one of PP candidates and the other of providers not eligible for QIO services. QIOs would select a set of providers to work with from the PP candidate pool. The PP candidate pool should be large enough that QIOs can meet their recruitment targets. QIOs' performance would be evaluated by comparing the outcomes for the *entire* PP candidate pool (not just those selected as PPs) to the entire ineligible pool. Having QIOs select providers to work with from among the PP candidates offers the advantage of QIOs being able to choose providers most in need of help or most likely to be helped by QIO intervention. In addition, comparing the entire PP candidate pool to the entire ineligible pool means that any observed differences must be due to QIOs' efforts with the PPs within the PP candidate pool, since the two pools are otherwise equivalent. In contrast, comparing only the PPs to the pool of ineligibles would lead us back to the situation of being unable to distinguish whether superior performance by PPs was due to QIOs' efforts or simply to QIOs' skill at picking "winners" (providers most likely to improve anyway). Compared to the first option, the disadvantages to this second option are increased data collection costs and diminished statistical power. Data costs are higher because of the need to gather information on larger numbers of providers, namely all providers in both the PP candidate and ineligible pools. Statistical power to detect impacts is also substantially diminished because PPs comprise only a fraction of the PP candidate pool. There is thus a tradeoff between furnishing QIOs a wide choice of providers from which to pick by offering them a very large PP candidate pool, and increasing statistical power by raising the proportion of PPs in the overall PP candidate pool [1]

Option one is preferable in terms of simplicity, data collection costs, and statistical power, but the ability of QIOs to exercise discretion in selecting PPs may also be an important consideration. Either option would improve both the contract and program evaluability, and

[1]Impact estimates (effects on participants) in option two would be calculated by dividing the treatment-control difference between the PP candidate and ineligible pools by the participation rate (Bloom xxxx). Sample sizes for a given minimum detectable effect increase inversely with the square of the participation rate. If the proportion of PPs in the PP candidate pool is only 25 percent, the sample size must be 16 times larger than that for option one to yield the same power.

thereby strengthen the rigor with which attribution can be made, for themes and subtheme components in future SOWs that require QIOs to work with PPs.

B. REPORTING RESULTS

In this section we first outline our approach to the overall synthesis of evaluation findings that the final report will need to undertake. We then briefly discuss how this evaluation fits into the context of previous studies of the QIO program, and describe the specific deliverable reports that the evaluation will produce.

1. Synthesizing Results

We repeat here the main evaluation research questions first presented in Chapter I.

1. What is the impact of the program on the quality of care for Medicare beneficiaries (either nationally or subnationally)?
   - How do program costs and benefits compare, and what is the cost-effectiveness of the program? What factors mediate costs and benefits, and cost-effectiveness?
   - Do impacts differ for underserved beneficiaries and non-underserved beneficiaries (has the program narrowed healthcare disparities)?
2. Assuming there are impacts, what works for whom, and in what circumstances (what are the *mechanisms* of impacts)?
3. How might the program be improved to provide greater value?
   - Can key activities be more standardized across QIOs in a way that would improve the impact?

However, as noted in Chapter I, the many themes of the 9th SOW in fact comprise multiple, loosely related interventions aimed at multiple providers, and that seek to influence multiple outcomes of ambulatory to acute to long term care. The evaluation of the 9th SOW thus constitutes multiple smaller evaluations, and it will be challenging to integrate the many findings. For example, it may turn out that one subtheme component focused on one type of provider and care setting yields extremely promising results, while other components from the same or different themes targeting different providers or care settings appear less successful. Furthermore, the strength of evidence for each of the components will vary as well, given how the comparison and impact estimation strategies had to be modified for each subtheme component. Thus, as in the development and presentation of the evaluation methodologies, our approach to synthesizing results will start by considering each of the research questions above for each subtheme component individually.

For each theme, we will first assess the proportion of outcomes subsumed by the theme that exhibit favorable impacts, the size and statistical significance of those impacts, and the susceptibility of the estimators to bias. We will revisit any measures of cost-effectiveness and

cost-benefit that we have been able to calculate for specific subtheme components, as described in Chapter III. We will then assess the extent to which the implementation of the theme followed the logic models presented in Chapter II and Appendix A and review our findings from the mechanisms and environment analyses described in Chapter IV.

We will then enter summaries of all of these subtheme component specific assessments into a series of matrices in which the rows are the subtheme components and the columns are summaries of the individual assessments listed above, namely—estimated impacts on different outcomes; size, statistical significance, robustness and underlying rigor of these impacts; measures of cost-effectiveness and cost-benefit; faithfulness to the logic models and to implementation as planned; and mechanisms/environment/provider findings. Since impact analyses, cost-effectiveness/cost-benefit analyses, and mechanisms analyses may not be feasible for all of the components, some of the cells may remain blank. Inspection and analysis of these matrices will help us to answer each research questions for each of the subtheme components.

Finally, we will consider whether we can build these individual subtheme component assessments into an overall assessment. In addition to the evaluability issues already discussed, the challenges are (1) the breadth and variety of the 9th SOW, and (2) the tradeoffs in attempting to provide an overall assessment.

a. Breadth and Variety of the 9th SOW

As noted, the 9th SOW is ambitious, seeking to improve quality of care in hospitals, all nursing homes, struggling nursing homes, primary care practices with EHRs, primary care practices serving minority beneficiaries, self-care among beneficiaries with diabetes, Medicare Advantage drug plans, Part D prescription drug plans, nephrology practices, and dialysis centers. The clinical care of Medicare beneficiaries treated in these settings differs tremendously, as do the nature of providers themselves (such as size, to name just one characteristic), problems with quality of care, and interventions to improve quality.

b. Tradeoffs and Challenges in Summarizing Results

Obviously, an overall assessment is straightforward if all component evaluations are either uniformly positive (for example, large, unambiguous favorable impacts across the board; clear cost-effectiveness; unmistakable mechanism, environment, and provider characteristic findings), or uniformly negative (complete and convincing lack of impacts). However, such a scenario is highly unlikely, because of the problems in evaluability and breadth of activities just discussed. It will be tempting to boil the wealth of findings from the matrices described earlier into a single, simple message (such as the 9th SOW "worked" or "did not work"). However, such a single message risks discarding an enormous amount of information; it might mask, for example, that a few things worked extremely well, while others looked promising but evidence for their effectiveness was weak (because of a lack of rigor in attribution of effects to QIO efforts). On the other hand, a complex list of findings qualified by numerous caveats is also not helpful to decisionmakers. Although the nature of specific tradeoffs must await the findings of our analyses, we will work with CMS to produce concise, policy relevant reports that fairly represent the complexity of results while providing clear guidance and recommendations.

2. The Current Evaluation in the Context of Previous Studies of the QIO Program

We briefly discuss here how the current evaluation relates to previous studies of the QIO program. As discussed in Chapter I, the Institute of Medicine's Report *Medicare's QIO Program: Maximizing Potential* articulated a broad set of recommendations, and NORC, under contract to ASPE, also developed recommendations pertaining to data availability and evaluation of the program. CMS has been working to address many of the recommendations in those reports, which were based on the experience in the 7th and beginning of the 8th SOW.

The Mathematica evaluation, by its existence and scope, meets the IOM recommendation for an external evaluation. The evaluation design meets several of the specifics the IOM recommended as well, including developing a method for attributing quality improvements to the QIO's intervention, designing a "mechanisms" analysis that examines the relative effectiveness of various types of interventions, including cost-effectiveness analysis, and careful assessment of the QIOs' role in quality improvement interventions relative to other players.

The extent to which the current evaluation relates to other IOM and NORC recommendations varies by the type of recommendation, with largest relevance to the IOM program management recommendations. The Mathematica evaluation will assess the success of CMS efforts to address these recommendations to a large degree, through QIO directors' reports about the clarity of the goals and objectives in their core contract and how they will be evaluated, the clarity and consistency of communications from CMS, the reasonableness of the timeframe for achieving goals, and the effect on their operations and value of contract modifications. We do not have plans to assess the QIO selection process or the incentives contained within QIO and QIOSC contracts, other foci of the program management recommendations.

Other types of IOM and NORC recommendations most often relate to program decisions that are not within the evaluation's goal of evaluating the 9th SOW QIO Program. For example, the IOM recommended that CMS initiate a comprehensive review of its data-sharing systems, processes and regulations to identify and correct practices and procedures…that restrict the sharing of data by the QIOs for quality improvement purposes or that inhibit prompt feedback to the QIOs and provider on provider performance. Similarly NORC recommended CMS identify opportunities for shortening data lags and preparing databases that are ultimately used to report to QIOs. These types of reviews are not part of the contracted evaluation. However, if these recommendations have not been addressed by CMS, and the issues that led to these recommendations are still a significant concern for QIOs in achieving their goals, they will likely resurface through our QIO survey and case studies. Details of each IOM and NORC recommendation and the relevance to the evaluation are presented in Appendix E.

3. Forthcoming Reports

The evaluation will produce several reports. These include a summary report of QIOs' attainment of the mid-course milestones in their contracts, and a report on findings from the evaluation's surveys of hospitals, nursing home,s and QIO staff. In late September of 2010 we will submit a detailed draft outline (including chapter headings and table shells or dummies) for the interim report that is due in early February of 2011. The February 2011 interim report will contain results of quantitative descriptive and impact analyses. The final evaluation report, due in

October 2011, will update the quantitative analyses of the February report with more recent data; present results of all of the qualitative components of the study, the mechanisms analysis, and the cost-effectiveness and cost-benefit analyses; and conclude with a synthesis of all analyses of the evaluation and future implications and recommendations. This schedule assumes that all of the QIO- and CMS-furnished data necessary for the evaluation are accurate and available in time for report analysis and preparation. Table V.1 summarizes the delivery schedule.

C. PROJECT TIMELINE

As noted Mathematica executed subcontracts with each individual QIO in the summer of 2009 in order to obtain legal access to QIO owned data. Mathematica has also working since early 2009 with relevant divisions within OCSQ and several CMS data contractors to gain access to data. Mathematica obtained the specialized SDPS computers necessary to physically access the SDPS/QIONet system through a VPN connection and was given VPN user accounts to the system at the end of 2009. Mathematica gained access in late April 2010 to the QIO owned PATRIOT data which identifies which providers are PPs, and includes QIOs' reports to CMS on QIO activities and interventions.

The remaining key activities and milestones in the evaluation are as follows (Figure V.1)

- **Medicare Claims-based QIO Program Data.** We also plan on using three sets of Medicare claims-based files that are created for the QIO program for other purposes (1) the quarterly analytic files, (2) the Program Progress Reports, and (3) the risk-adjusted mortality, readmission, and surgical outcome data for Hospital Compare.
  - The quarterly analytic files are processed from Medicare claims data for the QIO program by, or with input from Iowa Foundation for Medical Care (IFMC), Buccaneer Computer Systems and Services, Inc. (BCSSI), and Edaptive Systems LLC (the Program Management Business Requirements contractor or PMBR). These files comprise the denominator of Medicare beneficiaries eligible for the core prevention measures, the diabetes utilization measures, and the CKD measures. They also contain indicator flags for whether the beneficiary received these services (breast and colon cancer screening, and influenza and pneumococcal vaccination for core prevention; hemoglobin A1c testing, lipid testing, and eye exams for diabetes utilization; and urine microalbumin testing, ACE-I/ARB medication, and initial hemodialysis through an arterio-venous fistula for CKD). We were granted access to previous quarterly analytic files in late April 2010.We assume that these files and the forthcoming quarterly files have been processed correctly, and that future files will be available for our use in time for our deliverable schedule.
  - The Program Progress Reports (PPR) are summary reports based on the quarterly analytic files created for both CMS and the QIOs. We assume that these reports will be available for our use in time for our deliverable schedule.

TABLE V.1

UPCOMING DELIVERABLES FOR 9TH SOW EVALUATION

Deliverables	Description or Comment	Due Dates
Report of QIOs' achievement of their milestones		10 weeks after we are given access to 18-month scores determined by CMS
Survey report	Includes report on partner's experience of service by the QIOs and report on the survey of QIOs	24 weeks after OMB clearance (anticipated due date of December 21, 2010)
Preliminary draft outline (including chapter headings) and set of dummy tables	Report to outline data, end points and timeframes for February 2011 interim impact report	September 27, 2010
Final outline and set of dummy tables following receipt of CMS comments on draft outline		October 25, 2010
Draft interim impact report	The report will present quantitative descriptive and impact analyses of the most recent available data.	February 1, 2011
Final interim impact report		February 15, 2011
Draft final report	The draft final report will update quantitative results using the most recent data; present the mechanism, cost-benefit/cost-effectiveness, and qualitative results; and synthesize results from all components of the evaluation.	September 19, 2011
Final report		October 3, 2011

Note: This table lists only forthcoming reports. Other deliverables specified in Mathematica's contract, such as monthly conference calls, or assisting the QIO data contractor, are not shown. The schedule is contingent on correct CMS- and QIO-furnished data being available to us in time for analysis and reporting.

- The risk-adjusted mortality, readmission, and surgical outcome data for Hospital Compare are being created under the Production and Implementation of the CMS Hospital Outcomes and Efficiency Measures (PIHOEM) project by Mathematica under a separate contract to CMS. Many of the non-public files being produced by Mathematica are also being provided to Colorado Foundation for Medical Care (CFMC) to support its role as the Care Transitions QIOSC. We assume that these data will be available for our use in time for our deliverable schedule.

- **Survey Data Collection.** The Paperwork Reduction Act supporting statement for the surveys of QIOs, nursing homes, and hospitals, and for the partner interviews and case studies are under review by OMB and we await clearance. The current schedule calls for the fielding of these surveys around August through December of 2010.

- **Site visits and Beneficiary Focus Groups**. Assuming OMB clearance, these will be held from November 2010 through May 2011.

- **Site visits and Beneficiary Focus Groups**. These will be held during the spring and summer of 2010. An additional wave of qualitative data collection of 12 QIOs and their partners will take place in the fall of 2010 and spring of 2011.

- **Evaluation Analyses.** As noted above, we will be submitting three drafts of the Data Collection and Analysis Report in the fall of 2010 (in September, October, and November). Once CMS has approved the final analysis plans in November 2010, we will begin work on the final report.

FIGURE V.1
OVERVIEW OF THE SCHEDULE FOR THE EVALUATION

Year	2009							2010												2011										
Calendar Month	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N
Task/Deliverables																														
Second Meeting						▲																								
Monthly Conference Calls	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Monthly Progress Reports	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Design Report																														
Evaluation methodology development Conceptual framework and method Evaluation design report				△							▲																			
Data Collection Activities																														
Survey of QIOs																														
Survey of hospitals and nursing homes																														
Case studies and beneficiary focus groups																														
Other Reports																														
Condition-specific analyses to inform preparations for the 10th SOW		△	▲																											
Potential net financial impacts on Medicare of selected 9th SOW QIO interventions					△	▲																								
Validation of RDD assumptions for hospital SCIP/HF analysis using baseline data									▲																					
Interim report on QIOs' achievement of their milestones [a]															▲															
Draft detailed outline of interim impact report of February 2011 (including chapter headings and a set of dummy tables)																△														
Final detailed outline of interim impact report of February 2011																	▲													
Survey report of findings from surveys of nursing homes, hospitals, and QIOs [b]																			▲											
Draft interim impact report																				△										
Final interim impact report																				▲										
Final Reports																														
Draft final report																												△		
Final final report																													▲	
△ = draft deliverable																														
▲ = final deliverable																														
Note: above schedule assumes QIO- and CMS- furnished data are correct and available in time for analysis and reporting																														
[a] 10 weeks after we are given access to 18-month scores determined by CMS																														
[b] 24 weeks after OMB clearance (projected draft report due date of December 2010)																														

REFERENCES

Boukus, Ellyn, Alwyn Cassil, and Ann S. O'Malley. "A Snapshot of U.S. Physicians: Key Findings from the 2008 Health Tracking Physician Survey." Washington, DC: Center for Studying Health System Change, September 2009.

Centers for Medicare & Medicaid Services. *Quality Improvement Organization Manual*, Publication number 100-10. [http://www.cms.hhs.gov/Manuals/IOM/itemdetail.asp?filterType=none&filterByDID=-99&sortByDID=1&sortOrder=ascending&itemID=CMS019035&intNumPerPage=10], accessed September 8, 2009.

Centers for Medicare & Medicaid Services. "Reporting Hospital Quality Data for Annual Payment Update." [http://www.cms.hhs.gov/HospitalQualityInits/08_HospitalRHQDAPU.asp#TopOfPage], accessed September 9, 2009.

Centers for Medicare & Medicaid Services. "Use of the CARE Instrument in the QIO Program 9th Statement of Work—CARE Users Now Urged to Transition to New Formats for Collecting CARE Data." Fact Sheet, June 25, 2009. [http://www.cms.hhs.gov/QualityImprovementOrgs/downloads/CAREFSUpdate062509.pdf], accessed September 9, 2009.

Chromy, James R. "Sequential Sample Selection Methods." *Proceedings of the American Statistical Association, Survey Research Methods Section*, 1978, pp. 401 -406

Clarkwest, Andrew, Sue Felt-Lisk, Sarah Croake, and Arnold Chen. "Assessment of the Eighth Scope of Work of the Medicare Quality Improvement Organization Program." Princeton, NJ: Mathematica Policy Research. April 2009.

Cook. T. D., and V. C. Wong. Empirical tests of the Validity of the Regression Discontinuity Design. *Annales d'Economie et de Statistique* (in press) [http://www.northwestern.edu/ipr/publications/papers/cook_empirical_tests.pdf], accessed April 26, 2010

Dehejia, Rajeev H. and Sadek Wahba. "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *The Review of Economics and Statistics*, vol. 84, no. 1, February 2002, pp. 151–161.

DesRoches, Catherine M., Eric G. Campbell, Sowmya R. Rao, Karen Donelan, Timothy G. Ferris, Ashish Jha, Rainu Kaushal, Douglas E. Levy, Sara Rosenbaum, Alexandra E. Shields, and David Blumenthal. "Electronic Health Records in Ambulatory Care—A National Survey of Physicians." *New England Journal of Medicine*, vol. 359, no. 1, July 3, 2008 , pp. 50-60.

Fan, Jianqing and Irene Gijbels. *Local Polynomial Modelling and its Applications*. London: Chapman and Hall, 1996.

Felt-Lisk, Sue. "Design of Methods for Tasks 3 and 4, Phase II." Memorandum submitted to Bob Kambic, August 11, 2009.

Giambo, Pamela, Vasudha Narayanan, Josh Rubin, Lauren Shrader, and Sushama Rajapaksa. "Surveys of Quality Improvement: 2007 Provider Satisfaction Survey Methodology Report." Rockville, MD: Westat, Inc., November 29, 2007.

Gimm, Gilbert, Christine Yip, and Marian V. Wrobel. "Colorectal Cancer Screening Analysis." Memorandum to Robert Kambic. Washington, DC:: Mathematica Policy Research, July 30, 2009

Gold, Marthe R., Joanna E. Siegel, Louise B. Russell, and Milton C. Weinstein. *Cost-Effectiveness in Health and Medicine.* New York: Oxford University Press (USA), 1996.

Imbens, Guido and Karthik Kalyanaraman. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." Working Paper 14726. Cambridge, MA: National Bureau of Economic Research, February 2009. [http://www.nber.org/papers/w14726.pdf], accessed September 9, 2009.

Institute of Medicine. *Medicare's Quality Improvement Organization Program: Maximizing Potential*. Washington, DC: National Academies Press, 2006.

Jencks, S.F. Letter to the editor regarding "Quality Improvement Organizations and Hospital Care." *JAMA*, vol. 294, no. 16, October 26, 2005, p. 2028.

Jencks, S.F., E.D. Huff, and T. Cuerdon. "Change in the Quality of Care Delivered to Medicare Beneficiaries, 1998-1999 to 2000-2001." *JAMA*, vol. 289, no. 3, 2003, pp. 305-312.

Kaufman, Leonard, and Peter J. Rousseeuw. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: John Wiley & Sons, Inc., 1990.

Kovac, Martha, Sue Felt-Lisk, Arnold Chen, and John Hall. "Program Evaluation of the Eighth and Ninth Scope of Work Quality Improvement Organization Program: Supporting Statement for Paperwork Reduction Act Submission." Princeton, NJ: Mathematica Policy Research, January 12, 2010.

Leavitt, Michael O. "Report to Congress: Improving the Medicare Quality Improvement Organization Program – Response to the Institute of Medicine Study." Washington, DC: Centers for Medicare & Medicaid Services, 2006. [http://www.cms.hhs.gov/QualityImprovementOrgs/downloads/QIO_Improvement_RTC_fnl.pdf], accessed September 9, 2009

Lee, David S., and Thomas Lemieux. "Regression Discontinuity Designs in Economics." Working Paper 14723. Cambridge, MA: National Bureau of Economic Research, February 2009. [http://www.nber.org/papers/w14723], accessed September 9, 2009.

McBean, A. Marshall. "Improving Medicare's Data on Race and Ethnicity." Medicare Brief No. 15. Washington, DC: National Academy of Social Insurance, October, 2006. [http://www.nasi.org/usr_doc/medicare_brief_015.pdf], accessed September 9, 2009.

Narayanan, Vasudha, Pamela Giambo, Stephanie Fry, Sherman Edwards, Jennifer Kawata. "Surveys of Quality Improvement: 2007 Provider Satisfaction Survey Analytic Report." Rockville, MD: Westat, Inc., March 14, 2008.

National Committee for Quality Assurance. *The State of Health Care Quality 2008.* Washington, DC: NCQA, 2008.

Paez, Kathryn, Kate Stewart, and Suzanne Felt-Lisk. "Overview of Quality Improvement Initiatives." Washington, DC: Mathematica Policy Research, June 12, 2009.

Pham, Hoangmai H., Ann S. O'Malley, Peter B. Bach, Cynthia Saiontz-Martinez, and Deborah Schrag. "Primary Care Physicians' Links to Other Physicians through Medicare Patients: The Scope of Care Coordination." *Annals of Internal Medicine*, vol. 150, no. 4, February 17, 2009, pp. 236-242.

Robert Wood Johnson Foundation and University of Wisconsin Population Health Institute. "County Health Rankings." [http://www.countyhealthrankings.org/], accessed April 27, 2010.

Rollow, William, Terry R. Lied, Paul McGann, James Poyer, Lawrence LaVoie, Robert T. Kambic, Dale W. Bratzler, Allen Ma, Edwin D. Huff, and Lawrence D. Ramunno. "Assessment of the Medicare Quality Improvement Organization Program." *Annals of Internal Medicine*, vol. 145, 2006, pp. 342-353.

Rubin, Donald B. "Estimating Causal Effects of Treatments in Randomised and Nonrandomised Studies." *Journal of Educational Psychology*, vol. 66, 1974, pp. 688-701.

Schochet, Peter Z. "Technical Methods Report: Statistical Power for Regression Discontinuity Designs in Education Evaluations." Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, August 2008.

Schmitz, Bob, Michaella C. Morzuch, and Marian V. Wrobel "Analysis of Vaccination for Pneumococcus and Influenza." Memorandum to Robert Kambic. Washington, DC:: Mathematica Policy Research, July 30, 2009.

Snyder, Claire, and Gerard Anderson. "Do Quality Improvement Organizations Improve the Quality of Hospital Care for Medicare Beneficiaries?" *JAMA*, vol. 293, no. 23, 2005, pp. 2900-2907.

Sutton, J., L. Silver, L. Hammer, and A. Infante. "Toward an Evaluation of the Quality Improvement Organization Program: Beyond the 8th Scope of Work." Final Report. Washington, DC: U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, 2007.

U.S. Government Accountability Office. "Nursing Homes: Federal Actions Needed to Improve Targeting and Evaluation of Assistance by Quality Improvement Organizations." GAO-07-373. Washington, DC: GAO, May 2007.

van der Klaauw, Wilbert. “Breaking the Link Between Poverty and Low Student Achievement: An Evaluation of Title I.” *Journal of Econometrics*, vol. 142, no. 2, February 2008, pp. 731-756.

Waldo, Daniel R. “Accuracy and Bias of Race/Ethnicity Codes in the Medicare Enrollment Database.” *Health Care Financing Review*, vol. 26, no. 2, winter 2004, pp. 61-72.

Wrobel, Marian V. and Myles Maxfield. “Readmission Analysis.” Memorandum to Robert Kambic. Cambridge, MA: Mathematica Policy Research, July 30, 2009.

Wrobel, Marian V. and Myles Maxfield “Screening Mammography Analysis.” Memorandum to Robert Kambic. Cambridge, MA: Mathematica Policy Research, July 30, 2009.