Evaluation of HHS Delivery System Reform Efforts and Affordable Care Act Provisions: Consolidated Evaluation Design Recommendations

Final Report

August 31, 2012

Prepared for

Susan Bogasky, Project Officer
Assistant Secretary for Planning and Evaluation (ASPE)
Office of the Secretary
U.S. Department of Health and Human Services
Hubert H. Humphrey Building, Room 443F5
200 Independence Avenue, SW
Washington, DC 20201

Prepared by

The CDM Group, Inc.
Brandeis University
The Heller School for Social Policy and Management
The Schneider Institutes for Health Policy

Authors

Robert Mechanic, M.B.A.
Jennifer Perloff, Ph.D.
Darren Zinner, Ph.D.
Moaven Razavi, Ph.D.
Michael J. Keane, Dr.P.H.

The report was prepared under contract number: HHSP23320095629WC. Task order number: HHSP23337001T

Table of Contents

LIST OF EXHIBITS

Chapter 1: Executive Summary

The purpose of this project is to provide the Department of Health and Human Services (HHS) with recommendations for a coordinated evaluation design to tie results from the numerous Medicare, Medicaid and Dual Eligible delivery system reform initiatives and Affordable Care Act (ACA) provisions into an overall model of health system change. These delivery system reform models have the potential to improve quality, reduce costs, and engage patient-centered care. This project is a first step in an overall planning effort to build the foundation for future evaluations and meta-analyses of many diverse initiatives.

Our work included a review of numerous new programs that would be included in a coordinated evaluation, including:

- Pioneer Accountable Care Organization (ACO)
- Medicare Shared Savings Program
- Bundled Payment for Care Improvement
- State Multi-Payer Primary Care Demonstration
- FQHC Advance Primary Practice Demonstration
- Comprehensive Primary Care Initiative
- State Demonstrations to Integrate Care for Dual Eligible Individuals
- Hospital Value-Based Purchasing
- Medicare Readmission Reduction Initiative
- Community-Based Care Transitions Program
- Partnership for Patients
- Independence at Home Demonstration

This executive summary (**Chapter 1**) highlights the main findings of this report, summarized by chapter. It incorporates input from a Technical Expert Panel (TEP) that met on June 5 (see **Appendix)**, and is followed by an introduction **(Chapter 2)** that further describes the objectives of this project. Key chapters of the report are summarized below.

Summary of Chapter 3: Conceptual Analytic Framework

Our approach to a coordinated evaluation design for HHS delivery reform programs began with a conceptual analytic framework. The general approach can be summarized in the following manner:

$$\Delta \text{ Outcome}_{1..n} = A + bI_{1..n} + c\Delta I_{1..n} + dO_{1..n} + e\Delta O_{1..n} + fE_{1..n} + g\Delta E_{1..n} + INT + Err$$

I = Intervention
O = Organization
E = Environment
INT = Interaction Among Interventions
A = Constant
Err = Error Term

This approach aims to measure changes in outcomes that are influenced by the delivery system reform interventions (I) contained in the ACA. The impact of these interventions will vary depending on the characteristics and capabilities of the organizations (O) that participate in these new programs—not only their initial capabilities but also their ability to adapt (Δ) their existing processes and methods over time. Observed impacts will also be influenced by the market environments (E) in which the Centers for Medicare & Medicaid Services (CMS) implement each intervention. These environments will also change over time (Δ) based on the actions of key actors, including private insurers and state policymakers. Finally, there will be significant interactions (INT) among the interventions to the extent that multiple initiatives take place simultaneously in geographic markets and that specific organizations participate in multiple initiatives.

Our general approach is divided into four general categories: (1) national view, (2) cross-demonstration perspective, (3) target populations, and (4) feedback reporting. First we consider a "main effects" approach that uses national data sets (e.g., claims, encounters or public health data) to look at aggregate trends in spending and outcomes. This approach could compare outcomes between demonstration and non-demonstration areas. Although it would have strong statistical power, the national model would provide limited insight into the role of organizational characteristics on observed changes.

Second we consider the impact of specific combinations of demonstrations. At this level, the analysis can also begin to integrate organizational features to assess the specific drivers of observed effects. Third, we assess the impact of ACA demonstrations on specific target populations. This subgroup analysis is a natural extension of the cross-demonstration work described above. Here we focus on the overall and demonstration specific effects of different ACA activities on specific populations, such as dual-eligible beneficiaries, those with complex chronic illness, or those approaching end of life. Finally, we propose approaches for monitoring outcomes over time to keep policymakers informed and help demonstration sites implement changes more rapidly in response to performance feedback.

Summary of Chapter 4: Organization and Market Variables

This chapter offers recommendations on organization and market variables that are likely to be associated with quality and cost performance for health care organizations participating in delivery reform initiatives. Collection of consistent organizational and market variables across the different HHS delivery reform initiatives may help evaluators distinguish between the impacts of HHS/ACA programs while controlling for the impact of organizational and environmental factors on observed changes in quality and cost performance. It may also help policymakers assess how program impacts vary across different types of organizations. This information may help target future programs toward environments where they are most likely to be successful and to design new programs that may be more effective for certain types of organizations and environments.

Through searches of the peer-reviewed literature, interviews with researchers in the health management field, and syntheses of publicly available organizational surveys, this section presents 53 organizational variables across eight categories:

1. Organizational structure and service capacity
2. Governance structure
3. Financial characteristics
4. Information technology and data management
5. Clinical process improvement capabilities
6. Culture, leadership and teamwork
7. Patient centeredness
8. Local market characteristics and state policy environment

For each variable, we provide rationale for its inclusion, provide sample wording of a similar survey item, and compare the advantages and disadvantages of each variable. Additionally, we suggest relevant data sources for these variables, including public and private databases, other Federal and state programs (e.g., Electronic Health Records [EHR] Incentive Program, National Committee for Quality Assurance [NCQA] Patient Centered Medical Home Survey and Accreditation), claims, administered organizational surveys, patient satisfaction surveys, as well as qualitative research, case studies, and ethnographies. Finally, the section highlights hurdles in the availability and use of these organizational variables, citing a lack of strong empirical research linking their effect on health outcomes, accommodating variation among the types of organizations, and gathering data on comparison groups.

Summary of Chapter 5: Review of National and Regional Data Sets and Data Collection Recommendations

This analysis of data sets focuses on how they can best support ACA evaluation activities. Fee-for-service claims data are a mainstay for evaluating health reform initiatives, in part because they are national and consistently collected for all beneficiaries over time. However, claims have many limitations, including sparse clinical detail that limits assessment of health outcomes. A variety of methodological issues are also associated with using claims to measure spending, particularly when comparing regions over time. The chapter assesses the benefits of investing in activities like price standardization, improved risk adjustment, and integrating Medicaid and Medicare data because these efforts would enhance the ability to use payment data for monitoring and evaluation activities.

Additionally, the chapter discusses the potential benefits of providing rapid feedback through claims databases and other data collection techniques. There is growing demand from health care providers participating in ACA initiatives to receive timely claims data that would help them track patients as they move through the delivery system, as close to real time as possible. Many hospitals participating in the ACO or bundled payment demonstrations, for example, want to know which patients use institutional care or emergency department services after leaving the hospital. Although institutions are requesting claims data for this purpose, it is less than ideal because of claims lag time. Efforts to better understand the impact of claims lag times are important, but likely insufficient to address the need for a beneficiary tracking system. Initiatives like the Beacon Communities Demonstration, where communities are testing new ways to use health information technology, may offer better solutions. This notion of real time (or virtually real

time) data is important to individual sites, but can also support broader evaluation activities, particularly with regard to modeling short-term effects.

Summary of Chapter 6: Framework for Evaluating the ACA Health Reform Initiatives as a Group

This chapter describes a continuum of quantitative methods for a coordinated evaluation plan to quantify the marginal effects associated with individual ACA reforms, designed individually and collectively to encourage improvements in health care delivery that may lead to improved beneficiary outcomes. This is a challenging task since ACA initiatives are complex, target overlapping goals, and take on reform in a dynamic environment with competing or complementary state, local, and private efforts. As a result, a coordinated evaluation of multiple ACA initiatives requires multiple levels of analysis. Specifically, the design needs to consider the effect on populations, organizations and the environment. In addition, we need to ask not only *Did an intervention work (enough)?* but also *Why did it work?* This chapter addresses these issues through a multi-level design that begins with main effects and moves on to drill down analyses that consider specific causal pathways and the impacts on specific populations.

At its most basic level, a main effects analysis determines whether there were statistically significant changes in health spending or outcomes within communities, states, and the nation as a whole. It is possible at this level that positive and negative effects are pooled, showing no change. In terms of national priorities, this is important information. The major focus is to propose an approach capable of capturing the effect of multiple interventions on multiple outcomes. In other words, any main effect model has to fulfill the condition of many-to-many causality relations from interventions to outcomes.

Another key objective is to model geographic hot-spots where there is significant reform activity. The initial main effect model can be adapted in response to hypotheses about the potential interactions between distinct reform initiatives. By enhancing the construct, the model can quantitatively consider the synergetic or even competing/contradicting effects of simultaneous interventions taking place within certain organizations or regions. The next step is to extend the modeling effort into sub-regions through either an integrated version of the macro-level main-effect model with fixed or random effects of regions as clusters, or conducting a series of sub-group analyses based on customized replica of the macro level model for individual sub-regions of interest.

Informed by results provided by the main effect models, Chapter 6 also considers the next level of evaluation questions focusing on the organizational or environmental determinants of success or failure, as well as the sustainability of observed effects over different time horizons. The chapter proposes modeling short-, medium-, and long-term effects to better understand the drivers of rapid improvement. Along the same lines, we propose modeling attainment, defined as achieving a significant amount of change, rather than looking only for improvement.

Summary of Chapter 7: Operational Considerations

Chapter 7 focuses on the challenges of aggregating disparate information from government databases and of collecting additional data from providers. Documenting the need for baseline data, especially on organizational characteristics and capabilities, the section emphasizes the need to: (1) prepare a synthesis of organization-related variables that CMS is currently collecting from provider organizations, (2) create a national registry of health care organizations, (3) be judicious in the collection of organizational variables, (4) utilize qualitative research to identify the most critical organizational factors, and (5) move quickly to establish baseline data.

CMS can greatly enhance the field of health services research by helping to map physicians (and their associated claims) to their parent healthcare organization.

Chapter 2: Introduction

With the ongoing implementation of the ACA, HHS is preparing to expand health insurance coverage to 30 million Americans. While this process may be challenging, the approaches for expanding coverage are well understood. HHS is simultaneously examining strategies to control the growth in health spending and improve quality for beneficiaries in federally sponsored health care programs through a series of new initiatives that encourage reforms in the delivery of care. Considerably less is known about the potential impact of these new payment and delivery reform initiatives, and HHS may invest substantial resources to evaluate the effects of these new programs, including the following:

- Pioneer Accountable Care Organization (ACO)
- Medicare Shared Savings Program
- Bundled Payment for Care Improvement
- State Multi-Payer Primary Care Demonstration
- FQHC Advance Primary Practice Demonstration
- Comprehensive Primary Care Initiative
- State Demonstrations to Integrate Care for Dual Eligible Individuals
- Hospital Value-Based Purchasing
- Medicare Readmission Reduction Initiative
- Community-Based Care Transitions Program
- Partnership for Patients
- Independence at Home Demonstration

These programs vary considerably in their scope, size, and target population; evaluating any one requires substantial planning and coordination. Yet, the Assistant Secretary for Planning and Evaluation (ASPE) also needs to consider the collective impact of these initiatives as a group and to anticipate the information that may be needed by HHS several years from now when designing subsequent phases of payment and delivery system reform initiatives. In that light, the primary goals of a coordinated evaluation across all of the HHS delivery reform initiatives are to determine: (1) the collective effect on cost, quality, and access, (2) which interventions (if any) drove the observed effect and should be prioritized for expansion to new geographic areas and populations, (3) whether particular combinations of interventions have a synergistic effect on driving performance improvement, and (4) which types of organizations have performed well under delivery reform initiatives and are therefore most appropriate for targeting future programs or expansions.

Under the guidance of a TEP, this evaluation plan aims to cohesively assess the effects of ACA interventions to inform policy. Specifically, it seeks to evaluate and attribute findings from individual delivery system reform initiatives and synthesize multiple interventions in a coordinated evaluation plan to inform the improvement of overall delivery system reform. After developing a conceptual analytical framework, the report reviews organizational variables and data sources to aid in this effort, and discusses rapid cycle collection efforts to provide timely feedback to both participants and policymakers. Lastly, this report provides recommendations for a coordinated evaluation design, combining quantitative and qualitative analyses, to aggregate

and attribute findings from a variety of individual Medicare, Medicaid, and Dual Eligible delivery system reform initiatives.

The report is organized as follows. Chapter 1 provides a summary of each section, highlighting the main points and summarizing key takeaways. This section, Chapter 2, provides an introduction to the project to review overall goals and methodology. Chapter 3 (Task 3) provides the conceptual framework for the analysis plan, outlining methodological approaches and key patient, organizational, and environmental influences in the analysis. Chapter 4 (Task 4) reviews the health care management literature and suggests nearly 50 organizational variables and their potential data sources. Chapter 5 (Task 5) reviews existing data sources and availability. The purpose of Chapter 6 (Task 6) is to identify, examine, and recommend methods for a coordinated evaluation plan to assess the extent to which ACA reforms improve overall health system delivery. Chapter 7 (Task 7) outlines operational considerations with respect to data collection and analyses that may substantially affect the implementation of the recommendations from the previous tasks. Finally, the appendix summarizes key points from the TEP meeting on June 5, 2012, and provides a list of the acronyms used in this report.

Chapter 3: Conceptual Analytic Framework

Developing recommendations for an approach to a coordinated evaluation design for HHS delivery reform programs begins with a conceptual analytic framework. Our analytic framework is divided into three principal sections: (1) general approach, (2) profiling delivery system reform initiatives, and (3) program evaluation strategy.

3.1 GENERAL APPROACH

Our general approach to this project evolves from the schematic presented in **Exhibit 3-1** and can be summarized in the following manner:

$$\Delta \text{ Outcome}_{1..n} = A + bI_{1..n} + c\Delta I_{1..n} + dO_{1..n} + e\Delta O_{1..n} + fE_{1..n} + g\Delta E_{1..n} + INT + Err$$

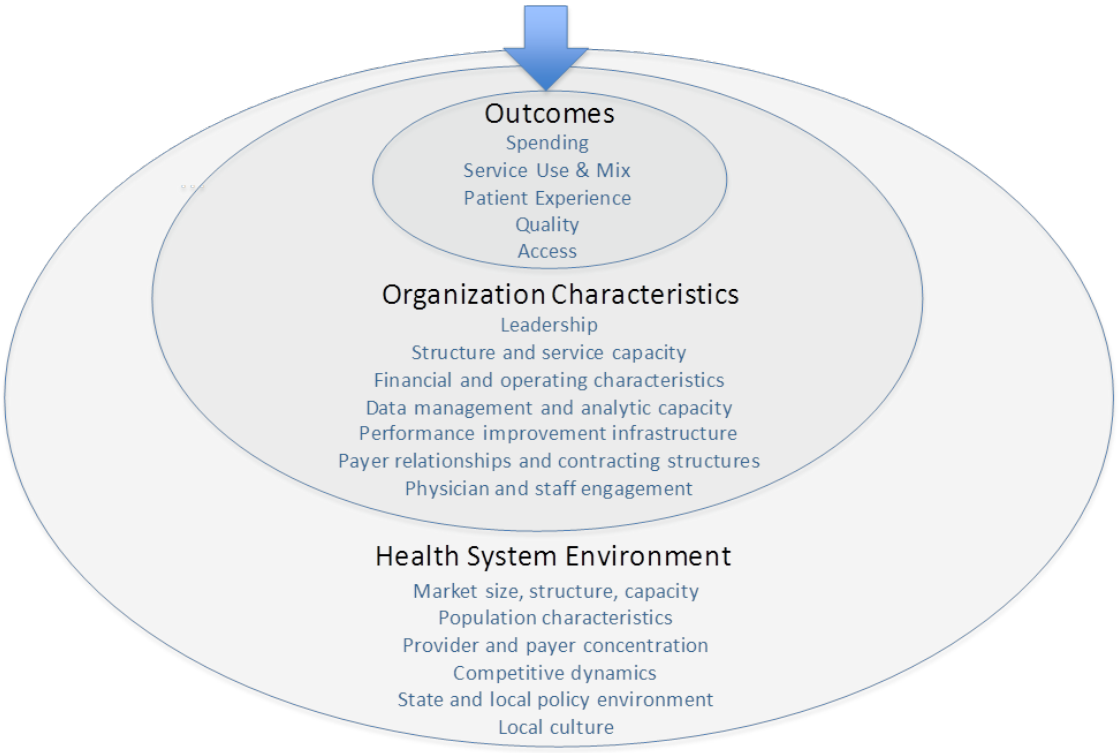
I = Intervention
O = Organization
E = Environment
INT = Interaction Among Interventions
A = Constant
Err = Error Term

Essentially, this approach aims to measure changes in outcomes that are influenced by the delivery system reform interventions (I) contained in ACA. The impact of these interventions may vary depending on the characteristics and capabilities of the organizations (O) that participate in these new programs—not only their initial capabilities but also their ability to adapt (Δ) their existing processes and methods over time. Observed impacts will also be influenced by the market environments (E) in which CMS implements each intervention. These environments will also change over time (Δ) based on the actions of key actors, including private insurers and state policymakers. Finally, there will be significant interactions (INT) among the interventions to the extent that multiple initiatives take place simultaneously in geographic markets and that specific organizations participate in multiple initiatives.

Therefore, the challenge of this project is to: (1) identify a consistent series of variables that meaningfully describe the organizational, market and policy conditions and dynamics that will affect the outcomes of the ACA's delivery reform initiatives, (2) present an approach for collecting these data, (3) assess the likely interactions among the ACA initiatives, and (4) prepare coordinated evaluation design recommendations based on these analyses.

Our primary approach to developing organizational and market variables, identifying viable data sources, assessing interactions, and developing coordinated recommendations is a combination of literature and document review, expert interviews, and consultation with a technical expert panel. The guidelines for this project required that we conduct our analysis using publicly available information. Although we interviewed some CMS staff as part of this project, the interviews were limited to programs where CMS had already awarded evaluation contracts. While we reviewed publicly available request for proposals for new CMS evaluations, we did not have access to materials for any awarded contracts.

Exhibit 3-1: Impact of ACA Delivery Reform Initiatives on Health System Change



3.2 PROFILING DELIVERY SYSTEM REFORM INITIATIVES

In moving from a conceptual framework to recommendations for a coordinated evaluation design, it was important to develop a clear understanding of the scope and nature of the new HHS delivery reform initiatives. We began by developing a delivery reform map that identified elements that were likely to influence the evaluation process across the different programs (see **Exhibit 3-2**). However, such a map is only a snapshot. During the course of this project, CMS announced a number of new program awards and many more details about the programs. We have updated these profiles throughout the project, but they are necessarily incomplete because of the dynamic nature of CMS program awards. Details of these profiles can be found in **Appendix A** and allows us to make a number of observations.

- **Providers—**A vast majority of the new delivery reform programs involve hospitals as the primary participant or as a core participant in a broader delivery network. Hospitals are the primary participants for the bundled payment pilot, hospital value-based purchasing, readmission reduction, and partnership for patients. Hospitals are likely to be central participants in the ACO and community care transitions program. Many hospitals may also be involved in state Patient-Centered Medical Home (PCMH) programs through their ownership of primary care practices.
- **Indirect impacts—**Providers that are not directly involved in the new programs may nonetheless be significantly affected. For example, improving coordination of post-acute care and reducing unnecessary spending in post-acute settings will be a major focus of the

bundled payment program. PCMH programs will have a major objective of reducing potentially avoidable hospital admissions and readmissions.

- **Geography—**Many ACA delivery reform programs are national in scope, but as voluntary programs, are likely to be concentrated in regions with certain characteristics—either those with high costs (i.e., significant opportunity for providers to achieve savings) or areas with high provider readiness (i.e., groups experienced with managed care). Once participants in the projects are announced, we expect it may lead to "hotspots" of activity. A key question is whether more detailed evaluations of hotspots may provide better information for policymakers than a broad evaluation across all geographies.
- **Project size/scope—**Programs like Medicare Shared Savings that cover millions of beneficiaries and tens of billions of Medicare payments will have greater impacts on aggregate outcomes. However, small programs, like Independence at Home, may have large impacts relative to their size. The evaluation methodology will have to consider such size and scope differences.
- **Timing—**Most of the programs for evaluation may begin in 2012 and 2013 and continue for 3 to 5 years. Some of the programs (i.e., Multipurpose Senior Services Program [MSSP]) are permanent and will continue beyond their initial contract period. Other programs (e.g., PCMH) extend and expand state programs that were in existence prior to the ACA. It is also important to consider the time that providers will require to ramp up their capabilities to perform successfully in these programs. This implies that the strongest effects of these programs may well occur beyond their original program period.
- **Beneficiary type**—While some programs are focused on discrete groups of Medicare beneficiaries (frail elderly, dual-eligibles, and those with multiple chronic conditions) others cover Medicare beneficiaries more generally. Nevertheless, even in broader programs like MSSP, organizations will focus resources on managing the most frail, complex, and high-risk patients. Their success in so doing may be important for achieving improvements in more aggregate outcome measures.

Exhibit 3-2: Delivery Reform Initiative Analytic Map

	Participating Providers
Medicare Shared Savings Regular, Pioneer Advanced Pmt.	1. Integrated Delivery Systems 2. Multi-specialty groups 3. PHOs 4. IPAs
Bundled Payment	1. Hospitals 2. PHOs 3. IDNs
Hospital VBP	1. Hospitals
Readmission Reduction	1. Hospitals (in bottom quartile)
State Dual Eligible	NA – but probably full range of providers
State PCMH	1. PCP practices 2. Multi-spec practices 3. Hospital clinics 4. FQHCs
Partnership for Patients	1. Hospitals & and hospital systems (large systems)
Community Care Transitions	1. Partnerships between hospitals and community based organization, CBOs)
Independence at Home	NA – Range of providers able to deliver home based primary care. (Probably hospitals in partnerships with others).

	Impacts	
	Direct	**Indirect**
Medicare Shared Savings Regular, Pioneer Advanced Pmt.	Integrated delivery systems	Hospitals; Medical specialists; Post-acute care providers
Bundled Payment	Hospitals PHOs	Non-affiliated physicians; Post-acute care providers
Hospital VBP	Hospitals	NA
Readmission Reduction	Hospitals	Post-acute care providers; Primary care providers
State Dual Eligible	Integrated delivery systems	
State PCMH	Primary care delivery systems	Hospitals, post-acute care providers, specialists, ancillary service providers
Partnership for Patients	Hospitals and hospital systems	NA
Community Care Transitions	Hospital and community based organization partnerships	Non-participating hospitals, post-acute care providers.
Independence at Home	Physicians and care-coordination partners	Hospitals, post-acute care providers.

Exhibit 3-2: Delivery Reform Initiative Analytic Map (continued)

	Geography
Medicare Shared Savings Regular, Pioneer Advanced Pmt.	National – need award data to determine geographic concentrations
Bundled Payment	National – need award data to determine geographic concentrations
Hospital VBP	National – All hospitals
Readmission Reduction	National – 25 percent of hospitals
State Dual Eligible	15 $1M design contracts to states. Hope over time to enroll as many as 25 states.
State PCMH	Statewide: MI, MN, VT Selected Areas: ME, NY, PA, NC, RI
Partnership for Patients	National – likely concentration based on hospital system location
Community Care Transitions	National – Concentrated in communities with appropriate partnerships
Independence at Home	National - Physician practices across the country

	Timing
Medicare Shared Savings Regular, Pioneer Advanced Pmt.	January 1, 2012 to January 1, 2015 with options to January 1, 2017 Initial contracts 3 years with 2 options Program duration = permanent
Bundled Payment	January 1, 2013 – January 1, 2016
Hospital VBP	2013 (1 percent withhold) to 2017 (2 percent withhold)– permanent
Readmission Reduction	2013 (1 percent) – 2014 (2 percent) – 2015 (3 percent) and beyond
State Dual Eligible	Target enrollment date of 1/1/2013 for 3-year program
State PCMH	3-year program. VT/NY/RI: 2011–13; NC/MN 2012–12; ME/ME/PA 2013–15
Partnership for Patients	Starting 1/1/12 - 2 years + one option period
Community Care Transitions	Start in 2012. Applications accepted on a rolling basis.
Independence at Home	Start 2012 and operates for 3 years

Exhibit 3-2: Delivery Reform Initiative Analytic Map (continued)

	Scope/Size
Medicare Shared Savings Regular, Pioneer Advanced Pmt.	Total dollars for attributed beneficiaries x number of ACOs x beneficiaries per ACO. 30 Pioneer @30,000 members + 100 ACO @ 15,000 members (@10,000/member) = $22.5 billion/yr
Bundled Payment	Average value of selected bundles x number of hospitals x 2-5 episodes/5-25 DRGs x 300 hospitals = Total dollars
Hospital VBP	2 percent of DRG by 2017 x all hospitals = max payout is a percent of DRG
Readmission Reduction	3 percent of DRG x 25% of hospitals
State Dual Eligible	Design contracts awarded to 15 states. Number of beneficiaries will depend on final state proposals and awards (forthcoming). CMS goal of targeting up to 2 million duals.
State PCMH	8 states. States estimate approximately 1,200 PCP practices and 1 million Beneficiaries. CMS estimates maximum of 700K – 800K beneficiaries.
Partnership for Patients	$218 million – improvement support grants, not 'patient revenue.' Awards to 26 "Hospital Engagement Networks" that according to CMS encompass 80 percent of total Medicare discharges.
Community Care Transitions	Participants enrolled on a rolling basis. Currently have 7 awarded sites. Will continue to award sites until reaching the funding ceiling ($500M paid based on a "per-eligible discharge rate" to participating sites). Anticipate awarding around 70 partnerships (with around 350 hospitals),
Independence at Home	Limited to 10,000 beneficiaries

	Beneficiary Characteristics
Medicare Shared Savings Regular, Pioneer Advanced Pmt.	Population based – all beneficiaries in target areas
Bundled Payment	Beneficiaries with target admissions (defined by DRG) at target hospitals
Hospital VBP	All beneficiaries with a hospital admission.
Readmission Reduction	Beneficiaries with a hospital admission
State Dual Eligible	Dual eligible beneficiaries in target states
State PCMH	Beneficiaries who receive care from a participating practice
Partnership for Patients	Beneficiaries with a hospitalization who are at risk for adverse outcomes
Community Care Transitions	High risk beneficiaries with a hospitalization
Independence at Home	Community based high need populations

3.3 EVALUATION STRATEGY

Our general approach to evaluation is divided into four general categories: (1) national view, (2) cross-demonstration perspective, (3) target populations, and (4) feedback reporting.

3.3.1 Level I: The National View

Our evaluation strategy will consider the impact of ACA demonstrations at multiple levels. As national legislation, our initial question is whether the demonstrations as a group have an observable impact on cost, quality, and access. This is a main effects analysis that assumes some theoretical coherence across all demonstrations. It is probably fair to characterize the theoretical underpinning of the ACA as a series of efforts to use changes in incentives—combined with expanded data availability and targeted technical assistance—to drive changes in delivery systems and ultimately the practice/culture of medicine. This means that an overall evaluation of ACA demonstrations needs to consider sustained changes in the practice of medicine in addition to observable changes in cost, quality, and access.

The basic analytic approach will involve the use of national data (e.g., claims, encounters, or public health data) to look at national trends. It may also be possible to compare outcomes between demonstration and non-demonstration areas. Changes in the delivery of care may be captured in claims (e.g., service delivery patterns), and though assessment of organizational changes in health care organizations participating in these programs that must be captured through surveys and qualitative data (e.g., site visits).

This main effects level needs to consider the impact on individuals and populations, as well as the impact on organizations. This is best done using mixed methods designs that integrate quantitative and qualitative analysis in a systematic way.

3.3.2 Level II: The Demonstration and Cross-Demonstration Perspective

One level down, we will consider the impact of individual demonstrations or specific combinations of demonstrations. Here, we will recommend that HHS look much more closely at specific evaluation designs for areas of congruence. Although we do not have access to the evaluation designs currently under development for these new programs, we would recommend that HHS abstract study designs for each demonstration to understand the use of experimental, quasi-experimental, and observational approaches. It will also be important to map primary and secondary outcomes. Efforts with similar outcomes or designs may lend themselves to quantitative comparisons like meta-analysis or pooling data. Results from demonstrations focused on readmissions reductions as an outcome, for example, could be combined to understand the complete range of effect sizes, rather than a simple national average.

At level II, the analysis can also begin to integrate organizational features to assess the specific drivers of observed effects. This presents a number of challenges, including limited measure of organizational and contextual factors, limited availability of organizational data nationally (not just in treatment areas), multiple treatment models within a given demonstration, and overlapping interventions within a given market area.

Untangling this complexity will require a detailed understanding of the context, organizational capacity, and hypothesized or desired change process (e.g., changes in financial incentives leading to changes in care coordination intended to reduce readmissions). Armed with this information, it then becomes possible to develop heuristic models that can be tested empirically or intervention typologies that become inputs into a dynamic modeling process.

At this stage, it is important to be aware of policy priorities and the ultimate goals of the evaluation. Narrowly focusing on program models, for example, may miss the broader capacity of organizational changes to support the ability to improve performance, mange financial risk, or administer new programs.

3.3.3 Level III: Target Populations and Drivers of Change

Finally, we will consider the impact of ACA demonstrations on specific target populations. This subgroup analysis is a natural extension of the cross-demonstration or cross-site work described above. Here we will focus on the overall and demonstration-specific effects of different ACA activities on specific populations, such as dual-eligibles, those with complex chronic illness, or those approaching end of life. The goal is to drill down and better understand which interventions are effective for which population subgroups. Given geographic and program design diversity, it will once again be important to consider the role of delivery system and contextual factors on population outcomes. For example, *Does HMO penetration have an impact on the design and outcomes of state dual eligible demonstrations?*

3.3.4 Level IV: Feedback Reporting

In addition to overlapping geographically, ACA demonstrations also have complex timing issues. Programs start at different times and many are likely to have different timelines before achieving a maximal effect. Monitoring outcomes over time has two potential benefits. In addition to keeping implementation staff and policymakers informed, demonstration sites themselves will benefit from understanding the changes in outcomes over time.

There are a number of challenges associated with doing this well. For example, claims-based measures face claims maturity issues—payment systems were not designed for surveillance and monitoring, so it is important to understand the type of 'noise' introduced by billing processes. Quantifying the billing bias may allow for the calculation of measures using younger data. It will also be important to provide demonstration participants with intermediate outcomes, such as information collected from site visits or consumer surveys. Market research may provide valuable clues about how small sample information can be used to inform decisionmaking.

Chapter 4: Organization and Market Variable Recommendations

4.1 OVERVIEW

In this chapter, we offer recommendations on organization and market variables that we believe are likely to be associated with quality and cost performance for health care organizations participating in HHS delivery reform initiatives including those created by the ACA. The overall goal of this project is to provide recommendations for a coordinated evaluation design that would tie results from the numerous Medicare, Medicaid and Dual Eligible delivery system reform initiatives and ACA provisions into an overall model of health system change.

The impact of the new interventions may vary depending on the characteristics and capabilities of the organizations that participate in these initiatives—not only their initial capabilities but also their ability to adapt processes and methods over time. The market environments in which CMS implements each intervention may also influence observed changes in quality and cost performance and these environments may also change over time based on the actions of key actors, including private insurers and state policymakers.

Collection of consistent organizational and market variables across the different HHS delivery reform initiatives will help evaluators distinguish between the impacts of the HHS/ACA programs while controlling for the impact of organization and environmental factors on observed changes in quality and cost performance. It may also help policymakers assess how program impacts vary across different types of organizations. This information may be helpful in targeting future programs towards environments where they are most likely to be successful and to designing new programs that may be more effective for certain types of organizations and environments with below average performance.

4.2 APPROACH

Our approach to developing organizational and market variable recommendations included the following steps. First we performed a literature search. Next, we reviewed a range of organizational survey instruments that have been used across a variety of studies. We reviewed several organizational readiness assessments that were provided to us by private organizations on the condition that we only use them as background information. We also interviewed individuals involved in health care organization survey research as well as health system executives (identified below) to obtain feedback about the types of variables they believe are most important for performance on cost and quality outcomes.

4.2.1 Literature Search

A full meta-analysis of the literature is beyond the scope of this project. But to gain an understanding of the field's breadth, key articles, and key academics, we conducted a search within the Information Sciences Institute (ISI) Web of Science under the term *health care organization* for the years 2000-2012 (n>16,300 articles). We further limited the results to the sub-categories of *health care science services* and *health policy services* (n=3,892 articles) to define the discipline of health care management and organization. We examined the articles in this set that were most cited and explored the methods by which they collected their data,

including national surveys. We also documented the most frequent authors, which we utilized to determine key informants to interview.

The literature review showed that the discipline is rife with attempts to implement novel health care interventions, many with only modest or partial success. In many cases, the lack of adoption success stems from a failure of cultural or organizational receptiveness, rather than a shortcoming of the technical aspects of the intervention itself (Berwick, 2003; Rogers, 1995). Attention must be paid to how individual, group/team, organization, and larger environment/system level variables affect reform efforts (Ferlie & Shortell, 2001), and how these factors relate to the specific characteristics of the intervention (Grol & Grimshaw, 2003).

The literature provides some guidance on market variables that are associated with quality and cost performance. Many analyses employ organizational variables as a control (e.g., organizational size or academic affiliation), but don't explicitly emphasize these components. Others specifically examine how managerial and market forces affect health care. While enumerating all of these studies would be too cumbersome to include here, some well-known examples include:

- Successful implementation of quality improvement efforts is dependent on culture and leadership of the organization (Shortell et al., 1995; Weiner, Shortell, & Alexander, 1997)
- Health outcomes are strongly related to the volume of activity in an organization in which enhanced organizational experience, learning curves, and focused routines lead to centers of excellence across many disease conditions (e.g., Birkmeyer et al., 2005 in cancer; Nguyen et al., 2004 in bariatric surgery).
- Provider capacity affects spending levels (i.e., higher relative supply of hospitals and specialist physicians is associated with higher spending in geographic markets; higher relative supply of primary care physicians is associated with lower spending; The Dartmouth Atlas Group)
- Health Maintenance Organization (HMO) penetration was significantly associated with lower hospital and Medicare costs in the 1990s (numerous articles).
- Health costs and quality can vary substantially between for-profit and not-for-profit hospitals (Jha et al., 2005), nursing homes (Harrington et al., 2001), or insurers (Himmelstein et al., 1999).

Yet for all these papers, the use of organizational and market variables to explain observations about health care performance has been haphazard. Much of this literature simply documents the presence of technology or processes in organizations that are thought to contribute to improved performance on spending and quality, but does not tie these variables to actual performance (e.g., Casalino et al., 2003). Alternatively, when research has documented superior performance for certain types of organizations, it is reported as a statistical association with coarse metrics of overall organizational size or structure rather than the impact of discrete organizational characteristics (e.g., Weeks et al., 2010).

Very few researchers have put these perspectives together, seeking to determine not just 'what works,' but to understand 'what works and why' (Damshroeder et al., 2009). Some researchers

are actively exploring this connection. For example, the U.S. Department of Veterans Affairs Quality Enhancement Research Initiative (QUERI) has begun incorporating organizational factors into their research plans. They emphasize the importance of understanding how organizational structure and processes can affect the ability to identify and implement best practices (Yano, 2008). However, these efforts are still in the early stages.

4.2.2 Interviews

We interviewed several prominent health services researchers to discuss how organizational and market variables are likely to influence health care outcomes and potential approaches to collecting this information. We spoke with the following individuals; their rank as the most frequent U.S.-based authors in the health care organization and management discipline is indicated in parentheses.

- Stephen Shortell, PhD., University of California, Berkeley (#1)
- Constance Horgan, PhD, (#5), Dominic Hodgkin PhD, (#6), Elizabeth Merrick PhD, (#7), Brandeis University (joint interview)
- Lawrence Casalino , MD, PhD, Weill Cornell University Medical College (#17)
- Sara Singer, PhD, MBA, Harvard School of Public Health (>#50)

We also interviewed several health care executives to ensure that these theories and approaches were meaningful to real-world decisionmakers.

- Francis J Crosson, MD, former Executive Director of The Permanente Federation and Senior Fellow, Kaiser-Permanente Institute for Health Policy
- Thomas Graf, MD, Chairman, Community Practice Service Line, Geisinger Health System
- Dana Safran, ScD, Senior Vice President. Performance Measurement & Improvement, Blue Cross/Blue Shield of Massachusetts

4.2.3 Survey Instruments

Lastly, we reviewed a range of survey instruments that health services research have developed to evaluate health care organizations. We focused on publicly available (or soon-to-be) surveys, which provide examples of how organizational variables have been systematically collected across multiple health care entities.

- **National Study of Physician Organizations and the Management of Chronic Illness (NSPO I**), Stephen M. Shortell et al., School of Public Health, University of California-Berkeley, 2000-2001.
- **Competing Values Framework (CVF),** developed for healthcare by Stephen Shortell et al., RAND Improving Chronic Illness Care Evaluation; adapted by the Veterans Health Administration. Source: 2004 VHA All Employee Survey.
- **A National Survey of Health Record Keeping among Physicians & Group Practices in the United States**, Catherine DesRoches et al., Institute for Health Policy, Massachusetts General Hospital, 2008.
- **Community Tracking Study, Survey of Physicians**, Center for Studying Health Systems Change, 2008.

- **Learning Organization Survey,** David A. Garvin, Amy C. Edmondson, and Francesca Gino, Harvard Business School, 2008.
- **Use of Electronic Health Records in U.S. Hospitals,** Ashish Jha et al., Department of Health Policy and Management, Harvard School of Public Health, 2009.
- **National Study of Small and Medium-Sized Physician Practices (NSSMPP),** Lawrence Casalino, Weill Medical College, Cornell University, 2009.
- **2009 Commonwealth Fund National Survey of Federally Qualified Health Centers**, Commonwealth Fund – HarrisInteractive, 2009.
- **2010 AHA Annual Survey,** American Hospital Association, 2010.
- **Survey of Risk-Based Contracting and Physician Compensation in Organized Delivery Systems**, Robert Mechanic and Darren Zinner, Heller School for Social Policy and Management, Brandeis University, 2011.
- **Medical Group Compensation and Financial Survey,** American Medical Group Association, 2012.
- **Relational Coordination Survey**, Jody Hoffer Gittel, Heller School for Social Policy and Management, Brandeis University.
- **Health Systems Integration Study Questionnaire**, Robin Gilles et al., Northwestern University, 1996.
- **National Survey of Physician Organizations III (NSPO III),** Stephen M. Shortell et al., School of Public Health, University of California-Berkeley, 2012 (DRAFT).

The literature review, interviews, and surveys described above are used as the foundation for our organization and market variable recommendations.

4.3 TAXONOMY OF PROVIDERS

New HHS delivery reform initiatives will directly or indirectly influence a full range of health care providers and organizations. **Exhibit 4-1** provides context for setting priorities about the types of providers and related organizational variables. This exhibit illustrates that the delivery reform initiatives target a wide range of provider categories, but that the majority of programs are focused on improving cost and quality performance across a continuum of care rather than for specific categories of service (e.g., hospital quality only). For example, in Model 2 of the CMS bundled payment program, applicants will be financially responsible for episodes of care that begin with admission to the hospital and include 30–180 days of post-acute care services. Since, the majority of spending for many DRG-based episodes occurs in the post-acute care setting (RTI International, 2011), participants in this program (many of which are hospitals) will have to establish new systems for coordinating services outside of the hospital.

In keeping with the HHS focus on care coordination across the continuum, we began our assessment of organizational variables with a focus on integrated delivery systems, multi-specialty physician groups, and contracting entities like physician-hospital organizations that can take responsibility for managing patient populations. We recognize that a wide range of organizations (hospitals, physician groups, post-acute care providers etc.) will participate in these

programs. However, specifying organizational variables for each type of provider is not possible within our charge of developing 25–40 variable recommendations.

We recognize that evaluators will frequently want to collect additional information from individual providers (e.g., hospitals) within integrated networks. Many participants in the HHS delivery reform programs are not integrated providers. But the ability of freestanding providers to successfully improve coordination across the continuum of care will depend on the nature of their contractual and informal relationships with other components of the delivery system. Therefore, although we have concentrated our recommended variables on physician-centered integrated delivery systems, many of the variables we propose can be adapted for freestanding providers.

Exhibit 4-1: Overview of Key Organizations Participating in HHS Delivery Reform Initiatives

Program	Focus of Performance Improvement	Primary Provider Organizations Targeted in Program
Shared Savings and Pioneer ACO	Continuum of Care	Integrated Delivery Systems (IDN) Multi-specialty physician groups Contracting groups (Physician-hospital organizations, independent practice associations)
Bundled Payment	Continuum of Care	Hospitals, PHOs, IDNs, Post-Acute Care Providers
Hospital Value-Based Purchasing (VBP)	Hospital care expanding to continuum of care	Hospitals
Readmission Reduction	Continuum of Care	Hospitals (in bottom quartile)
State Dual Eligible	Continuum of Care	Range of provider organizations depending on state program.
Multi-payer Advanced Primary Care Practice Demonstration	Continuum of Care	Primary care physician practices Multi-specialty practices Hospital clinics
FQHC Advanced Primary Care Practice Demo.	Continuum of Care	Federally Qualified Health Centers
Partnership for Patients	Hospital care	Hospitals and hospital systems
Community-Based Care Transitions	Continuum of Care	Hospitals working with community based organizations (CBOs)
Independence at Home	Continuum of Care	Range of providers able to deliver home based primary care. (Likely hospitals in partnerships with others).

4.4 ORGANIZATIONAL AND MARKET VARIABLE RECOMMENDATIONS

We have developed a taxonomy in which the organizational and market variable recommendations are divided into eight categories. This section offers a brief discussion of each category. Detailed recommendations are provided in **Exhibit 4-2** where for each item, we provide a definition and an example of how a data request might be worded,

drawn from existing surveys or the literature. The rationale for selecting individual variables is provided in **Exhibit 4-3**. Principal advantages and disadvantages are listed in **Exhibit 4-4**. We have chosen the following categories to organize our variable recommendations.

1. Organizational structure and service capacity
2. Governance structure
3. Financial characteristics
4. Information technology and data management
5. Clinical process improvement capabilities
6. Culture, leadership and teamwork
7. Patient centeredness
8. Local market characteristics and state policy environment

4.4.1 Organizational Structure and Service Capacity

General organizational variables are needed to characterize the organizations participating in the various HHS delivery reform programs. This includes the type of organization, size, and composition of providers, services, and health care personnel. Size variables may affect program implementation efforts by allowing organizations to more easily absorb new overhead expenses (e.g., new personnel, administration, and coordination programs) across the organization. The scope of an organization and its ability to control or integrate services across the continuum of care may improve groups' ability to manage care transitions. An important aspect of organizations is the degree of integration and alignment with physicians. Therefore variables that indicate the model of physician affiliation with the organization and methods of compensating physicians are critical.

4.4.2 Governance

The organization's ownership model (public, non-for-profit, for-profit) may influence the level and nature of investments that organizations are willing and able to direct towards delivery reform initiatives as well as the priority they place on achieving certain outcomes. The composition of the board and the level of participation by different stakeholders (physicians, managers, community members, patients) will influence the willingness of organizations to try and implement changes.

4.4.3 Financial, Payer, and Reimbursement Model Characteristics

The payer mix and reimbursement contracts of health care organizations are an indicator of their experience with managed care and alternative (non-fee-for-service) reimbursement models, experience that may influence their performance under HHS delivery reform programs. Understanding each organization's market share is also important, as a dominant market share may lead to an emphasis on revenue enhancement through private insurer price negotiation rather than reducing the cost of delivering medical services. Finally, organizational profitability and capital reserves are an indication of resources available to invest in the infrastructure necessary for improving delivery system performance.

4.4.4 Information Technology and Data Management

One common attribute of successful delivery systems is their ability to use clinical and financial data to measure performance, standardize processes, and increase the quality, efficiency, and reliability of care through timely feedback of information to clinicians, managers, and patients. To improve care delivery, an organization must first have a strong understanding of its baseline processes and outcomes. Under the axiom, *you manage what you measure*, health information technology allows an organization to develop reports on the care of specific providers, clinics, or groups, and for individual patients or subsets of clinically-relevant patients (i.e., a patient registry). With this data, organizations can identify areas of high variance, implement programs to standardize care processes based on clinical evidence, and develop programs and incentives to reward quality, efficiency, and patient-centeredness. This category of variables includes information about the existence of infrastructure, like electronic medical records, enterprise-wide data repositories, and analytic software tools. More importantly it will include variables that measure the functionality of these tools, including discrete categories of clinical decision support.

4.4.5 Clinical Process Improvement Capacity

This category of variables examines the formal commitment of organizations to a management method for achieving improved performance. It is designed to capture specific investments in technologies, processes, and personnel to increase quality, support high-risk patients, and improve handoffs and care transitions.

4.4.6 Culture, Leadership, and Teamwork

New HHS initiatives require improved coordination across settings—whether among divisions of an organization or across external health care partners. Doing so effectively will require effective work processes and provider relationships. The success of organizations in these pilots may depend upon how well the leadership and frontline workers within organizations can adapt. To that end, we have included several variables on organizational culture and leadership, seeking to examine whether the organization emphasizes a supportive learning environment and has established modes of working that allow for experimentation and analysis of those experiments. In short, we seek to measure the extent to which these groups are *learning organizations.*

4.4.7 Patient Centeredness

New HHS delivery reform programs have emphasized the need for patient-centered care. Therefore, we have recommended several variables that indicate organizational efforts to give patients better access to information and engage them in shared decisionmaking. These reflect desired attributes of the patient-centered medical home.

4.4.8 Local Market Characteristics

Organizational performance on Medicare and Medicaid initiatives will be influenced by local market factors that either reinforce or confound the incentives and objectives of the HHS delivery reform programs. The three principal factors in this category are provider market

conditions, insurer market conditions, and state policy environment. We expect provider consolidation in local markets to have an inverse relationship to the rate of performance improvement. Providers that can easily raise prices face far less financial pressure for restructuring activity. In contrast, local insurance markets consisting of strong plans that are also committed to payment innovation will reinforce the financial incentives created by the HHS programs. Finally, aggressive state policy can critically affect observed rates of spending growth and changes in quality. We believe it is critical that program evaluators be cognizant of the impact of state health insurance exchanges. States that enroll large populations of both subsidized and commercial enrollees and that direct their exchanges to take aggressive positions on limiting growth in health insurance premiums will substantially affect the incentives that flow through to the delivery system.

4.5 DATA SOURCES

The organizational variables listed in **Exhibit 4-2** can be generated from a variety of sources. In this section, we describe several ways in which these variables can be populated, discussing the pros and cons of each method. In **Exhibit 4-5**, we match each variable to the most likely source.[1]

Publicly Available Datasets: Several existing public databases already contain aspects of the variables listed in **Exhibit 4-2**. These data are economical because they have already been collected, cleaned, and (often) vetted. AHRQ's Nationwide Inpatient Sample, for example, includes information on hospital payer mix, patient demographics, and hospital characteristics, such as ownership size and teaching status. The 2009 survey includes 1,050 hospitals from 44 States. AHRQ's Healthcare Cost and Utilization Project (HCUP) maintains similar surveys of emergency rooms and ambulatory surgical centers.

However, most public data sets are focused on specific categories of providers and do not contain information about organizations that offer or contract for a wide range of services across the continuum of care. For example, the AHRQ's Medical Expenditure Panel Survey (MEPS) may have information about certain healthcare organizations, including hospitals, office-based providers, home-care providers, pharmacies, and other caregivers. The MEPS Medical Provider Component, which helps estimate the exact cost of care, includes questions about the index patient's insurance and the use of capitation, but it is unlikely to have a large enough sample in any given provider group to allow CMS to draw conclusions. Similarly, CMS Hospital Compare provides information on hospitals' process of care, outcomes of care for select conditions (e.g., heart attacks, pneumonia), readmission rates, patient safety measures, and patients' perceptions of care. In the process, these institutions are required to maintain a patient registry for these conditions. However, these data are collected for Medicare patients only and some variables are available only for hospitals that have submitted this information voluntarily.

[1] It should be noted that the nature of this project inherently biases toward specific sources of data. By requesting specific organizational variables that are associated with implementation success and health outcomes, the exercise discriminates against more open-ended, qualitative methods like case studies and ethnographies. Moreover, we have focused our attention on prior efforts that had collected organizational information across multiple health care groups, under the assumption that any future CMS work would need to be conducted simultaneously and economically. Thus, when choosing the variables, we based our recommendations on previous surveys and other coordinated data collection efforts, further biasing the results to specific data sources.

Private Datasets: Data for several variables in **Exhibit 4-2** can be found in private datasets. The American Hospital Association's (AHA's) annual survey of hospitals provides information on size and type of facilities. HealthLeaders-Interstudy provides managed care penetration at the state and local level. The American Medical Group Association and Council for Accountable Physician Practices regularly survey their members on aspects of care coordination and market influences. These data sets vary considerably in their scope, cost, and availability to outside researchers. We know of no data sets that provide comprehensive information about organizations that offer or contract for a wide range of services across the continuum of care.

Data Collected for Other Federal and State Programs: Beyond the data collected for individual program evaluations, the Federal Government collects information in conjunction for accreditation and evaluation of other programs. In addition, individual states also systematically register providers, facilities, and other groups (a complete list of state resources is beyond the scope of this project). These have a strong advantage in that providers are often required by law to comply and thus are comprehensive of all health care groups. It is unknown how readily researchers could access these data for purposes beyond its original charter. We highlight two programs here that are especially relevant to the variables in Exhibit 2:

- **EHR Incentive Program**: As a consequence for meaningful use requirements, CMS has annual data on the implementation and functionality of electronic health records within physician offices and hospitals. For example, to be certified as a system, the program requires these providers to meet goals of health information technology systems in terms of data recording (e.g., active medications, height/weight, automated order-entry, patient problem lists), preliminary decision support (e.g, formulary drug checks), and communications (e.g., electronic summaries to patients, summary care record to other providers, immunization summaries to public health agencies). Many of the specific requirements are similar to variables listed in the *Information Technology and Data Management* variable subgroup.
- **NCQA Patient Centered Medical Home (PCMH) Survey and Accreditation**: The PCMH survey queries provider groups on a host of issues similar to those in **Exhibit 4-2**. These include: use of data for population management, ability to identify high-risk patients, coordination with facilities and care transitions, implementing/demonstrating continuous quality improvement, measuring/reporting physician performance, and use of certified EHR technology.

Claims: Medicare claims data can be used to help determine patient characteristics and initial health status. Medicare claims are received from specific types of providers and do not contain information about provider affiliation (e.g., provider identifiers for all physicians and facilities associated with a single organization). In fact, no such database of provider affiliations exists, although CMS will likely compile such information for applicants or participants in its various delivery reform programs. This may be helpful in linking outcomes to organizations but will provide only limited information about organizational characteristics.

Organizational Surveys: Organizational surveys have the benefit of allowing researchers to craft specific items to answer the questions at hand. Often, these surveys can be self- or

electronically-administered, greatly reducing the cost and personnel involved in data collection. Common questionnaires can be standardized across provider groups, allowing for precise comparisons across organizations on an exact set of responses. Standardization, however, also creates an important drawback for surveys because they may not be adaptable to certain settings or changes in settings over time. Generating a common survey instrument may require over-generalizing constructs so that they are appropriate for multiple types of organizations, potentially missing key variables or context that are most relevant for any given group. Finally, surveys require the researchers to garner sufficient response rates—a task that can demand significant time and resources.

- **Executive-level surveys**: For many variables, questions can be targeted to senior or executive leadership. Indeed, for questions regarding insurer contracts or market share, only senior-level executives will have that information. Executive-level surveys greatly speed and simplify data collection efforts, allowing for relatively small sets of questionnaires to represent a very large organization. While this makes survey administration more economical, these surveys also encompass several sources of bias. For instance, senior leaders or organizations working with CMS may answer questions that put them in a favorable light (self-response bias). Additionally, the persons who answer the survey are being asked to represent the entire organization, which may not always be accurate (responder bias). Follow-up interviews with executives can help researchers better understand the magnitude of some of these biases, as well as create a venue to explore the context of individual organizations. But such methods increase on the cost and time to complete the data collection.
- **Employee-level surveys**: For some organizational variables, especially culture and leadership domains, single- or small-group responses will not provide adequate information. For example, the Competing Values Framework requires multiple respondents at each organization to respond—a methodology that may be difficult to implement systematically across CMS programs. These large-scale employment surveys are significantly more expensive and time-consuming to administer.

Qualitative Research, Case Studies, and Ethnographies: Indepth, qualitative research has the advantage of providing a holistic examination of the complex mix of factors that promote successful implementation of HHS initiatives. It allows for a more nuanced approach, allowing for contextual factors to emanate organically without being influenced by rigidly defined, standardized variables. However, these methods are expensive and time-consuming, especially in terms of personnel costs. Researcher bias cannot be avoided, and the results are difficult to replicate. Further, qualitative research, without p-values and rigorous methodologies, may not be convincing to quantitative researchers or organizations like the Congressional Budget Office that are responsible for scoring the impact of program initiatives.

Other Sources: Some of the variables in **Exhibit 4-2** cannot be fully determined from the categories listed above. Most prominently, the subcategory of *Patient Centeredness* can only be approximated through surveys or interviews of the organization's employees. Researchers who truly want to measure this concept would also want to conduct patient surveys about their use of shared-decisionmaking, open-scheduling, and electronic access to communications and EHRs. Additionally, the data sources for many of the variables in the subcategory of *Local*

Market Characteristics and State Policy Environment do not exist yet. New research may be required to understand how state goals and local community groups aid or hinder in the implementation of HHS initiatives.

4.6 ORGANIZATIONAL VARIABLES: RESEARCH AND DATA GAPS

The variables, rationale, and data sources in this chapter provide our assessment of existing research and of the opportunity for developing organizational variables likely to be associated with performance for providers participating in the new HHS delivery reform initiatives. However, the overarching goal of this project is larger than the research that we have used to develop our recommendations. This section briefly explores critical gaps in the research and data regarding organizational variables and their relationship to organizational performance.

Lack of Empirical Research Linking Organizational Variables to Performance. The principal objective of HHS delivery reform initiatives is to improve health care system performance on the three-part aim: individual health care, population health, and health care spending (Berwick, Nolan, and Whittington, 2008). However, the literature lacks systematic evaluation of the empirical relationship between organizational variables and performance on broad measures of the triple aim. There is a broad literature examining the impact of discrete interventions in discrete settings on particular outcome measures. But this literature does not translate well to the broad organizational performance improvements envisioned by the ACA. New HHS programs like the Pioneer ACO program, however, will provide rich new opportunities for analyzing these relationships in greater detail.

A broader concept of organizational characteristics thought to be associated with broader measures of performance is that of clinical integration, defined as:*..the coordination of health services across providers, functions, activities, processes and settings in order to achieve maximal value for person for whom the system has assumed responsibility.* (Conrad and Shortell, 1996). Clinical integration and care coordination are central to many individual HHS programs. Measures of clinical integration exist (e.g., Gilles et al., 1996) but they have not been tested widely across a range of organizations nor have they been empirically tied to outcomes.

Accommodating Variation in Types of Organizations: CMS initiatives within the ACA affect many different types of organizations, including hospitals, integrated delivery systems, organized physician groups, solo practitioners, and even some state agencies. A myriad of hybrid combinations of these organizations exists: a hospital with integrated post-acute care facilities, physician practices with home-health partnerships, etc. In many cases, it may be challenging to fit idiosyncratic health care organizations into common definitions and to craft a single survey or set of questions that are appropriate for all sites.

Data Availability: Academic researchers have the freedom to choose research questions for which there are strong, unassailable data to help answer them. Hence, the literature has many more studies on hospital dynamics for which there are good data than on other venues of health care, such as outpatient, post-acute, and the physician sector.

Similarly, there are limited measures of the internal dynamics of health care organizations. Structure does not always equal function; medical groups may compensate physicians with a salary, but this does not guarantee that they have the incentive to change their practice patterns. More research is needed to explore the culture, leadership, and relational coordination among members of a health care organization. And it may require significant resources to fully understand these dynamics.

Lastly, more research will be needed on the impact of external forces in driving successful implementation efforts, in particular, how the actions of private insurers, state policies, and community groups affect organizational behavior and performance.

Gathering Data on Comparison Groups: To accurately evaluate delivery reform initiatives, researchers may need health care data on groups who chose not to participate. Gaining access to health outcomes may be difficult, but this may be more of a challenge for the organizational and managerial variables. Many of the data sources mentioned in this chapter focus attention only on participating groups through surveys or qualitative case studies and interviews. While non-participating organizations could also be surveyed, determining the right control group and inducing their (even limited) participation may be challenging.

Exhibit 4-2: Organizational Structure and Capacity Variables

1. General Organizational Structure & Service Capacity (10)

Variable, Definition	Example
1. Type of Organization: General definition of the organization (e.g., integrated delivery system, hospital, multispecialty medical group, single-specialty medical group etc.)	*[CTS Survey of Physicians]* Please check the box that best describes where you work: a solo practice; a two-physician practice; a group practice with three or more physicians; a group or staff model HMO; a community health center; a hospital run by state, county, or city government; a hospital run by a private for-profit or non-profit organization; a medical school or university; some other setting.
2. Organizational Experience/History: How long has the organization been in existence; previous experience with payment reform	*[NSPO I]* How long has the oldest practice unit, which is now all or part of your group, been in existence?
3. Location: General location of the facilities (inner city, urban, suburban, rural)	*[NSPO I]* Which designation best describes the metropolitan area or communication surrounding the largest or primary location of your practice: central city, urban, suburban, small city, rural?
4. Control of Continuum of Care: Extent to which the organization owns or partners with other types of care delivery sites (e.g., hospital, specialty physician groups, skilled nursing facilities, ambulatory surgical centers, post-acute care facilities, rehabilitation facilities, other)	*[Brandeis CAPP]* Does your medical group own or operate any of the following entities: acute care hospital, skilled nursing facility, ambulatory surgery center, home health (please provide number of each and number of beds where applicable)?
5. Integration with Third-Party Payer Whether the organization has an owned or affiliated insurance plan.	*[Brandeis CAPP]* What proportion of your organization's total patient revenue is paid by an owned or affiliated insurance plan?
6. Size - Number of Physicians: Number of FTE physicians, by type (e.g., primary care, specialist, hospitalist)	*[NSPO III]* Approximately how many of the physicians in your medical group, across all its locations, are: family physicians, general internists, general practitioners, cardiologists, endocrinologists, other.
7. Size - Number of Other Clinicians: Number of all FTE non-physician clinical staff, including nurses, nurse practitioners, and physician assistants, but may also include dieticians, health educators, and other providers.	*[NSPO I]* Please indicate the total number of the following health professionals (full or part-time) working in your medical group across all locations: nurses, nurse practitioners, physician assistants.
8. Size - Utilization: Estimation of annual volume based on the number of units of medical services (e.g., inpatient admissions, patient visits, etc.)	*[AHA Survey]* Report the number of inpatient admissions, inpatient days, outpatient visits.

Exhibit 4-2: Organizational Structure and Capacity Variables (continued)

1. General Organizational Structure & Service Capacity (10) (continued)

Variable, Definition	Example
9. Physician Affiliation Model: Percent of physicians who are employed, formally contracted, informally affiliated	*[AHA Survey]* Report the number of physicians with privileges at your hospital by: total employed, total individual contract, total group contract, not under contract.
10. Physician Compensation Model: Percent of employed/non-employed physician compensation that is salary, production-based (i.e., RVU), or performance-based	*[Brandeis-CAPP]* For [employed, non-employed] [primary care/specialist] physicians, what percent of compensation is based on the following factors: salary, production, efficiency/resource use, quality metrics, patient satisfaction, other?

2. Governance Structure

Variable, Definition	Example
1. Ownership/Organization: Entity or group that controls the policies and/or strategies of the organization, including **f**or profit vs. not-for-profit status	*[AHA Survey]* Control – Indicate the type of organization that is responsible for establishing policy for overall operation of your hospital: Federal Government; government nonfederal (state, county, city, hospital district); nongovernment not-for-profit (church-operated, other not-for-profit); investor-owned for profit (individual, partnership, corporation). *[AMGA Survey]* What is the legal organization of your medical group: business corporation, professional corporation, general partnership, limited liability partnership, not-for-profit corporation or foundation?
2. Board Members & Board Representation: What kinds of personnel make up the board, including physicians, other healthcare organizations, insurers, and consumers?	*[NSPO I]:* For the governing body of your medical group (NOT the owner of your practice, i.e. hospital or health system), please indicate: the total number of board positions, the number of [primary care/specialist] physicians on board, the number of meetings per year.

Exhibit 4-2: Organizational Structure and Capacity Variables (continued)

3. Financial, Payer and Reimbursement Model Characteristics (6)

Variable, Definition	Example
1. Size - Total Patient Services Revenue: Financial size of the organization in terms of patient revenue. Also may include profitability or operating margins in a defined fiscal period.	*[AMGA Survey]* What is your estimated annual medical net revenue? *(later asks groups to calculate an organizational profit/loss statement)* *[Brandeis-CAPP]* Please estimate your total Net Patient Service Revenue based on the medical budget for designated physician group.
2. Market Share: Percent of care delivered (e.g., hospital beds/services, professional services) in organization's primary service area	*[Brandeis-CAPP]* What is the organization's market share in its primary service area for: professional services, hospital services?
3. Payer Mix: Percentage of the organization's revenue from government payers, private insurance, and owned-affiliated insurers	*[AHA Survey]* Report the total facility gross and net revenue by: Medicare FFS, Medicare managed care, Medicaid FFS, Medicaid managed care, Medicaid DSH, other government, self-pay, third-party payers, all other
4. Payer Contracts/Reimbursement Mechanisms: Percentage of reimbursement contracts based on fee-for-service, shared savings, capitation, or other risk-based payments	*[Brandeis-CAPP]* What percent of organization's patient revenues paid under the following payment mechanisms: fee-for-service, pay-for-performance under FFS, episode payments, shared savings, partial capitation, global capitation, other.
5. Patient Mix: The sociodemographic and eligibility characteristics of the patients the organization treats, and may also include patients' initial health status to the extent that it is a risk-adjustment (and not outcome of care).	*[CTS Survey of Physicians]* Approximately what percentage of your patients belong to the following groups: African-American or Black, Hispanic or Latino, Asian or Pacific Islander, Native American, or Alaska Native? What percentage of your patients do you have a hard time speaking with or understanding because you speak different languages?
6. Major Payers: Extent to which revenue is concentrated in a few health plans with the ability to materially affect financial incentives through alternative payment models.	*[Brandeis-CAPP]* Please list the three health insurance plans that account for the largest share of your patient service revenue and the approximate percentage for each plan.

Exhibit 4-2: Organizational Structure and Capacity Variables (continued)

4. Information Technology and Data Management (6)

Variable, Definition	Example
1. EHR Availability: Basic question regarding the presence of an electronic health record	*[Brandeis-CAPP]* Has your organization implemented a common electronic medical record? Fully/partially?
2. EHR Documentation/ Meaningful Use: Assess the level of documentation of patient data, especially as it applies to meaningful use statutes?	*[Jha et al, EHR Survey]* Does your hospital have a computerized system for: patient demographics, physician notes, nursing assessments, problem lists, medication lists, discharge summaries, advanced directives (i.e. DNR)?
3. EHR Order Entry/ Results Management: Assess the level of functionality of the electronic health record within the organization	*[Jha et al, EHR Survey]* Does your hospital have a computerized system for [computerized provider order entry/ results viewing] for: labs, radiology, diagnostic tests, consultant reports?
4. EHR Adoption: Assess the level of penetration within the organization of personnel who use electronic health records	*[Brandeis-CAPP]* What percentage of treating physicians uses your organization's electronic medical record? Routinely order medications electronically? Routinely order lab or other tests electronically?
5. EHR Decision Support: Assess the sophistication of the organization's electronic health record to provide quality care	*[Jha et al, EHR Survey]* Does your hospital have a computerized system for decision support, including clinical guidelines, clinical reminders, drug allergy alerts, drug-drug interaction alerts, drug-lab interaction alerts, drug dosing support?
6. Data Management Capabilities: The extent to which the organization has invested and developed data analytic capabilities, including disease registries and practice variation analyses.	*[Brandeis-CAPP]* Has your organization [fully, partially, or not] implemented: enterprise-wide data warehouse and analytic software, patient disease registries, practice variation analysis?

Exhibit 4-2: Organizational Structure and Capacity Variables (continued)

5. Clinical Process Improvement Capabilities (6)

Variable, Definition	Example
1. Management Methodology: To what extent (if any) does the organization have a central management method for process improvement (e.g., Lean, Six Sigma, TPS, plan-do-study-act)?	*[Brandeis/CAPP]* Has your organization implemented a defined management methodology for process improvement (e.g. lean manufacturing)? Which methodology does your organization use? Approximately what percent of [physician/non-physicians] are trained in the methodology? Have participated in a performance improvement event?
2. Electronic Access to Data to Coordinate Care: The extent to which a provider can electronically evaluate and coordinate care, either through the organization's EHR or through partnerships with other providers	*[NSPO III]* Approximately what percentage, if any, of the physicians in your medical group have electronic access to [patients' emergency room visits, discharge summaries, laboratory results, pharmacy records]?
3. Performance Feedback: The extent to which physicians are given feedback on the quality and cost of care they provide patients	*[Brandeis/CAPP]* Do you [measure, report] physician performance in the following areas: production, efficiency/resource use, quality metrics, patient satisfaction, other? How frequently do physicians receive performance feedback?
4. Care Coordination Staff: Does the organization invest specific resources in personnel who coordinate care, including primary care teams or care coordinators?	*[NSPO III]* Does your medical group have any non-physician staff, for example, nurses, dieticians, or health educators, who have time set aside to meet with and/or call patients to help educate them and manage their disease?
5. Care Transition Programs: The extent to which the organization has developed formal processes (either internally or with partners) to coordinate care between ambulatory, hospital, and post-acute facilities.	*[NSPO III]* Does your medical group participate in formal organized care transition program, which improves transitions of care from hospital discharge to home care, nursing home care, or follow-up with the patient's primary care physician or specialist?
6. Care Improvement Programs: What is the extent to which the organization has implemented specific programs intended to address unnecessary spending.	*[Brandeis/CAPP]* To what extent is your organization working on [far along, getting started, planning, not considering] the following initiatives: a) reducing avoidable hospital admissions; b) reducing avoidable hospital readmissions; c) high-risk patient management programs; d) preferred relationships with efficient specialists, hospitals, post-acute care facilities; e) reducing variation for defined episodes of care; f) performance-based physician compensation; g) physician leadership training; h) regular physician peer-group meetings; i) pharmaceutical management.

Exhibit 4-2: Organizational Structure and Capacity Variables (continued)

6. Culture, Leadership, and Teamwork (8)

Variable, Definition	Example
1. Organizational Culture: Description what the organization and its managers value as a predictor of quality improvement implementation, employee and patient satisfaction, and team functioning. Often categorized into "team culture," "hierarchical culture," "entrepreneurial culture," and "rational culture."	*[Competing Values Framework]* My facility is a very dynamic and entrepreneurial place. People are willing to stick their necks out and take risks; My facility is a very formalized and structured place. Bureaucratic procedures generally govern what people do.
2. Leadership: Employee's assessment of the behavior of the organization's leaders, signaling institution's true (versus nominally espoused) values, culture, and processes	*[Learning Organization Survey]* My managers encourage multiple points of view; My managers provide time, resources, and venues for identifying problems and organizational challenges.
3. Quality of Communication: Staff assessment of the frequency, timeliness, and accuracy of communication with key clinical and administrative staff.	*[Relational Coordination Survey]* How **frequently** do people in each of these groups communicate with you about [focal work process or client population]?
4. Shared Goals: Staff assessment of whether their goals are consistent with those of key clinical and administrative staff	*[Relational Coordination Survey]* Do people in these groups **share your goals** regarding [focal work process or client population]?
5. Shared Knowledge: Staff assessment of whether they receive sufficient information from key clinical and administrative staff to support effective work	*[Relational Coordination Survey]* Do people in these groups **know** about the work you do with [focal work process or client population]?
6. Safety Climate: The ability of the organization to encourage openness/psychological safety, learn from errors and near misses, and find time for reflection	*[Learning Organization Survey]* People in this unit are usually comfortable talking about problems and disagreements; Despite the workload, people in this unit find time to review how the work is going.
7. Concrete Learning Processes and Practices: The extent to which the organization invests processes and resources into creating a learning organization, including information collection, analysis, and education and training	*[CTS Survey of Physicians]* Does the hospital where most of your patients are treated have a system for reporting medical errors, in which the person reporting the error remains anonymous? *[Learning Organization Survey]* This unit experiments frequently with new ways of working; This unit has a formal process for conducting and evaluating experiments or new ideas.
8. Employee Tenure/Turnover: The length of tenure (or frequency of turnover) in an organization increases (or exacerbates) the ability to retain strong organizational knowledge and infers a strong (or weak) work-environment and organizational culture	

Exhibit 4-2: Organizational Structure and Capacity Variables (continued)

7. Patient Centeredness

Variable, Definition	*Example*
1. Patient Centered Medical Home: The extent to which the organization practices care according to the patient-centric tenets of the PCMH	*[NSPO III]* Has your medical group received recognition as a Patient-Centered Medical Home from the National Committee for Quality Assurances (NCQA)? What level of recognition have you received?
2. Patient Access to Medical Records: The extent to which patients can access or contribute to their medical record	*[NSPO III]* Does your medical group allow patients to [view, make changes to or update] their medical record online? Does your practice use the electronic medical record to provide patients with clinical summaries of each office visit?
3. Shared Decisionmaking: The extent to which the organization formally incorporates patient input into treatment decisions	
4. Open Scheduling: The extent to which patients have access to make or edit appointments with the organization's providers	*[NSPO III]* Approximately what percentage, if any, of physicians in your medical group use the "advanced access" or "open access" scheduling method in an effort to offer same-day appointments to virtually all who want to be seen, regardless of the reason for which they want to be seen?
5. Email/Extended Communication: Does the organization communicate with patients and other providers via email?	*[DesRoches et al. EHR Survey]* Please indicate how frequently you communicate by email with each of the following: patients about medical issues, other physicians in your practice about patient care, other staff in your practice about patient care, other physicians who are not in your practice about patient issues.

Exhibit 4-2: Organizational Structure and Capacity Variables (continued)

8. Local Market Characteristics and State Policy Environment (10)

Type of Variable	Variable examples (Sources)
1. Relative Health Care Spending (price-adjusted). Areas with relatively higher health spending may have more opportunity to reduce spending through effective delivery reforms.	Medicare spending per beneficiary (Dartmouth Atlas) Average family health insurance premium (MEPS)
2. Provider Market Concentration. Provider concentration indicates the ability of providers to negotiate desired private sector price increases, thus lowering the relative incentives to reduce spending or modify historical service use patterns.	Herfindahl index for hospitals
3. I**nsurer Market Concentration**. Insurer concentration indicates the ability of insurers to limit private sector price increases, thus increasing relative incentives to reduce spending or modify historical service use patterns (in ways that may affect patterns of care provided to Medicare and Medicaid patients)	Percentage of commercial enrollees in the three largest health insurance plans
4. Insurance Plan Type. Penetration of HMO/POS products is related to use of capitation and other alternative payment models as provider organizations are more willing to accept risk under benefit design that includes a requirement for referral authorization prior to specialty care	Percentage of total commercial enrollment in HMO/POS product Percentage of total Medicare beneficiary in HMO product Percentage of Medicaid beneficiaries in HMO product
5. Private Insurer Payment Innovation. Indication of the extent to which private insurers are implementing delivery reform initiatives that complement those of Medicare and Medicaid	What percentage of commercially insured enrollees is enrolled in performance-based reimbursement contracts? *Note: there are no sources for this information on an area-specific basis at present*.
6. Community Engagement. Extent to which there is local community organization across sectors to address causes and remedies of high healthcare spending and/or inadequate quality	Is there a local multi-stakeholder coalition/collaborative focused on controlling cost? Is there a local multi-stakeholder coalition/collaborative focused on improving quality?

Exhibit 4-2: Organizational Structure and Capacity Variables (continued)

8. Local Market Characteristics and State Policy Environment (10) (continued)

Type of Variable	Variable examples (Sources)
7. State Policy – Health Care Spending Goals. Indicates states willingness and capacity to address issues of total health care spending	Does the state measure overall state health care spending? Has the state established specific goals for annual growth in overall state health care spending? Does the state have a mechanism for enforcing specific goals for growth in overall health care spending?
8. State Policy – Health Insurance Exchange. Indication of the extent to which state is willing to use health insurance exchange authority to actively promote lower cost health insurance options. These efforts would likely prove complementary to HHS delivery system reform programs.	What proportion of local residents purchase health insurance coverage through an American Health Benefits Exchange? Medicaid recipients Subsidized populations Commercial members (individual) Commercial enrollees (small group) Commercial enrollees (large group)
9. State Policy – Health Regulation. Indication of the extent to which state has or is willing to use health insurance regulatory authority to actively promote lower cost health insurance options. These efforts would likely prove complementary to HHS delivery system reform programs.	Does the state review health insurance premiums? Individual market Small group market Large group market Has the state set limits on annual growth in health insurance premiums?
10. State Policy: Transparency. Indicator of the extent to which the state is promoting public accountability and increased competition through publication of performance measures.	Does the state collect data on prices negotiated between private insurers and health care providers? Does the state publish data on prices negotiated between private insurers and health care providers? *[NSPO III]* Are data on [patient satisfaction/ experience, clinical quality, cost] within your medical group or its physicians publicly reported by health plans or other external entities?

Exhibit 4-3: Rationale for Organizational Structure and Capacity Variables

1. General Organizational Structure & Service Capacity

Variable, Definition	Rationale
1. Type of Organization: General definition of the organization (e.g., integrated delivery system, hospital, multispecialty medical group, single-specialty medical group, etc.)	Need for general classification of participating organizations
2. Organizational Experience/History: How long has the organization been in existence; previous experience with payment reform	
3. Location: General location of the facilities (inner city, urban, suburban, rural)	Reflects general differences in spending levels and availability of services among urban, suburban and rural locations.
4. Control of Continuum of Care: Extent to which the organization owns or partners with other types of care delivery sites (e.g., hospital, specialty physician groups, skilled nursing facilities, ambulatory surgical centers, post-acute care facilities, rehabilitation facilities, other)	Reflects size and scale of organization and potential to integrate or coordinate services across owned or closely affiliated providers.
5. Integration with Third-Party Payer Whether the organization has an owned or affiliated insurance plan.	Reflects access to staff and infrastructure that can provide data analytic and actuarial services that can help organizations measure performance.
6. Size - Number of Physicians: Number of FTE physicians, by type (e.g., primary care, specialist, hospitalist)	Indicates scale and scope of organization, Scale can help implementation efforts by allowing investments in overhead resources in human and financial capital to be efficiently spread across the organization.
7. Size - Number of Other Clinicians: Number of all FTE non-physician clinical staff, including nurses, nurse practitioners, and physician assistants, but may also include dieticians, health educators, and other providers.	Indicates level of support available to physicians by skilled non-physician clinicians. May related to ability to care for populations more efficiently.
8. Size - Utilization: Estimation of annual volume based on the number of units of medical services (e.g., inpatient admissions, patient visits, etc.)	Indicates scale and scope of organization, Scale can help implementation efforts by allowing investments in overhead resources in human and financial capital to be efficiently spread across the organization.
9. Physician Affiliation Model: Percent of physicians who are employed, formally contracted, informally affiliated	Physician affiliation can be associated with acceptance of or alignment with performance improvement initiatives
10. Physician Compensation Model: Percent of employed/non-employed physician compensation that is salary, production-based (i.e., RVU), or performance-based	Indicates the extent to which physicians are compensated based on production or rewarded for objective measures of quality, efficiency or patient satisfaction

Exhibit 4-3: Rationale for Organizational Structure and Capacity Variables (continued)

2. Governance Structure

Variable, Definition	Rationale
1. Ownership/Organization: Entity or group that controls the policies and/or strategies of the organization, including for profit vs. not-for-profit status	
2. Board Members & Board Representation: What kinds of individuals make up the board? Are physicians and patients represented? Are participants from across the continuum of care represented?	Broad representation of constituents in governance may reflect willingness to focus on aligning services across the continuum of care and embracing patient-centered reform initiatives.

3. Financial, Payer and Reimbursement Model Characteristics

Variable, Definition	Rationale
1. Size - Total Patient Services Revenue: Financial size of the organization in terms of patient revenue. Also may include profitability or operating margins in a defined fiscal period.	Indicates scale and scope of organization, Scale can help implementation efforts by allowing investments in overhead resources in human and financial capital to be efficiently spread across the organization.
2. Market Share: Percent of care delivered (e.g., hospital beds/services, professional services) in organization's primary service area	
3. Payer Mix: Percentage of the organization's revenue from government payers, private insurance, and owned-affiliated insurers	Percent of revenue from government payers may indicate importance for the organization of success in new delivery reform programs.
4. Payer Contracts/Reimbursement Mechanisms: Percentage of reimbursement contracts based on fee-for-service, shared savings, capitation, or other risk-based payments	Historical experience with capitation or other risk-based payment models may indicates greater preparedness for CMS performance-based payment models. The extent that revenue is tied to performance on health care spending may indicate ability to perform successfully in HHS delivery reform programs.
5. Patient Mix: The sociodemographic and eligibility characteristics of the patients the organization treats, and may also include patients' initial health status to the extent that it is a risk-adjustment (and not outcome of care).	Changes in patient mix will affect performance on spending and quality outcomes.
6. Major Payers: Extent to which revenue is concentrated in a few health plans with the ability to materially affect financial incentives through alternative payment models.	

Exhibit 4-3: Rationale for Organizational Structure and Capacity Variables (continued)

4. Information Technology and Data Management

Variable, Definition	Rationale
1. EHR Availability: Basic question regarding the presence of an electronic health record	Availability of a shared EHR platform can help organizations coordinate care across providers
2. EHR Documentation/ Meaningful Use: Assess the level of documentation of patient data, especially as it applies to meaningful use statutes	Measure of functionality of the electronic health record
3. EHR Order Entry/ Results Management: Assess the level of functionality of the electronic health record within the organization	Measure of functionality of the electronic health record
4. EHR Adoption: Assess the level of penetration within the organization of personnel who use the electronic health record's capabilities.	Degree of penetration of EHR use in the organization provides an indication of clinical integration which may be associated with ability to improve performance
5. EHR Decision Support: Assess the sophistication of the organization's electronic health record to provide quality care.	Automated decision support can reduce unnecessary variance in care delivery.
6. Data Management Capabilities: The extent to which the organization has invested and developed data analytic capabilities, including disease registries and practice variation analyses	Organizations with ability to measure and analyze their own performance may be more likely to succeed in performance improvement.

Exhibit 4-3: Rationale for Organizational Structure and Capacity Variables (continued)

5. Clinical Process Improvement Capabilities (6)

Variable, Definition	Rationale
1. Management Methodology: To what extent (if any) does the organization have a central management method for process improvement (e.g., Lean, Six Sigma, TPS, plan-do-study-act)?	Indication of degree of commitment to a formal method for performance improvement
2. Electronic Access to Data to Coordinate Care: The extent to which a provider can electronically evaluate and coordinate care, either through the organization's EHR or through partnerships with other providers	Indicative of clinical integration across the continuum of care that may improve capacity to improve performance
3. Performance Feedback: The extent to which physicians are given feedback on the quality and cost of care they provide patients	Indicative of clinical integration across the continuum of care that may improve capacity to improve performance
4. Care Coordination Staff: Does the organization invest specific resources in personnel who coordinate care, including primary care teams or care coordinators?	Indicative of clinical integration across the continuum of care that may improve capacity to improve performance
5. Care Transition Programs: The extent to which the organization has developed formal processes (either internally or with partners) to coordinate care between ambulatory, hospital, and post-acute facilities.	Indicative of clinical integration across the continuum of care that may improve capacity to improve performance
6. Care Improvement Programs: Extent to which the organization has implemented specific programs intended to address unnecessary spending.	Indicative of clinical integration across the continuum of care that may improve capacity to improve performance

Exhibit 4-3: Rationale for Organizational Structure and Capacity Variables (continued)

6. Culture, Leadership, and Teamwork

Variable, Definition	Rationale
1. Organizational Culture: Description of what the organization and its managers value as a predictor of quality improvement implementation, employee and patient satisfaction, and team functioning. Often categorized into "team culture," "hierarchical culture," "entrepreneurial culture," and "rational culture."	Strong organizational culture, history, and mission focuses thinking, shared ethics, and purpose.
2. Leadership: Employee's assessment of the behavior of the organization's leaders, signaling institution's true (versus nominally espoused) values, culture, and processes.	Willingness to embark on performance improvement initiatives and ability to implement them successfully is enhanced by strong physician leadership, robust management structure, and clear reporting lines of communication.
3. Quality of Communication: Staff assessment of the frequency, timeliness, and accuracy of communication with key clinical and administrative staff	Indication of the ability of team members to work together and effectively implement new initiatives
4. Shared Goals: Staff assessment of whether their goals are consistent with those of key clinical and administrative staff	Indication of the ability of team members to work together and effectively implement new initiatives
5. Shared Knowledge: Staff assessment of whether they receive sufficient information from key clinical and administrative staff to support efficient, effective completion of their work.	Indication of the ability of team members to work together and effectively implement new initiatives
6. Safety Climate: The ability of the organization to encourage openness/psychological safety, learn from errors and near misses, and find time to reflection.	Indication of organization's emphasis on a culture of improvement over a culture of blame
7. Concrete Learning Processes and Practices: The extent to which the organization invests processes and resources into creating a learning organization, including information collection, analysis, and education and training	Indication of organization's commitment to ongoing positive change and to support employees in effecting this change
8. Employee Tenure/Turnover: The length of tenure (or frequency of turnover) in an organization	The rate of turnover increases (or exacerbates) the ability to retain strong organizational knowledge and infers a strong (or weak) work-environment and organizational culture

Exhibit 4-3: Rationale for Organizational Structure and Capacity Variables (continued)

7. Patient-Centeredness

Variable, Definition	Rationale
1. Patient Centered Medical Home: The extent to which the organization practices care according to the patient-centric tenets of the PCMH	Goals and objectives of patient-centered medical home are fundamentally based on the three-part-aim and therefore fully consistent with HHS delivery reform efforts
2. Patient Access to Medical Records: The extent to which patients can access or contribute to their medical record	Indication of organizational commitment to patient-centeredness
3. Shared Decisionmaking: The extent to which the organization formally incorporates patient input into treatment decisions	Indication of organizational commitment to patient-centeredness
4. Open Scheduling: The extent to which patients have access to make or edit appointments with the organization's providers	Indication of organizational commitment to patient-centeredness
5. Email/Extended Communication: Does the organization communicate with patients and other providers via email?	Indication of organizational commitment to patient-centeredness

8. Local Market Characteristics and State Policy Environment (10)

Type of Variable	Rationale
1. Relative Health Care Spending (Price-Adjusted)	Areas with relatively higher health spending may have more opportunity to reduce spending through effective delivery reforms. Local market conditions (relative health care costs, market concentration) can reduce or encourage willingness and attention on reform initiatives. Private insurer contracts and state/local efforts with health care groups can augment the intentions of CMS to help drive organizational change.
2. Provider Market Concentration	May indicates the ability of providers to negotiate desired private sector price increases thus lowering the relative incentives to reduce spending or modify historical service use patterns.
3. Insurer Market Concentration	May indicate the ability of insurers to limit private sector price increases thus increasing the relative incentives to reduce spending or modify historical service use patterns (in ways that may affect patterns of care provided to Medicare and Medicaid patients).

Exhibit 4-3: Rationale for Organizational Structure and Capacity Variables (continued)

8. Local Market Characteristics and State Policy Environment (10) (continued)

Type of Variable	Rationale
4. Insurance Plan Type. Penetration of HMO/POS products.	May be related to use of capitation and other alternative payment models or potential for expansion of these models, as provider organizations are more willing to accept risk under benefit design that includes a requirement for referral authorization prior to specialty care.
5. Private Insurer Payment Innovation	Indication of the extent to which private insurers are implementing delivery reform initiatives that complement those of Medicare and Medicaid.
6. Community Engagement	Extent to which there is local community organization across sectors to address causes and remedies of high healthcare spending and/or inadequate quality.
7. State Policy—Health Care Spending Goals	Indicates states willingness and capacity to address issues of total health care spending
7. State Policy—Health Insurance Exchange	Indication of the extent to which state is willing to use health insurance exchange authority to actively promote lower cost health insurance options. These efforts would likely prove complementary to HHS delivery system reform programs.
8. State Policy—Health Regulation	Indication of the extent to which state has or is willing to use health insurance regulatory authority to actively promote lower cost health insurance options. These efforts would likely prove complementary to HHS delivery system reform programs.
9. State Policy—Transparency	Indicator of the extent to which the state is promoting public accountability and increased competition through publication of performance measures.

Exhibit 4-4: Advantages and Disadvantages of Organizational Structure and Capacity Variables

1. General Organizational Structure & Service Capacity

Variable	Pros	Cons
1. Type of Organization	Classification of organizations is needed	Significant variability in organization characteristics within classifications
2. Organizational Experience	May be indicative of organizational stability, development of systems, and presence of common culture in integrated organizations (e.g.multi-specialty groups)	Probably less relevant for individual institutions (e.g., hospitals, nursing homes); may only be a proxy for other variables.
3. Location	Readily available, may have specific policy implications for rural facilities	None
4. Control of Continuum Of Care	Important variables for understanding integration Easy to validate	Often unavailable in public databases and will require survey of organizations; may change over time
5. Integration With Third-Party Payer	Good indicator of access to data analytic and actuarial resources	Relatively few organizations are presently integrated with third party payer
6. Size - Number of Physicians	Necessary variable for size/scale	Requires survey of organization
7. Size - Number of Other Clinicians	Indication of extent that physicians are leveraged with extenders	Requires survey of organization
8. Size - Utilization	Necessary variable for size/scale	Requires survey of organization
9. Physician Affiliation Model	Important measure of integration	Requires survey of organization
10. Physician Compensation Model	Important measure of financial incentives faced by individual physicians	Requires survey of organization

2. Governance Structure (2)

Variable	Pros	Cons
1. Ownership/Organization	Frequently used as a control variable in analysis of healthcare facility costs	Publically available for certain organizations (e.g. facilities) but may require survey for others (e.g. physician groups)
2. Board Composition	Indication that key stakeholders are involved in organizational decisionmaking	Requires survey of organization

Exhibit 4-4: Advantages and Disadvantages of Organizational Structure and Capacity Variables (continued)

3. Financial, Payer and Reimbursement Model Characteristics (5)

Variable	Pros	Cons
1. Size - Total Patient Services Revenue	Necessary variable for size/scale	Readily available for certain delivery system elements (e.g., hospitals) but may require survey for other elements (e.g. physician groups)
2. Market Share	Indicator of organization's ability to dictate terms to suppliers and purchasers	May be either positively or negatively correlated with performance improvement
3. Payer Mix	Indicates proportion of organization's revenue dedicated to public vs. private payers, proportion that may be risk-based	Does not provide specific details of the contracts or their financial incentives for coordinated care
4. Payer Contracts/Reimbursement Mechanisms	Indicates organization's experience and current use of performance based contracts	Requires survey of organization Contracting organizations (i.e., PHOs, IPAs) will be able to report on contracts that it manages but will not have information on all sources of revenue for contracted physicians and other components of the contracted network
5. Patient Mix	CMS has ready access to information about Medicare beneficiary characteristics Spending levels and other performance measures vary based on patient characteristics	None
6. Major payers	Could provide indication of potential for expansion of private sector performance based contracts	Requires additional information about initiatives underway at local private payers

Exhibit 4-4: Advantages and Disadvantages of Organizational Structure and Capacity Variables (continued)

4. Information Technology and Data Management

Variable	Pros	Cons
1. EHR Availability	Precursor to EHR adoption	Does not provide insight into how EHR is used
2. EHR Adoption	Strong gauge of potential of organization to take advantage of performance improvement facilitated by EHR	Does not provide insight into how EHR is used
3. EHR Documentation/ Meaningful Use	Availability of documented measures	Documented measures will not include full range of EHR capabilities of interest
4. EHR Order Entry/ Results Management	Specific EHR capabilities may be associated with performance	Does not provide insight into extent specific EHR capability is used
5. EHR Decision Support	EHR decision support is very likely associated with performance	Does not provide insight into extent specific EHR capability is used
6. Data Management Capabilities	Essential capacity for monitoring and improving performance	Does not provide insight into quality of data analysis and reporting within the organization or use of reporting by managers and clinicians

Exhibit 4-4: Advantages and Disadvantages of Organizational Structure and Capacity Variables (continued)

5. Clinical Process Improvement Capabilities (6)

Variable	Pros	Cons
1. Management Methodology	Indication of degree of commitment performance improvement.	Process measure. Difficult to quantify the extent to which the methodology is applied or its effectiveness. Requires survey to collect information.
2. Electronic Access to Data to Coordinate Care	Indicative of clinical integration across the continuum of care that may improve capacity to improve performance.	Process measure. Difficult to quantify the extent to which the data are used or their effectiveness. Requires survey to collect information.
3. Performance Feedback	Indicative of clinical integration that may improve capacity to improve performance.	Process measure. Difficult to quantify the extent to which the data are used or their effectiveness. Requires survey to collect information.
4. Care Coordination Staff	Indicative of clinical integration across the continuum of care that may improve capacity to improve performance	Process measure. Difficult to quantify the effectiveness of these staff or extent of their activities. Requires survey to collect information.
5. Care Transition Programs	Indicative of clinical integration across the continuum of care that may improve capacity to improve performance	Process measure. Difficult to quantify the effectiveness of these programs or extent of their activities. Requires survey to collect information.
6. Care Improvement Programs	Indicative of clinical integration that may improve capacity to improve performance	Process measure. Difficult to quantify the effectiveness of these programs or extent of their activities. Requires survey to collect information.

Exhibit 4-4: Advantages and Disadvantages of Organizational Structure and Capacity Variables (continued)

6. Culture, Leadership, and Teamwork

Variable	Pros	Cons
1. Organizational Culture	Essential element of performance	Difficult to measure accurately. Requires substantial investment in physician and employee surveys.
2. Employee Turnover	Documents consistency of personnel, allowing for retained organizational knowledge	Difficult to measure, requiring historical analysis of human resources/employee survey
3. Leadership	Essential element of performance	Difficult to measure accurately. Requires substantial investment in physician and employee surveys.
4. Quality of Communication	Essential element of performance	Difficult to measure accurately. Requires substantial investment in physician and employee surveys.
5. Shared Goals	Essential element of performance	Difficult to measure accurately. Requires substantial investment in physician and employee surveys.
6. Shared Knowledge	Essential element of performance	Difficult to measure accurately. Requires substantial investment in physician and employee surveys.
7. Safety Climate	Essential element of performance	Difficult to measure accurately. Requires substantial investment in physician and employee surveys.
8. Concrete Learning Processes and Practices	Essential element of performance	Difficult to measure accurately. Requires substantial investment in physician and employee surveys.

Exhibit 4-4: Advantages and Disadvantages of Organizational Structure and Capacity Variables (continued)

7. Patient Centeredness

Variable	Pros	Cons
1. Patient Centered Medical Home	Consistent with the 3-part aim Certification data available from NCQA	Difficult to evaluate the scope or quality of PCMH programs without qualitative assessment.
2. Patient Access to Medical Records	Indication of organizational commitment to patient-centeredness	Process measure. Impact on outcomes in not known.
3. Shared-decision making	Indication of organizational commitment to patient-centeredness	Difficult to evaluate the scope or quality of shared decision making initiatives without qualitative assessment.
4. Open Scheduling	Indication of organizational commitment to patient-centeredness	Process measure. Impact on outcomes in not known.
5. Email/Extended Communication	Indication of organizational commitment to patient-centeredness	Process measure. Impact on outcomes in not known.

Exhibit 4-4: Advantages and Disadvantages of Organizational Structure and Capacity Variables (continued)

8. Local Market Characteristics and State Policy Environment

Type of Variable	Pros	Cons
1. Relative Health Care Spending (Price-Adjusted)	Easily available– Dartmouth Atlas or Medicare claims data	None
2. Provider Market Concentration	Easy to measure for hospital inpatient services and certain other institutional services	Much more difficult to measure for physician services
3. I**nsurer Market Concentration**	Easy to measure from publicly available sources	None
4. Insurance Plan Type	Data are available from private services like Interstudy	Possibly less valuable as an indicator with growth of shared savings payment models that do not require HMO-type benefit design.
5. Private Insurer Payment Innovation.	Important indicator of private sector payment incentives for performance improvement	Requires detailed primary data collection through case study type approaches Changing at rapid rate requiring annual updating
6. Community Engagement	Presence of regional health care collaboratives /community quality improvement initiatives (e.g. RWJF AF4Q) are readily available	Determining effectiveness of regional initiatives requires detailed research that is subjective in nature
7. State Policy – Health Care Spending Targets	Indicator of state government pressure on market participants to constrain costs	Only a few states currently contemplating such targets and fewer have determined enforcement mechanisms
8. State Policy – Health Insurance Exchange	Indicator of state willingness to utilize health insurance exchange authority to actively promote better functioning insurance markets	Many possible exchange configurations—would require qualitative assessment of potential for affecting health care quality and cost
9. State Policy – Health Insurance Regulation.	Indicator of state willingness to use health insurance regulatory authority to actively promote lower cost health insurance options	Would require some qualitative assessment of potential for affecting health care quality and cost
10. State Policy: Transparency	Indicator of state willingness to publish health care provider prices and quality measures in order to create stronger impetus for performance improvement	Would require some qualitative assessment of the extent of transparency created and potential for affecting health care quality and cost

Exhibit 4-5: Primary Data Sources for Organizational Variables

	Publicly-Available Data	Privately-Held Datasets	Data from Fed/State Programs	Medicare Claims	Exec-Level Org Surveys	Employee Level Org Surveys	Qual/ Case Studies/ Ethnog.	Comment
1. General Organizational Structure & Service Capacity								
1. Type of Organization	x				x		x	
2. Organizational Experience/History					x		x	
3. Location	x			x	x			
4. Control of Continuum of Care					x			May be available from websites, or claims w/provider affiliations
5. Integration With Third-Party Payer					x			
6. Size – Number of Physicians		x			x			Ex. AHA survey
7. Size – Number of Other Clinicians		x			x			
8. Size – Utilization	x				x			
9. Physician Affiliation Model					x		x	
10. Physician Compensation Model					x	x	x	
2. Governance Structure								
1. Ownership/Organization	x	x			x		x	May be available from websites
2. Board Members & Board Representation					x		x	
3. Financial, Payer and Reimbursement Model Characteristics								
1. Size – Total Patient Services Revenue					x			
2. Market Share					x			
3. Payer Mix	x	x			x			
4. Payer Contracts/Reimbursement Mechanisms					x			
5. Patient Mix				x				
6. Major payers					x			

Exhibit 4-5: Primary Data Sources for Organizational Variables (continued)

	Publicly-Available Data	Privately-Held Datasets	Data from Fed/State Programs	Medicare Claims	Exec-Level Org Surveys	Employee Level Org Surveys	Qual/ Case Studies/ Ethnog.	Comment
4. Information Technology and Data Management								
1. EHR Availability			x		x	x	x	EHR Incentive Program
2. EHR Documentation/ Meaningful Use			x		x	x	x	EHR Incentive Program, PCMH Standards
3. EHR Order Entry/ Results Management			x		x	x	x	EHR Incentive Program, PCMH Standards
4. EHR Adoption						x		Employee-level surveys provides extent HIT is distributed and used
5. EHR Decision Support			x		x	x	x	EHR Incentive Program
6. Data Management Capabilities			x		x		x	EHR Incentive Program
5. Clinical Process Improvement Capabilities								
1. Management Methodology			x		x		x	PCMH Standards asks about CQI programs
2. Electronic Access to Data to Coordinate Care			x		x	x	x	EHR Incentive Program, PCMH Standards
3. Performance Feedback			x		x	x	x	PCMH Standards
4. Care Coordination Staff			x		x	x	x	PCMH Standards
5. Care Transition Programs			x		x	x	x	PCMH Standards
6. Care Improvement Programs			x		x	x	x	PCMH Standards

Exhibit 4-5: Primary Data Sources for Organizational Variables (continued)

	Publicly-Available Data	Privately-Held Datasets	Data from Fed/State Programs	Medicare Claims	Exec-Level Org Surveys	Employee Level Org Surveys	Qual/ Case Studies/ Ethnog.	Comment
6. Culture, Leadership, and Communication								
1. Organizational Culture						x	x	Requires employee-level surveys, interviews
2.Employee Turnover						x		Requires employee-level surveys, interviews
2. Leadership						x	x	Requires employee-level surveys, interviews
3. Quality of Communication						x	x	Requires employee-level surveys, interviews
4. Shared Goals						x	x	Requires employee-level surveys, interviews
5. Shared Knowledge						x	x	Requires employee-level surveys, interviews
6. Safety Climate					x	x	x	Targeted surveys of quality, safety personnel
7. Concrete Learning Processes and Practices						x	x	Requires employee-level surveys, interviews
7. Patient Centeredness								
1. Patient Centered Medical Home			x		x		x	PCMH Standards
2. Patient Access to Medical Records	x		x		x		x	PCMH Standards
3. Shared Decisionmaking			x		x	x	x	Employee-level surveys provide extent SDM is distributed and used
4. Open Scheduling			x		x	x	x	PCMH Standards
5. Email/Extended Communication			x		x	x	x	PCMH Standards

Exhibit 4-5: Primary Data Sources for Organizational Variables (continued)

	Publicly-Available Data	Privately-Held Datasets	Data from Fed/State Programs	Medicare Claims	Exec-Level Org Surveys	Employee Level Org Surveys	Qual/ Case Studies/ Ethnog.	Comment
8. Local Market Characteristics and State Policy Environment								
1. Relative Health Care Spending (Price-Adjusted)	x	x		x				Claims analysis, Dartmouth Atlas, other sources
2. Provider Market Concentration	x	x						Private vendors or calculated for public data
3. Insurer Market Concentration	x	x						Available from private vendors or calculated for public data
4. Insurance Plan Type, Managed Care Penetration		x						Ex: Interstudy
5. Private Insurer Payment Innovation								No data source currently available
6. Community Engagement								None, requires *de novo* study of local groups
7. State Policy—Health Care Spending Goals								None, requires *de novo* study of state policies
7. State Policy—Health Insurance Exchange								None, requires *de novo* study of state policies
8. State Policy—Health Regulation								None, requires *de novo* study of state policies
9. State Policy—Transparency								None, requires *de novo* study of state policies

4.7 REFERENCES

Birkmeyer NJO; Goodney PP; Stukel TA; et al. Do cancer Centers designated by the National Cancer Institute have better surgical outcomes? *Cancer* Feb 1 2005; 103(3): 435-441.

Berwick D. Disseminating Innovations in Health Care. *JAMA*, April 16 2003; 289(15); 1969-1975.

Berwick DM, Nolan TW, Whittington J. The Triple Aim: Care, Health, and Cost. *Health Affairs*, May 2008 vol. 27 no. 3 759-769.

Casalino L, Giles RR, Shortell SM. External incentives, information technology, and organized processes to improve health care quality for patients with chronic diseases. *JAMA*, 2003; 289(4): 434-441.

Conrad DA, Shortell SM, Integrated health systems: Promise and performance. *Frontiers of Health Services Management*, 1996, 13(1).

Damschroder LJ, Aron DC, Keith RE et al. Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science,* 2009, 4(50): 1-15.

Devers KJ, Shortell SM, et al. Implementing organized delivery systems: An integration scorecard. *Health Care Management Review*, 1994, 19(3):7-20.

Ferlie EB, Shortell SM: Improving the quality of health care in the United Kingdom and the United States: A framework for change. *Milbank Quarterly*, 2001, 79:281-315.

Gilles RR, Shortell SM et al. Conceptualizing and measuring integration: Findings from the Health Systems Integration Study. *Journal of Healthcare Management*, Winter 1993; 38(4).

Gittell, Jody Hoffer. *High Performance Healthcare: Using the Power of Relationships to Achieve Quality, Efficiency and Resilience.* New York: McGraw-Hill, 2009.

Grol R; Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet*, Oct 11 2003; 362(9391): 1225-1230.

Harrington C; Woolhandler S; Mullan J; et al. Does investor ownership of nursing homes compromise the quality of care? *American Journal of Public Health,* Sep 2001; 91(9): 1452-1455.

Himmelstein DU; Woolhandler S; Hellander I; et al. Quality of care in investor-owned vs not-for-profit HMOs. *JAMA,* July 14 1999; 282(2): 159-163.

Jha AK; Li ZH; Orav EJ; et al. Care in US hospitals—The Hospital Quality Alliance. *NEJM*, July 21 2005; 353(3): 265-274.

Nguyen NT; Paya M; Stevens M; et al. The relationship between hospital volume and outcome in bariatric surgery at academic medical centers. *Annals of Surgery,* Oct 2004; 240(4): 586-593.

Rogers EM. *Diffusions of Innovations*. 4th Ed. NY: Free Press, 1995.

RTI International. *Post-Acute Care Episodes: Expanded Analytic File.* Prepared for Assistant Secretary for Planning and Evaluation (ASPE), U.S. Department of Health and Human Services, June 2011. Accessed August 15, 2012 at: http://www.innovations.cms.gov/Files/x/Bundled-Payments-for-Care-Improvement-Chart-Book-2011.pdf

Shortell SM, Obrien JL, Carman JM et al. Assessing the impact of continuous-quality-improvement, total-quality-management—Concept versus implementation. *Health Services Research*, June 1995; 30(2): 377-401.

Tollen L, Enthoven A, Crosson FJ, Taylor N, Audet AM, Schoen C, Ross M. *Delivery System Reform Tracking:*

A Framework for Understanding Change. The Commonwealth Fund – Issue Brief 1510, July 2011. Accessed August 15,2012 at: http://www.commonwealthfund.org/~/media/Files/Publications/Issue%20Brief/2011/Jun/1510_Tollen_delivery_sys_reform_tracking_ib_v2.pdf.

Weeks WB, Gottlieb DJ, Nyweide DJ, et al. Higher health care quality and bigger savings found at large multispecialty medical groups. *Health Affairs,* May 2010 (29):991-997.

Weiner BJ; Shortell SM; Alexander J. Promoting clinical involvement in hospital quality improvement efforts: The effects of top management, board, and physician leadership. *Health Services Research*, Oct 1997; 32(4): 491-510.

Yano, EM. The role of organizational research in implementing evidence-based practice: QUERI Series. *Implementation Science*, May 29, 2008; Vol 3, Article 29.

Chapter 5: Data Availability and Collection Recommendations

5.1 INTRODUCTION

Fee-for-services (FFS) claims are a mainstay in assessing the cost of care for Medicaid and Medicare beneficiaries, in part because they are national and consistently collected for all beneficiaries over time. However, they are far from ideal. Claims do not fully capture patient severity or provide any information on clinical outcomes (aside from mortality which is available in the beneficiary summary file). Further, FFS claims exclude those in health plans. In the case of Medicare, this is about 25 percent of all beneficiaries at any given point in time. Starting in 2012, Medicare health plans will be submitting standardized encounters. At some point, this information should become consistent enough to use for evaluation and monitoring (perhaps 2014). Unlike claims, encounters do not include information on costs. However, if you switch to Medicaid, the vast majority of beneficiaries are in managed care plans, leading to a potentially significant selection bias if you only use FFS data.[2]

The current claims system does cover pharmacy, but only for Medicare beneficiaries enrolled in a Part D plan (about 50 percent). There are also a variety of methodological issues associated with measuring costs, particularly when you are comparing regions over time. For example, there is not a generally acceptable method for price standardization, and risk adjustment models are only able to capture a small portion of differences based on severity. The claims data system is always evolving, but its primary purpose is payment, not recording health information.

Although claims data may remain important, many other data sources may help provide information on outcomes, consumer experience, utilization, organizations, and market areas. The ideal data set for assessing ACA demonstrations is national (i.e., collected consistently across states), with broad representation of consumers or delivery system entities and repeated measures over time. Very few, if any, data sets match this ideal. Thus, we have to think creatively about how to meet current and future research needs with existing data. It is possible that new data collection efforts are necessary to fully assess the effectiveness of ACA demonstrations with an eye towards replication. This is particularly true when it comes to organizational and market area variables.

5.2 APPROACH

To address this question, we have reviewed data sources that have been used historically for evaluating CMS demonstrations, and collected information on other national or regional data sources that could be harnessed in future work. The analysis was based on three sources of information: a web-based review of national agencies that collect health or social data (see **Exhibit 5-1**), a review of data sources used in published and unpublished evaluations of CMS demonstrations and related research, and consultation with two reference librarians who specialize in national data sets. For regional and local data sources, we followed references in the literature. Unlike the national data, it was more difficult to find a single source that reviewed and

[2] The assumption is that sicker beneficiaries remain the FFS program, while the healthy majority are enrolled in managed care products. There has been some limited research on this.

covered the many customized data collection efforts that support health care management and evaluations.

Exhibit 5-1: Sources of Information on Data Sets

- Agency for Healthcare Research and Quality (AHRQ)
- Centers for Disease Control and Prevention (CDC)
- Centers for Medicare & Medicaid Services (CMS)
- National Center for Health Statistics
- National Center for Chronic Disease Prevention & Health Promotion
- National Bureau of Economic Research
- U.S. Census Bureau

Each data source was reviewed in terms of measures, geographic coverage, sampling strategy, and time design. For this particular task, measures were initially conceptualized in broad categories including: health outcomes, access, cost, process of care or utilization, quality, beneficiary experiences, healthcare systems environment, and healthcare organizations. For a small number of data sets, we did a more indepth analysis that looked at specific variables, such as total cost of care or caregiver burden. In terms of geographic coverage, we defined data sets as national or regional. For the time design category, we sought to understand how frequently the data was available, including time series that support rapid feedback all the way to data sets that skip several years between data collections.

There are many different ways to categorize datasets, including the type or source of information (e.g., claims/encounters, surveys, clinical information from Electronic Health Records (EHRs) or registries) or the type of available measures (e.g., spending, quality, health outcomes, patient experience, organizations, and environment). Two other important dimensions are geographic representativeness and frequency of data collection. It is also valuable to consider the links between data sources. For example, *Can provider information be linked to beneficiaries or surveys linked to claims?*

The section below reviews our findings by type of measures supported by a given data set. **Exhibit 5-2** provides an overview of specific data collection efforts. As some data sets support multiple types of measures, we have tried to note that where relevant.

5.3 DATA AVAILABILITY AND COLLECTION RECOMMENDATIONS

5.3.1 Spending

Claims-based data are a step removed from patient medical records. That said, they are based on a patient's health condition(s) and service utilization. As payment data, they lack the clinical subtlety to convey a fully accurate picture of the patient's health conditions or the care received (including care purchased outside the Medicaid or Medicare reimbursement system). Claims data also provide no information in the way of health outcomes (aside from mortality, which is often noted in beneficiary summary files). Another potential issue with claims data involves the

gradual movement of the U.S. health care system from per-unit (i.e., fee-for-service) to per-patient/population (i.e., bundled payments based on episodes of illness, fully capitated rates, and global budgets). Over time, fee-for-service may be relevant to fewer and fewer people.[3]

Beneficiaries who are dually eligible for Medicaid and Medicare are particularly difficult to track in claims data. Right now, to see the complete history for these individuals, you have to merge Medicaid and Medicare data. Although Medicare claims data are relatively standardized, Medicaid data has state-level variations in coding structures and other differences that need to be handled in the merging process. There is no easy way to do this, although a handful of organizations have developed state or regional data sets of dually eligible beneficiaries to support policy analysis or research. To the extent that dually eligible beneficiaries are enrolled in managed care plans, their data is not available in claims at all. Improving the link between Medicaid and Medicare is the first step in making data on dually eligible beneficiaries more widely available. It may also be beneficial to have one or two organizations maintain a selective (i.e., limited number of states) dually eligible data set that is made available to policymakers and researchers.

Many other data sets that include cost information are national. This includes, for example, the Medical Expenditure Panel Survey (MEPS; total spending per person) or the Centers for Disease Control and Prevention's (CDC's) National Nursing Home Survey (NNHS; nursing home expenditures). Although these data sets can be measured for overall trends, they cannot be used for comparing specific geographic areas. Spending information is also available through private datasets like MarketScan, a longitudinal compilation of claims for more than 40 million covered lives from several larger employers.[4] Unlike the CMS FFS claims, MarketScan includes managed care encounters and pharmacy claims for those who have the benefit. MarketScan tends to include a younger, somewhat healthier population compared to those captured in public sector claims. This and similar data sets[5] can be used to fill out the picture of what is happening in a given health care market area, allowing examination of unintended consequences.

All-payer data sets are a relatively new phenomena that also provide an opportunity to observe spending. There are currently 9 states with active all-payer data sets, and 5 more in the process of being implemented. There are an additional 17 states with an interest in creating an all-payer data set, but most of those efforts are nascent.

Even when you have paid claims, it can be difficult to use the cost information. In Medicare, for example, the allowed amounts reflect variation in regional payments. To compare across areas, therefore, you need to impose some type of price standardization. Since price standardization is rarely perfect, it can introduce noise or error into the measure of costs. When treatment effects are small, the error from price standardization can mask treatment effects. Risk adjustment is also important for fair price comparisons. One way to extend the usefulness of current claims data is to continue developing methodologies that support analysis of cost and utilization from these sources. Further work on defining costs, such as episode costs rather than annual expenditures per beneficiary, would also be useful.

[3] Note, this is already an issue with Medicaid, where the vast majority of beneficiaries for a given state can be enrolled in managed care.
[4] There are also a number of similar sets coming onto the market right now, including many large insurance companies that are selling access to managed care encounters for research.
[5] There is a growing number of private market data sets from different insurers and some efforts to pool across insurers.

5.3.2 Quality

Many different types of data can support the development and scoring of quality measures. For example, a large number of the Healthcare Effectiveness Data and Information Set (HEDIS) quality measure can be claims-based. This is also true of the Agency for Healthcare Research and Quality (AHRQ) Patient Quality Indicators (PQIs) and Patient Safety Indicators (PSIs). At the same time, HEDIS measures require clinical data—like blood glucose levels or other test results—and there are a number of efforts underway to develop new electronic health record (EHR)-based measures. NCQA (who administers HEDIS), and other private entities are increasingly making health-plan-level quality information available for research and evaluation. This is data that is not publicly available, but can be purchased. It also only includes those organizations that use NCQA as a vendor. In the case of something like the NCQA Medical Home survey, the number of participants may be large, but it is not universal.

The American Medical Association (AMA) Physician Consortium for Performance Improvement (PCPI) has developed a series of quality measures based on voluntarily reported Current Procedural Terminology (CPT)-2 codes (an alphabetical listing of clinical conditions and topics with which the measures and codes are associated). This is an example of an innovative effort to use the claims processing system to gather new information. Unfortunately, reporting is voluntary and the uptake and use of CPT-2 codes has been slow.

A third source of information to support quality measurement is local and state public reporting efforts. The Massachusetts Health Quality Partners (MHQP), for example, provides HEDIS quality measures for approximately 150 physician groups in the state. This is a voluntary reporting effort, with insurers submitting information to MHQP. The information is complemented with consumer experience data that is also collected by the insurers. As part of the process, participating organizations standardized their patient experience measurement to gather consistent information. This type of collaborative public reporting has been undertaken in other market areas and represents both a source of information and a model for pooling data.

5.3.3 Patient Experience

Beyond clinical quality (processes of care and patient/population health outcomes), we are interested in how patients experience the health care system. An assessment of patient experience is essentially asking: *How patient-centered is the health care system?* Patient-centeredness encompasses clinical quality, but also pays particular attention to how accessible and convenient the health care system is (e.g., *Does the provider have weekend office hours and make phone or email check-ins available? How long do patients have to wait for an appointment and how long do patients have to wait to see a provider after arriving at the appointment? Do providers realize and act upon the cultural and linguistic heterogeneity of their communities and practices? How satisfied are patients with the overall experience?*

Of the data sources we reviewed, the AHRQ's Consumer Assessment of Healthcare Providers and Systems (CAHPS) provides the most comprehensive assessment of these more *high-touch* elements of patient experience. There is both an ambulatory and inpatient (Hospital Consumer Assessment of Healthcare Providers [HCAHPs]) version, with additional versions, like HCAHPs for Accountable Care Organizations (ACOs), in the works. CAHPS is used, either in part or in whole, by other data sources we reviewed (e.g., MEPS). One thing to consider carefully with this

measure is the underlying sampling design. On the inpatient side, for example, hospitals need to report up to 300 cases, regardless of size. Some CMS evaluations of ACA initiatives require additional data collection for treatment sites to capture specific sub-groups, but parallel data is generally not available for comparison sites. One recommendation would be to expand the sampling frame to include more beneficiaries. This would provide increased flexibility for evaluation studies, particularly when the emphasis is on a specific condition or population sub-group.

5.3.4 Health Outcomes

A variety of data sets seek to capture information on health outcomes. For example, the National Health and Nutrition Examination Survey (NHANES) and the Behavioral Risk Factor Surveillance System (BRFSS) provide annual (i.e., point-in-time) national and state-level data, respectively, on health risk factors and preventive behaviors. Additionally, longitudinal studies of population health—for example, the National Children's Study, which follows children from before birth to 21 years of age and the National Longitudinal Study of Adolescent Health, which follows a cohort that is currently between the ages of 28 and 36—can provide insight into the impact of health care payment and delivery system initiatives across the lifespan. While all of these data sets could be used for a national assessment of trends, it is harder to use them to compare specific regions or market areas.

Assuming that the health of the individual patient—or, in aggregate, a patient population—is most relevant for health policy decisionmaking, high-quality clinical data (i.e., care received and health outcomes attained) derived from patients' medical records are very valuable. Absent direct collection of clinical information (e.g., NHANES) or review of patients' medical records (e.g., National Hospital Ambulatory Medical Care Survey [NHAMCS]), many of the data sources presented in **Exhibit 5-2** rely on self-report from patients, their caregivers, and providers to track clinical processes of care and health outcomes. Not surprisingly, the data sources we reviewed overwhelmingly assessed processes of care, likely owing to their discrete and temporally immediate short-term nature. When health outcomes are assessed, they are typically a generic evaluation of a patient's health status. This immediate checkpoint is useful, because it is predictive of long-term clinical health outcomes. However, it should be included alongside other more granular health outcomes data that can more confidently be linked to clinical processes of care.

5.3.5 Organization Characteristics

Moving from the individual patient (i.e., micro) to the organizations in which patients find themselves (i.e., meso), we should consider what organizational characteristics can be more easily associated with doing more (e.g., structure [and composition], capacity, financial and operating characteristics, data management and analytic capacity, and payer relationships and contracting structures) versus doing better (e.g., leadership, performance improvement infrastructure, and clinician and staff engagement). While these two overarching qualities of the health care system are not perfectly exclusive, the focus, historically, has overwhelmingly been on doing more. It is both desirable and possible for the two strategies to complement one another.

Among the data sources we reviewed, data are more readily available for more characteristics. For example, in NHAMCS, we can discern practice size and other characteristics, including if practices are single or multi-specialty, types of practitioners, some of the practice capabilities (e.g., lab

testing), and basic to intermediate electronic medical record capacity and usage. We can also know, based on provider self-report, how providers are paid (e.g., fee-for-service through capitation) and through what mechanisms (e.g., public payer, private payer, or self-pay). Several other data sources contain similar information (e.g., the National Hospital Care Survey, NNHS, and the National Survey of Residential Care Facilities [NSRCF]). Data on the better characteristics are, at this point in time, substantially more limited. Patient-centered medical home assessment tools may prove extremely useful for a somewhat rough understanding of *how* providers work together—for example, if a care team exists and what its composition is—as well as levels of clinician and staff engagement with the payment and delivery interventions within their practices. As the health care delivery system becomes more relationship-centered—most importantly between patients and clinicians and among clinicians with varying educations (e.g., MDs, NPs, PAs, RNs), experiences, and within diverse settings—evaluations will need to take account of the evolution of those relationships and their impact on clinical quality. The same can be said for leadership (which is not the same as ownership) especially if it is leaders who are primarily responsible for engaging their clinicians and staff in new payment and delivery models as well as performance improvement.[6]

In thinking about how to evaluate the impact of organizations on patient/population outcomes, there are three emergent considerations. First, the fields of organizational science and management science have a great deal of expertise to offer, both theoretically/conceptually and methodologically, in trying to better understand the complexity of organizations and the people who work within them. Second, researchers will need to think about how organizational characteristics can be operationalized and captured consistently. Third, qualitative methods will prove extremely useful in understanding how individual clinicians and health care organizations are transforming a system of *more* into a system of *better*.

5.3.6 Health System Environment

Beyond health care organizations lies the health environment, which encompasses both population health and economic, political, and sociocultural determinants and features of health and the health care system. For the purposes of health services research, the data source that most readily provides comprehensive information on the macro health care system is the Area Resource File (ARF). The ARF provides county-level data on the health care workforce (e.g., supply and distribution), health care facilities (e.g., facility type, bed size, admissions, and inpatient days), and census-year population characteristics.

The Kaiser Family Foundation's State Health Facts provides some data regarding managed care penetration and competitiveness of individual and small group markets, and AMA publishes a yearly update of competition in health care markets for 48 states and 368 metropolitan areas.[7] The Federal Trade Commission is another likely source of data for provider (especially hospital) mergers.

[6] Partnership for Patients and the Institute for Healthcare Improvement: http://www.ihi.org/explore/CMSPartnershipForPatients/Pages/default.aspx

[7] https://catalog.ama-assn.org/Catalog/product/product_detail.jsp?productId=prod1940016

5.3.7 Recommendations

Based on our review of national and regional data sets, we have the following recommendations to support the evaluation of ACA initiatives:

1. Support measure development and methodological work that facilitates the use of claims data to measures expenditures.
2. Build and sustain a national sample of dual-eligible beneficiaries.
3. Continue to develop and standardize Medicaid data for cross-state analysis.
4. Where possible, build on the strengths of existing national data collection activities to improve geographic representations by expanding sampling frames. This notion could also be expanded to include more frequent collection of data elements that support rapid feedback reporting.
5. Continue to support the development of her-based outcome measures and processes for pooling these measures across disparate information systems.

For the most part, these issues are well known and in many cases there are efforts underway to address them. Thinking about how to strengthen or extend existing data assets may be more efficient than launching new data collection activities. The one exception, as pointed out in Chapter 4, is organizational and environmental variables. This is an area where nationally available data is extremely limited. This information is important to understand the drivers of change both within and across initiatives.

Also worth noting, there is growing pressure from sites implementing ACA initiatives to use claims data to track patients as they move through the care delivery system, in as close to real time as possible. Many hospitals participating in the ACO or Bundled Payment (BP) demonstrations, for example, want to know which patients use institutional care or emergency department services after leaving the hospital. Information on ambulatory care visits is also valuable for understanding whether or not an individual is on a positive or negative care trajectory. Although institutions are requesting claims data for this purpose, it is a less than ideal source because of claims lag time. Efforts to better understand the impact of claims lags times is important, but most likely does not address the need for a beneficiary tracking system. Initiatives like the Beacon Communities demonstration (funded by The Office of the National Coordinator for Health Information Technology (ONC) within the Office of the Secretary for HHS) where communities are testing new ways to use health information technology to provide care, may offer better solutions. Medicare Advantage plans have also developed innovative ways to share billing information with providers to support integrated care. Medicare Advantage has also taken on the very real problem of providers that do not have EHRs, developing systems to convert paper records into electronic formats. This notion of real time (or virtually real time) data is important to individual sites, but can also support broader evaluation activities, particularly for modeling short-term effects.

Exhibit 5-2: National and Regional Data Sources

Source	Description/ Strengths	Levels of Aggregation	Years or Time Design	Sample Size	Health Status/ Outcomes	Health system Environment	Health Care Organization	Process of Care/ Utilization	Access	Cost
Medicare Claims (Standard Analytical Files)[8]	Available for: outpatient, inpatient, physician (i.e., Part B), home health agency, hospice, durable medical equipment	Individual beneficiary, can be rolled up to the provider level	Ongoing	47 million beneficiaries; 9.4 million beneficiaries in the 20 percent sample; 2.5 in the 5 percent sample				x (utilization insomuch as claim = actual utilization and not what is coded)		x
Medicaid Analytic eXtract (MAX) files[9]	Significant data delay: 2010 data due out Winter 2012-2013. Medicaid data (FFS and MCO) from all 50 states + DC with person-level, inpatient, long-term-care, "other services," and prescription drugs. Can be linked to NCHS (CDC) data as well as Medicare Enrollment Database (identify duals)	Individual beneficiary, can be rolled up to the provider level	Ongoing	62 million beneficiaries				x		x

[8] http://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/IdentifiableDataFiles/StandardAnalyticalFiles.html
[9] http://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MAXGeneralInformation.html

Exhibit 5-2: National and Regional Data Sources (continued)

Source	Description/ Strengths	Levels of Aggregation	Years or Time Design	Sample Size	Health Status/ Outcomes	Health system Environment	Health Care Organization	Process of Care/ Utilization	Access	Cost
National Health Interview Survey (see also the California Health Interview Survey and other state surveys); CDC	The Family component collects demographic information on each family member in the house and also information on health status and limitations, injuries, healthcare access and utilization, health insurance, and income and assets.	National (with state versions)	Cross-sectional survey of house-holds conducted every year	Approximately 35,000 households containing about 87,500 persons	x					
Behavioral Risk Factor Surveillance System (BRFSS)	Preventive health practices and risk behaviors that are linked to chronic diseases, injuries, and preventable infectious diseases in the adult population.	50 states, the District of Columbia, Puerto Rico, the U.S. Virgin Islands, and Guam.	Data collected monthly	One adult interviewed per household	x					
NCQA Quality Compass	The online database features up to 3 years of performance trending of HEDIS and CAHPS® measures for publicly reporting plans.	Health Plans that use HEDIS	Up to 3 years of trend data	Depends on measure and health plan				x		

Exhibit 5-2: National and Regional Data Sources (continued)

Source	Description/ Strengths	Levels of Aggregation	Years or Time Design	Sample Size	Health Status/ Outcomes	Health system Environment	Health Care Organization	Process of Care/ Utilization	Access	Cost
Community Health Center UDS	The Uniform Data System (UDS) tracks a variety of information, including patient demographics, services provided, staffing, clinical indicators, utilization rates, costs, and revenues. UDS data are collected from grantees and reported at the grantee, state, and national levels.	All FHQCs	Data are reported annually in the first quarter of the year.	20.2 million patients				x		
Medical Expenditure Panel Survey (MEPS)	MEPS includes a household and employer (insurance) component. The household survey collects information on health conditions, health status, service utilization, access to care, satisfaction, insurance coverage, and sources of payment for all members of the household. The survey uses a panel design which includes 5 interviews over the course of 2 years.	National	Two full calendar years	Approxi-mately 13,875 (number of families) and 34,920 (number of persons) as of 2009.				x	x	x

Exhibit 5-2: National and Regional Data Sources (continued)

Source	Description/ Strengths	Levels of Aggregation	Years or Time Design	Sample Size	Health Status/ Outcomes	Health system Environment	Health Care Organization	Process of Care/ Utilization	Access	Cost
Healthcare Cost and Utilization Project (HCUP)	State level statistics on inpatient stays, ED visits by payer	National, state, and all-payer healthcare data	Annual	NIS contains data from approxi-mately 8 million hospital stays from roughly 1,000 hospitals; KID contains a sample of over 3 million discharges for children age 20 and younger from more than 3,500 U.S. community hospitals				x		
Family Evaluation of Hospice Care (FEHC) Survey	Among information gathered are caregivers' perceptions of the hospice's performance and patient's experience in the following areas: patient comfort and emotional support, coordination of care, decisionmaking, information sharing, respect for the patient, and emotional support to the family.	National	Quarterly	Approxi-mately 29,292 surveys in 2 quarters.				x		

Exhibit 5-2: National and Regional Data Sources (continued)

Source	Description/ Strengths	Levels of Aggregation	Years or Time Design	Sample Size	Health Status/ Outcomes	Health system Environment	Health Care Organization	Process of Care/ Utilization	Access	Cost
National Cancer Data Base (NCDB)	The NCDB contains standardized data elements on patient demographics, patient insurance status, tumor site, stage and morphology, comorbidities, first course of treatment, disease recurrence, and survival information. In addition, the NCDB contains information on patient ZIP Code and county of residence, which is used to incorporate area-based sociodemographic characteristics. Selected characteristics of the reporting health care facility are also collected.	National	Ongoing data collection	70 percent of all newly diagnosed cases of cancer in the U.S.	x					
National Health and Nutritional Examination Survey (NHANES) (CDC)[10]	Cross-sectional annual survey of 5,000 participants. In-person interview, physical examination, and laboratory tests. Some years of data collection can be linked to Medicare and Medicaid data.	National	Yearly	5,000 people in each in different counties	x			x	x	

[10] http://www.cdc.gov/nchs/nhanes/about_nhanes.htm

Exhibit 5-2: National and Regional Data Sources (continued)

Source	Description/ Strengths	Levels of Aggregation	Years or Time Design	Sample Size	Health Status/ Outcomes	Health system Environment	Health Care Organization	Process of Care/ Utilization	Access	Cost
National Hospital Ambulatory Care Medical Survey (NHACMS) (CDC)[2]	Cross-sectional annual survey. Data derived from random sample of visits - within a randomly assigned four-week reporting period - from emergency departments and outpatient departments of non-institutional general and short-stay hospitals (ambulatory surgical centers included as of 2010)	National	4 week reporting period/ yearly	One or more races (up to 5) for each sampled visit. Each year, 120-170 outpatient department visits and 160-240 emergency department visits. Each year has, on average, 140-300 outpatient department visits and 200-240 emergency department visits.	x (status only)	x	x	x	Possibly	Possibly
National Hospital Care Survey (NHCS) (CDC)[11]	New (as of 2011) cross-sectional annual survey. Inpatient data derived from (unclear of sampling) hospital-based settings including ambulatory surgical centers, both hospital-based and free-standing. Data will be able to be linked with Medicare and Medicaid data.	National	Annual		x (status more likely)		x (seems to be a particular focus/strength of this effort)	x		

[11] http://www.cdc.gov/nchs/nhcs.htm

Exhibit 5-2: National and Regional Data Sources (continued)

Source	Description/ Strengths	Levels of Aggregation	Years or Time Design	Sample Size	Health Status/ Outcomes	Health system Environment	Health Care Organization	Process of Care/ Utilization	Access	Cost
National Home and Hospice Care Survey (NHHCS) (CDC)[12]	Cross-sectional annual survey. (Note: survey administration is inconsistent/ intermittent) Agency and patient-level data derived from two-stage sampling: stratified to select agencies and random to select patients (within agency).	National	Periodic, most recent is 2007	Total number of agencies that participated in the 2007 NHHCS is 1,036; data available on 9,416 current home health patients and hospice discharges from these agencies	x (status more likely)	?	x	x (processes of care)		

[12] http://www.cdc.gov/nchs/nhhcs/about_nhhcs.htm

Exhibit 5-2: National and Regional Data Sources (continued)

Source	Description/ Strengths	Levels of Aggregation	Years or Time Design	Sample Size	Health Status/ Outcomes	Health system Environ-ment	Health Care Organiza-tion	Process of Care/ Utiliza-tion	Access	Cost
National Nursing Home Survey (NNHS) (CDC)[13]	Cross-sectional annual survey. (Note: survey administration is inconsistent/intermit-tent) Nursing home and patient-level data derived from two-stage sampling: stratified then proportional random sampling (by bed size) to select nursing homes and random (? - unclear) to select patients (within home).	National	Periodic	For the 2004 NNHS, 1500 nursing homes were selected. Of these, 283 refused to participate and 43 were considered out of scope. A total of 14,017 residents were sampled from the responding facilities. Of these, 8 were out of scope and 502 refused.	x		x	x		x (charges based on nursing home report)
National Survey of Residential Care Facilities (NSRCF) (CDC)[14]	New (2010) cross-sectional survey (no info on 2011 or 2012). Residential care facility and patient-level data derived from two-stage sampling: stratified then proportional random sampling (by bed size) to select facilities and proportional (by bed size) random to select up to 6 patients (within facility).	National	One year (2010)	2,302 facilities and data available on 8,094 residents from these facilities	x		x	x	Possibly	x (charges based on facility report)

[13] http://www.cdc.gov/nchs/nnhs/about_nnhs.htm
[14] http://www.cdc.gov/nchs/nsrcf/about_nsrcf.htm

Exhibit 5-2: National and Regional Data Sources (continued)

Source	Description/ Strengths	Levels of Aggregation	Years or Time Design	Sample Size	Health Status/ Outcomes	Health system Environ-ment	Health Care Organiza-tion	Process of Care/ Utiliza-tion	Access	Cost
National Survey of Family Growth (NSFG) (CDC)[15]	Cross-sectional survey (Note: survey administration is intermittent, but consistent/ongoing) Data derived from random sample of U.S. households using in-person interviewing General note: It does not appear that the data elements are collected consistently within families (i.e., for each family member)	National	Annual	5,000 interviews annually	x		Possibly	x	x	Possibly
Longitudinal Studies of Aging (LSOA) (CDC)	Multi-cohort study (cohort 1 = 1984 and cohort 2 = 1994; unclear if additional cohorts to follow) of adults ≥70 years old. Cohorts interviewed multiple times during 6-year period (interview schedule does not appear to be systematic) Data derived from nationally representative sample of non-institutionalized adults ≥70 years old. First cohort interviews conducted in person, but second cohort via telephone. Aspirations to link data to Medicare claims in future cohorts (unclear if there will be 2004 and 2014 cohorts).	National	Periodic time series	9,447 NHIS participants of age 70 and over.	x		Possibly	x	x	Possibly

[15] http://www.cdc.gov/nchs/nsfg/about_nsfg.htm

Exhibit 5-2: National and Regional Data Sources (continued)

Source	Description/ Strengths	Levels of Aggregation	Years or Time Design	Sample Size	Health Status/ Outcomes	Health system Environ-ment	Health Care Organiza-tion	Process of Care/ Utiliza-tion	Access	Cost
Market Scan	Commercial claims and encounters	Participating employers	Ongoing	Nearly 150 million beneficiaries included since the start in 1995				x		x
Area Resource File (ARF)	Compilation of publicly available statistics on health care capacity and providers at the county level. Data set also includes population descriptors on age and poverty status	County, state	Annual	All counties in the U.S.		x				
The Community Tracking Study (CTS) (RWJF)[16,17]	Cross-sectional study (Note: survey administration is intermittent, but consistent/ongoing) Data derived from national representative sample of communities through telephone interviews and site visits (limited over time to 12 communities) with both individuals and physicians.	Target communities	Every two years	12 communities	x		x	x	x	x

[16] http://www.icpsr.umich.edu/icpsrweb/content/HMCA/community-tracking-study.html
[17] http://www.hschange.com/index.cgi?data=12

Chapter 6: Coordinated Evaluation Design Recommendations

6.1 OVERVIEW

The primary purpose of this chapter is to identify, examine, and recommend methods for a coordinated evaluation design to assess the extent to which new HHS delivery system reforms and ACA provision reforms affect health spending, quality, patient experience, and overall performance of the health care delivery system. A related goal is to make recommendations to facilitate the comparison of Medicare, Medicaid, and Dual Eligible delivery system reform initiatives that have different methods for achieving overlapping goals. This is very challenging because ACA initiatives are complex and occur in a dynamic environment with competing or complementary state, local, and private efforts. Organizational readiness to participate in these activities is also an important feature, with implications for how results can be generalized beyond any given pilot.

One strategy for dealing with the complexity of overlapping initiatives is to find ways to break down the complexity into simpler, more heuristic models. For example, it is possible to think of cost, quality, and access as a function of organizational capabilities and incentives. In this simple model, changes in one or both of these factors (as stimulated by ACA initiatives and other environmental factors) lead to changes in outcomes. This heuristic model provides a framework for understanding a range of intervention strategies across different initiatives.

Even with a simplified model, there is still a strong desire to test causal relationships: Did X cause Y (e.g., *Did a change in financial incentives lead to reduced readmissions?*). The idealized evaluation design uses random assignment and strict control over access to treatment to address this type of question. However, in the context of most ACA initiatives, randomization is not possible. Without randomization, identifying a strong counterfactual is essential, particularly when it comes to individual site evaluations.

Thinking across ACA initiatives, however, requires moving beyond observed spending and quality measures to also consider factors affecting implementation at the organizational level—such as organizational context, institutional willingness to participate, and fidelity to an evidence- or theory-based intervention under investigation. All of this contributes to a greater understanding of policy or intervention levers that can be used to create change.

Finally, we need to consider sustainability. We can no longer afford one-time fixes that produce a short-term outcome that is quickly washed away by broader underlying trends that put upward pressure on utilization and cost (with little noticeable improvement in quality). The ultimate success of the ACA initiatives as a group will be measured by systematic changes across multiple delivery systems to produce new ways of delivering care. If, for example, you subscribe to Weisbrod's (1991) hypothesis that changing technology drives increased costs, then the success of ACA is defined by transforming how new technologies are used.

Regardless of whether you subscribe to the Weisbrod hypothesis or another theory of why health care costs have been rising, the success of the ACA ultimately rests on sustained, broad-based change, such as new community standards of medicine—meaning a new culture and process of adopting, diffusing, and utilizing new technology. Single sites, or even single markets, cannot accomplish this level of change in isolation. It will require a culture of change involving broad value-consciousness within and across markets. The ultimate success of ACA initiatives rests on this type of macro-level, sustained change.

Thus, we would propose that an evaluation framework for ACA initiatives include:

- Clear delineation of the causal model that underlies either a single initiative or a set of initiatives
- A main effects analysis that analyzes changes in outcomes at a broad national level
- Extending the main effect model to sub-regions through one of the scenarios of fixed/random effect constructs within the main effect model, or by designing a replica of the main effect model for the sub-region also known as sub-group analysis
- Drill down analysis that focuses on the drivers of success
- Testing for sustained effects over time

The following sections define a cross-initiative perspective and elaborate on each dimension of the proposed framework.[18]

6.2 SPECIFIC DIMENSIONS IN THE ACA INITIATIVE EVALUATION FRAMEWORK

6.2.1 Defining a Cross-Initiative View

Each ACA initiative considered in this report is a complex combination of delivery system reforms and incentive payments. Some are national; others are more limited demonstrations in states or selected market areas. Each initiative has its own timeline–with some new efforts and some that build on existing efforts that have been underway for years. Clearly there is a tremendous amount of variability when it comes to evaluating a single initiative, let alone thinking across multiple ACA initiatives as a group.

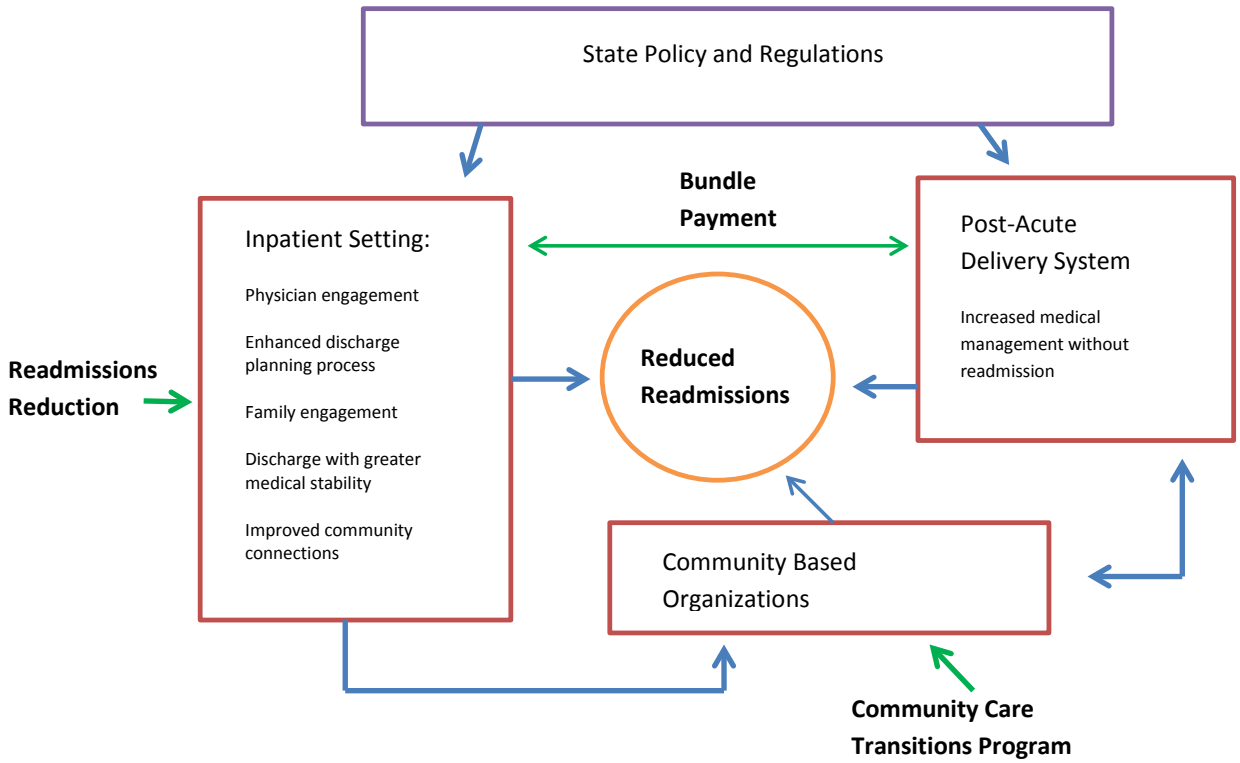
For purposes of this review, the emphasis is placed on the cross-initiative perspective—*How do we look across a set of activities and understand what is going on?* This view builds on the existing evaluation work that is already taking place within CMS. In other words, program-level evaluation and monitoring activities are inputs into this cross-initiative effort.

How do you look across initiatives? One way is to create discrete combinations of initiatives that are focused on a common outcome. A good example of this is readmissions reduction—several initiatives focus on this outcome using different incentives or targeting different parts of the delivery systems.

[18]Another important consideration is level of evidence –Within a policy context, different decisionmakers may require different evidence to expand a demonstration (e.g., budget neutrality, positive health impacts). Who is the target audience for evaluation results and what do they consider valid evidence of an effect?

As shown in **Exhibit 6-1**, the readmissions reduction initiative directly targets hospitals, in this case penalizing excess readmissions for certain conditions. The Bundled Payment (BP) initiative also involves hospitals, but creates strong incentives for those hospitals to directly engage post-acute providers.[19] Finally, the Community-based Care Transitions Program (CCTP) brings in community-based agencies to help keep beneficiaries stable outside of institutional settings. One demonstration offers a bonus (difference between bundled payment and costs) for reducing readmissions rates, another creates a penalty. The third demonstration pays for a new service in a different part of the delivery system. In combination, the three initiatives may have an additive effect or any one may be effective alone—we don't know. Untangling this type of question is the goal of cross-initiative evaluation.

Exhibit 6-1: Schematic of Relationship between ACA Initiatives that Target Readmissions Reeducation



By pooling information (both quantitative and qualitative) across logical groupings of initiatives, there is an opportunity to test different hypotheses and assess the overall effect of all initiatives. Some outcomes, like costs, may have almost all initiatives, while others outcomes may be more specialized.

Another option for combining initiatives is a common evaluation strategy, such quasi-experimental designs with pre-post comparisons. The assumption here is that treatment effects can be pooled for detailed quantitative analysis of common effects. Although there may be more

[19] Note, for Model 3, the bundle begins in the post-acute setting, not the inpatient setting. In either case, the two systems need to work together.

analytic options when two evaluation designs use common methods, the policy relevance of pooling this information may be less clear. As a result, the sections that follow focus on groups of initiatives with common outcomes.

6.2.2 Clear Delineation of the Causal Model

An important aspect of defining evaluation questions is understanding the theoretical model or best practice that underlies the initiative. Although most ACA initiatives focus on reducing costs, improving quality, and increasing access, each ACA initiative takes on the delivery system from a different perspective. Even in overlapping demonstrations there are different drivers targeting different parts of the delivery system, as demonstrated in the readmissions example above. A theory-driven evaluation approach will consider the entire model, testing for direct and indirect effects of hypothesized causal pathways.

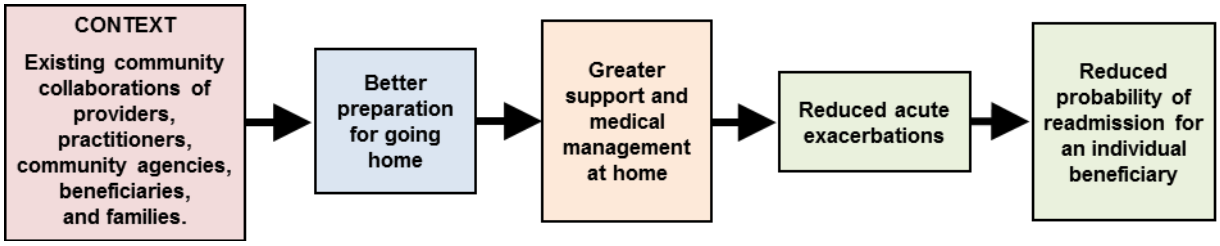
Focusing specifically on CCTP, there is tremendous variation in the specific intervention models across sites (with approximately 70 sites in the initial phase). An important question across all 70 sites is whether or not they are able to reduce readmissions without increasing other adverse events (e.g., ED visits). However, expansion of CCTP, which creates a new group of providers able to received CMS payments, requires knowledge of which types of organizations are capable of effectively receiving and utilizing these payments. This becomes an important question which could easily be missed if the evaluation design focused exclusively on reduced readmissions.

This idea of crafting questions that capture what you need to know from a policy perspective is even more important at the cross-initiative level than the individual initiative level, because the additive effects only emerge at this higher level. Using this approach, you can develop and test theoretically important assumptions about change.

6.2.2.1 Outcomes

Although we tend to think about outcomes as independent concepts, they are in fact derived from the theoretical causal model—the more specific the model, the more specific the outcome. Once again, consider readmissions. This can be defined as both the rate of readmissions for each admission (i.e., readmissions/admissions) or a population rate of readmissions (readmissions/ 1,000 Medicare beneficiaries). Both are valid measures, but one may have greater theoretical value when it comes to reducing costs or increasing quality.

The theoretical model across multiple ACA initiatives is likely to include a number of processes that can be defined as distinct causal chains.



Each casual chain can be used to identify intermediate and ultimate outcomes. In some situations, it may not be possible to measure or observe the ultimate outcomes (farthest right point in the causal chain). In this situation, the goal is to move from right to left along the causal chain until you find an outcome that can be measured. The closer you are to the ultimate outcome, the stronger the evidence of the desired effect. However, the causal model provides a link between intermediate and ultimate outcomes, building evidence that the ultimate outcome is possible.[20]

It may also be desirable to look across multiple outcomes, understanding the synergistic relationship between, for example, cost and quality. In both cases, building composite or latent outcomes may facilitate analysis.

Another important aspect of selecting outcomes relates to the evaluation questions themselves. Given the fact that the nation continues to experience relentless increases in healthcare costs, it is no longer enough to ask: *Did the initiatives produce change?* Instead, the fundamental evaluation question can be rephrased to: *Did the initiative produce enough change?* This shift in emphasis is important because it moves the conversation from measures of improvement to a real discussion of measures of attainment.

More specifically, improvement measures any change in the outcome of interest—*Is the value at time 2 better than the value at time 1?* However, attainment is a measure of how close you get to a given target. This could be a real value, like the top performer last year or a theoretical target. The idea of attainment builds in the concepts of enough change to make a difference. This may have to be calculated for cross-initiative evaluation activities if it is not part of the underlying, initiative-specific evaluation materials.

6.2.3 Main Effects Analysis

As suggested by the proposed framework, the evaluation of multiple ACA initiatives as a group requires analysis at multiple levels within the system. Specifically, the design needs to consider the effect on populations, organizations, and the environment. In addition, we need to ask not only: *Did it work (enough)?* but *Why did it work?* These questions can be addressed through a multi-level design that begins with main effects and then moves on to drill down analyses that consider specific causal pathways or the impact on specific populations.

Population Health Management lends a useful perspective for thinking about main effects. For respective populations that are treated by ACA interventions, the analysis seeks to identify reliable changes that are consistent with hypothesized effects— in nature, direction, and magnitude. The respective populations could be defined according to the interventions, in other words, an intent-to-treat evaluation approach. At its most basic level, this main effects analysis helps determine if anything observable happened for whole communities, states, or even the nation as a whole. It is fully possible at this high level that positive and negative effects are pooled, showing no change. In terms of national priorities, this is important information.

[20] Note, previous research can often be used to validate the relationships between concepts in a causal chain. However, to the extent that ACA initiatives draw on new causal processes, there may not be a strong body of literature linking short-term or intermediate outcomes to ultimate outcomes.

As an example of this approach, you could posit a series of useful variables that indicate health-related success, such as hospital days/1000 for CHF or COPD. You could then track such variables and compare populations over time (the difference in difference) and defining success. Based on that criterion, you can proceed to subgroups or examination of the causes of success.

6.2.3.1 Modeling the Collective Effect of Multiple ACA Initiatives

The main effects can be expended to consider the impact of multiple ACA initiatives. One strategy to accomplish this capitalizes on variation in the distribution of initiatives to identify an overall treatment effect. The modeling assumes that the unit of analysis will be organizations (e.g., hospitals), or geographic areas (e.g., states), where multiple ACA initiatives are taking place. This approach assumes patients are assigned to only one ACA intervention and the share of patients in each organization/area is known. There were two specific concerns raised during the TEP meeting with regard to the main effect models. The major concern was the construct of such a model in capturing the effect of multiple interventions on multiple outcomes. In other words, any main effect model has to fulfill the condition of many-to-many causality relations from interventions to outcomes.

The other TEP concern was how to model so called "hot-spot" impacts. In other words, any main effect construct has to take into consideration the synergetic or even competing/contradicting effects of simultaneous interventions taking place in certain organizations or regions.

The following sections describe a series of possible scenarios and how they would be tested analytically. Throughout these models, we explain how we have analytically addressed the two aforementioned TEP concerns.

6.2.3.2 Hypothesis I: Independency of Interventions

Under this hypothesis, each ACA intervention has an independent effect that is not impacted by other interventions in any way. Although this may seem unrealistic, it is possible in some cases that two specific initiatives are orthogonal, targeting different patient populations. This is possible in situations where CMS wants to avoid double dipping (or multiple incentive payments) for a single beneficiary. However, it is harder to believe that organizational changes associated with one initiative do not spill over to other aspects of care delivery.

If we believe in independency of interventions, the main effect model will control for intervention variables using a set of dummy variables, one per ACA intervention. For example, if we were about to model the effect of Hospital Value-Based Purchasing (HVBP), readmission reduction, and BP, we would employ three dummy variables—one per ACA intervention—and assign them the value 1 for all organizations/regions that have implemented any of them. An organization with all three interventions would get three ones for three dummies and an organization with none would get three zeros. The latter organization is in fact a control organization in our example. The equations could be written as:

$$Outcome\ 1 = \alpha + \beta\ Org\ Vars + \gamma\ Intervention\ Dummies + \lambda\ other\ control\ vars + \mathcal{E}$$

$$Outcome\ n = \alpha + \beta\ Org\ Vars + \gamma\ Intervention\ Dummies + \lambda\ other\ control\ vars + \mathcal{E}$$

Assuming *n* is the number of ACA interventions, we need to estimate *n* separate equations—one per outcome variable. The term $\mathcal{E}$ represents the unobserved portion of impact on the outcome variable. Each individual model would perhaps defer in the set of organizational variables as well as other control variables according to theories and the literature. For example, the set of organizational factors that could impact cost as an outcome variable could differ from those that affect access or quality. So far, the proposed construct which implies separate equations for separate outcomes has addressed the many-to-many relationship emphasized by the TEP members. In the next set of hypotheses, we explain how the TEP's argument around hot-spotting phenomenon could be addressed.

6.2.3.3 Hypothesis II: Synergy Between Specific Types of Interventions; or Between all Interventions

This hypothesis suggests that ACA interventions lead to substantial effects on overall organizational structure and culture, resulting in changes in provider practice as well as inter-organizational collaboration for better care coordination. Under this model, due to synergetic effects, the collective effect size of multiple interventions will be greater than sum of the individual intervention effects over all interventions.

The common mathematical modeling approach in this case would be to add the interaction terms between those interventions deemed to fulfill the synergetic effect assumptions. This is operationalized by taking the model for *hypothesis I* and adding all relevant interaction terms. Equations could be revised as:

$$Outcome\ 1 = \alpha + \beta\ Org\ Vars + \gamma Intervention\ Dummies + \mu Interaction\ between\ Intervention\ Dummies + \lambda\ other\ control\ vars + \mathcal{E}$$

...

$$Outcome\ n = \alpha + \beta\ Org\ Vars + \gamma Intervention\ Dummies + \mu Interaction\ between\ Intervention\ Dummies + \lambda\ other\ control\ vars + \mathcal{E}$$

Assuming the previous example with three ACA interventions, we need to add three additional dummy variables, one per pair-wise interactions (HVBP*ReadmReduc;

HVBP*BndlPay;BndlPay*ReadmReduc).

A more complete version of this hypothesis suggests that not only the presence of certain interventions is needed to achieve the synergetic effects, but also a minimum threshold of penetration is required for individual interventions to have a big enough effect size on organizational variables. This enhancement, above and beyond straightforward interaction effects for pair-wise ACA interventions, is in fact proposed to examine a simple version of the TEP mentioned hot-spotting phenomenon. The enhanced hypothesis suggests that the synergetic effect of three interventions (HVBP, readmission reduction, and BP) in Hospital A where patients are divided equally between the three interventions is greater than Hospital B, where 100 percent of patients are enrolled in HVBP and only 20 percent are subject to readmission reduction or bundled payment. Similarly the synergetic effects in Hospitals A and B are both greater than in Hospital C where there is no enrollment in bundled payment and the share is 80 percent HVBP and 20 percent readmission reduction.

While it may be expensive and challenging to precisely define the ingredient needed to achieve desired organizational change, there is a mathematical approach for this enhanced hypothesis that involves the calculation of the appropriate concentration index among all ACA interventions implemented in any given organization.

Among many concentration measures, the Herfindahl-Hirschman Index (HHI) offers a reasonably good picture of penetration for multiple ACA interventions in each organization. The HHI is calculated by squaring the patient population share of each ACA intervention being implemented in a given organization and then summing the resulting numbers over all interventions. The calculation formula is:

$$H = \sum_{i=1}^{N} s_i^2$$

In the formula, s_i is the patient population share of ACA intervention, i is the organization, and N is the total number of interventions in the organization. For example, in a hospital with three ACA interventions with equal patient shares of 33 percent each, the HHI is $0.33^2 + 0.33^2 + 0.33^2 = 0.11 + 0.11 + 0.11 = 0.33$. However if one of the interventions dominates the hospital and the shares become 5 percent and 10 percent and 85 percent, the concentration index will increase to 0.74, suggesting a substantial increase in concentration in favor of one dominant intervention.

In the same hospital, if the number of interventions increases from 3 to 5 all with equal share, then the index drops to 0.2, suggesting a more equal share. The HHI ranges from $1/N$ to one, where N is the number of ACA interventions in the organization. Lower values suggest a high degree of integration across initiatives and higher values suggested less integration.

By adding the HHI for each participating provider on top of the full set of individual ACA intervention dummy variables and interaction effect dummy variables, we will be able to test the enhanced revision of the *Hypothesis II* within our evaluation model. The equations will be:

Outcome 1 = α + β Org Vars + γIntervention Dummies + μInteraction between Intervention Dummies + η HHI + λ other control vars + Ɛ
...
Outcome n = α + β Org Vars + γIntervention Dummies + μInteraction between Intervention Dummies + η HHI + λ other control vars + Ɛ

We can take the modeling effort of synergetic effects one step further by differentiating between contributions from certain ACA interventions deemed to be more fundamental in reforming the organizational structure. In theory, not all ACA initiatives bring about the same level of organizational change needed to boost the effect of other interventions. To model this theoretical situation, we can assign weights to initiatives' patient populations so that essential interventions receive a higher degree of importance in calculation of HHI. Since this deviates from the standard definition, we call this measure 'prioritized HHI.'

As more interventions are implemented, we expect a negative coefficient for the HHI or prioritized HHI variable. Note, by definition, more interventions means less concentration

hence lower HHI, therefore a negative coefficient suggests that a synergetic effect exists among ACA interventions (in other words, negative is good if you believe there is an interactive effect).

There is one important limitation to bear in mind with this type of modeling. In theory, if a comprehensive set of interaction terms for all possible combinations of paired interventions are already in the model, the HHI variable could potentially introduce co-linearity (correlation between independent variables). The co-linearity will increase if the synergetic effect of all interventions could be fulfilled by a few certain ACA interventions, meaning that the other interventions have trivial synergetic effects once the major ACA initiatives are taken into account. Any model that combines interaction effects and synergetic effects has to be carefully designed to avoid such technical issues.

6.2.3.4 Hypothesis III: Crowding-out Effect of Multiple Interventions

Instead of a synergistic or positive effect from multiple ACA initiatives, it is also possible that multiple interventions would lead to competing priorities and administrative burden. The logical conclusion of this argument would contradict the assumptions under *Hypothesis II*. Therefore we can also hypothesize a crowding-out effect under the condition where an organization is stressed by undertaking too many interventions. This is more serious for organizations with fewer resources. Crowding-out effects, particularly in the short-term, could undermine the valuable efforts put into multiple interventions, even worse for those interventions that might target conflicting objectives.

To test the crowding-out effect hypothesis, we can modify the functional form of the model for *Hypothesis II* for its HHI variable. In this case, we impose a nonlinear functional form to our equation, allowing appearance of crowding-out effects beyond a turning point. A simple quadratic form for our HHI variable seems a good solution. The quadratic form requires the presence of HHI and HHI squared (HHI^2) in our model. We expect an inverted U shape curve in our model estimations if the crowding-out effect competes with the collective effects of many interventions in organizations. It implies a negative coefficient for the variable HHI^2 and a positive coefficient for HHI variable. The equations would be:

$$Outcome\ 1 = \alpha + \beta\ Org\ Vars + \gamma Intervention\ Dummies + \mu Interaction\ between\ Intervention\ Dummies + \eta\ HHI + \rho\ HHI^2 + \lambda\ other\ control\ vars + \mathcal{E}$$

...

$$Outcome\ n = \alpha + \beta\ Org\ Vars + \gamma Intervention\ Dummies + \mu Interaction\ between\ Intervention\ Dummies + \eta\ HHI + \rho\ HHI^2 + \lambda\ other\ control\ vars + \mathcal{E}$$

Exhibit 6-2 shows two hypothetical examples for *Hypothesis II* and *Hypothesis III*. By definition, as the number of interventions increases and patients get distributed among more and more interventions, the HHI index decreases. HHI equals 1 is the extreme case when the entire organization is dedicated to implement only one ACA intervention. For *Hypothesis II,* where there is no crowding-out effect, the synergetic effect increases from minimum to maximum with no limitation in terms of number of interventions (from 10 percent to 100 percent in our example). For *Hypothesis III,* however, crowding-out effect implies that the collective effect increases from minimum to its maximum when a certain number of ACA

interventions take place beyond which competing and conflicting additional interventions will reduce the collective effect to a lower point. **Exhibits 6-3 and 6-4** correspond to numbers in **Exhibit 6-2** for *Hypothesis II* and *Hypothesis III* respectively.

Exhibit 6-2: Hypothetical Examples for Effect Sizes Under Endless Synergetic Effect Assumption Versus Collective Effect With Crowding-Out Effect

Penetration of Interventions	HHI for multiple ACA interventions	Hypothesis II: Endless Synergetic Effect (%)	Hypothesis III: Collective Effect w Crowding-Out Effect (%)
Only one dominant intervention	1.00	10	10
	0.95	15	15
	0.90	20	25
	0.85	30	35
	0.80	35	40
A few interventions	0.75	40	50
	0.70	45	60
	0.65	50	70
	0.60	55	80
	0.55	60	90
	0.50	65	100
Multiple interventions	0.45	70	100
	0.40	70	100
	0.35	75	90
	0.30	80	80
	0.25	85	70
Too many interventions	0.20	90	60
	0.15	95	50
	0.10	95	40
	0.05	100	30

Exhibit 6-3: Increase in Effect Size due to Synergetic Effects from Multiple ACA Interventions

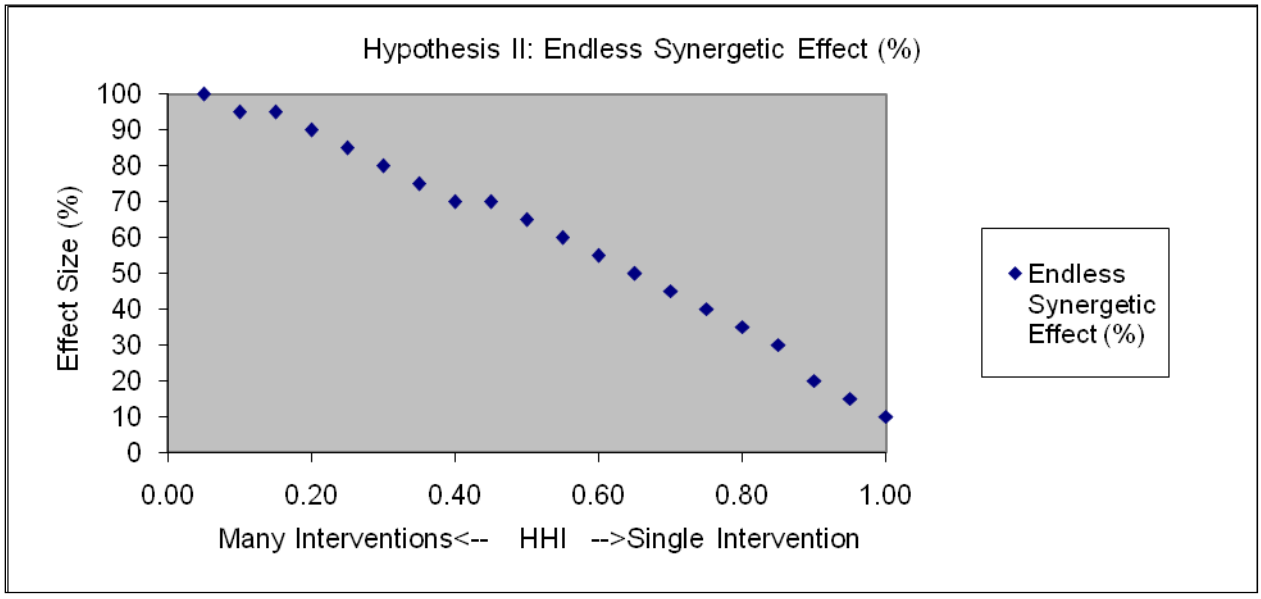
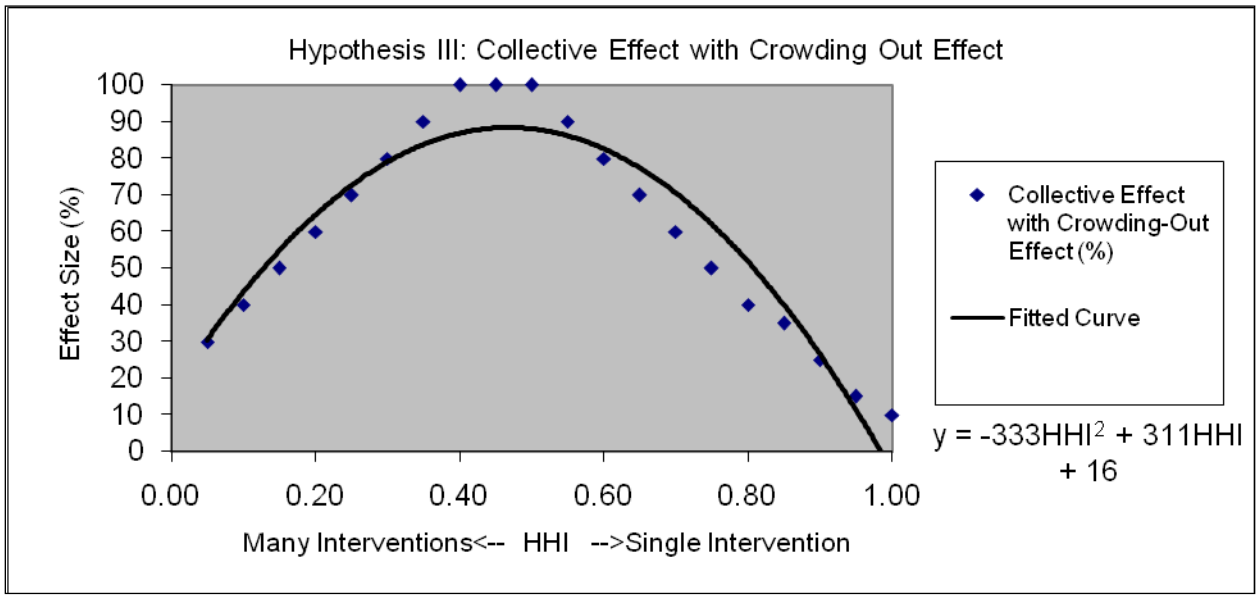


Exhibit 6-4: Turning Point in Effect Size due to Crowding-Out Effects From Too Many Concurrent ACA Interventions



A revision of *Hypothesis III* could be considered when none of the ACA interventions have conflicting goals or their concurrent implementation implies so. In other words, we can assume that crowding-out effect in this case is solely due to resource constraints. Since larger organizations tend to be both more resourceful and better situated to take advantage of economy of scale in implementation, we can use the hospital size or another indicator of resources and interact it with the HHI in our model.

6.2.4 Extending the Main Effect Analysis into Sub-Regions

Evaluators might be interested in examining the effects of ACA within specific sub-regions. We propose two solutions as an extension of our main effect model.

1. First, we can analyze all sub-regions at once by adding regional dummy variables to the main model. For example we can treat the Midwest region as the reference group and add n-1 dummy variables for other regions. However, if we believe that there is a regional effect above and beyond a simple change in the model's intercept, we can use more advanced econometric techniques to estimate the fixed or random effects due to clustering of organizations within known sub-regions. The advantage of a full-fledged random effect model is to decompose the variations in the outcome variable into region effects and explanatory variable effects. In other words, we can differentiate sources of variation both between and within regions.

2. The second choice is designing a replica of the main effect model for each sub-region, looking at each region independently. This method is also known as sub-group analysis. The advantage is that you can customize the regional model based on available data, adding explanatory variables that may be unique to a given area. However, it can be difficult to generalize the region-specific findings into other regions.

In extending the main effect models to sub-regions, the evaluator needs to consider the pros and cons of each scenario depending on the specific objectives of the analysis.

There may also be interest in sub-populations, rather than sub-regions. This can be handled analytically just like sub-regions. Instead of state or region indicators, population membership indicators would be used. If a given population is likely to have a unique experience (e.g., dual eligibles), the separate sub-group models provide more flexibility around the inclusion of specialized covariates.

6.2.5 Drill Down Analysis

Once a main effect is discovered, drill down analysis is designed to find out why or how this effect was achieved. This can be done quantitatively, by further extending the concepts used for sub-regions described above. However, in many cases, key constructs cannot be measured, either because standard assessment tools don't exist or because the mechanisms are new or poorly understood. In these cases, qualitative data, most likely in the form of selective case studies, is necessary.

In terms of a quantitative approach, we can take the concept of an indicator variable for a given initiative and replace it with an indicator variable for a hypothesized cause. In this approach, the broad flags indicating an initiative are replaced by a richer and more detailed set of explanatory variables that capture potential drivers of the observed effect. For example, rather than have flags for readmissions reduction, BP, and CCTP in the model, included would be enhanced discharge planning, medication reconciliation, para-professional home visit, ambulatory follow-up within 7 days of discharge, and so on. The idea is to model, possibly using a series of structural equations, the process or processes that are in effect.

Two potentially serious concerns with this approach include statistical power (particularly in a multi-level model) and the ability to measure all relevant constructs. Another concern is breaking apart the initiatives and the drivers. It is theoretically possible to model both together, and with a large enough sample size, this may be desirable. However, it is also possible to look at drivers outside of the specific initiative or initiatives that gave rise to their existence. This is really a different or more detailed view of the same thing. Once you know that enhanced discharge planning is a consistent driver of readmissions reduction across geographic locations, you can work backwards to understand the cause of this intermediate effect.

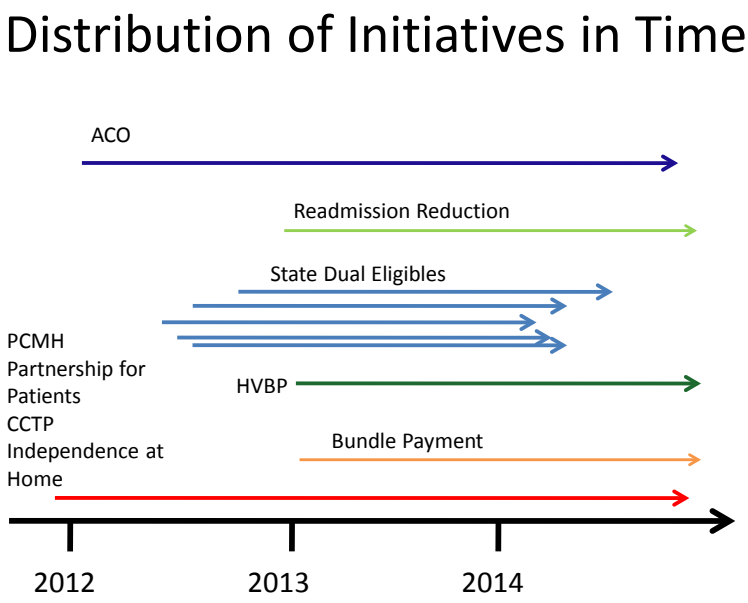
In many cases, statistical modeling will not be an option because of data constraints. In these situations, case studies or comparative case studies are a valuable tool for understanding complex causal processes. One key to making this work is a strong sampling plan for sites. Traditional models select a subset of top performers and bottom performers for indepth comparative analysis. Other options include nontreatment sites or the addition of a random sample of all sites. This work can be enhanced by value free coding methodologies, where coders may not know which sites are in the treatment group.

Strong mixed methods research includes a plan for integrating qualitative and quantitative analyses. There may be opportunities, for example, to convert qualitative findings into discrete variables. Another approach is to allow quantitative findings to drive hypothesis testing during the qualitative phase. There are other mixed methods design options as well, but the key is a plan on how to combine results.

6.2.6 Test for Sustained Effects over Time

Time plays an important role when it comes to understanding and observing the effects of ACA initiatives. First, there is the temporal ordering of the initiatives themselves. As indicated in **Exhibit 6-5,** many initiatives were launched in 2012, but by 2013, the density of initiatives may be higher, possibly peaking in 2014. In many areas, multiple initiatives may be in different phases of development at any point in time. Thus, analysis designed to look for interactive effects between initiatives needs to consider where in the timeline you are looking. Comparing the beginning of one initiative with the end of another initiative may not be the optimal point for seeing significant relationships.

Exhibit 6-5: Distribution of Initiatives in Time



Time plays another important role in terms of shaping how quickly we know what works. Although we tend to think about treatment effects as a single point estimate, they in fact have a trajectory that changes over time. Early effects may be weak, gaining momentum until the initiative reaches full implementation. Certain organizational or environmental factors may affect this trajectory, and it may be valuable to know what factors drive more rapid success. This can be tested empirically using a variety of different longitudinal analytic techniques such as survival analysis or growth curve modeling. Survival analysis captures time to an event (e.g. post-discharge ambulatory follow-up). Growth curve modeling (a form of random effects modeling), models the rate of change in a continuous variable, say readmissions, care transitions, or total cost.

Modeling time to an effect is quite similar to using achievement as an outcome rather than improvement. The basic idea is to find pathways that produce desired changes more quickly than other pathways. This acknowledges the urgency of producing meaningful change as quickly as possible.

Rapid cycle feedback during an individual initiative can be seen as a derivative of the same approach—rather than model the drivers of the underlying outcome process (a retrospective approach), rapid cycle efforts seek a prospective answer. At the cross-initiative level, this type of effort would require very careful mapping of expected and observed effects at a specific point in time. Implementing this would be quite challenging, given data and methodological limitations.

We also don't want to forget the tail of the distribution or long-terms effects. It is possible that desired effects are not observable until later than expected or that the effects strengthen over longer time horizons (e.g., 12–36 months, rather than 6–12 months). The notion of sustainability also comes into play here. As suggested by our framework, meaningful changes persist over time

and have an impact on the overall trajectory, not just one-time impacts when the initiative is new. This is important to test along with sort and medium-term effects.

6.3 RECOMMENDATIONS

This chapter has presented a framework for thinking about the evaluation of multiple ACA initiatives. The framework represents a starting point for designing specific evaluations that will focus on outcomes or policy questions of interest across multiple initiatives.

The basic approach begins with a theoretical modeling process to understand which organizations and incentives are the targets of relevant initiatives. This approach gives special attention to areas of intersection (*What should happen when BP and CCTP occur together?*). This model gives rise to a series of causal pathways or hypotheses that can be tested empirically.

The framework goes on to suggest that treatment effects be considered first at a national level and then at increasing levels of geographic disaggregation, such as the regional or state level. Both types of models represent different ways to look at the main effect of the initiatives. We have proposed a specific analytic approach designed to capture the interactive effect of multiple initiatives in a single area.

A second step in the analysis focuses on areas where we observe positive effects and drills down to identify the drivers of those effects. Here, we drop the initiatives as an independent construct, and instead use specific organizational or market area features, including the presence of incentives and penalties to predict treatment effects. These drill-down analyses are designed to open up the black box and identify causal mechanisms. Some of this work can be done empirically, but given data limitations, it is also important to use in-depth site visit information to untangle these complex processes. Secondary analysis of initiative specific case studies may be useful in this regard.

Finally, all of these modeling and analytic activities need to carefully consider time. Not only is there an interest in rapid feedback, but we also need to test for sustained effects over long time horizons.

Based on this framework, we recommend:

- A theory driven approach
- Main effects analysis accompanied by drill down analysis where significant effects are observed
- Consider measuring attainment, not just improvement
- Carefully model the effect of time, considering short-, medium-, and long-term outcomes.

6.4 SPECIFIC METHODOLOGICAL ISSUES FOR INITIATIVE SPECIFIC EVALUATIONS

Selection Bias

Comparison Groups

The primary purpose of a comparison group is to provide a counter factual—*What would have happened if the intervention had not taken place?* This is an important step any quasi-experimental design. In the case of ACA initiatives, however, it difficult to find areas that are truly comparable to the 'early adapters' who are able to participate initiatives that may require taking on risk, establishing new billing relationships or investing in care redesign, among other things. Thus, there may be limited choices for geographic based matching. Another option is synthetic comparison groups, based on propensity score matching or another matching technique, that are focused on finding similar beneficiaries to those participating in the initiatives. An even stronger design may include both approaches. One strategy would be the use of multiple comparison groups with complementary strengths and non-overlapping weaknesses.

When it comes to looking across multiple ACA initiative, the comparison group becomes, to some extent, a given, assuming this decision has already been made by the evaluation contactor in conjunction with CMS.

Comparison groups also serve a valuable role when it comes to rolling out or replicating an intervention or strategy. Understanding what is unique about treatment sites is often easier in comparison to other areas. This is important when it comes to replication. In this type of situation, qualitative data can be valuable for capturing dimensions that do not yet have standardized assessments. Qualitative data is also important for capturing barriers to participation among those who do not even apply to be part of the initiative.

Stopping rules

As discussed above, there is urgency to understanding what work to improve cost, quality and patient experience. This raises the question: *When do you have enough evidence to stop and expand a demonstration?* There is a narrow literature on stopping rules within the field of clinical research which includes trade-offs between the benefits (expected) and harms (unexpected) effects of a given treatment. Experts at CMS and Hunter College are working to expand these ideas from clinical trials to quasi-experimental designs.

Something to consider is an intermediate step between a demonstration and a full role out. Experts in the alternative quality contract, a value based purchasing strategy, and care transitions both described processes by which new models rolled out gradually, allowing organizations time to prepare. This graduated process revealed organizational readiness issues that could be resolved and improved and the initiatives increasingly reach out to less prepared or sophisticated physician groups or community organizations. This refers back to the notion that outcomes have a lifecycle of their own.

Chapter 7: Operational Considerations

7.1 OVERVIEW

In this chapter, we outline operational considerations for HHS to consider when implementing the recommendations in Chapters 4, 5, and 6. This chapter primarily discusses issues regarding the aggregation of information from Government databases and processes for collecting additional data from providers. While we cover some of these issues in the prior sections of this report, here we encapsulate key operational issues discussed in our Technical Expert Panel (TEP) meeting on June 5, 2012.

- **Create a national registry of health care organizations**. Although there are many databases with registries of individual providers, there are no good national or regional sources of information illustrating how providers are interconnected into systems. CMS is compiling good data about the composition of organizations in the Pioneer ACO and Medicare Shared Savings program. It may be especially important to future research efforts to have similar information about organizations that are not currently in these programs. The most important information in such a registry would include the physician identifiers that are associated with different types of organizations that will take responsibility for the care of beneficiaries. This includes: integrated delivery systems, physician group practices, contracting organizations (independent practice associations and physician-hospital organizations), and hospitals (i.e., hospital-employed physicians).
- **Move quickly to establish baseline organizational data**. A high proportion of new HHS delivery reform initiatives are beginning in 2012 and 2013. CMS has claims data to provide baseline information for spending and selected quality measures. Although it will have some organizational variables, it must move quickly to identify and begin collecting those it does not possess. Although much of the evaluation will assess how incentives change outcomes, it will also be valuable to look at changes in organizations that occur in response to incentives created by HHS delivery reform programs. Understanding such changes will help HHS and CMS in efforts to identify and spread successful practices.
- **Prepare a synthesis of all organization-related variables CMS is currently collecting from provider organizations**. This includes data that organizations are required to provide to CMS before they enter new programs and data that may be required to support program monitoring and evaluation. Such a synthesis may help HHS harmonize the collection of organizational variables across programs and to identify gaps in organizational variables that may require special data collection efforts.
- **Be judicious in the collection of organizational variables**. CMS and HHS must be cognizant of the cost and reporting burden of collecting organizational variables. CMS' direct data collection may best be focused on structural, capacity, information, and operational process variables. Members of the TEP emphasized that culture and leadership of health care organizations are critical for performance, but are much more difficult to measure and may be more sensitive for the Government to collect. They

suggested that using an independent academic institution to collect this type of information would be more acceptable to stakeholders.

- **Utilize qualitative research to determine the most critical organizational factors.** While quantitative analysis will tell HHS what happened, qualitative research is essential to help understand why or how it happened. Most of the CMS evaluations under development include site visits and case studies. These efforts can be used to identify and prioritize the organizational variables that provider organizations believe are most important and that are reasonable to provide through survey or other means. This indepth information can also help untangle complex causal processes that are not yet well understood.

7.2 COLLECTING BASELINE DATA AND CREATING AN INVENTORY OF ORGANIZATIONAL INFORMATION

Measuring changes in spending, quality, and delivery system organization resulting from HHS delivery reform initiatives and the ACA requires understanding the state of the world before these initiatives have been fully implemented. It is critical to collect baseline data on spending trends, the composition of organizations, and the nature the incentives facing them (including those from state and private insurer programs). The success of physician and hospital organizations in ACA programs will be correlated to their baseline capabilities. Many programs are planned for full implementation by 2013 and 2014, so timing is critical; HHS should consider strategies to collect organizational information as quickly as possible.

A first step would be to create an inventory of organizational information that is currently or planned to be collected through existing program evaluations (such an inventory was beyond the scope of this project). To augment this effort, several external data sources are suggested in Chapter 4. However, variables may likely be incomplete and inconsistent across evaluations, requiring HHS to proactively collect data directly from organizations. Several strategies of data collection of organizational variables are offered in Chapter 4. Members of the TEP emphasized this, but also cautioned about the scope of new data collection efforts. Policymakers need to balance the need for more variables against the burden of reporting. They also need to balance the higher response rate from a government-administered survey that can compel compliance against a less threatening data collection effort that might be administered by an academic group on the government's behalf.

Lastly, this data collection effort must be maintained and repeated over time to better understand changes in organizational variables and incentives. Organizational change is typically slow and difficult to implement, and organizational learning may accumulate over several years. One TEP member suggested funding data collection of these types of variables well past any official end-date of the demonstration. As described in Chapter 6, some analyses may use changes in organizational characteristics as the main dependent variable.

7.3 MAPPING PROVIDERS TO ORGANIZATIONS

Health and cost outcomes can be calculated at the beneficiary level and these patients can be readily assigned to physicians or health care organizations using claims data. However, there is no reliable public data source that links physician to organizations. HHS can greatly help the coordinated evaluation of ACA by investing in methodology that maps physicians to

organizational entities. These would include integrated delivery systems, physician group practices, contracting organizations (independent practice associations and physician-hospital organizations), and hospitals. Several potential methods for compiling this information have been suggested. One is through requiring physicians to identify their affiliation to a group practice on the claim form. Another is through determining whether some of this information is available through existing Federal data sources, private health plans, or for-profit data aggregators like IMS Health.

One area that HHS is actively exploring is the organization charts contained within the Medicare Provider Enrollment, Chain, and Ownership System (PECOS). More research is needed to evaluate the potential for PECOS, especially in its ability to link National Provider Identification (NPI) numbers to their organizational parent. Potentially, PECOS can also link other types of providers, including physician assistants, certified clinical nurse specialists, nurse practitioners, clinical psychologists, certified nurse midwives, and clinical social workers.

7.4 QUALITATIVE CASE STUDIES

Fundamentally, ACA delivery reform initiatives work under the hypothesis that a change in incentives will drive changes in organizational factors, which may lead to changes in meaningful outcomes (see Chapter 6). Some of these outcomes, such as cost and volume of services, may be more straightforward to measure through claims and other analyses. However, other outcomes, especially those that outline the progression of activities hypothesized above, may require case studies and qualitative analysis. This qualitative research will be crucial; no single mathematical analysis will be able to provide definitive answers to all the research questions posed by a coordinated evaluation. While quantitative analysis may be needed to support the CMS, Actuary, and Congressional Budget Office scoring, HHS may also need studies that help give nuance to the numbers. An analysis that combines empirical analysis with well-structured qualitative research may provide a more compelling rationale about the successes and failures of the programs.

Qualitative research in this arena can accomplish many goals. Case studies can help identify best practices and allow more advanced groups to share best practices with the groups that are fast-followers. It can provide stories of success that keep participants and the public motivated. It may also help identify important lessons on the edges of the distribution. While mathematical analyses look for the average effect, qualitative research can provide more information about specific subpopulations or types of organizations. For example, *Do these programs improve care for chronic-care patients or rural beneficiaries that are struggling to find coordinated care? Do we see a large affect among physician groups and hospitals who have not yet started on delivery reform?*

To provide more rigor, HHS should consider incorporating qualitative comparative analyses that seek to make causal inferences by comparing the characteristics from a limited number of organizations under study.[21] These qualitative analyses can approach more well-controlled studies, especially if the researchers are blinded to the organization's actual performance and

[21] Benoît Rihoux, "Qualitative Comparative Analysis (QCA) and Related Systematic Comparative Methods Recent Advances and Remaining Challenges for Social Science Research," *International Sociology* September 2006; Vol 21(5): 679–706.

independently assess the characteristics associated with successful implementation of ACA programs. Finally, qualitative studies can help refine what to collect in broader surveys and focus attention to specific, key variables in quantitative analyses. For example, many of the organizational variables recommended in Chapter 4 are correlated, so that there may be diminishing returns associated with collecting all of them.

7.5 ACCOMMODATING OVERLAP

No patient will be officially in more than one demonstration. Currently, CMS is developing a database that identifies which patients are primarily enrolled. However, these patients will undoubtedly touch other providers who are in other demonstrations. And some national initiatives, such as Hospital Readmissions policy or the Hospital VBP initiative, will affect every beneficiary who has an inpatient admission, regardless of whether they are also enrolled in an ACO or patient-centered medical home. *How will a coordinated evaluation plan handle this overlap, where multiple interventions may be occurring within the same organization or geographic region?*

To some extent, policymakers may look past these overlaps, since they do not affect the overarching question central to a coordinated evaluation: *Did the ACA affect health care costs, quality, and access?* The answer to this question is not affected by whether a patient was affected by one or more sub-initiative, just by the fact that the beneficiary's care was improved by the implementation of the ACA itself.

As mentioned in Chapter 6, policymakers may wish to conduct analyses of geographic regions and study the changes to the local health system where multiple programs are taking place. This area-specific approach will focus on the effect of cumulative changes that have a high degree of ACA activity.

By conducting such hot-spotting analyses, researchers can identify bundles of programs that are correlated with improved outcomes. True delivery system reform is likely to emanate from organizations that are focusing on more than one type of intervention (accepting risk-based contracts, implementing health information technology, changing physician compensation structures). Just as medical societies promulgate packages of best practices to treat specific diseases, policymakers could identify and emphasize specific bundles of interventions that are associated with more efficient, coordinated care.

However, these overlaps may limit the ability to answer important sub-questions—*Which initiatives were responsible for the lion's share of the overall effect? What distinguishes organizations that successfully implemented these programs?* More research may be needed to explore how beneficiaries are assigned to specific, potentially overlapping programs.

Appendix

Notes from the Technical Expert Panel Meeting, June 5, 2012

OVERVIEW

This appendix summarizes the day-long meeting of the Technical Expert Panel (TEP), which was held on June 5, 2012 in the offices of the Assistant Secretary for Planning and Evaluation, HHS in Washington DC. While not an exact transcription, main points are summarized here for reference and grouped by theme; they have subsequently been incorporated into chapters of the Final Report. The agenda and list of participants are listed at the end of this appendix, along with a list of the acronyms used in this report.

SUMMARY POINTS FROM THE TECHNICAL EXPERT PANEL

Conceptual Approach to a Coordinated Evaluation

- The task assumes two fundamentally different questions: 1) Did the sum total of ACA programs achieve better health, access, and/or cost reductions (e.g., *Did it work?*); and 2) What factors were responsible for achieving these outcomes (e.g., *Why did it work*?). These two questions should be kept separate, as the methods for answering each of may differ dramatically.
- The Secretary, Actuary, and Congressional Budget Office may need hard numbers and specific health outcomes to gauge the success of these programs. Total cost of care and multi-dimensional quality metrics will be important to measure.
- The goal of this project is to determine what is the best way to tell a credible, convincing story with numbers? Evaluating the second question may help establish the validity of the first, even if the methodology isn't as scientifically rigorous. This is how business would likely evaluate such a bold strategy.

Quantitative Analyses

- A coordinated evaluation is complicated by the presence of multiple outcomes, multiple interventions, as well as multiple versions of the same intervention. Additionally, with rapid feedback and designed flexibility, the intervention may not be static over time. Most of these interventions will not have randomization or well-defined control groups, but that is more of an issue for individual program evaluations.
- The answer to the primary question (*Did it work?*) should be conducted in a population-based, intent-to-treat analysis.
- Many ACA initiatives assume that a change in outcomes may occur with a necessary change in organizational capabilities and incentives. Interventions may be combined if they are aimed at achieving changing similar incentives or organizational structures.
  - Intervention ΔIncentives → ΔOrganizational Factors → ΔOutcomes

  - Because of this, it may also be important to conduct sub-analyses that investigate how organizational factors have changed over time. Analyses should consider the multiple levels within the data: beneficiaries within organizations within interventions. Determining what changes are exogenous or endogenous to the intervention may be challenging, but it may give an indication of the effect of the ACA on delivery system reform.
  - Be wary of endogeneity of case mix and coding in any model. Coding practices are subject to significant changes in response to financing and organizational reforms. Coding practice changes will be reflected in case-mix, counterfacting any real changes in population health and co-morbidity. Thus, Interventions → Coding → Case-Mix → Outcome. Most of the time, the intermediate variables are less prone to this problem.

- Quantitative analyses will likely look at changes over time, whether through a differences-in-differences or interrupted-time-series methodology. It is critically important to understand and measure the nature of capabilities and incentives at baseline.
- Many factors will contaminate a time-series look at ACA interventions: the recession, private market influences, changes in supplemental coverage, state health policy decisions.
  - A right model construct with wise use of dummy variables can treat states like Massachusetts that are subject to many reforms simultaneously. Something like concurrent reform dummy variables applied to a few states with many reforms.

Qualitative Analyses

- Qualitative analyses get at the second question (*Why did it work?*).
- Methodological rigor can be maintained by sending teams to evaluate sites that are blinded to the organization's performance. They can independently identify patterns of high and low performers and then correlate that with actual performance data (see Elizabeth Bradley, Yale University)
- Qualitative Comparative Analysis uses a small number of case studies (e.g., 5-30) to characterize variables associated with each case and regress them on outcomes. It helps to establish correlations between performance and organizational factors (see the work of Charles Ragin).
- Case studies and field interviews of sites are labor-intensive, but they can be limited. Typically, only a few site visits are needed to adequately inform the variables to then collect in larger survey of organizations. The precise number cannot be predicted ahead of time.
- Surveys, case studies, and other data collection efforts can be conducted by HHS or outsourced to academics. Policymakers may need to balance the benefits and risks of each strategy (compliance, burden of the data collection, expense, etc.). The principal investigators for the National Survey of Physician Organizations estimate that it takes about 40 minutes to conduct an exhaustive survey of physician organizations (with skip patterns).

Organizational Variables

- Organizational data are very important. However, the number of organizational variables collected should be parsimonious, based on the limited degrees of freedom in many demonstrations and the burden of additional reporting to government and private programs.
- Currently, there is no census of medical groups or hospital-owned physicians. The U.S. health system lacks a standardized attribution census for physicians. It is always hard to identify provider organizations for individual physicians. IMS Health has some of these data, but they are expensive. CMS can greatly help the health management field by mapping physicians to organizations, either through claims data (PECOS) or through other mechanisms (i.e., health-information technology adoption), and updating such mapping annually.
- It is important to understand which organizational variables are fixed and which are proxies for something more fundamental (e.g., culture).
- Culture and leadership are the least defined variables on the list presented. They are important but also the most complicated, least understood, and expensive variables for data collection and analysis. The Veterans Health Administration has been conducting these types of studies and may have recommendations for how to implement such surveys.
- Individual organizations may already be conducting many of the same kinds of cost and quality analyses for their internal purposes. HHS can help coordinate the sharing of best practices by hosting conferences and meetings.

Other Comments

- Organizational variables need to consider other non-physician providers who are also important contributors to the health care team.
- Need to evaluate beneficiaries and organizations across the spectrum. The ACA may have the biggest impact on the struggling beneficiaries at the ends of the illness distribution, or the health groups that haven't really started to change yet.
- While patient satisfaction is important and represented in the variables, it is also critical to gauge physician- and nurse-satisfaction, turnover, and morale. These give an indication of whether providers can maintain these improvements.
- Rapid cycle evaluations are crucial because most of the reforms have been already started and spending is in its warm-up phase. Rapid cycle evaluation would effectively contribute to dissemination of best practices. If we find a mechanism through which all participating organizations receive the success and failure stories of their peers, it may add to the efficiency and speed of implementation.
- Need to consider other options beyond the arbitrary $p<=0.05$ criteria. Many health care organizations may go ahead with a project if there is an 80 percent chance of success, or a test that showed that the probability of saving money is greater than 50 percent.

Evaluation of HHS Delivery System Reform Efforts and Affordable Care Act Provisions:

Technical Expert Panel Agenda
Washington, D.C., June 5, 2012

I: Charge for the Day (10.00 – 10.30)

This session will review the overarching objective of this project – to provide recommendations for a coordinated evaluation design that would tie results from numerous Medicare, Medicaid, and Dual Eligible delivery system reform initiatives and ACA provisions into an overall model of health system change. This session will briefly review the range of HHS delivery reform initiatives included in the project and outline meeting objectives, which will focus on: 1) enhancing and strengthening designs for individual program evaluations; 2) identifying common outcome measures and covariates for the evaluations; and 3) methods for synthesizing results from multiple demonstrations

II: Case Study and Agenda (10.30 – 10.45)

This session will briefly review the ACA initiatives that we will use as an illustrative case study for discussing methodological issues in evaluation design. After reviewing how specific methodological issues might apply to a specific program, we will expand the discussion to assess whether the technical expert panel's conclusions could be generalized across multiple HHS delivery reform initiatives (e.g., ACO, Bundled Payments, Partnership for Patients, medical home, dual eligibles etc.).

III: Defining the Intervention and Organizational Context (10.45 – 12.00)

The impact of new interventions will vary depending on the characteristics and capabilities of the organizations that participate in these initiatives and their local market environments. This session will review the team's recommendations for organizational and market variables that may be associated with quality and cost performance and discuss how best to collect common organizational metrics across multiple programs.

IV: Approaches to Evaluation (12.30 – 2.00)

This session will examine ways to augment individual evaluations to better support a coordinated evaluation design. This session will explore methodological issues in the evaluation of programs, including:

1. Study design (randomization, interrupted time series, regression discontinuity)
2. Comparison groups (geographical matching, selection bias, program overlap)
3. Evidence & stopping rules (feedback reporting, continuous/rapid evidence generation)

V: Design Challenges for Comparisons Across Programs (2.00 – 3.00)

Beyond the shared analytical hurdles within each initiative, this session will explore methodological and design issues for creating a coordinated evaluation across the diverse set of programs. Potential frameworks could include: estimating effect sizes from each individual evaluations, pooling/clustering demonstrations of common study designs, utilizing comparative effectiveness analysis, conducting micro-simulations and dynamic modeling, and employing qualitative and mixed methods analysis

TEP MEMBERS			
Name	**Title**	**Organization**	**email**
Jim Burgess, Ph.D	Professor of Health Policy and Management	Boston University	jfburges@bu.edu
Larry Casalino, MD, Ph.D	Professor	Cornell University	lac2021@med.cornell.edu
Michael Chernew, Ph.D	Professor	Harvard Medical School	chernew@hcp.med.harvard.edu
Partha Deb, Ph.D.	Professor, Senior Advisor	Hunter College, CMMI	partha.deb@hunter.cuny.edu
Thomas Graf, MD	Vice President	Geisinger Health System	TRGRAF@geisinger.edu
Kate Koplan, M.D.	Director of Medical Management	Atrius Health	Kate_Koplan@atriushealth.org
Steve Sheingold, Ph.D	Division Director	Assistant Secretary for Planning and Evaluation, HHS	Steven.Sheingold@HHS.GOV
Will Shrank, M.D., MSHS	Director, Rapid-Cycle Evaluation Group	Center for Medicare and Medicaid Innovation	William.Shrank@cms.hhs.gov
Christopher Tompkins, Ph.D	Associate Professor	Brandeis University	tompkins@brandeis.edu
BRANDEIS UNIVERSITY STAFF			
Robert Mechanic, M.B.A.	Senior Fellow	Brandeis University	mechanic@brandeis.edu
Jennifer Perloff, Ph.D.	Scientist	Brandeis University	perloff@brandeis.edu
Darren Zinner, Ph.D.	Scientist	Brandeis University	dzinner@brandeis.edu
Moaven Razavi, Ph.D.	Senior Research Assoc	Brandeis University	mrazavi@brandeis.edu
CDM STAFF			
Kathryn Herron	President and CEO	CDM Group	kathryn.herron@cdmgroup.com
Candice Grayton	Consultant	CDM Group	Candice.Grayton@cdmgroup.com
PRIMARY ASPE/CMS CONTACTS			
Susan Bogasky		Assistant Secretary for Planning and Evaluation, HHS	Susan.Bogasky@hhs.gov
Jun Li		Center for Medicare and Medicaid Innovation	Jun.Li@cms.hhs.gov

Glossary of Acronyms

ACA	Affordable Care Act
AHRQ	Agency for Healthcare Research and Quality
CDC	Centers for Disease Control and Prevention
CMS	Centers for Medicare & Medicaid Services
MEPS	Medical Expenditure Panel Survey
NNHS	National Nursing Home Survey
FFS	Fee-for-services claims
HEDIS	Healthcare Effectiveness Data and Information Set
PQIs	Patient Quality Indicators
PSIs	Patient Safety Indicators
EHR	Electronic health record
NCQA	National Committee for Quality Assurance
AMA	The American Medical Association
PCPI	Physician Consortium for Performance Improvement
CPT-2	Current Procedural Terminology codes
MHQP	Massachusetts Health Quality Partners
CAHPS	Consumer Assessment of Healthcare Providers and Systems
HCAHPs	Hospital Consumer Assessment of Healthcare Providers
ACO	Accountable Care Organizations
NHANES	National Health and Nutrition Examination Survey
BRFSS	Behavioral Risk Factor Surveillance System
NHAMCS	National Hospital Ambulatory Medical Care Survey
ARF	Area Resource File
AHA	American Hospital Association
BP	Bundled Payment
ISI	Information Sciences Institute
HMO	Health Maintenance Organization
QUERI	Quality Enhancement Research Initiative
HCUP	Healthcare Cost and Utilization Project
MEPS	Medical Expenditure Panel Survey
ONC	Office of the National Coordinator for Health Information Technology within The Office of the Secretary for the Department of Health and Human Services
MAX	Medicaid Analytic eXtract files
UDS	Uniform Data System
FEHC	Family Evaluation of Hospice Care Survey
NCDB	National Cancer Data Base
NHCS	National Hospital Care Survey
TEP	Technical Expert Panel
CCTP	Community-based Care Transitions Program
HVBP	Hospital Value-Based Purchasing
HHI	Herfindahl-Hirschman Index
PECOS	Medicare Provider Enrollment, Chain, and Ownership System (PECOS)