December 16, 2013

Measurement, Monitoring, and Evaluation of State Demonstrations to Integrate Care for Dual Eligible Individuals

Aggregate Evaluation Plan

Prepared for

Normandy Brangan
Rapid-Cycle Evaluation Group
The Center for Medicare and Medicaid Innovation
Centers for Medicare & Medicaid Services
Mail Stop WB-06-05
7500 Security Blvd.
Baltimore, MD 21244

Submitted by

Edith G. Walsh, PhD, Project Director
RTI International
1440 Main Street, Suite 310
Waltham, MA 02451-1623

RTI Project Number 0212790.003

RTI Project Number
0212790.003

Measurement, Monitoring, and Evaluation of State Demonstrations to Integrate Care for Dual Eligible Individuals

Aggregate Evaluation Plan

Authors

RTI International
Edith Walsh, PhD
Wayne Anderson, PhD
Angela Greene, MS, MBA
David Kaiser, MA
Galina Khatutsky, MS
Brieanne Lyda-McDonald, MSPH
Melissa Morley, PhD
Rebecca Perry, MSc
Kevin Smith, MA
Joshua Wiener, PhD

American Institutes for Research
Andrew Thomson, BS

Muskie School of Public Service, University of Southern Maine
Maureen Booth, MA
Julie Fralich, MBA

National Academy for State Health Policy
Diane Justice, MA

Urban Institute
Lisa Dubay, PhD
Timothy Waidmann, PhD

Project Director
Edith G. Walsh, PhD

Federal Project Officer
Normandy Brangan

RTI International
1440 Main Street, Suite 310
Waltham, MA 02451-1623

December 16, 2013

Contents

Tables

Number **Page**

EXECUTIVE SUMMARY

The Medicare-Medicaid Coordination Office (MMCO) and Innovation Center at the Centers for Medicare & Medicaid Services (CMS) have created the Financial Alignment Initiative to test integrated care models for beneficiaries who are dually eligible for Medicare and Medicaid (Medicare-Medicaid enrollees). The goal of the Financial Alignment Initiative is to develop person-centered care delivery models integrating the full range of medical, behavioral health, and long-term services and supports (LTSS) for Medicare-Medicaid enrollees.

CMS contracted with RTI International to monitor the implementation of the demonstrations and to evaluate their impact over time on beneficiary experience, quality, utilization, and cost. This report describes the *Aggregate Evaluation Plan* that will guide the overall evaluation. RTI will develop separate State Evaluation Reports for each individual State participating in the Financial Alignment Initiative. The activities described in this report may be revised if modifications are made to the demonstrations or if other circumstances change during the demonstration period. Although this document will not be revised to address all changes that may occur, the annual and final evaluation reports will note areas where the evaluation as executed differs from this evaluation plan.

Section 1—Introduction

The goals of RTI's evaluation are to monitor each State's demonstration implementation, evaluate the impact of these demonstrations on the beneficiary experience, monitor unintended consequences, and monitor and evaluate the demonstrations' impact on a range of outcomes for the eligible populations as a whole and for subpopulations (e.g., people with mental and/or substance use disorders, LTSS recipients). To achieve these goals, RTI will collect qualitative and quantitative data from States quarterly; analyze Medicare and Medicaid enrollment, claims, and encounter data, and data from the Nursing Home Minimum Data Set (MDS); conduct site visits and interviews with staff involved in the demonstration, beneficiary focus groups, and key stakeholder interviews; and incorporate relevant findings from beneficiary surveys.

RTI will report preliminary information to CMS in an initial 6-month implementation report and quarterly data reports to CMS and States. RTI will also integrate this information into annual State-specific and cross-cutting reports and a final evaluation report. The key research questions and data sources for the evaluation are summarized in ***Table ES-1***.

The principal focus of the evaluation will be at the State level. CMS has engaged an operations support contractor to monitor fulfillment of the demonstration requirements outlined in the Memoranda of Understanding (MOUs), contracts, and final agreements, including Medicare-Medicaid Plan (MMP)-level monitoring in capitated States. RTI will integrate that information into the evaluation as appropriate.

Table ES-1
Research questions and data sources

Research questions	Stakeholder interviews and site visits	Beneficiary focus groups	Claims and encounter data analysis	Demonstration statistics[1]
1) What are the primary design features of each State's demonstration and how do they differ from the State's previous systems?	X	X	—	X
2) To what extent did each State implement its demonstration as designed? What factors contributed to successful implementation? What were the barriers to implementation?	X	—	—	X
3) What impact do these demonstrations have on the beneficiary experience overall, by State and for beneficiary subgroups? Do beneficiaries perceive improvements in how they seek care, choice of care options, how care is delivered, personal health outcomes and quality of life?	X	X	—	X
4) What impact do the demonstrations have on cost and is there evidence of cost savings in each State? How long did it take to observe cost savings in each State? How were these savings achieved in each State?	—	—	X	—
5) What impact do these demonstrations have on utilization patterns in acute, long-term, and behavioral health services, overall, by State, and for beneficiary subgroups?	X	X	X	X
6) What impact do these demonstrations have on health care quality overall, by State, and for beneficiary subgroups?	—	—	X	X
7) Does the demonstration change access to care for medical, behavioral health, long-term services and supports (LTSS) overall and for beneficiary subgroups, by State? If so, how?	X	X	X	X
8) What policies, procedures, or practices implemented by each State in its demonstration can inform adaptation or replication by other States?	X	X	—	X
9) What strategies used or challenges encountered by each State in its demonstration can inform adaptation or replication by other States?	X	X	—	X

— = not applicable

[1] Demonstration statistics refer to data that States, CMS, or other entities will provide regarding topics including enrollments, disenrollments, grievances, appeals, and number of health and/or medical homes. States will be providing quarterly data updates through the State Data Reporting System. States will also submit reports summarizing findings of any external quality review organization analyses, beneficiary surveys, and other quality monitoring activities required by CMS or undertaken by the States during the demonstration period.

Section 2—Demonstration Implementation

The evaluation of demonstration implementation will be based on case study methods and quantitative data analysis of enrollment patterns and is designed to answer the following overarching research questions:

- What are the primary features of each State demonstration, and how do they differ from the State's previous system available to the demonstration-eligible population?
- To what extent did each State implement the demonstration as proposed?
- Which States were able to fully implement their intended proposals?
- Were certain models more easily implemented than others?
- Were the demonstrations more easily implemented for certain subgroups?
- What factors contributed to successful implementation?
- What were the barriers to implementation?
- How have beneficiaries participated in the ongoing implementation and monitoring of the demonstrations?
- What strategies used or challenges encountered by each State can inform adaptation or replication by other States?

Demonstration Design Features. RTI will examine how each States' strategies and demonstration design features translate at the plan or practice level. ***Table ES-2*** lists the design features, and examples of key components of design features, that RTI will monitor. Table 3 in Section 2 of this report contains a more complete listing.

Table 4 in Section 2 of this report provides a comprehensive list of implementation tracking elements that RTI will monitor for each design feature. Examples include State efforts to build MMP and provider core competencies for serving beneficiaries with various disability types; State requirements for coordination and integration of clinical, LTSS, and behavioral health services; documentation of coordination activities between MMPs and community-based organizations; phase-in of new or enhanced benefits, and methods to communicate them to eligible populations; and strategies for expanding beneficiary access to demonstration benefits.

As part of the implementation evaluation, the design features will be used in descriptive and comparative analyses across States. Additionally, the design features will be used in quality, utilization, access to care, and cost analyses to identify demonstration characteristics associated with better outcomes.

Table ES-2
Demonstration design features and examples of key components
(see Table 3 of this report for more detail)

Integrated delivery system

- Primary care, including medical homes and health homes
- LTSS
- Behavioral health services
- Developmental disability services

Integrated delivery systems supports

- Care team composition
- Health IT applied throughout the demonstration (at State level, by MMPs, at provider level or other)

Care coordination/case management

- Assessment process
- Service planning process
- Care management stratification process

Benefits and services

- Scope of services/benefits
- New or enhanced services

Enrollment and access to care

- Integrated enrollment and access to care
- Provider accessibility standards
- Opt out, disenrollment, and auto assignment policy

Beneficiary engagement and protections

- State policies to integrate Medicare and Medicaid grievances and appeals
- Quality management systems

Financing and payment elements

- Financing model: capitation or managed fee for service
- Incentives
- Shared savings

LTSS = Long-term services and supports; MMP = Medicare-Medicaid Plan; IT = information technology.

RTI will also track implementation progress indicators using data that States report quarterly through the State Data Reporting System (SDRS—see Section 4 of this report for more information), and other data that RTI obtains. ***Table ES-3*** presents examples of progress indicators that RTI will track through the SDRS and include in preliminary quarterly monitoring reports to CMS and the States.

Table ES-3
Examples of SDRS progress indicators

Indicator
No. of individuals…
— eligible to participate in the demonstration
— currently enrolled in the demonstration
— passively enrolled in the demonstration
— who opted out of the demonstration prior to enrollment
— who voluntarily disenrolled from the demonstration
— whose enrollment in the demonstration ended (e.g., death, loss of eligibility)

SDRS = State Data Reporting System.

Data Sources. RTI will use a variety of data sources to assess the implementation of each demonstration, including State documents (e.g., MOUs, waivers, contracts, State Plan Amendments); quarterly SDRS data submissions; and interviews with State agency staff, stakeholders, and coordinated care organizations/providers conducted during two site visits to each State and by telephone.

Section 3—Beneficiary Experience

The impact each State demonstration has on beneficiary experience is an important focus of the evaluation. RTI will monitor and evaluate the experiences of beneficiaries, their families, and caregivers to assess how closely the demonstrations meet CMS's goal of designing person-centered care delivery models. RTI will address the following research questions:

- What impact do these demonstrations have on beneficiary experience overall, by State, and for beneficiary subgroups?
- What factors influence the beneficiary enrollment decision?
- Do beneficiaries perceive improvements in their ability to find needed health services?
- Do beneficiaries perceive improvements in their choice of care options, including self-direction?
- Do beneficiaries perceive improvements in how care is delivered?
- Do beneficiaries perceive improvements in their personal health outcomes?
- Do beneficiaries perceive improvements in their quality of life?

RTI's framework for evaluating beneficiary experience is influenced by work conducted by the Center for Health Care Strategies (CHCS) on the elements of integration that directly affect beneficiary experience for Medicare-Medicaid enrollees. Table 6 in Section 3 of this report aligns key elements identified in the CHCS framework with the demonstration design features described previously, and ***Table ES-4*** provides examples of these elements that the evaluation will monitor.

Table ES-4
Examples of assessing beneficiary experience by beneficiary impact
(see Table 6 of this report for more detail)

Integrated delivery system
Beneficiaries have choice of medical, behavioral, and LTSS services and providers.
Beneficiaries are empowered and supported to make informed decisions.
Beneficiaries report that LTSS and behavioral health are integrated into primary and specialty care delivery.
Beneficiary burden is reduced through elimination of duplicative tests and procedures.
Delivery systems supports
Beneficiaries report that providers are knowledgeable about them and their care history.
Beneficiaries report adequacy of discharge and referral instructions.
Beneficiaries report that providers follow up after visits or discharge.
Care coordination/case management
Assessment process integrates/addresses health, behavioral health, and LTSS.
Beneficiaries report that they actively participate in the assessment process as do their medical providers.
The system facilitates timely and appropriate referrals and transitions within and across services and settings.
Benefits and services
Beneficiaries are aware of covered and enhanced benefits and use them.
The demonstration covers important services to improve care outcomes that are not otherwise available through the Medicaid or Medicare program.
Enrollment and access to care
Beneficiaries have choices and assistance in understanding their enrollment options, and they report ease of disenrollment.
Beneficiaries can access the full range of scheduled and urgent medical care, behavioral health services, and LTSS.
Beneficiaries report improved quality of life as a result of access to the full range of services.
Beneficiaries have access to multilingual and culturally sensitive providers.
Beneficiary engagement and protections
Beneficiaries understand their rights, and they get assistance in exercising their rights and protections.
Beneficiaries are treated fairly, are informed of their choices, and have a strong and respected voice in decisions about their care and support services.
Beneficiaries have easy access to fair, timely, and responsive processes when problems occur.
Finance and payment
Beneficiary experience is taken into account when awarding provider and plan incentives.

LTSS = Long-term services and supports.

NOTE: Data sources for assessing beneficiary experience include stakeholder interviews, beneficiary focus groups, survey questions, demonstration data, and interviews with State agency staff on demonstration implementation.

Data Sources. RTI will solicit direct feedback from beneficiaries through *focus groups* to gain insight into how the initiative affects them. There will be four focus groups in each State with 8 to 10 individuals in each group. Based on a State's enrolled population, each focus group will include a cross-section of individuals.

RTI will review and include in the evaluation as appropriate the results of *beneficiary surveys* administered by the State, CMS, or other entities. RTI will not directly administer beneficiary surveys as part of the evaluation and is not requiring States to administer beneficiary surveys for this evaluation. Several States have proposed to administer a beneficiary survey as part of their demonstrations. RTI will work with States, CMS, or other entities to incorporate these data into the evaluation.

Per the capitated model demonstration requirements outlined in MOUs and three-way contracts, Healthcare Effectiveness Data and Information Set (HEDIS), Health Outcomes Survey (HOS), and Consumer Assessment of Healthcare Providers and Systems (CAHPS) must be reported consistent with Medicare requirements. The operations support contractor for CMS will administer a CAHPs survey in managed fee-for-service (MFFS) States. RTI will work with States and CMS to acquire the results of these surveys and incorporate them into the evaluation as appropriate.

RTI will also conduct key *stakeholder interviews* to better understand the level of beneficiary engagement with the demonstration, its perceived impact on beneficiary outcomes, and any unintended consequences. RTI will conduct interviews with members of beneficiary groups whose stakeholders are served by a State's demonstration, such as members of consumer advisory groups, beneficiary rights advocates, and public guardian groups.

Finally, RTI will use *other data* collected from States during site visits, reports, and other materials developed by States, the SDRS, and data obtained from CMS or other entities to assess the beneficiary experience. Data of particular interest include, but are not limited to, the following:

- Complaint, appeal, and grievance data from CMS, demonstration Ombuds programs, or other entities, as available.
- Disenrollment and opt-out rates, where appropriate.
- Information about waiting lists or lags in accessing services, which will provide useful indications of where the system lacks capacity, as a topic for discussion during site visits or focus groups.

RTI will explore whether specific demonstration features can be identified in each State that affect beneficiary experience and, where possible, how those features also affect evaluation findings through quantitative analyses of quality, utilization, and costs.

Section 4—State Data Reporting System (SDRS)

The SDRS will be RTI's tool for collecting and storing information about each State's demonstration design and progress, and monitoring and reporting on demonstration progress by individual States and the Financial Alignment Initiative as a whole. The SDRS will store model summary, implementation tracking, and demonstration impact and outcomes data to be used in the quarterly preliminary reports for CMS and States as well as in annual aggregate and State-specific reports.

Data Stored in the SDRS. The *model summary* data describe each State's demonstration model, such as geographic areas, new services offered, and eligible populations (see Table 10 of this report for more detail). RTI will input this information into the SDRS and will update it only if there are changes to a State's demonstration.

States will input *implementation tracking* data into the SDRS on a quarterly basis. States will report numerical Progress Indicators data, including those presented in ***Table ES-5*** and in Table 11 of this report. States will also enter data via the Tracking Elements by Design Feature subsection of the SDRS, which includes Yes/No and free text responses. ***Table ES-6*** (and Table 12 of this report) present examples. See Section 5 of this report for more information.

Table ES-5
SDRS data collection: Progress indicator elements

Indicator[1]
Eligibility
Total number of beneficiaries who are eligible to participate in the demonstration[2]
Enrollment
Total number of beneficiaries who are enrolled in the demonstration[2]
Number of beneficiaries who are newly enrolled in the demonstration[2]
Number of newly enrolled beneficiaries who were automatically (passively) enrolled in the demonstration[2]
Number of beneficiaries who opted out or chose not to enroll in the demonstration without ever being enrolled[2]
Disenrollment
Number of beneficiaries who voluntarily disenrolled (i.e., made a choice to disenroll) from the demonstration[2]
Number of demonstration enrollees whose eligibility for the demonstration ended involuntarily (e.g., moved out of area, lost Medicaid eligibility, were incarcerated)[2]
Demonstration service area
Whether demonstration is currently operating statewide vs. in specific counties or geographic areas (and provide list of counties served, if in specific geographic areas)
Specific to capitated model demonstrations
Number of three-way contracts with MMPs
New CMS initiatives in the demonstration area that may affect Medicare-Medicaid enrollees
Specific to demonstrations that use health homes
Number of health homes participating in the demonstration
Number of enrollees served by health homes
Specific to demonstrations using medical homes
Number of medical homes participating in the demonstration
Number of enrollees served by medical homes

LTSS = long-term services and supports; MFFS = managed fee for service; MMP = Medicare-Medicaid plan; NCQA = National Committee for Quality Assurance; SDRS = State Data Reporting System.

[1] All indicators may not apply to all States (e.g., for some MFFS States, beneficiaries who are eligible for the demonstration are the same as beneficiaries who are enrolled in the demonstration).

[2] Progress indicators that will be presented in quarterly reports to CMS and the States.

Table ES-6
SDRS data collection: Examples of tracking elements by design feature
(see Table 12 of this report for more detail)

Integrated delivery systems
New policies or administrative procedures for improving the integration of primary care, long-term services and supports (LTSS), and behavioral health services under the demonstration
Changes in reporting requirements for any of the entities involved in the demonstration
Integrated delivery systems supports
Training or capacity-building activities to build core competencies of demonstration MMPs/providers in serving demonstration populations
Activities to help primary care providers transform care delivery
Care coordination/case management
New State policies/guidelines regarding care coordination/case management, promoting the adoption of electronic health records, etc.
Benefits and services
New or expanded services/benefits for demonstration participants
Enrollment and access to care
Activities to increase beneficiary enrollment
Major issues and challenges implementing the demonstration and solutions developed
Beneficiary engagement and protections
Activities to engage stakeholders in policy development or oversight of the demonstration
Activities to engage enrollees, families, or advocates in policy development or oversight of the demonstration
Quality management and measurement
Tracking of new quality indicators
Receiving data from MMPs/providers to support new quality indicators
Financing and payment
Changes in payment methodology for MMPs and providers
Data development
Timing of the most recent MSIS or T-MSIS data file submissions
Whether MMPs experienced any problems submitting encounter data to CMS (for capitated models)
Successes related to the demonstration not covered by other questions

MMP = Medicare-Medicaid Plan; MSIS = Medicaid Statistical Information System; SDRS = State Data Reporting System; T-MSIS = Transformed Medicaid Statistical Information System.

RTI will input *demonstration impact and outcomes* data into the SDRS on a quarterly basis. This content will be generated by RTI from Medicaid Statistical Information System (MSIS)/Medicare fee-for-service (FFS) claims; encounter data from MMPs, Medicaid managed care organizations, and Medicare Advantage plans; Nursing Home MDS analysis; and other data that may be available. Data availability will affect when specific analysis results will be reported.

Section 5—Quantitative Analyses

RTI will conduct quantitative analyses for individual States and will include them in annual reports. The final evaluation report will also include an aggregate analysis to learn more about the effects of different State demonstration design features on quality, utilization, and costs. Different analytic approaches are required for MFFS States versus capitated model States in terms of data requirements, analytic issues, and outcome variables. This section of the report discusses the overall approach to identifying demonstration group and comparison group beneficiaries. For State-specific details on identifying demonstration and comparison groups, see the State-specific evaluation design reports.

Research Approach. RTI will use an intent-to-treat (ITT) approach for the quantitative analyses, comparing the eligible population for each State's demonstration with a similar population that is not affected by the demonstration (i.e., a comparison group). Under the ITT framework, outcome analyses will include all beneficiaries eligible for the demonstration in the demonstration States, including those who opt out, participate but then disenroll, are eligible but are not contacted by the State or participating providers, and those who enroll but do not engage with the care model, and a group of similar individuals in the comparison group. This approach diminishes the potential for selection bias and highlights the effect of the demonstrations on all beneficiaries in the demonstration-eligible population. RTI will compare the characteristics of those who enroll with those who are eligible but do not enroll in the care model and will conduct analyses to further explore demonstration effects on demonstration enrollees, acknowledging that selection bias must be taken into account in interpreting the results.

Demonstration Groups. To identify the population eligible for a State's demonstration, States will submit demonstration evaluation (finder) files to RTI on a quarterly basis that will include information on enrolled beneficiaries as well as all beneficiaries eligible for the demonstration. RTI will use this information to identify the characteristics of each State's eligible beneficiaries. Section 5 of this report provides more detail on the content of the demonstration evaluation (finder) files.

Comparison Groups. In this evaluation design, the comparison group provides an estimate of what would have happened to the demonstration group in the absence of the demonstration. Thus, the comparison group members should be similar to the demonstration group members in terms of characteristics, including health care and LTSS needs, and should reside in areas similar to the demonstration areas in terms of the health care system and the larger environment. Identifying the comparison group members will entail two steps: (1) selecting the geographic area from which the comparison group will be drawn and (2) identifying the individuals who will be included in the comparison group.

RTI will determine whether an in-State comparison group is possible for each demonstration State. An in-State comparison group will only be a potential option for demonstrations

implemented in a subset of the State, rather than statewide, where the demonstration and nondemonstration areas are similar, and where the nondemonstration areas contain sufficient numbers of beneficiaries. For States where an in-State comparison group is not possible, we will develop an out-of-State comparison group or, potentially, a comparison group that includes both in-State and out-of-State areas.

To construct a comparison group from out-of-State areas, or a combination of in-State and out-of-State areas, we will limit potential comparison areas to those not participating in the Financial Alignment Initiative. We expect to draw out-of-State comparison groups using areas from multiple comparison States. The areas included in the comparison group will be determined by the closeness of the match with the demonstration areas, and the size of the Medicare-Medicaid enrollee population in the comparison area. The goal will be to identify a comparison group at least as large as the demonstration group to ensure a sufficient sample to support sensitivity analyses around the choice of comparison groups. The first annual report will document decision rules for choosing the comparison area.

To identify comparison areas, RTI will conduct a statistical distance analysis to assess the similarity of a demonstration region with each of its potential comparison areas. The process entails the following three steps:

Step 1. Identify characteristics that will be used to compare demonstration and comparison areas that reflect State-level policies prior to the demonstration. Example characteristics include the following:

- Medicare spending per Medicare-Medicaid enrollee
- Medicaid spending per Medicare-Medicaid enrollee
- Nursing facility users per Medicaid beneficiary age 65 and over
- Home and community-based services (HCBS) users per Medicaid beneficiary age 65 and over
- Personal Care users per Medicaid beneficiary age 65 and over
- Medicare Advantage penetration
- Medicaid managed care penetration per full-benefit Medicare-Medicaid enrollee

Step 2. Compute statistical distance scores for each demonstration area and potential comparison Metropolitan Statistical Areas (MSAs) to measure the similarity between the demonstration population and a potential comparison population. The smaller the distance score, the more similar the States are with respect to the selected characteristics. This step will occur when the final list of demonstration States is confirmed, so the most recent data available can be used and the full range of potential comparison areas is known. More detail is provided in Section 5 of this report.

Step 3. Select comparison areas by identifying the comparison areas with the smallest statistical distance scores. The number of areas to be selected will depend on their combined population of Medicare-Medicaid enrollees. RTI will also consider other factors, including the timeliness of a State's MSIS submissions.

To identify the individuals within the comparison geographic areas to include in the comparison group, RTI will estimate propensity scores and weight comparison group beneficiaries so that the distribution of selected characteristics looks like the demonstration group. For this evaluation, the propensity score is an estimate of the probability that a beneficiary is in the demonstration group conditional on a set of observed characteristics. The following characteristics are examples of those that may be included in the propensity model:

- Beneficiary characteristics such as age, sex, MSIS eligibility information on socioeconomic status, prior Medicare and Medicaid expenditures, LTSS/HCBS, hierarchical condition category (HCC) risk scores, and end stage renal disease (ESRD) status. These data will be obtained from Medicare and Medicaid files and will include encounter data or per member per month (PMPM) payments where appropriate and available.
- MSA-level characteristics from Census Bureau databases, and the Area Resource File (ARF) such as health care providers/100,000 population, morbidity/mortality, and urbanicity.
- State-level policy factors, such as the proportion of long-term services and supports spending that is for HCBS (rather than for facility-based care), Medicaid nursing facility eligibility criteria, and implementation of Health Home State Plan Amendment (SPA; except for within-State comparison groups).

RTI will estimate a logistic model by regressing group status (demonstration vs. comparison pool) on the set of individual and area characteristics to determine the propensity scores for demonstration and comparison group beneficiaries. Comparison group members will then be weighted by their predicted propensity score to ensure that the comparison group reflects the distribution of characteristics in the demonstration population.

The comparison areas will be determined within the first year of implementation of each demonstration, in order to use the timeliest data available. The comparison group members will be determined retrospectively at the end of each demonstration year, allowing us to include information on individuals newly eligible or ineligible for the demonstration during that year. The groups will be refreshed annually to incorporate individuals who become eligible for the demonstration over time. Section 5 of this report provides more detail on this process.

The demonstrations under the Financial Alignment Initiative will be implemented during a period when several other CMS demonstrations and initiatives are occurring. As part of our analytic framework using multivariate analyses in the last year of the evaluation, RTI will work to identify any demonstration impact beyond that resulting from other demonstrations and programs, as appropriate.

Data Sources. RTI will use several data sources for the quantitative analysis. To identify beneficiaries eligible for the demonstration, a State will provide demonstration evaluation (finder) files described earlier and in Section 5 of this report. From CMS, RTI will use

Medicare enrollment data, claims, and Nursing Home MDS for the predemonstration and demonstration periods, where appropriate, to create a beneficiary-level file with summary variables on Medicare utilization and payment by service type (e.g., inpatient, skilled nursing, home health). Medicaid claims data will be used to construct service use patterns, particularly for services not covered by Medicare—notably, facility-based long-term care, HCBS waiver services, and behavioral health services. Because the evaluation uses an intent-to-treat design and includes a predemonstration period, RTI will use CMS encounter data from Medicare Advantage plans, nondemonstration Medicaid managed care plans, and MMPs to capture utilization of beneficiaries receiving services through managed care. RTI will also obtain CMS data on prescription drug PMPM payments for beneficiaries from the monthly plan payment files at CMS, and potentially Part D reconciliation costs directly from the CMS payment group, to support analysis of Part D costs.

Section 5 of this report provides more detail on how the analytic files will be constructed, including possible challenges.

Section 6—Analysis Overview

Quantitative analyses of quality, utilization, and cost will be performed for each demonstration State and in the aggregate for the Financial Alignment Initiative as a whole. Sections 7, 8, and 9 of this report describe the quality, utilization, and cost measures that will be examined. This section outlines the methods that RTI will use to analyze those data. The timing of the quantitative analyses will depend on data availability (as discussed in Section 5 of this report); the methods outlined here may be modified to incorporate any changes that may occur in States unrelated to the demonstration, such as the effects of other demonstrations or State-specific policy changes.

Monitoring. RTI will track quarterly changes in individual States for selected beneficiary experience, quality, utilization, access to care, and cost measures for the demonstration group using pre- and postperiod data analyzed by RTI and stored in the SDRS. RTI will use available Medicare and Medicaid data each quarter to calculate means, counts, and proportions for selected measures. RTI will also analyze available Nursing Home MDS data to calculate facility admission rates. The monitoring analysis will be used to develop the quarterly reports for CMS and the States.

Individual State Descriptive Analyses. RTI will conduct individual-State descriptive analyses at the end of each demonstration period[1] that focus on beneficiary experience, utilization, access to care, cost, and quality measures, as well as changes across time or subgroups of interest within each demonstration period. Examples of measures include total costs for Medicare and Medicaid separately, primary and specialty care utilization rates,

[1] Demonstration period as defined in each State's MOU.

rates of avoidable hospitalization and inappropriate readmissions, counts of hospital and nursing facility admissions and length of stay, rates of HCBS use, and mortality. More information on the measures to be reported can be found in Sections 7 to 9 of this report. RTI will present means, counts, and proportions; and statistical tests of means and counts for the 2-year baseline period, each demonstration period, and the comparison group. RTI will also provide results comparing beneficiary subgroups by age groups, subpopulations, and other important characteristics to inform CMS and States about improvements over time. The results of these analyses will be presented in an annual report for each State.

Impact Analyses Within States. RTI will assess the overall impact of the demonstration on quality, utilization, and cost measures using a difference-in-differences method with a comparison group for the final evaluation report for each State. This multivariate analysis will be done after the demonstration is complete to allow for sufficient claims run-out for the demonstration State and the comparison areas, to avoid over- or underestimates of results. Under the difference-in-differences method, pre- and postdemonstration changes in the outcome measures (e.g., utilization, quality, cost measures) for the demonstration group will be compared with the pre- and postexperience of a comparison group. This methodological framework will also be applied to each of the quality, utilization, access to care, and cost measures that will be tracked within each State over time. These analyses will use linked Medicare and Medicaid claims and encounter data for the predemonstration and demonstration period. RTI will finalize the specific outcome measures for the difference-in-differences analyses after the demonstration has concluded to ensure that comparable, high-quality data are available for the demonstration and comparison States.

More details on the multivariate analyses, including the regression equations that will be used, can be found in Section 6 of this report.

Sensitivity Analyses. RTI will test the sensitivity of the impact estimates for State demonstrations because the validity of the difference-in-differences approach depends in large part on the assumption that changes over time in the comparison group are a reasonable counterfactual for what would have happened to the demonstration group. One such decision that RTI will test is the choice of comparison groups. As part of efforts to check the consistency of the impact estimates, RTI will compare the findings from the core models with estimates based on assumptions, such as different combination of States for out-of-State comparison groups and different propensity-score models, for a few States. Consistency in the estimates across models will provide more confidence in the reliability of the impact estimates.

Additional Analyses. RTI will use the Nursing Home MDS to analyze additional changes in patterns of facility-based LTSS quality and use. RTI will evaluate admission rates, acuity upon admission, and selected quality measures for both short-stay (i.e., skilled nursing facility users) and long-stay facility residents. RTI will also conduct an analysis of encounter

data coding patterns, because capitated payments to MMPs under the demonstration may be affected by changes in coding intensity during the demonstration. These analyses will examine the extent to which changes in coding intensity observed in demonstration States compare with nondemonstration States or a predemonstration period.

Because enrollment is voluntary, RTI will compare the characteristics of those who enroll with those who are eligible but do not enroll in the care model and conduct analyses to further explore demonstration effects on demonstration enrollees, acknowledging that selection bias must be taken into account in interpreting the results.

Aggregate Impact Analyses. After the final multivariate analyses have been performed for individual State evaluations, the RTI team will conduct aggregate analyses to examine changes resulting from various State demonstration design features on quality, utilization, and cost outcomes. The goal of this type of analysis is to look across States that have implemented similar design features (e.g., MFFS or capitated payment model, demonstration design, contract vs. noncontract States, Health Home SPA States) and disentangle the relative effectiveness of the demonstration design choices, whereas the individual State evaluations will estimate the impact of a chosen *set* of demonstration design choices relative to the status quo in the State.

This part of the evaluation will address several research questions—for example:

- Which demonstration model (MFFS or capitated) has achieved greater savings?
- Are there differences in key outcomes (e.g., quality, utilization, expenditure types) that can be attributed to the type of financial alignment model used?
- Do the effects achieved by alternative integrated care models occur equally fast? Or does one model (MFFS vs. capitated) achieve gains more quickly than the other?
- Does the approach to enrollment (e.g., passive enrollment) affect access to care and costs?
- How does the relative degree of care management intensity and diversity across services affect outcomes?
- Do these effects vary across subgroups of beneficiaries, such as those using LTSS?

RTI will carefully consider which States to include in these meta-analyses in order to provide thoughtful conclusions. Some States will be moving beneficiaries from MFFS to capitated approaches, some from capitated approaches to better-integrated capitated approaches. The populations eligible for the demonstration also differ across States. We will include States with similar approaches or populations in appropriate meta-analyses.

Section 6 of this report provides more detail on the current plan for how the aggregate analyses will be conducted, including how the variables will be constructed and the model equation.

Section 7—Utilization and Access to Care

The impact of these demonstrations should result in changes in service use, in annual utilization patterns, and in specific patterns of care. Of particular interest is the impact across subpopulations of Medicare-Medicaid enrollees, and whether any observed impact is short term or continuous. Research questions regarding utilization include the following:

- What is the impact of the State demonstrations on utilization patterns during the course of the demonstration?
- What is the impact on hospital and nursing facility admission rates, potentially avoidable hospitalization utilization rates by setting, and on LTSS utilization rates? What is the impact of the demonstration on hospital and nursing facility length of stay?
- Do demonstrations change the balance between HCBS and nursing facility use, the types of Medicare-Medicaid enrollees who use these services, and utilization rates by type of HCBS such as personal care? Do Medicare-Medicaid enrollees receive more HCBS as a result of the demonstrations?
- Is any impact short term (e.g., lasting only for 1 year before returning to predemonstration level, increasing over time, reaching a plateau after a year or 2)?
- Does the observed impact vary by health condition or other beneficiary characteristics?
- Will case management or care coordination lead to lower hospital admission rates or, if admitted, shorter lengths of stay and shorter nursing facility and home health care episodes?
- Are demonstration group members using fewer inpatient services and more ambulatory services?
- Is the impact greater for more medically complex (multiple chronic condition), high-cost (top 10 percent) enrollees?

In addition, State demonstrations are expected to improve access to services, which should be evident through changes in utilization patterns of certain services. Research questions pertaining to access to care are as follows:

- Access to medical care: Do demonstration participants experience increases in the mean number of primary care visits and increased visit rates by specialty type?
- Access to LTSS: Does acuity on admission to nursing facilities increase? Do discharge rates back to the community from nursing facilities increase? Is there an increase in the proportion of HCBS users self-directing care?
- Access to behavioral health services: Does the mental health outpatient utilization rate increase? Does the outpatient substance use disorder service utilization rate increase?

Monitoring. To monitor States' progress during the demonstration, we will calculate high-level measures for each State to identify changes in utilization over time. Examples of utilization and access to care measures are listed in ***Table ES-7***. Various inpatient and

emergency room measures that RTI plans to include in quarterly reports are described in more detail in Section 8. RTI will also identify a range of key utilization and access to care measures to include in quarterly, annual, and final evaluation reports. For each utilization type, these measures will usually be expressed as visits per 1,000 eligible beneficiaries and users as a percentage of the demonstration-eligible population.

Table ES-7
Examples of utilization and access to care measures for the SDRS
(see Section 4 of this report for more information on the SDRS and Section 7 for more information on utilization and access to care)

Utilization measures[1]	Access-to-care measures
■ Inpatient acute ■ Inpatient psychiatric ■ Emergency room ■ Skilled nursing facility ■ Nursing facility (long stay) ■ Outpatient (primary care) ■ Outpatient behavioral health (mental health; AOD) ■ Home health ■ State Plan personal care ■ Waiver personal care ■ Other HCBS (home health, other waiver services)	■ Number of physician visits ■ Acuity on admission to nursing facilities ■ Discharge rate back to the community from nursing facilities ■ Any outpatient utilization for severe and persistent mental health conditions ■ Any substance use disorder treatment utilization

AOD = alcohol or other drugs; HCBS = home and community-based services; SDRS = State Data Reporting System.

[1] The final set of measures will be determined based on the timing and availability of data. Utilization and access-to-care measures will be calculated by RTI using data provided by CMS.

Individual State Descriptive Analyses. For annual reports, we will measure utilization rates of Medicare- and Medicaid-covered services for each State, using unlinked data to identify the effects of State demonstrations on the type and level of service use, ranging along a continuum from facility-based care to care provided at home (***Table ES-8***). Both Medicare and Medicaid data will be used for this analysis.

RTI will calculate average utilization rates at predemonstration and at the beginning, middle, and end of each demonstration. Use rates for each State will be stratified by HCC scores, health status measures, or similar measures. We will adjust for hospitalizations in the prior year using categorical HCC scores. Statistical tests will be used to test for significant differences in use across years and between subpopulations within a State.

Table ES-8
Service categories for reporting utilization measures
(see Table 16 of this report for more information)

Service type	Medicare only	Medicaid only	Medicare and Medicaid
Inpatient	—	—	X
Emergency room	—	—	X
Skilled nursing facility	X	—	—
Nursing facility (long-term stay)	—	X	—
Other facility-based[1]	—	—	X
Outpatient[2]	—	—	X
Outpatient behavioral health (mental and substance use disorder)	—	X	—
Home health	—	—	X
HCBS (PAS, waiver services)	—	X	—
Dental	—	—	X

— = not applicable. HCBS = home and community-based services; PAS = personal assistance services.

[1] Includes long-term care hospital, rehabilitation hospital, State mental health facility stays.

[2] Includes visits to physician offices, hospital outpatient departments, rehabilitation agencies.

Impact Analyses. As discussed above, in the final year of the evaluation RTI will conduct multivariate difference–in-differences analyses to evaluate the impact of individual State demonstrations relative to their selected comparison groups (see Section 6 of this report for a detailed description of the multivariate impact analyses). Examples of outcome variables in the multivariate analyses include rates and lengths of short- and long-term nursing facility stays, number of primary care provider (PCP) visits, number of specialty physician visits, and rates and number of months of personal assistance services and HCBS waiver services. Any inpatient analyses other than rates of overall inpatient use will be discussed in the section on quality measures.

One key strategy for reducing costs without compromising quality of care is to improve care coordination by reducing fragmentation and redundancies in services. RTI will develop analyses to address this issue, such as analyzing patterns of primary versus specialty care. Measures for assessing fragmentation of care for LTSS and behavioral health services will also be explored after reviewing candidate measures.

Section 8—Quality of Care

For purposes of monitoring and conducting impact analyses, RTI has selected a set of quality measures that are largely utilization based (see ***Table ES-9***). Many of these

measures are available through claims and encounter data that RTI will obtain from CMS. RTI expects these data to be available for descriptive quarterly reporting for demonstration States, and for final impact analyses for both demonstration and comparison groups. Some measures, such as HEDIS, have standardized definitions that will allow RTI to monitor the results across all capitated demonstration States on an annual basis.

Table ES-9
Evaluation quality measures
(see Table 18 in this report for definitions and specifications for these measures)

Measure concept	State model (capitated or MFFS)
RTI team calculations based on data obtained from CMS	
30-day all-cause risk-standardized readmission rate	Capitated, MFFS
Influenza immunization	Capitated, MFFS
Pneumococcal vaccination for beneficiaries 65 and older	Capitated, MFFS
Ambulatory care sensitive condition admissions—overall composite (AHRQ PQI #90)	Capitated, MFFS
Ambulatory care sensitive condition admissions—chronic composite (AHRQ PQI #92)	Capitated, MFFS
Preventable ED visits	Capitated, MFFS
ED visits, excluding those resulting in inpatient admission or death	Capitated, MFFS
Admissions with primary diagnosis of a severe and persistent mental illness or substance use disorder	Capitated, MFFS
Follow-up after hospitalization for mental illness	Capitated, MFFS
Screening for clinical depression and follow-up	Capitated, MFFS
Cardiac rehabilitation following hospitalization for cardiac event	Capitated, MFFS
Percent of high-risk long-stay NF residents with pressure ulcers	Capitated, MFFS
Screening for fall risk	Capitated, MFFS
Initiation and engagement of alcohol and other drug dependence treatment	Capitated, MFFS
HEDIS data obtained from CMS	
Adult BMI assessment	Capitated
Annual monitoring for patients on persistent medications	Capitated
Antidepressant medication management	Capitated
Breast cancer screening	Capitated
Care transition record transmitted to health care professional	Capitated
Comprehensive diabetes care—selected components	Capitated
Controlling high blood pressure	Capitated

BH = behavioral health; BMI = body mass index; ED = emergency department; HEDIS = Healthcare Effectiveness Data and Information Set; LTSS = long-term services and supports; MFFS = managed fee for service; NF = nursing facility.

To conduct a thorough examination of quality of care, the evaluation will supplement the measures from Table ES-9 with the following:

- Additional State-specific quality measures that will be finalized within the first 6 months of each State's implementation.
- Quality of life, satisfaction, and access-to-care information derived from the evaluation as discussed in Section 3 and in Section 7.
- HEDIS measures that are required of MMPs and outlined in each State's MOU.
- Beneficiary surveys, such as HOS and CAHPS surveys, that MMPs are required to report to CMS.
- CAHPS surveys administered in MFFS States.
- Measures developed by RTI from publicly available sources, such as the Area Resource File (ARF), that define the health care environment in each beneficiary's residential area. These may reflect variation in the supply of available providers or general economic conditions that may apply to health care markets.

Monitoring. The quality measures used for quarterly monitoring will be reported for demonstration States and not comparison groups, because comparison groups will not be identified until after the end of each demonstration year. The measures will be standardized, however, across States to the extent possible and will be useful for monitoring trends over time within a State and across the demonstration.

Individual State Descriptive Analyses. Because States have developed different approaches to integration and may target specific groups or services, RTI will develop some measures unique to individual States within the first 6 months of each demonstration to supplement the core evaluation measures. RTI will also incorporate the quality measures that States and MMPs are required to report to CMS for this demonstration, and listed in each State's MOU, into the evaluation. Although these measures will not be available for comparison areas, they will provide insight into the quality of care that beneficiaries receive while in the demonstration.

Impact Analyses. RTI will use the evaluation measures that it calculates as dependent variables in multivariate regression analyses in the final evaluation report to identify factors contributing to quality outcomes. The analytic methods for quality measures will follow the same template as described in Section 6 (Analysis Overview) with some refinements. For example, the methods will need to be refined depending on whether the outcome is binary or continuous.

Section 9—Cost

The evaluation will use the same basic descriptive and regression-based techniques as outlined in Section 6 of this report (Analysis Overview) to analyze cost. It will examine how costs are associated with the variety of services that beneficiaries receive, including

medical, behavioral health, and LTSS. The research questions regarding cost analyses include the following:

- Do the demonstrations reduce costs?
- If so, how were the demonstrations able to reduce the costs of Medicare-Medicaid enrollees compared with the comparison group?
- How do the demonstrations differentially affect expenditures for beneficiaries at risk for having high costs?

Monitoring. RTI will identify high-level cost measures that can be calculated for all States to monitor changes over time. For MFFS demonstration States, RTI will provide per-capita or per-user costs for categories of services (e.g., inpatient, outpatient, long-term nursing facility, mental health) from claims to understand how costs change quarterly. For capitated demonstration States, costs to the Medicare and Medicaid programs are the PMPM rates paid to MMPs, combined with capitated or FFS costs for those who opt out or disenroll. RTI plans to include Part D PMPM and any PMPM reconciliation data provided by CMS. Accounting for all of these types of costs is important because of the cost implications of possible selection bias.

Individual State Descriptive Analyses. RTI will measure predemonstration and demonstration spending on Medicare-Medicaid enrollees for both Medicare and Medicaid, and present descriptive cost analyses in quarterly and annual reports. RTI will also present in annual reports the costs for various subpopulations, such as demographic groups, LTSS users, and beneficiaries with intellectual and developmental disabilities, ESRD, and other chronic conditions. RTI will test for differences across demonstration periods in each State.

For MFFS States, RTI will also assess costs for the service types shown in ***Table ES-10***. For capitated model States, RTI anticipates that service-level spending will not be available in the encounter data reported by MMPs, so the utilization analysis described in Section 7 will be used to understand the impact of the demonstration by type of service. Other factors, such as changes in coding intensity, could also play a role in demonstration costs, and RTI will consider such factors in its analysis.

Impact Analyses. RTI will estimate the demonstrations' impact on Medicare and Medicaid costs, using regression-based techniques to learn what factors contribute to cost savings or increases. The goals are to learn whether certain types of demonstration approaches save more money than others, or whether costs are lower in the demonstration group than the comparison group for certain subgroups. CMS is also interested in which types of services (e.g., inpatient, HCBS) contribute the most to cost differences between the demonstration and comparison groups as State demonstrations promote changes in utilization patterns through care management.

Table ES-10
Service categories and associated data sources for reporting cost in MFFS States

Service type	Encounter data (Medicare Advantage)	Medicaid only (FFS)	Medicare and Medicaid (FFS)
Inpatient	—	—	X
Emergency room	—	—	X
Nursing facility (short rehabilitation stay)	—	—	X
Nursing facility (long-term stay)	—	X	—
Other facility-based[1]	—	—	X
Outpatient[2]	—	—	X
Outpatient behavioral health (mental and substance use disorder)	—	X	—
Home health	—	—	X
HCBS (State Plan PAS, waiver services)	—	X	—
Dental	—	X	X
Prescription drug PMPM	—	—	X
Managed care PMPM	X	—	—

— = not applicable; FFS = fee for service; HCBS = home and community-based services; PAS = personal assistance services; PMPM =per member per month payments.

[1] Includes long-term care hospital, rehabilitation hospital, State mental health facility stays.

[2] Includes visits to physician offices, hospital outpatient departments, rehabilitation agencies.

In addition to cost analyses for all Medicare-Medicaid enrollees eligible for the demonstration, MFFS demonstration States should be expected to reduce total costs for high-cost beneficiaries. Demonstration and comparison group beneficiaries will be stratified to identify the groups of beneficiaries that have traditionally been the most expensive service users in the demonstration State. High-cost beneficiaries may include those with multiple comorbidities, severe and persistent mental illness (SPMI), LTSS users, or prior inpatient and/or skilled nursing facility stays. RTI also will conduct cost analyses exploring demonstration effects on demonstration enrollees, acknowledging that selection bias must be taken into account in interpreting the results.

For capitated model demonstrations, RTI will estimate cost savings accruing to the Medicare and Medicaid programs separately twice during the demonstration, using a regression-based approach and a comparison group. To determine annual total costs (overall and by payer) for these analyses, RTI will aggregate the Medicare and Medicaid PMPM payments paid to the MMPs, Medicare Advantage plans, and Medicaid managed care organizations, and the FFS costs for the eligible population that are not enrolled in the demonstration. RTI will include Part D PMPM and any PMPM reconciliation data provided by CMS in the final

assessment of cost impact to ensure that all data are available. The methodology will be reviewed and approved by the CMS Office of the Actuary. For MFFS States, savings will be calculated after each demonstration period using an actuarial methodology for performance payment purposes. This methodology has been presented to CMS in a separate memorandum. The evaluation will also use a regression-based approach to examine savings at the end of the demonstration. The assumptions underlying the two methods will be as consistent as possible.

Section 10—Subpopulations and Health Disparities

It is important to understand whether the demonstrations have differential effects on subpopulations as defined by disability type, or demographic or clinical characteristics, such as cognitive status, clinical complexity, and residence (community-residing or in a residential setting).

The overarching research questions for subpopulations are as follows: Does the demonstration have an impact on the quality of care, service utilization patterns, and the beneficiary experience for subpopulations and the costs incurred for their services, and do these effects differ from those on the overall population of Medicare-Medicaid enrollees? To answer these questions, four specific research issues will be reflected in the qualitative protocol development and the quantitative analyses:

- How do the demonstrations, as implemented by the different States, address the unique needs of the subpopulations? Are there special initiatives designed to meet the needs of these populations (e.g., special care coordination efforts, new services for people with severe and persistent mental illness, or nursing facility diversion programs)? Do the demonstration States successfully implement what they proposed? Do the models that focus on subpopulations work better than those that are designed for more general populations?
- Do the demonstrations reduce expenditures and improve beneficiary experience, quality of care, and health outcomes for subpopulations? What is the effect on service use?
- Do the demonstrations reduce or eliminate undesirable disparities (e.g., by race or ethnicity) in access to care, beneficiary experience, health care utilization, expenditures, quality of care, and health outcomes?
- To the extent that the demonstrations have positive outcomes for subpopulations, what features of the demonstration account for these outcomes?

RTI will work with CMS to identify high-priority, policy-relevant populations to analyze for each State. Possible subpopulation groups of interest include racial and ethnic groups, people living in rural or inner-city areas, younger people with disabilities, people age 65 and older, people with SPMI, people with developmental disabilities, users of LTSS, and high-cost beneficiaries.

Below are examples of how the evaluation will be targeted to beneficiaries with behavioral health conditions and individuals residing in nursing facilities.

Beneficiaries with Behavioral Health Conditions. The evaluation team will conduct subanalyses for the population of individuals with behavioral health conditions, defined to include severe and persistent mental illness and substance use disorders. These analyses will evaluate the impact of the demonstrations on quality, utilization, and access to care for medical, LTSS, and behavioral health services, and will also examine qualitative data gathered through interviews, focus groups, and surveys. Examples of the range of measures that will be examined for beneficiaries with behavioral health conditions include outpatient behavioral health services; HCBS services; new long-term nursing facility admissions for beneficiaries with SPMI; access to a full range of scheduled and urgent medical care, behavioral health services, and LTSS; beneficiary reports of improved quality of life as a result of access to the full range of services; beneficiary choice of medical, behavioral, and LTSS services and providers; beneficiary reports on satisfaction with their life; care coordination assessment processes that integrate/address health, behavioral health, and LTSS; hospitalizations for beneficiaries with SPMI; outpatient visits after hospitalization for mental illness; and initiation and engagement of alcohol and other drug dependence treatment. Results of descriptive analyses will be presented in annual reports. The final evaluation reports will include multivariate analyses.

Nursing Facility Residents. By aligning Medicare and Medicaid incentives, the demonstrations have an opportunity to improve quality of care in nursing facilities, reduce potentially avoidable hospitalizations of nursing facility residents, and, through rebalancing efforts, prevent, delay, or shorten facility stays. Conversely, if demonstration providers seek to achieve savings by negotiating lower-cost contracts with nursing facilities, lower quality of care could result. The evaluation will analyze nursing facility admission rates, acute-care utilization (e.g., physician visits, hospitalizations, emergency room use) and cost patterns for individuals receiving short-term skilled nursing facility care and for long-stay residents. In addition, we will use the Nursing Home MDS to evaluate the level of impairment or acuity of new nursing facility entrants to evaluate the extent to which the demonstrations are maintaining frail individuals in the community, and monitor selected nursing facility quality measures. We will monitor trends in nursing facility admissions and quality within the demonstration States (or regions within a State) and analyze demonstration impact in comparison with facilities in comparison States or regions, using multivariate techniques.

Section 11—Next Steps

We will present the results of our analyses in a series of deliverables, including quarterly reports to CMS and States, annual reports, and a final evaluation report for each State as well as a final aggregate evaluation report (***Table ES-11***). ***Table ES-12*** summarizes the sources of data that the evaluation team will use to monitor demonstration progress and

evaluate the outcomes of the demonstrations. It provides an overview of the data States will be asked to provide, evaluation activities in which State staff will participate, and data the evaluation team will access from CMS.

Table ES-11
Deliverable timeline for monitoring and evaluation activities

Deliverable	Timeline	Data included
State-Specific Evaluation Design Plans	Summer 2013 through 2014, on a rolling basis.	The State-specific evaluation design plans detail the application of the overall research design for each State given the characteristics of each State's demonstration.
State-Specific Initial Reports	Reporting on the first 6 months of implementation in each State.	Based on qualitative data collected through site visits, interviews, or other State reporting, these reports will provide information to CMS and each individual State about early implementation experience.
Quarterly Reports to CMS and States	Quarterly, beginning the quarter after the State-specific initial 6-month report.	These reports will include preliminary information on enrollment, disenrollment, quality, utilization, and cost measures for ongoing monitoring in each State. Initially, they will include data from the SDRS and predemonstration Medicare and Medicaid data as available. Later reports will include more information as the data become available.
Annual State-Specific and Aggregate Reports	Annually, for each of the demonstration performance periods.	These reports will summarize and update preliminary information in quarterly reports to CMS and States and provide context for the analysis. They will also include descriptive analysis of quality, utilization, and cost measures and qualitative information collected during site visits, focus groups, and telephone interviews. All beneficiaries eligible for the demonstration will be included in the annual analysis. Savings will be calculated at least twice during the demonstration for capitated model States using a regression-based methodology: once during the demonstration and once after the end of the demonstration, for the final evaluation report. Savings will also be calculated annually for MFFS States using an actuarial methodology, for performance payment purposes.
Final State-Specific Evaluation Reports and Final Aggregate Evaluation Report	After the demonstration period has ended.	The final State-specific reports and the final aggregate evaluation reports will contain multivariate analyses to provide a comprehensive understanding of the effects of the demonstration interventions on quality, utilization, and cost. The final report will also include cost-savings calculations and qualitative information collected during site visits, focus groups, and telephone interviews.

MFFS = managed fee for service; MMP = Medicare-Medicaid Plan; SDRS = State Data Reporting System.

Table ES-12
Information sources for the evaluation of the Financial Alignment Demonstrations

RTI will obtain data from:	Type of data
CMS	■ Encounter data (Medicare Advantage, Medicaid, and MMP) ■ HEDIS measures ■ Results from HOS and CAHPS surveys ■ Medicare and Medicaid fee-for-service claims ■ Medicare Part D costs and dual-eligibility status ■ Medicare Advantage Prescription Drug (MARx) data ■ Nursing Home data (MDS) ■ CMS-HCC and RXHCC risk scores ■ Demonstration quality measures that States are required to report to CMS (listed in MOUs) ■ Demonstration quality measures that health plans are required to report to CMS (listed in three-way contracts or other guidance) ■ Other administrative data as available
State	■ Detailed description of State's method for identifying eligible beneficiaries ■ File with monthly information identifying beneficiaries eligible for the demonstration (can be submitted monthly or quarterly)[1] ■ State Data Reporting System (SDRS; described in detail in Section 4 of the *Aggregate Evaluation Plan*) quarterly submissions of demonstration updates, including monthly statistics on enrollments, opt-outs, and disenrollments ■ Participation in key informant interviews and site visits conducted by RTI team ■ Results from surveys, focus groups, or other evaluation activities (e.g., EQRO reports) conducted or contracted by the State,[2] if applicable ■ Other data State thinks would benefit this evaluation, if applicable
Other sources	■ Results of focus groups conducted by RTI subcontractor (Henne Group) ■ Grievances and appeals ■ Other sources of data, as available

CAHPS = Consumer Assessment of Healthcare Providers and Systems; EQRO = external quality review organization; HCC = hierarchical condition category; HEDIS = Healthcare Effectiveness Data and Information Set; HOS = Health Outcomes Survey; MDS = Minimum Data Set; MMP = Medicare Medicaid Plan; MOU = Memorandum of Understanding; RXHCC = prescription drug hierarchical condition category.

[1] These data, which include both those enrolled and those eligible but not enrolled, will be used (in combination with other data) to identify the characteristics of the total eligible and the enrolled populations. More information is provided in Section 5 of this report.

[2] States are not required to conduct or contract for surveys or focus groups for the evaluation of this demonstration. However, if the State chooses to do so, the State can provide any results from its own independent evaluation activities for incorporation into this evaluation, as appropriate.

Section 1. Introduction

1.1 Purpose

The Medicare-Medicaid Coordination Office (MMCO) and Innovation Center at the Centers for Medicare & Medicaid Services (CMS) have created the Financial Alignment Initiative to test integrated care models for beneficiaries who are dually eligible for Medicare and Medicaid (Medicare-Medicaid enrollees). The goal of the Financial Alignment Initiative is to develop person-centered care delivery models integrating the full range of medical, behavioral health, and long-term services and supports (LTSS) for Medicare-Medicaid enrollees, with the expectation that integrated delivery models would address the current problems associated with the lack of coordination of Medicare and Medicaid benefits, financing, and incentives. At the time of this report, 23 States were working with CMS to develop their demonstrations: 14 contract States (those previously awarded Federal funding to design and develop integrated demonstrations for Medicare-Medicaid enrollees) and 9 noncontract States.

CMS contracted with RTI International to monitor the implementation of the demonstrations and to evaluate their impact over time on beneficiary experience, quality, utilization, and costs under the evaluation, titled *Measurement, Monitoring, and Evaluation of State Demonstrations to Integrate Care for Dual Eligible Individuals*. This *Aggregate Evaluation Plan* presents detailed plans for the work involved in this project, including identification of evaluation implications for the models, the data needs for the State-specific and cross-cutting analyses in all aspects of the evaluation, and selection of data elements and outcomes most relevant and feasible to incorporate into quarterly data reports to CMS and States and into annual and final evaluation reports. State-specific evaluation design plans will be developed for each State selected to participate in the Financial Alignment Initiative. The State-specific plans will tailor the plans presented in this report to each State's particular intervention and eligible populations.

This report is a guide for the evaluation, subject to future modifications. Although this document will not be revised to address all changes that may occur, the annual and final evaluation reports will note areas where the evaluation as executed differs from this evaluation plan.

1.2 Current State Plans

The States are currently in different phases of demonstration development, working with CMS to finalize their demonstration designs. ***Table 1*** is based on recently available information about the States' demonstration proposals.

Table 1
State proposal models and scope (as of September 2013)

State[1]	Contract state	Financial alignment model(s)	Eligible population	Geographic area
California	Contract	Capitated	Full-benefit Medicare-Medicaid enrollees, aged 21 and older in 8 counties. Beneficiaries not eligible include residents of certain rural zip codes, ICF-DDs or Veteran's Homes of California; individuals with an ESRD diagnosis residing in certain counties; individuals who qualify for Medicaid as medically needy who are not continuously certified; those with other health coverage; and current enrollees in PACE or the AIDS Healthcare Foundation or in the following 1915(c) waivers: Nursing Facility/Acute Hospital, HIV/AIDS, Assisted Living, and In Home Operations.[2]	Regional
Colorado	Contract	MFFS	Full-benefit Medicare-Medicaid enrollees of all ages. Beneficiaries not eligible include those in existing State Medicaid managed care programs.	Statewide
Connecticut	Contract	MFFS	**Model 1**: Full-benefit Medicare-Medicaid enrollees, aged 18 and older. Beneficiaries not eligible include enrollees in health homes for individuals with SPMI, and Model 2 enrollees.	Statewide
			Model 2: Full-benefit Medicare-Medicaid enrollees, aged 18 and older. Beneficiaries not eligible include enrollees in health homes for individuals with SPMI.	Regional
Idaho	Noncontract	Capitated	Full-benefit Medicare-Medicaid enrollees aged 21 and over.	Statewide
Illinois[3]	Noncontract	Capitated	Full-benefit Medicare-Medicaid enrollees 21 and older. Beneficiaries who use HCBS developmental disability services, HCBS waiver, or ICF/MR services; have comprehensive third-party coverage; or are on spend-down are not eligible.	Regional
Iowa	Noncontract	MFFS	Full-benefit Medicare-Medicaid enrollees of all ages eligible for participation in a Medicaid health home.	Statewide

(continued)

Table 1 (continued)
State proposal models and scope (as of September 2013)

State[1]	Contract state	Financial alignment model(s)	Eligible population	Geographic area
Massachusetts[3]	Contract	Capitated	Full-benefit Medicare-Medicaid enrollees, aged 21–64 upon enrollment. Enrollees in a 1915(c) waiver, residents of an intermediate care facility for individuals with intellectual disabilities (ICF/IID), individuals enrolled in a Medicare Advantage plan, PACE, an Employer Group Waiver Plan (EGWP) or other Employer-Sponsored Plans, or plans receiving a Retiree Drug Subsidy (RDS), or individuals participating in the CMS Independence at Home (IAH) demonstration are not eligible.	Counties with at least one MMP
Michigan	Contract	Capitated	Full-benefit Medicare-Medicaid enrollees of all ages. PACE enrollees and spend-down beneficiaries are not eligible.	Regional
Minnesota[3]	Contract	Other	Full-benefit Medicare-Medicaid enrollees in the Minnesota Senior Health Options (MSHO) program.	Statewide
Missouri	Noncontract	MFFS	Full-benefit Medicare-Medicaid enrollees of all ages eligible for participation in a health home and who reside in the community (non–facility-based).	Statewide
New York[3]	Contract	Capitated	**Model 1:** Full-benefit Medicare-Medicaid enrollees who are aged 21 and older and who require community-based long-term services and supports for more than 120 days or who require Medicaid-covered facility-based long-term services and supports. The following people are not eligible for enrollment: people who receive services through the Office of People with Developmental Disabilities (OPWDD); people who receive inpatient services in an Office of Mental Health (OMH) facility; and people on the Traumatic Brain Injury Waiver.	Regional
			Model 2: Full-benefit Medicare-Medicaid enrollees aged 21 and older, who are not residents of an Office of Mental Health facility, and who are receiving services from the State Developmental Disability system.	Regional
North Carolina	Contract	MFFS	Full-benefit Medicare-Medicaid enrollees 21 and older. Enrollees are not eligible in the months they are using specialty behavioral health plan services.	Statewide

(continued)

Table 1 (continued)
State proposal models and scope (as of September 2013)

State[1]	Contract state	Financial alignment model(s)	Eligible population	Geographic area
Ohio[3]	Noncontract	Capitated	Full-benefit Medicare-Medicaid enrollees 18 and older. Beneficiaries on spend down; with third-party creditable coverage; receiving ICF-IDD or IDD waiver services; participating in PACE; or who are CMS Independence at Home demonstration participants are not eligible.	Regional
Oklahoma	Contract	MFFS	**Model 1:** Full-benefit Medicare-Medicaid enrollees of all ages, not enrolled in demonstration Model 2 or 3.	Statewide
		MFFS	**Model 2:** Full-benefit Medicare-Medicaid beneficiaries aged 45 and older.	Regional
		Capitated	**Model 3:** Full-benefit Medicare-Medicaid beneficiaries aged 45 and older.	Regional
Rhode Island	Noncontract	Capitated	Full-benefit Medicare-Medicaid enrollees, aged 21 and older.	Statewide
South Carolina	Contract	Capitated	Full-benefit Medicare-Medicaid enrollees aged 65 and older, in non-facility-based settings. Those enrolled in PACE are not eligible.	Statewide
Texas	Noncontract	Capitated	Full-benefit Medicare-Medicaid enrollees 21 and older required to enroll in STAR+ PLUS. Those residing in nursing facilities for 4 months or longer; residents of ICF/MRs; and enrollees in some 1915(c) waivers are not eligible.	Regional
Vermont	Contract	Capitated	Full-benefit Medicare-Medicaid enrollees of all ages. Those enrolled in PACE Vermont are not eligible.	Statewide
Virginia[3]	Noncontract	Capitated	Full-benefit Medicare-Medicaid enrollees, aged 21 and older. Individuals in State mental hospitals, State Hospitals, ICF/MR facilities, Residential Treatment Facilities, and long-stay hospitals; in some HCBS waivers;[4] in hospice; receiving the ESRD Medicare benefit; receiving other comprehensive group or individual health insurance coverage; in Money Follows the Person (MFP) Program; in PACE; or in the Independence at Home demonstration are not eligible.	Regional

(continued)

Table 1 (continued)
State proposal models and scope (as of September 2013)

State[1]	Contract state	Financial alignment model(s)	Eligible population	Geographic area
Washington	Contract	MFFS[3]	**Model 1:** Full-benefit Medicare-Medicaid enrollees of all ages eligible for Medicaid health home services.	Statewide, except for capitated counties
		Capitated	**Model 2:** Full-benefit Medicare-Medicaid enrollees of all ages. Enrollees in Medicare Advantage or PACE; who receive hospice services; or who reside in a State Residential Habilitation Center are not eligible.	Regional
Wisconsin	Contract	Other	Full-benefit Medicare-Medicaid enrollees residing in nursing facilities in long-term FFS Medicaid stays.	Regional

ESRD = end stage renal disease; FFS = fee for service; HCBS = Home and Community Based Services; ICF-IDDs = intermediate care facilities for individuals with developmental disabilities; ICF/IDs = Intermediate Care Facilities for Individuals with Intellectual Disabilities; ICF/MR = intermediate care facilities for people with mental retardation; MFFS = managed fee for service; MMP = Medicare-Medicaid Plans; PACE = Program of All-Inclusive Care for the Elderly; SPMI = severe and persistent mental illness; STAR+ PLUS = a Texas Medicaid managed care program.

[1] As of April 2013, Hawaii had requested a 2015 demonstration start date, which is not currently an option for the Financial Alignment Demonstrations. Updates will be made if circumstances change.

[2] Also not eligible are individuals enrolled in a prepaid health plan that is a nonprofit health care service plan with at least 3.5 million enrollees statewide, that owns or operates its own pharmacies and that provides medical services to enrollees in specific geographic regions through an exclusive contract with a single medical group in each specific geographic region in which it operates to provide services to enrollees.

[3] Denotes that at the time of this report, the State had a Memorandum of Understanding (MOU) with CMS to establish the parameters of the initiative. Note that only the managed fee-for-service demonstration in Washington has a finalized MOU.

[4] The specific waivers are Individual and Family Developmental Disabilities Support; Intellectual Disabilities; Technology Assisted Waiver; Day Support; or Alzheimer's Assisted Living.

SOURCE: Each State's State Demonstration to Integrate Care for Dual Eligible Individuals proposals to CMS, with the following exceptions: California, Illinois, Massachusetts, Ohio, Virginia, and Washington (MFFS model only). Information for these six states is drawn from each State's Memorandum of Understanding with CMS.

1.3 Research Questions

Nine major research questions of the evaluation are presented in ***Table 2,*** along with possible data sources. The data sources will be described in more detail later in this report. As part of the evaluation design process, we have assessed each State's capacity to report the information required for this evaluation and have highlighted potential issues that may affect aspects of this evaluation plan.

1.4 Structure of Report

The sections of this report are organized by the following evaluation design areas: implementation; beneficiary experience; the State Data Reporting System (SDRS); an overview of the quantitative design; and analyses of quality, utilization, access to care, and costs. Evaluation design considerations regarding subpopulation analyses, which will be implemented as appropriate in each aspect of the analysis (implementation, beneficiary experience, quality, utilization, and costs), are also discussed. Each section presents additional research questions for the evaluation design area and methods for addressing the questions. The final section of this report will discuss the process and next steps for the evaluation.

Table 2
Research questions and data sources

Research questions	Stakeholder interviews and site visits	Beneficiary focus groups	Claims and encounter data analysis	Demonstration statistics[1]
1) What are the primary design features of each State's demonstration and how do they differ from the State's previous systems?	X	X	—	X
2) To what extent did each State implement its demonstration as designed? What factors contributed to successful implementation? What were the barriers to implementation?	X	—	—	X
3) What impact do these demonstrations have on the beneficiary experience overall, by State and for beneficiary subgroups? Do beneficiaries perceive improvements in how they seek care, choice of care options, how care is delivered, personal health outcomes and quality of life?	X	X	—	X
4) What impact do the demonstrations have on cost and is there evidence of cost savings in each State? How long did it take to observe cost savings in each State? How were these savings achieved in each State?	—	—	X	—
5) What impact do these demonstrations have on utilization patterns in acute, long-term, and behavioral health services, overall, by State, and for beneficiary subgroups?	X	X	X	X
6) What impact do these demonstrations have on health care quality overall, by State, and for beneficiary subgroups?	—	—	X	X
7) Do these demonstrations change access to care for medical, behavioral health, long-term services and supports (LTSS) overall and for beneficiary subgroups, by State? If so, how?	X	X	X	X
8) What policies, procedures, or practices implemented by each State in its demonstration can inform adaptation or replication by other States?	X	X	—	X
9) What strategies used or challenges encountered by each State in its demonstration can inform adaptation or replication by other States?	X	X	—	X

— = not applicable.

[1] Demonstration statistics refer to data that States, CMS, or other entities will provide regarding topics including enrollments, disenrollments, grievances, appeals, and number of health or medical homes. States will be submitting quarterly data updates through the State Data Reporting System (***Section 4***). States will also submit reports summarizing findings of any external quality review organization analyses, beneficiary surveys, and other quality monitoring activities required by CMS or undertaken by the States during the demonstration period.

This page intentionally left blank

Section 2. Demonstration Implementation

2.1 Overview

We will monitor and evaluate implementation of each demonstration, using qualitative and quantitative approaches. The qualitative component of the evaluation will consist of collecting baseline information, monitoring demonstration implementation, and analyzing system changes in each State. The electronic State Data Reporting System (SDRS) described in ***Section 4*** will incorporate both qualitative and quantitative information to monitor and evaluate demonstration implementation.

We have established RTI evaluation teams for each State that will focus on collecting data, monitoring the information submitted to the SDRS, and establishing and maintaining lines of communication with each demonstration State. These teams have been communicating with Medicare-Medicaid Coordination Office staff during the demonstration design phase, and throughout the development of State-specific evaluation designs. The teams have also worked directly with each State to collect information about the State's data capabilities relevant to developing individual and aggregate evaluation plans. The protocols described in this section have been informed by currently available information about the States' demonstration designs. We will use this information to develop site visit questions specific to the models and populations being served in the individual State demonstrations. In addition to individual State-specific reports, we will conduct annual aggregate analyses of demonstration implementation and include aggregate sections in the final evaluation report.

2.2 Implementation Research Questions

The evaluation of the implementation process is designed to answer the following overarching research questions about the State demonstrations:

- What are the primary features of each State demonstration and how do they differ from the State's previous system available to the demonstration-eligible population?
- To what extent did each State implement the demonstration as proposed?
- Which States were able to fully implement their intended proposals?
- Were certain models more easily implemented than others?
- Were the demonstrations more easily implemented for certain subgroups?
- What factors contributed to successful implementation?
- What were the barriers to implementation?

- How have beneficiaries participated in the ongoing implementation and monitoring of the demonstrations?
- What strategies used or challenges encountered by each State can inform adaptation or replication by other States?

2.3 Methodology

2.3.1 Research Design

To understand and assess demonstration implementation and systems change, we will use case study methods and qualitative and quantitative data analysis. Prior to demonstration implementation, we will establish a profile of each State's preexisting service delivery system for Medicare-Medicaid enrollees to enable us to identify key elements that will be modified through the demonstration and measure the effects of those changes. Using case study methods, we will conduct descriptive analyses of key demonstration design features for each State, and comparative analyses across States. Finally, demonstration design variables will be included in quality, utilization, access to care, cost, and beneficiary experience qualitative and quantitative analyses that will identify demonstration characteristics associated with comparatively better outcomes.

2.3.2 Monitoring Implementation of the Demonstration by Key Demonstration Design Features

Our analysis of the implementation of each State demonstration will be organized by key demonstration design features. This framework will be used to define our areas of inquiry, structure the demonstration implementation variables we track, organize information from our data collection sources, and outline our annual report. For this task, our goal is to frame analysis at the level of State policy or practice and examine how States' strategies and demonstration features translate at the plan or practice level.

Table 3 illustrates the key components of each demonstration design feature that we will analyze as part of the implementation evaluation.

Table 3
Demonstration design features and key components

Design feature	Key components
Core components of integrated delivery systems (how the delivery system is organized/integrated; interrelationships among the core delivery system components)	■ MMPs ■ Primary care, including medical homes and health homes ■ LTSS ■ Behavioral health services ■ Developmental disability services ■ Integration functions that bridge delivery systems and roles of community-based organizations
Integrated delivery systems supports	■ Care team composition ■ Health IT applied throughout the demonstration (at State level, by MMPs, at provider level or other) ■ Data (Medicare claims or encounter data) and other feedback to MMPs, medical/health homes, other providers (by the State or other entities) ■ Primary care practice support (e.g., coaching, learning collaboratives, training)
Care coordination/case management (by subpopulation and/or for special services) ■ Medical/primary ■ LTSS ■ Behavioral health services ■ Integration of care coordination	■ Assessment process ■ Service planning process ■ Care management stratification process ■ Support of care transitions across settings ■ Communication and hand-offs between care coordinators/case managers and providers
Benefits and services	■ Scope of services/benefits ■ New or enhanced services ■ Excluded services ■ Service authorization process
Enrollment and access to care	■ Integrated enrollment and access to care ■ Provider accessibility standards ■ Marketing/education protocols ■ Enrollment brokers ■ Beneficiary information and options counseling ■ Opt-out, disenrollment, and auto-assignment policy ■ Assignment/referrals to providers, health homes, medical homes ■ Phased enrollment of eligible populations ■ Workforce development for worker supply and new functions
Beneficiary engagement and protections	■ State policies to integrate Medicare and Medicaid grievances and appeals ■ Quality management systems ■ Ongoing methods for engaging beneficiary organizations in policy decisions and implementation ■ Approaches to capture beneficiary experience, such as surveys and focus groups
Demonstration financing model and methods of payment to plans and providers	■ Financing model: capitation or managed fee for service ■ Entities to which States are directly making payments ■ Innovative payment methods to MMPs and/or to providers
Elements of payments to MMPs and providers	■ Incentives ■ Shared savings ■ Risk adjustment

IT = information technology; LTSS = long-term services and supports; MMP = Medicare-Medicaid Plan.

2.4 Implementation Tracking Elements for Each Demonstration Design Feature

Through document review and interviews with State and Federal officials, we will identify and document the preexisting service delivery system for Medicare-Medicaid enrollees in each State. This will enable us to identify key elements that each State intends to modify through the demonstration and measure the effects of those changes. Using a combination of case study methods, document review, and telephone interviews, we will conduct a descriptive analysis of each State's key demonstration design features.

We will analyze how States are carrying out their plans and track any changes they make to their demonstration designs as implementation proceeds. We will identify both planned changes (e.g., phasing in new populations or benefits) and operational and policy modifications the States make based on changing circumstances. Finally, we anticipate that in some instances, changing State policy environments will trigger alterations to the State's original demonstration design. We will collect data through the SDRS to track implementation. We will follow up with State demonstration staff through quarterly telephone calls to clarify State entries as needed and to generally keep abreast of demonstration developments. We will use site-visit interviews to learn more about what factors are facilitating or impeding progress or leading to revisions in demonstration implementation for the participating States.

During site-visit interviews and our ongoing communications with States, we will collect detailed information on how they have structured care coordination and the extent to which they are prescriptive in setting care coordination expectations in contracts with Medicare-Medicaid Plans (MMPs), health homes, medical homes, and other care coordination entities. In managed fee-for-service demonstration models, we will identify the roles and functions of health homes, medical homes, or any other entities with which the State has vested responsibility for primary/medical care coordination and their linkages with long-term services and supports (LTSS) and behavioral health services. In capitated demonstration models, we will analyze the scope of care coordination responsibilities assigned to MMPs, the extent to which they conduct these functions directly or through a contract, and internal structures established to promote service integration. We will also examine how care coordination activities conducted under the demonstration by MMPs compare with approaches used by States in their capitated programs serving other populations.

For both demonstration model types, we will assess whether States have designed and implemented care coordination to be consistent with principles of person-centeredness. We will also examine the following: (1) how services are integrated and coordinated across providers and settings; (2) requirements for assessments and service planning; (3) the locus, scope, and authority of care coordinators and how these vary for different sets of services; (4) the degree to which States provide informatics infrastructure support to

providers, networks, or MMPs; and (5) protocols for sharing information among care coordinators and providers on changes in an individual's status, such as hospital or nursing facility admissions, medication reconciliations, and care plan modifications. Person-level care coordination data (i.e., which individuals receive what type of care coordination) may be available in some States. To the extent that States are able to report on specific aspects of care coordination services provided to individual Medicare and Medicaid enrollees, we will examine these data.

We will analyze the design and implementation of each State's approach to care coordination, and, to the extent possible, develop a method for categorizing across all States the intensity and scope of mandated functions. This framework will be developed based on information collected during the implementation of the demonstrations through the SDRS, document reviews, and interviews.

The evaluation team will track implementation of the following elements of each demonstration's design features. These elements will be monitored through the SDRS, which includes a combination of measurable data elements (e.g., number of MMP contracts; new waivers or State Plan Amendments [SPAs]) and descriptions of progress with implementation of the particular feature (e.g., practice-level reporting; adoption of electronic health records [EHRs]). ***Table 4*** shows the types of demonstration implementation elements we will track using State submissions to the SDRS, quarterly calls with State demonstration staff, other interviews, and site visits.

Table 4
Implementation tracking elements by demonstration design feature

Design feature	Tracking elements
Integrated delivery system	■ Three-way contracts with MMPs ■ Documentation of coordination activities between MMPs and community-based organizations ■ New waiver authorities submitted for the demonstration and approved ■ Emergence of new medical homes and health homes ■ Strategies for integrating primary care, behavioral health, and LTSS (as documented in State policies, contracts, or guidelines) ■ Recognition and payment for services by nontraditional workers ■ Innovative care delivery approaches adopted by the demonstration
Integrated delivery systems supports	■ Ongoing learning collaboratives of primary care providers ■ Support with dissemination and implementation of evidence-based practice guidelines (e.g., webinars for providers, learning collaboratives) ■ Decision-support tools provided or supported by State (e.g., practice-level reporting on QIs) ■ State efforts to build MMP and provider core competencies for serving beneficiaries with various disability types ■ Provision of regular feedback to MMPs and providers on the results of their performance measures

(continued)

Table 4 (continued)
Implementation tracking elements by demonstration design feature

Design feature	Tracking elements
Care coordination	■ Adoption of person-centered care coordination practices ■ State systems for collecting data on care coordination use ■ As available, care coordination activities directed to individual enrollees ■ State requirements for assessment and service planning ■ State requirements for coordination and integration of clinical, LTSS, and behavioral health services ■ State approaches to stratify care coordination intensity based on individual needs ■ State requirements for care transition support, medication reconciliation, notification of hospitalizations; State actions to facilitate adoption of EMR and EHR ■ Use of informatics to identify high-risk beneficiaries
Benefits and services	■ Phase-in of new or enhanced benefits, and methods to communicate them to enrollees and potential enrollees ■ Adoption of evidence-based practices and services (e.g., use of chronic disease self-management programs by practices, fall prevention programs, other)
Enrollment and access to care	■ State efforts to provide integrated consumer information on enrollment, benefits, choice of MMP/provider ■ Options counseling and information provided by Aging and Disability Resource Centers and State Health Insurance Assistance Programs ■ Initiatives to increase enrollment in the demonstration ■ Strategies for expanding beneficiary access to demonstration benefits ■ Emergence of new worker categories/functions (e.g., health coaches, community care workers)
Beneficiary engagement and protections	■ Strategies implemented to engage beneficiaries in oversight of the demonstration ■ Quality management strategy, roles, and responsibilities ■ Implementation of quality metrics ■ Adoption of new State policies for beneficiary grievances and appeals based on demonstration experience
Financing and payment	■ Revisions to the demonstration's initial payment methodology, including risk-adjustment methodology ■ Risk-mitigation strategies ■ Performance incentive approaches ■ Value-based purchasing strategies

EHR = electronic health records; EMR = electronic medical records; LTSS = long-term services and supports; QIs = quality improvement initiatives; MMP=Medicare-Medicaid Plan.

To the extent that the operations support contractor engaged by CMS for these demonstrations is evaluating and monitoring implementation at the MMP level for capitated demonstrations, we will integrate that information into our measurement and reporting (though the principal focus of the evaluation of capitated models will be at the State level, rather than the MMP level).

2.5 Progress Indicators

In addition to tracking implementation of demonstration design features in each State, we will also track progress indicators based on States' data reported quarterly through the SDRS (see ***Section 4*** for more information) and other data provided by RTI and other entities. ***Table 5*** presents examples of progress indicators we will track through the SDRS. These data-driven indicators will be presented in standardized tables and graphs for quarterly monitoring reports to CMS and the States.

Table 5
Examples of SDRS progress indicators

Indicator
No. of individuals…
— eligible to participate in the demonstration
— currently enrolled in the demonstration
— passively enrolled in the demonstration
— who opted out of the demonstration prior to enrollment
— who voluntarily disenrolled from the demonstration
— whose enrollment in the demonstration ended (e.g., death, loss of eligibility)

SDRS = State Data Reporting System.

2.6 Data Sources

The evaluation team will use a variety of data sources to assess whether State demonstrations were implemented as proposed; identify modifications to the initial demonstration design features made during implementation; document changes in the time frame or phase-in of key elements; and determine factors that facilitated implementation or presented challenges. These data sources include the following:

- **State policies and State specifications for provider and plan agreements:** The evaluation team will review a wide range of State-developed documents that specify each State's approach to implementing its demonstration in order to develop a baseline profile of its predemonstration and demonstration delivery system. Review of States' agreements with CMS articulated through the demonstration Memoranda of Understanding (MOUs), waivers, contracts, and State Plan Amendments will further enhance our understanding of each State's approach.
- **Demonstration data (collected via the SDRS):** On a quarterly basis, we will collect data from States to inform ongoing analysis and feedback to States and CMS throughout the demonstration. The data will be incorporated into quarterly data reports to States and CMS and annual and final evaluation reports. Specifically, we will collect data to track implementation variables such as development of new policies, as presented in ***Table 4*** and progress indicators that are mostly numeric counts of key demonstration elements presented in ***Table 5***. These demonstration data may include specific information provided by CMS or other entities and incorporated into the SDRS.

- **State agency staff, stakeholders, selected coordinated care organizations/ providers:** There will be at least two site visits to each demonstration State. The first one will occur within 6 months of the demonstration start. Using two-person site visit teams, we will obtain perspectives from individuals we interview on progress to date, internal and external environmental changes, other CMS initiatives the State is participating in, reasons each State took a particular course, and current successes and challenges. In addition to the site visits, we will develop a schedule of quarterly telephone interviews with various individuals involved with the demonstration in each State.

In addition to consumer advocates, as discussed in ***Section 3.4.2, Key Stakeholder Interviews,*** candidates for interviews may include the following:

- Representatives from Implementation Council
- Representatives from CMS–State Contract Management Team
- State officials, such as:
  - Secretary of State health and human services agency
  - State Medicaid Director
  - Chief Medical Officer
  - Chief Quality Officer
  - Demonstration Project Director
  - LTSS Program Director
  - State Medicaid agency Finance Manager
  - State Aging and Disabilities Director
  - State Behavioral Health Director
  - State Developmental Disabilities Director
  - State representatives from the State demonstration's advisory committee
- Directors of MMPs (specific to each State)
- Directors of health home/medical home providers (specific to each State)
- Representatives from entities providing options counseling for the demonstration
- Representatives from the demonstration Ombuds program

The site-visit interview protocols used in the evaluation will contain a core set of questions that allow us to conduct a cross-State evaluation as well as questions that are specific to each State's model. Questions tailored to those we interview in each State will be developed and added to the core set to inform the State-specific implementation evaluation as the State's model is finalized. A site-visit interview protocol with core questions is provided in ***Appendix A***.

The RTI team will contact each State in advance to obtain input on the appropriate individuals to interview. We will work with the State to schedule the site visits and the on-site interviews to best meet the needs of those individuals.

2.7 Analytic Methods

For each State, evaluation of the demonstration implementation will be presented in the initial 6-month report, quarterly data reports to States and CMS, and annual State-specific evaluation reports, and integrated into a final aggregate report that examines implementation experiences across similar demonstrations and across all demonstrations, as appropriate. Quantitative and qualitative data will be collected and reported quarterly through the SDRS. We will integrate that information with qualitative data that we will collect through site visits and telephone interviews with State agency staff and other individuals and include these data in the 6-month, annual, and final evaluation reports, as appropriate. These data will provide context for interpreting the impact and outcomes related to beneficiary experience, quality, utilization, and costs and enable us to analyze (1) the changes each State has made to the preexisting delivery systems serving Medicare-Medicaid enrollees, (2) challenges each State met, and (3) approaches that can inform adaptation or replication by other States.

The activities to capture demonstration implementation may be revised if modifications are made to the demonstrations or if data sources are not available as anticipated. If modifications to this evaluation plan are required, they will be documented in the annual and final evaluation reports.

This page intentionally left blank

Section 3. Beneficiary Experience

3.1 Overview

The impact each State demonstration has on the beneficiary experience is an important focus of the evaluation. We will monitor and evaluate the experiences of beneficiaries, their families, and caregivers to assess how closely the demonstrations meet CMS's stated goal of designing person-centered care delivery models. The beneficiary experience will be monitored and evaluated using multiple approaches, including use of beneficiary focus groups; monitoring of beneficiary engagement activities, grievances, and appeals; interviews with key stakeholders, including relevant advocacy organizations; claims data analyses of key quality, utilization, and access to care measures; and review of the results of beneficiary surveys conducted for this demonstration by the State, CMS, or other entities.

We will also use interview, survey, and focus group data to provide context and assist in interpreting other quantitative evaluation findings on quality, utilization, and costs. The methods and data sources to be used include the following:

- the beneficiary voice, as documented through focus groups, consumer and advocacy stakeholder interviews, feedback from demonstration Ombuds programs, and results of any surveys that may be conducted by States, CMS, or other entities;
- State demonstration data, reported via the State Data Reporting System (SDRS) and from other sources (e.g., data on enrollments, disenrollments, appeals, grievances stakeholder engagement activities);
- claims and encounter data to evaluate beneficiary access to care and utilization of services and outcomes for key quality measures; and
- interviews with State agency staff involved in the demonstration.

Table 6 (described in more detail below) shows the range of topics and data sources we will use to monitor and evaluate the beneficiary experience. Using these multiple methods and data sources, we will assess the experience of beneficiaries, their families, and caregivers to determine the impact of the demonstration on their quality of life, health outcomes, access to needed services, the integration and coordination of services across settings and delivery systems, provider choice, beneficiary rights and protections, and the delivery of person-centered care. In the process, we will identify what has changed for beneficiaries since their enrollment in the demonstration and the impact of the demonstration on beneficiaries' health and quality of life.

This section of the evaluation plan focuses specifically on the methods we will use to monitor and evaluate beneficiaries' experiences from the perspective of the beneficiaries themselves, through the use of focus groups with beneficiaries and interviews with consumer and advocacy groups. We also discuss information about demonstration statistics we will obtain from the States and from results of beneficiary surveys—either those administered by the States themselves or by CMS or other entities. ***Section 2, Demonstration Implementation***, describes topics we will monitor and document through

interviews with demonstration staff and document reviews, including consumer protections and other demonstration design features intended to enhance the beneficiary experience. In the next section, ***Section 4, The State Data Reporting System***, we describe specific data we will collect from States to track information that reflects aspects of the beneficiary experience such as, disenrollment patterns. ***Sections 6*** to ***10, Analysis Overview; Utilization and Access to Care; Quality of Care; Cost;*** and ***Subpopulations and Health Disparities*** discuss how we will use claims and encounter data to establish baseline information about the beneficiaries enrolled in the demonstration, and how we will use these data to inform our understanding of the demonstrations' impact on their access to care and health outcomes.

3.2 Research Questions

Specifically, we will address the following research questions in this section:

- What impact do these demonstrations have on beneficiary experience overall, by State, and for beneficiary subgroups?
- What factors influence the beneficiary enrollment decision?
- Do beneficiaries perceive improvements in their ability to find needed health services?
- Do beneficiaries perceive improvements in their choice of care options, including self-direction?
- Do beneficiaries perceive improvements in how care is delivered?
- Do beneficiaries perceive improvements in their personal health outcomes?
- Do beneficiaries perceive improvements in their quality of life?

Through beneficiary focus groups and key stakeholder interviews (i.e., consumer and advocacy group members), we also will explore whether we can identify specific demonstration features in each State that affect beneficiary experience and, where possible, how those features also affect evaluation findings from quantitative analysis of quality, utilization, and costs.

3.3 Approach

This mixed-methods evaluation will combine qualitative information from focus groups and key stakeholder interviews with quantitative data related to beneficiary experience derived from the RTI SDRS (***Section 4***) and any surveys that States and other entities choose to conduct, where available. Qualitative data will be obtained directly from a beneficiary or beneficiary representative through focus groups, and interviews. We will apply a narrow definition of *representative* to include only family members, advocates, or members of organizations or committees whose purpose is to represent the interest of beneficiaries and who are not service providers or do not serve in an oversight capacity for the initiative. Although RTI will not have uniform baseline qualitative beneficiary experience data for the predemonstration period, beneficiaries will be asked about their experience before the demonstration and how it may have changed during the course of the demonstration.

Results of focus groups conducted by individual States prior to the demonstration may also be reviewed and, where appropriate, incorporated into the evaluation of a State's demonstration.

Our framework for evaluating beneficiary experience is influenced by work conducted by the Center for Health Care Strategies (CHCS) on how to design an integrated system for Medicare-Medicaid enrollees that strengthens the beneficiary experience. CHCS has identified the essential elements of integration that directly affect beneficiary experience, including the care process and quality of life: comprehensive assessment of need; person-centered plan of care; multidisciplinary care teams; family/caregiver involvement; a comprehensive provider network; strong home and community-based services (HCBS) options; robust data sharing and communications systems; and aligned financial incentives (Lind and Gore, 2010).

Table 6 aligns key elements identified in the CHCS framework with the demonstration design features described in ***Section 2***. We have added elements to the CHCS framework to more directly address the impact of State demonstrations on health outcomes and quality of life. Others have been modified to reflect that not all Medicare-Medicaid enrollees require intensive services, as would be suggested by the original CHCS language used when describing comprehensive assessments and multidisciplinary care teams. For each key element, we identify the impact on beneficiary experience and detail the data sources that we anticipate using to get the information.

As shown in ***Table 6***, we will solicit direct feedback from beneficiaries served through the demonstrations to determine how closely their experience compares to the measures we will use in the evaluation (i.e., improvements in personal health outcomes, beneficiaries' quality of life, how beneficiaries seek care, choice of care options, and how care is delivered). We will solicit this feedback both through focus groups and by reviewing the results of beneficiary surveys administered or required by the State (e.g., Consumer Assessment of Healthcare Systems and Providers [CAHPS], the Participant Experience Survey, or questions we recommend be included in States' beneficiary surveys, if States are conducting such surveys). We will include topics specific to the demonstration model and supplement our understanding of direct beneficiary experience with key stakeholder interviews, a review of demonstration statistics, claims, and encounter data analysis, and interviews with States on demonstration implementation. In ***Table 7***, we highlight some of the topics we will monitor and evaluate using demonstration statistics and claims or encounter data analysis. See ***Sections 6*** to ***10*** for a discussion of the quality, utilization, access to care, and cost measures we plan to examine as part of the overall evaluation of the demonstrations' impact on beneficiary outcomes, including for subpopulations. ***Appendix B*** (discussed later in this section) is an outline of the stakeholder interview protocol and ***Appendix C*** (also discussed later in this section) is a draft of a focus group protocol.

Table 6
Methods for assessing beneficiary experience by beneficiary impact

Direct measure	Key stakeholder interviews	Beneficiary focus groups	Recommended survey question[1]	Demonstration data[2]	Interviews with State agency staff on demonstration implementation
Integrated delivery system					
Choice					
Beneficiaries have choice of medical, behavioral, and LTSS[2] services.	X	X	X	X	X
Beneficiaries have choice of medical, behavioral and LTSS providers.	X	X	X	X	X
Beneficiaries have choice to self-direct their care.	X	X	—	X	X
Beneficiaries are empowered and supported to make informed decisions.	X	X	—	—	—
Provider network					
Beneficiaries report that providers are available to meet routine and specialized needs.	X	X	X	X	—
Beneficiaries report that LTSS and behavioral health are integrated into primary and specialty care delivery.	X	X	—	X	—
Beneficiary engagement					
Beneficiaries consistently and meaningfully have the option to participate in decisions relevant to their care.	X	X	X	X	—
There are ongoing opportunities for beneficiaries to be engaged in decisions about the design and implementation of the demonstration.	X	X	—	—	X

(continued)

Table 6 (continued)
Methods for assessing beneficiary experience by beneficiary impact

Direct measure	Key stakeholder interviews	Beneficiary focus groups	Recommended survey question[1]	Demonstration data[2]	Interviews with State agency staff on demonstration implementation
Streamlined processes					
Beneficiaries can easily navigate the delivery system.	X	X	—	X	—
Reduced duplication of services					
Beneficiary burden is reduced through elimination of duplicative tests and procedures.	—	—	—	X	—
Enrollment and access to care					
Enrollment					
Beneficiaries have choices and assistance in understanding their enrollment options.	X	X	—	X	X
Beneficiaries report ease of disenrollment.	X	X	—	X	—
Rate of beneficiaries who opt out of enrolling into demonstration.	—	—	—	X	—
Rate of disenrollment from the demonstration, by reason.	—	—	—	X	—
Access to care					
Beneficiaries can access the full range of scheduled and urgent medical care, behavioral health services, and LTSS.	X	X	—	X	—
Beneficiaries report improved quality of life due to access to the full range of services.	X	X	X	—	—
Beneficiaries report that waiting times for routine and urgent primary and specialty care are reasonable.	X	X	—	X	—

(continued)

Table 6 (continued)
Methods for assessing beneficiary experience by beneficiary impact

Direct measure	Key stakeholder interviews	Beneficiary focus groups	Recommended survey question[1]	Demonstration data[2]	Interviews with State agency staff on demonstration implementation
Health outcomes					
Beneficiary health rating	—	—	X	—	—
Quality of life					
Days free from pain.	—	—	X	—	—
Beneficiaries get the social and emotional supports they need.	—	X	X	—	—
Beneficiaries report that they are satisfied with their life.	—	X	X	—	—
Cultural appropriateness					
Beneficiaries have access to multilingual and culturally sensitive providers.	X	X	—	X	X
Beneficiaries report that written and oral communications are easy to understand.	X	X	—	X	—
Delivery systems supports					
Data sharing and communication					
Information is available and used by beneficiaries to inform decisions.	X	X	—	—	X
Beneficiaries report that providers are knowledgeable about them and their care history.	X	X	—	X	—
Beneficiaries report adequacy of discharge and referral instructions.	X	X	—	X	X
Beneficiaries report that providers follow up after visits or discharge.	X	X	—	X	—
Beneficiaries understand their options to specify that personal health data not be shared.	X	X	—	X	—

(continued)

Table 6 (continued)
Methods for assessing beneficiary experience by beneficiary impact

Direct measure	Key stakeholder interviews	Beneficiary focus groups	Recommended survey question[1]	Demonstration data[2]	Interviews with State agency staff on demonstration implementation
Care coordination					
Assessment of need					
Assessment process integrates/addresses health, behavioral health, and LTSS.	X	X	—	X	X
Medicare providers actively participate in the assessment process.	—	X	X	—	—
Beneficiaries report active participation in the assessment process.	X	X	—	X	—
Person-centered care					
Care is planned and delivered in a manner reflecting a beneficiary's unique strengths, challenges, goals, and preferences.	X	X	—	X	—
Beneficiaries report that care managers have the skills and qualifications to meet their needs.	—	X	X	—	—
Beneficiaries report that providers listen attentively and are responsive to their concerns.	X	X	X	X	—
Coordination of care					
The system facilitates timely and appropriate referrals and transitions within and across services and settings.	X	X	X	X	—
Beneficiaries have supports and resources to assist them in accessing care and self-management.	X	X	—	X	—
Beneficiaries report ease of transitions across providers and settings.	X	X	X	X	—

(continued)

Table 6 (continued)
Methods for assessing beneficiary experience by beneficiary impact

Direct measure	Key stakeholder interviews	Beneficiary focus groups	Recommended survey question[1]	Demonstration data[2]	Interviews with State agency staff on demonstration implementation
Family and caregiver involvement					
Beneficiaries have the option to include family and/or caregivers in care planning.	X	X	—	X	—
The family or caregiver's skills, abilities, and comfort with involvement are taken into account in care planning and delivery.	X	X	—	X	—
Benefits and services					
Awareness of covered benefits					
Beneficiaries are aware of covered benefits.	X	X	—	X	—
Availability of enhanced benefits					
The demonstration covers important services to improve care outcomes that are not otherwise available through Medicaid or Medicare program.	—	—	—	X	X
Flexible benefits are available to meet the needs of beneficiaries.	—	—	—	X	X
Awareness of enhanced benefits					
Beneficiaries are aware of enhanced benefits and use them.	X	X	—	X	—
Beneficiary safeguards					
Beneficiary protections					
Beneficiaries understand their rights.	X	X	—	X	—
Beneficiaries are treated fairly, are informed of their choices, and have a strong and respected voice in decisions about their care and support services.	X	X	—	X	—

(continued)

Table 6 (continued)
Methods for assessing beneficiary experience by beneficiary impact

Direct measure	Key stakeholder interviews	Beneficiary focus groups	Recommended survey question[1]	Demonstration data[2]	Interviews with State agency staff on demonstration implementation
Complaints, grievances, and appeals Beneficiaries have easy access to fair, timely, and responsive processes when problems occur.	X	X	—	X	—
Number and type of beneficiary complaints, grievance, and appeals.	—	—	—	X	—
Advocacy/member services Beneficiaries get assistance in exercising their rights and protections.	X	X	—	X	—
Finance and payment					
Provider incentives Beneficiary experience is taken into account when awarding provider and plan incentives.	X	—	—	—	X
Rate of auto-assignment (if available).	—	—	—	X	—
Rate of change of PCP requests (if available).	—	—	—	X	—

— = not applicable; LTSS = long-term services and supports; HCBS = home and community-based services; PCP = primary care provider.

[1] The evaluation team will recommend questions to add to surveys that are already being conducted by the States or CMS. RTI is not conducting beneficiary surveys, and is not requiring States to begin a new survey for the purposes of this demonstration.

[2] Drawn from State Data Reporting System and RTI analysis of administrative data, Consumer Assessment of Healthcare Systems and Providers (CAHPS) or Health Outcomes Survey (HOS) results, or from other beneficiary surveys conducted for the demonstration by the States or other entities.

We will analyze our findings by subpopulation. When we can recruit sufficient numbers of individuals from the subpopulations of interest to participate in the focus groups, we will also analyze our focus group findings about beneficiary experience to determine whether differences exist by subpopulations.

Table 7
Demonstration statistics on quality, utilization, access to care measures of beneficiary experience

Rate of auto-assignment to plan/demonstration (if available)
Rate of disenrollment from the demonstration by reason[1]
Rate of beneficiaries who opt out of enrolling into demonstration
Number and type of beneficiary complaints, grievance, and appeals
Use of preventive services[1]
Nursing facility admissions and readmissions[1]
Emergency room use[1]
Hospital admission and readmission rates[1]
Follow-up care after hospital discharge[1]

[1] See ***Sections 4, 7,*** and ***8*** for discussion of specific measures.

3.4 Data Sources

As shown in ***Table 6***, we will rely on five major data sources to assess beneficiary experience. In this section, we describe our plan for using focus group and stakeholder interviews; results of beneficiary surveys planned by States, CMS, or other entities; State demonstration data entered into the SDRS; and interviews with demonstration staff.

3.4.1 Focus Groups

We will conduct four focus groups in each demonstration State to gain insight into how the initiative affects beneficiaries. To ensure that we capture the direct experience and observations of those served by a State's demonstration, focus groups will be limited to demonstration enrollees, their family members, and informal caregivers. ***Table 8*** shows our current plan for the composition and number of focus groups.

Some States (e.g., Massachusetts and California) have conducted their own focus groups prior to implementing their demonstrations, and may continue to do so during the demonstration implementation. We will take into account each State's schedule as we plan our focus groups and will use findings from the States' own focus group activities to inform the content of the focus group discussion guides that the evaluation team will use. Preliminary focus group topics that we anticipate covering include beneficiaries' understanding of the demonstration, rights, options, and choices (e.g., plan, primary care provider [PCP]); reasons that beneficiaries choose to enroll and disenroll; their benefits;

concerns or problems encountered; experience with care coordination; access to primary and specialty care, and beneficiary-reported impact of participation on health outcomes and quality of life. The timing for conducting focus groups within each State will be influenced by our assessment of whether there is more to be learned about the experience of beneficiaries shortly after initial enrollment into the demonstration versus beneficiary perceptions of the demonstration's effectiveness later in the demonstration. If the latter, we will conduct focus groups at least 1 year after implementation so that beneficiaries have had a substantial amount of experience with the demonstration. We will make the decision regarding timing of the focus groups in conjunction with CMS.

Table 8
Purpose and scope of State focus groups

Primary purpose	To understand beneficiary experience with the demonstration and, where possible, to identify factors and design features contributing to their experience.
Number	Four focus groups per demonstration State.
Composition	Each focus group includes 8–10 individuals who may be beneficiaries or family members or caregivers representing beneficiaries. Based on a State's enrolled population, a cross-section of individuals will be included in each focus group. In cases where the demonstration design is focused on one or a few subpopulations of Medicare-Medicaid enrollees, we may limit the focus groups in that State to individuals representing one or two subgroups. These may include but are not limited to beneficiaries who have ■ developmental disabilities, ■ severe and persistent mental illness, ■ substance use disorders, ■ multiple chronic conditions, or who ■ use long-term services and supports.

We will recruit focus group participants from eligibility and enrollment files. In doing so, we will identify enrolled beneficiaries reflecting a range of eligibility, clinical, and demographic characteristics. Although we may consult with States regarding logistics, State demonstration staff will not be actively involved in the recruitment process. Our subcontractor, the Henne Group, will use a structured approach for screening potential participants and obtaining their agreement to participate. If there appear to be high initial rates of opting out or disenrollment from the demonstration in some States, we will consider convening focus groups with beneficiaries who have chosen to opt out or disenroll, to understand their decisions. We will work closely with the demonstration staff in each State to make the process for recruiting focus group members as smooth as possible for beneficiaries, such as selecting an accessible site and ensuring transportation and any needed special accommodations and supports to allow for full participation. Focus group

recruitment and all focus group arrangements will be conducted with a particular awareness of the subpopulations we will examine in each State.

For example, if the eligible population is Medicare-Medicaid enrollees 21 to 64 years of age, we expect a high proportion of individuals with severe and persistent mental illness and severe physical disabilities, which will make accessibility considerations particularly important. If the eligible group is Medicare-Medicaid enrollees aged 65 and over, we expect a high proportion of beneficiaries with cognitive impairment and multiple chronic conditions; therefore they may not be able to attend the groups themselves, which will necessitate representation by caregivers.

Diversity is another dimension where we can expect variability by State. For example, in California, a significant proportion of Medicare-Medicaid enrollees represents various minority groups and ethnicities, and includes a large proportion of persons with limited English proficiency (LEP). We will investigate the prevalence of non-English-speaking beneficiaries in the State and determine whether to conduct any of the focus groups in languages other than English.

Several other subpopulations will require tailored approaches to focus group recruitment and arrangements. Persons in nursing facilities usually have high levels of frailty and cognitive impairments, necessitating the use of proxies to represent them in focus groups. Participation in focus groups by beneficiaries with significant chronic conditions and/or severe and persistent mental illness may also require similar approaches to recruitment such as special transportation assistance and proxy participation to ensure adequate representation of these subpopulations. To support successful focus group recruitment, we may provide incentives (e.g., gift cards) for focus group participants, if this is consistent with current Federal guidelines.

A preliminary focus group protocol is presented in ***Appendix C***. This protocol includes a core set of questions that we plan to cover in all focus groups, and may be modified based on final decisions about focus group composition, content, and our understanding of issues raised during demonstration implementation. We will also modify the focus group protocol as appropriate to add questions that will be relevant for a given State's population. However, we will include a subset of questions for all focus groups across all State demonstrations. Before beginning each focus group, and after obtaining informed consent, we will obtain basic demographic and background information from focus group participants, using a short demographic questionnaire. The evaluation team will use this information to describe the individuals who participate in the discussion (e.g., age, racial makeup, gender). To the extent possible we will also use these data to aid in reporting and interpreting focus group results.

3.4.2 Key Stakeholder Interviews

The RTI evaluation team will conduct key stakeholder interviews with consumer and advocacy groups in each State, either in person as part of scheduled site visits or by telephone, with major beneficiary groups whose stakeholders are served by the demonstrations. The purpose of these interviews will be to assess the level of beneficiary engagement and experience with the demonstration and its perceived impact on beneficiary outcomes. For purposes of assessing beneficiary experience, key stakeholder interviews will not include service providers.

Table 9 identifies the overall purpose and scope of the key informant interviews, as well as the types of groups from which interviewees will be selected. We will finalize the list of key stakeholders following discussions with demonstration staff in each State, a review of events and issues raised during the development of the demonstration, and the composition of enrollment by subpopulations.

Table 9
Purpose and scope of key stakeholder interviews

Primary purpose	**Baseline:** Assess understanding of and satisfaction with demonstration design; expectations for the demonstration; perceived concerns and opportunities across demonstrations. **Throughout demonstrations:** Spot improvements and issues as they emerge and assess factors facilitating and impeding positive beneficiary experience, including positive health outcomes and quality of life, across demonstrations. **Final year:** Assess extent to which expectations were met; major successes and challenges; lessons learned from beneficiary's perspective across demonstrations.
Number and frequency	**Baseline:** Up to eight telephone interviews in each State within 6 months after implementation. **Throughout demonstrations:** Up to eight telephone or in-person interviews in each demonstration State each year. **Final year:** Up to eight telephone or in-person interviews in each demonstration State.
Subpopulations	Key stakeholder interviews will be held with beneficiary groups whose stakeholders are served by a State's demonstration. These may include, but are not limited to, members of consumer advisory groups, beneficiary rights advocates, and public guardian groups.

An outline of the key stakeholder interview at baseline is presented in ***Appendix B***. We will revise this draft as we obtain more information about the State demonstrations and the issues that arise during planning/design phases and early implementation. Modifications may be made to the protocols to include State-specific questions that will allow the evaluation team to address issues unique to an individual State.

3.4.3 Beneficiary Surveys

The evaluation team will not directly administer beneficiary surveys as part of the evaluation and is not requiring States to administer beneficiary surveys for purposes of the evaluation. We will ask that States share with the evaluation any relevant findings from beneficiary surveys that States, CMS, or other entities may be conducting independently. We can also provide recommendations regarding beneficiary survey selection or measure selection relevant to evaluating the beneficiary experience if a State or other entity is conducting a survey.

Several States have proposed to administer a beneficiary survey under their demonstrations. Others may include beneficiary surveys as part of the required external quality review of any Medicaid risk-based managed care program. Medicare-Medicaid Plans (MMPs) will be required to participate in the Health Outcomes Survey (HOS) and report Healthcare Effectiveness Data and Information Set (HEDIS) measures, including Medicare CAHPS. Our understanding is that the operations support contractor for this demonstration may administer a beneficiary survey in States not otherwise conducting a survey—for example, a version of the CAHPS survey in the managed fee-for-service (MFFS) States. The RTI Evaluation Team will work with States and CMS, or other entities as appropriate, to strengthen opportunities for comparing beneficiary survey findings across States. For example:

- The ideal scenario would be for all survey sponsors to use a standard instrument. Although some States have required their plans to use the nationally recognized CAHPS survey, final decisions have not been made on which survey within the suite of CAHPS instruments will be used. The newly published version of CAHPS for assessing patient-centered medical homes (PCMH) seems to be the most appropriate of the existing CAHPS instruments in that it assesses practice performance in areas relevant to the demonstration such as access to care, comprehensiveness, self-management support, shared decision-making, coordination of care, and information about care and appointments.
- If it is not possible to reach agreement on a common instrument, we can propose a limited set of standard questions across all demonstrations for inclusion in the States' beneficiary surveys, such as quality of life measures, to allow for comparison across States. Table 6 identifies areas for which we could recommend common questions across all demonstrations.
- A common sampling protocol would ensure that survey findings are statistically significant for the demonstration population. Under the demonstration, MMPs will be required to report data specific to their demonstration enrollees from HEDIS as well as HOS and CAHPS surveys.
- We will ask that States share reports that aggregate and summarize results of survey efforts by States and CMS or other entities. We will review these reports to determine whether the sample selected for these surveys is representative of and adequate to make estimates for the demonstration population and whether there are data relevant to the beneficiary experience that can be included in our evaluation.

3.4.4 Demonstration Data

We will use data about the demonstration that we collect from States during site visits, from reports and other materials developed by States, through the State Data Reporting System and data obtained from CMS or other entities to assess the beneficiary experience. Data of particular interest include the following:

- Complaint, appeal, and grievance data from CMS, demonstration Ombuds programs, or other entities, as available.
- Disenrollment and opt-out rates, where appropriate.
- Information about waiting lists or lags in accessing services will provide useful indications of where the system lacks capacity, as a topic for discussion during site visits or focus groups.
- Rate of change in Primary Care Provider (PCP) assignment, if available.

The above quantitative indirect measures will be collected for all Medicare-Medicaid enrollees served under the demonstration and will be analyzed by subpopulations where available, given that these measures may not apply to all models. In addition, if States plan to monitor quality using their own set of quality metrics, to the extent relevant, we will use findings from these State-calculated metrics to augment our assessment of beneficiary experience and outcomes in a given State.

3.4.5 Interviews with State Demonstration Staff

In addition to key stakeholder interviews conducted with beneficiary stakeholder groups, we will address issues of beneficiary engagement and feedback during our interviews with key State demonstration staff. These interviews, which are described in ***Section 2***, will provide another perspective on how States communicate and work with beneficiaries during the design and implementation of their demonstrations.

3.5 Analytic Methods

Our analysis will assess beneficiary experience and determine, where possible, how it is affected by financial model and demonstration design features. We will also examine whether and how beneficiary experience varies by subpopulations. The Henne Group will audio-record all focus groups, subject to approval of the group members, and the audio-recordings will be transcribed. Key stakeholder interview and focus group transcripts will be imported and analyzed using QSR NVivo 9, a qualitative data analysis software, to identify emergent themes and patterns regarding beneficiary experiences during the demonstration and issues related to the evaluation research questions. A structured approach to qualitative analysis in NVivo 9 will allow us to (1) identify themes that emerge across focus groups and key stakeholder interviews within and across States and (2) compare and contrast those themes by subpopulation within and across States. We will also compare our qualitative

findings by the type of demonstration model—capitated or MFFS—to assess whether particular demonstration design features are likely to affect beneficiary experience.

Most demonstration data will be collected and tracked through the SDRS and analyzed using descriptive statistics reported quarterly through that system. We will also request summary statistics and reports from States on beneficiary experience surveys conducted by the State. Information from site visits and site-reported data beyond those described specifically in this section also are expected to inform analysis of beneficiary experience research questions. The findings will be grouped into the beneficiary experience domains defined in ***Section 3.3***.

The activities to capture beneficiary experience may be revised if modifications are made to the demonstrations or if data sources are not available as anticipated. If modifications to this evaluation plan are required, they will be documented in the annual and final evaluation reports.

Section 4. State Data Reporting System

4.1 Overview

The State Data Reporting System (SDRS) will be RTI's tool for collecting and storing up-to-date information about each State's demonstration design and progress, capturing data elements on a quarterly basis, and monitoring and reporting on demonstration progress by individual States and the demonstration as a whole. For the demonstrations, a quarter is a 3-month period, with the first quarter beginning the day of demonstration implementation. Because these quarters do not necessarily coincide with calendar quarters, the reporting periods will differ by State, depending on the implementation dates. The SDRS includes three types of data collection activities:

1. Model summary: The RTI evaluation team for each State will enter data describing each State's demonstration model into the SDRS. The data will be static, updated only if there are changes to a State's demonstration model. That information will be included as background information (i.e., a model summary) in the SDRS-generated quarterly reports to CMS and the States.
2. Implementation tracking data: Data will be collected from the demonstration States on a quarterly basis; some of this information will be presented in the quarterly reports for CMS and the States to track States' progress and some of this information will be used to describe demonstration implementation in the annual aggregate and State reports.
3. Demonstration impact and outcomes: Data will be collected from RTI analyses of claims, encounter, and assessment (Minimum Data Set [MDS]) data on quality, utilization, and cost measures, as available, as well as additional information provided by CMS or other entities, and entered into the SDRS by RTI. Some of this information will be presented in the quarterly reports for CMS and the States to track States' progress, and some of this information will be used in the annual aggregate and State reports.

An important consideration in the system's design is the need for complete information weighed against the reporting burden on the States. Wherever possible, RTI will enter data into the SDRS without additional State reporting. Where specific data can be collected only with State involvement, every effort has been made to require only those data elements deemed most relevant to the evaluation. Additional data elements, such as grievances and appeals data, will be supplied by CMS or other entities based on their monitoring activities. Any data supplied by CMS or other entities will be entered into the SDRS by RTI staff.

RTI is developing a user guide and training materials, and will conduct webinars for the States prior to implementation of the SDRS. We will maintain a Help Desk to address any technical issues States encounter in using the system. States will have contact information for the RTI evaluation team for their State available in the SDRS and user guide if there is a need to address other nontechnical questions.

4.2 Methodology

4.2.1 Data Collection

To inform the evaluation design, the RTI evaluation team for each State has communicated extensively with the individual demonstration States to assess which data elements could be collected by the evaluation team without the State demonstration's involvement (i.e., through timely Medicaid Statistical Information System [MSIS] submissions). For other targeted, high-priority data elements, the evaluation teams for each State and the State officials conducted data readiness reviews to assess the current availability, timeliness, and ease of collection for data elements that may need to be reported directly by the States through the SDRS. RTI is currently finalizing the Model Summary and Demonstration Implementation sections; the Impact and Outcomes section will be finalized after further discussion with CMS and other entities, as appropriate.

Data summarizing the demonstration design for each State's model will be derived from the Memoranda of Understanding (MOUs) between CMS and each demonstration State. Once the MOU for each State is released, RTI will finalize the model description for each State to incorporate into the Model Summary section of the SDRS.

States will be asked to submit brief text-based responses into the SDRS each quarter about topics including demonstration progress, successes, and challenges. They will be asked to identify any reports available about their demonstration, including analyses completed by external quality review organizations and beneficiary survey or focus group results completed by the State. The RTI evaluation teams for each State will contact the States to follow up on these entries and to request any relevant available reports.

4.2.2 System Design

We are designing a system that consists of three distinct types of data entry and reporting, which will closely follow the evaluation domains, as follows: (1) model summary; (2) demonstration implementation tracking by design features; and (3) demonstration outcomes, consisting of elements to track quality, utilization, access to care, and costs. States implementing more than one financial alignment model will need to enter data for these models separately.

The SDRS will be able to generate quarterly reports for dissemination to CMS and the States; only the RTI team will be able to generate these reports. The reports will include charts summarizing key information across States at a given time and charts using quarterly data for longitudinal tracking of outcomes, such as quality, utilization (including access to care), total costs, total costs by payer for managed fee-for-service (MFFS) models, and costs for selected services.

4.2.3 System Structure and Elements

Currently, the three sections of the SDRS contain a total of about 130 data elements, including 21 static elements in the Model Summary to be entered by RTI evaluation teams for each State and about 45 items in the Demonstration Implementation section to be entered directly by the State. We are planning for approximately 40 to 50 numerical fields for RTI to enter in the Impact and Outcomes section, which is under development. RTI will finalize the list of database elements and wording for each, in consultation with CMS.

4.2.3.1 Model Summary

The data for the Model Summary section will be static: that is, the data for this section will be entered by the RTI evaluation team for each State once and should not require updating unless key design features change (e.g., financial alignment model type, geographic expansion, more services folded into the capitation rate). RTI will use each State's MOU to develop the overall summary of the model. This information will be included as a summary at the beginning of each State quarterly report to CMS and States.

RTI will control data entry into the Model Summary section. Each quarter, States will be asked to indicate whether any major demonstration design changes have occurred since the previous quarter. A "yes" response will present the user with a form to indicate which elements have changed and to add a description of the change. Changes entered by States will not overwrite the original content so that we can maintain historical data. The RTI evaluation teams for each State will be notified of any changes identified by States so they can research and resolve these changes as appropriate. RTI evaluation teams will contact the States when necessary to follow up for more information. ***Table 10*** presents data elements for the Model Summary.

4.2.3.2 Demonstration Implementation

This component will be updated quarterly with data uploaded into the SDRS by the States. States will be directed to a set of instructions on how to enter data and update data entries. The instructions will include details on how implementation quarters are defined and what to do if particular information is missing. The link to the user guide and Help Desk contact information will also be provided.

Table 10
Model summary description elements

State name
CMS contract status
No. of models under the demonstration
Model name
Financial alignment model (capitated, managed FFS or other)
Geographic area
Implementation date (i.e., when the State demonstration starts serving beneficiaries)
Eligible population (including subpopulations of focus for the demonstration)
Populations not eligible to enroll
Enrollment targets or caps
Description of phased enrollment
New services offered
Carved-out services
Enrollment and disenrollment methods
Payment methodology used for the MMPs or MFFS providers
Contact information for State demonstration staff

FFS = fee-for-service; MFFS = managed fee for service; MMP = Medicare-Medicaid Plan.

The Demonstration Implementation section is divided into several subparts; Progress Indicators and Tracking Elements by Design Feature are the two main subsections. The Progress Indicators subsection will ask the State to enter numerical data in monthly increments. For example, if a State were making a quarterly SDRS update for a quarter running from January through March, that State would enter data for each month of that quarter (January, February, March). Data elements in the Progress Indicators subsection are presented in ***Table 11***. The Tracking Elements by Design Feature subsection asks the States to check off Yes/No responses and provide text descriptions for items related to the reporting period. Data elements for the Tracking Elements by Demonstration Design Feature subsection are presented in ***Table 12***.

4.2.3.3 Demonstration Impact and Outcomes

The data for this section will be updated quarterly by the RTI team; States will not need to submit data for these measures. This content will be generated by RTI from MSIS/Medicare fee-for-service (FFS) claims, Medicaid managed care organization and Medicare Advantage plan encounter data and Medicare-Medicaid Plan (MMP) encounter data; Nursing Home MDS analysis; and other data that may be available from other entities. Availability of the data will drive quantitative analyses and affect when specific analysis results will be reported. For MFFS demonstrations, we expect the first few quarterly reports to include predemonstration

data from all available sources, but demonstration data may be limited to Medicare claims and MDS data because of anticipated lag time for MSIS data (the source for Medicaid claims and any Medicaid managed care encounters) and for Medicare Advantage encounter data. For capitated model demonstrations, the content of the first few quarterly reports also will depend on the timeliness of MMP data. Results incorporating additional data sources will be included in the quarterly reports to CMS and the States as they become available. Generally, measures will be the same across all States implementing the same financial alignment model (e.g., capitation or managed FFS).

Table 11
SDRS data collection: Progress indicator elements

Indicator[1]
Eligibility
Total number of beneficiaries who are eligible to participate in the demonstration[2]
Enrollment
Total number of beneficiaries who are enrolled in the demonstration[2]
Number of beneficiaries who are newly enrolled in the demonstration[2]
Number of newly enrolled beneficiaries who were automatically (passively) enrolled in the demonstration[2]
Number of beneficiaries who opted out or chose not to enroll in the demonstration without ever being enrolled [2]
Disenrollment
Number of beneficiaries who voluntarily disenrolled (i.e., made a choice to disenroll) from the demonstration [2]
Number of demonstration enrollees whose eligibility for the demonstration ended involuntarily (e.g., died, moved out of area, lost Medicaid eligibility, were incarcerated)[2]
Demonstration service area
Whether demonstration is currently operating statewide vs. in specific counties or geographic areas (and provide list of counties served, if in specific geographic areas)
Specific to capitated model demonstrations Number of three-way contracts with MMPs
New CMS initiatives in the demonstration area that may affect Medicare-Medicaid enrollees
Specific to demonstrations that use health homes
Number of health homes participating in the demonstration
Number of enrollees served by health homes
Specific to demonstrations using medical homes
Number of medical homes participating in the demonstration
Number of enrollees served by medical homes

LTSS = long-term services and supports; MFFS = managed fee for service; MMP = Medicare-Medicaid Plan; NCQA = National Committee for Quality Assurance; SDRS = State Data Reporting System.

[1] All indicators may not apply to all States (e.g., for some MFFS States, beneficiaries who are eligible for the demonstration are the same as beneficiaries who are enrolled in the demonstration).

[2] Progress indicators that will be presented in quarterly reports to CMS and the States.

Table 12
SDRS data collection: Tracking elements

Demonstration design feature
Integrated delivery systems
New waiver authorities or State Plan Amendments related to the demonstration (pending, approved, and not approved)
New policies or administrative procedures for improving the integration of primary care, long-term services and supports (LTSS), and behavioral health services under the demonstration
Changes in reporting requirements for any of the entities involved in the demonstration
Integrated delivery systems supports
Training or capacity-building activities to build core competencies of demonstration MMPs/providers in serving demonstration populations
Activities to help primary care providers transform care delivery
Reports on performance to MMPs/providers
Policies and procedures
New State policies/guidelines regarding care coordination/case management, promoting the adoption of electronic health records, etc.
New procedures to track/report data on care coordination/case management
Benefits and services
New or expanded services/benefits for demonstration participants
Enrollment and access to care
Whether enrollment targets were set and met
Whether enrollment caps were set and met
Education and outreach activities among eligible beneficiaries who are not enrolled
Activities to increase beneficiary enrollment
Major issues and challenges implementing the demonstration and solutions developed
Enrollee engagement
Activities to engage stakeholders in policy development or oversight of the demonstration
Activities to engage enrollees, families, or advocates in policy development or oversight of the demonstration
Quality Management and Measurement
Modifications to quality management approach
Tracking of new quality indicators
Receiving data from MMPs/providers to support new quality indicators
Financing and payment
Changes in payment methodology for MMPs and providers
Data development
Timing of the most recent MSIS or T-MSIS data file submissions
New State survey or evaluation reports available
Whether MMPs experienced any problems submitting encounter data to CMS (for capitated models)
Successes related to the demonstration not covered by other questions
Successes related to the demonstration not covered by other questions

MMP = Medicare-Medicaid Plan; MSIS = Medicaid Statistical Information System; SDRS = State Data Reporting System; T-MSIS = Transformed Medicaid Statistical Information System.

Table 13 provides examples of the quality, utilization, access to care, and cost measures that RTI will calculate and enter into the Demonstration Impact and Outcomes section. A subset of these measures will be presented in quarterly reports to CMS and States as data are available.

4.2.4 System Reporting Capabilities

The SDRS will generate reports for internal use and quarterly reports for CMS and the individual States.

Reports for internal use by the RTI evaluation teams for each State will depict missing data, completeness of data, and timeliness of data report. These reports will highlight data elements that are either missing for any given quarter or are late for any given quarter. These reports will be run routinely after each quarterly submission period is closed. Evaluation team leads for each State will follow up with the appropriate State staff to obtain the data and check to see whether the data have been uploaded.

We will also develop data extraction and data export/download capabilities for internal use by the RTI evaluation teams for each State. At various times during this evaluation (such as the annual report periods, prior to site visits, and during the final evaluation phase), it will be useful for the evaluation team to download and extract some or all of a State's data elements for a specified period in order to upload such data into other software (such as SAS) that would be more appropriate or convenient for detailed analyses.

The State-specific and aggregate quarterly reports to CMS and the States will automatically be generated by the SDRS.

4.2.5 System Technical Development

The data elements collected in the SDRS will be aggregate-level data only. No beneficiary-level data will be entered or stored in the SDRS.

The database design allows each State to have its own independent set of data fields, which will provide flexibility to track common measures across States or financial alignment models. The system will store the entries for each data field from each quarterly submission.

The State tables (the term for each State's set of data elements) and the aggregate tables will be written in a Structured Query Language (SQL) database and will have a Microsoft.NET-based, web-enabled data entry front end, plus some basic reporting for monitoring purposes.

Table 13
Examples of demonstration impact and outcome measures for the SDRS

Quality measures	Utilization measures[1]	Access to care measures	Cost measures[2]
■ 30-day all-cause risk-standardized readmission rate ■ Ambulatory care sensitive condition admissions—overall composite (AHRQ PQI #90) ■ Ambulatory care sensitive condition admissions—chronic composite (AHRQ PQI #92) ■ Preventable ED visits ■ Follow-up after hospitalization for mental illness ■ Cardiac rehabilitation following hospitalization for AMI, angina, CABG, PCI, CVA ■ Percent of high-risk residents with pressure ulcers (long stay)	■ Inpatient acute ■ Inpatient psychiatric ■ Emergency room ■ Skilled nursing facility ■ Nursing facility (long stay) ■ Outpatient (primary care) ■ Outpatient behavioral health (mental health; AOD) ■ Home health ■ State Plan personal care ■ Waiver personal care ■ Other HCBS (home health, other waiver services)	■ Number of primary care visits ■ Acuity on admission to nursing facilities ■ Discharge rate back to the community from nursing facilities ■ Any outpatient utilization for severe and persistent mental health conditions ■ Any substance use disorder treatment	■ Average cost PMPM (Medicare, Medicaid, and total)[3] ■ Inpatient acute ■ Inpatient psychiatric ■ Emergency room ■ Skilled nursing facility ■ Nursing facility (long stay) ■ Outpatient ■ Outpatient behavioral health (severe and persistent mental illness; AOD) ■ Home health ■ HCBS (Home health, State Plan personal care; waiver personal care; other waiver services) ■ Durable medical equipment

AMI = acute myocardial infarction; AOD = alcohol or other drugs; CABG = coronary artery bypass graft; CVA = cerebrovascular accident; ED = emergency department; HCBS = home and community-based services; MDS = minimum data set; PCI = percutaneous coronary intervention; PMPM = per member per month; SDRS = State Data Reporting System.

NOTE: These measures will be calculated by RTI using data provided by CMS.

[1] Each utilization measure will be presented using rates per 1,000 eligible beneficiaries and users as a percentage of eligible beneficiaries.

[2] Each cost measure will be presented as dollars per beneficiary per month.

[3] These are the only cost measures presented for capitated models; all cost measures will be presented for managed FFS models.

To create a streamlined reporting system, each State will have a dedicated, State-specific simple data entry screen accessible through a web interface. The only requirement for the States to complete quarterly entries will be access to a web browser. The data entry form will have field validation rules (format, allowable range, etc.), and inconsistencies will be flagged. However, it is not possible to develop data checks for open-ended fields. These fields will have a limit only on the number of characters that is possible to enter, developed according to the amount of information anticipated in that field. The RTI evaluation teams for each State will examine all entries and contact the State if follow-up is required.

States will enter data into the SDRS manually. Given that States will report a relatively small number of data elements per quarter, manual entry of the aggregate numbers by the State should not be excessively burdensome. Further, based on our preparatory work with demonstration States to date, we anticipate that most or all participating States already have substantial data analysis capability to participate in this evaluation. Many States reported having software, integrated data warehouses, or consultant services that presumably could generate many of the desired quarterly aggregate numbers relatively easily. However, the completeness of quarterly reports to CMS and the States will ultimately depend on data availability.

The data uploaded by RTI for the Demonstration Impact and Outcomes will use a batch upload protocol for each State table. RTI would send or upload a tab-delimited text file (or even an Excel file) that would extract each data element from the batch file and populate the appropriate fields in each State's table for the quarter.

Access to the SQL database (and to customized data entry screens and to State-specific tables) will be controlled through basic user authentication schemes (e.g., user names, passwords, access roles) assigned to the appropriate designated key staff in each State. In addition, RTI evaluation team members will have role-based log-on access to the State tables and to the higher-level aggregated tables. Because there will be no beneficiary-level data in any of these tables, and thus, no Personally Identifiable Information (PII), this database will be classified as low security risk, for which we will employ all appropriate and reasonable controls and precautions.

States will be able to access the State tables for Progress Indicators to upload data previously missing from their quarterly data submissions. However, States will need to notify RTI evaluation team leads for their State if corrections are required for data that have already been submitted. Evaluation team leads for each State will have permission to edit, update, and refresh data from previous quarters. Controls will be employed to prevent any data elements from being deleted. Audit trails will record who has made edits to the data. Daily backups will be maintained.

4.3 Training and Orientation for States

We have prepared training materials and a user guide, and will conduct training via webinars to ensure that State staff understand the SDRS and the expectations for data entry. We plan to offer webinar training dates scheduled closely following after the State demonstration implementation dates to orient State staff, make sure they can log in, and understand the screens and what their role is in entering data. For efficiency, we will train as many States as possible simultaneously. The total number of training sessions will depend on how many States implement the demonstration around the same time. Part of the training/orientation will be devoted to defining terms, so we have consistency among all States' entries. RTI is also developing a glossary of definitions to be included in the user guide. All terms will be defined on a screen with a mouse-over function and will also provide a link to the electronic version of the user guide. The recorded training webinar will be available for review as needed.

4.4 Timeline

The content of the Model Summary and Demonstration Implementation sections of the system have been developed and are in the process of being finalized. We have developed screenshots of the user interface. We have conducted cognitive testing of the Demonstration Implementation section with four States and have made edits to the items based on the feedback received.

The Demonstration Impact and Outcomes section is currently in the final stages of development. Given the iterative nature of the work, developing the data capture and reporting system with the database setup requires several months: a few months for database design (in process), a few months for implementation of the data capture, and additional months to set up reporting.

As part of the report development process, mock quarterly data will be entered into the database to ensure that we are collecting the necessary data to produce the quarterly reports to CMS and the States and the annual reports. At this time, we will also develop database structures to maintain aggregate data, as well as develop and test the processes for importing and aggregating common data elements from the individual State sources.

We anticipate potential modifications to the data elements captured. We also anticipate corresponding modifications to the user interfaces that support the submission of these data elements. The database and user interface are being designed to accommodate such potential changes. We will also make corresponding updates to reports, should the data elements change.

The activities to develop and maintain the SDRS may be revised if modifications are made to the demonstrations or if data sources are not available as anticipated. If modifications to this evaluation plan are required, they will be documented in the annual and final evaluation reports as appropriate.

Section 5. Quantitative Analyses

5.1 Introduction

This section of the report outlines the research design, the data sources, the analytic methods, and the outcome variables (quality, utilization, and cost measures) that the RTI team will focus on in evaluating the impact of State demonstrations. The individual analysis plans for each demonstration outcome follow in subsequent sections. These analyses will be conducted for individual States, and the final evaluation report will also include an aggregate analysis to learn more about the effects of different State demonstration design features on quality, utilization, and cost. This section also addresses differences in the analytic approach required for managed fee-for-service (MFFS) States versus capitated model States in terms of data requirements, analytic issues, and outcome variables. We discuss the approach to identifying the demonstration group in each State as well as for identifying comparison group beneficiaries. State-specific details on identifying demonstration and comparison groups can be found in the State-specific evaluation design reports.

RTI and its subcontractor Actuarial Research Corporation (ARC) are also developing the methodology for calculating annual Medicare and Medicaid cost savings for MFFS States. This methodology has been presented to CMS in a separate memorandum. Annual MFFS State savings will be calculated using an actuarial method, whereas the evaluation analyses will use a regression-based approach. The assumptions underlying the two methods will be as consistent as possible. These issues are addressed in the memorandum describing the annual actuarial savings calculation methodology for MFFS States.

The evaluation analyses include both descriptive and multivariate analyses to learn more about differences in key outcome variables within States over time and across States controlling for beneficiary and market characteristics. The analyses discussed in this section will become part of quarterly reports to CMS and States, annual State reports, and the final evaluation report.

5.2 Research Design—Major Considerations

5.2.1 Defining the Evaluation Framework

The RTI team will use an intent-to-treat (ITT) approach for the quantitative analyses conducted for this evaluation, comparing the eligible population under each State's demonstration with a similar population that is not affected by the demonstration (i.e., a comparison group). In this section, we discuss the intent-to-treat approach and our rationale for this decision.

ITT refers to an evaluation design in which all Medicare-Medicaid enrollees eligible for the demonstration constitute the evaluation sample, regardless of whether they actively participated in demonstration models. We recommend this design because we believe that CMS's primary interest lies in the effect of the demonstrations on all beneficiaries in the demonstration-eligible population—not just the effects on those who enroll or engage in the care model. Thus, under the ITT framework, outcome analyses include all beneficiaries eligible for the demonstration in the demonstration States, including those who opt out, or participate but then disenroll; are eligible but are not contacted by the State or participating providers to enroll in the demonstration or care model; and those who enroll but do not engage with the care model; and a group of similar individuals in the comparison group (discussed further below).

5.2.2 Identifying Demonstration Group Members

Under the ITT design, we must be able to identify all members of the population eligible for the demonstration, regardless of their enrollment status in the demonstration. That is, data are needed on all of the Medicare-Medicaid enrollees who would be eligible for enrollment in the demonstration design. Thus, identifying the eligible population for each State's demonstration will require information on the State's eligibility criteria and data on the characteristics of Medicare-Medicaid enrollees in the State on those dimensions (e.g., age, presence of chronic conditions, geographic location). RTI also will use this information in developing the comparison groups for each demonstration.

To identify the population eligible for a State's demonstration, the evaluation will need from each State a "finder" file that includes Medicare Health Insurance Claim Numbers (HICNs), Medicaid IDs, and other identifiers for all Medicare-Medicaid enrollees in each State's eligible population. Obtaining HICNs and Medicaid IDs from States will be necessary throughout the demonstration to identify beneficiaries newly eligible for the demonstration and those beneficiaries who have left the demonstration or are no longer eligible for the demonstration. Both MFFS and capitated model States will submit demonstration evaluation (finder) files to RTI on a quarterly basis, and will include information on enrolled beneficiaries as well as all beneficiaries eligible for the demonstration within their States. This information will be used to identify the characteristics of each demonstration State's eligible beneficiaries for RTI's use in selecting an appropriate comparison group as described in the sections that follow.

More details on the demonstration evaluation (finder) file are in ***Section 5.5, Data Sources*** below.

5.2.3 Identifying Comparison Group Members

In our evaluation design, the comparison group provides an estimate of what would have happened to the demonstration group in the absence of the demonstration—known as the

counterfactual. Thus, the comparison group members should be similar to the demonstration group members in terms of their characteristics and health care and long-term services and supports (LTSS) needs, and should reside in areas that are similar to the demonstration State in terms of the health care system and the larger environment. For this evaluation, identifying the comparison group members will entail two steps: (1) selecting the geographic area from which the comparison group will be drawn and (2) identifying the individuals who will be included in the comparison group.

The comparison groups will be used both for the evaluation and annual actuarial MFFS cost savings calculations. The comparison groups will be refreshed annually to incorporate new entrants into the eligible population as new individuals become eligible for the demonstration over time.

5.2.3.1 Determining the Geographic Basis for the Comparison

a. In-State Comparison Groups

Given the significant differences in health care systems across States, including significant differences in State Medicaid programs, our priority will be to determine whether an in-State comparison group is possible for each demonstration State. An in-State comparison group will only be a potential option in States that are implementing their demonstration in a subset of areas in the State, rather than statewide, and where the areas included in the demonstration and the areas excluded from the demonstration are similar.

In assessing the feasibility of an in-State comparison group for a demonstration State, we will compare the characteristics of the areas included in the demonstration and those not included in the demonstration on a range of population characteristics (e.g., age, income, racial/ethnic mix) and market characteristics (e.g., provider supply), as well as size of the population that meets the criteria for the eligible population under the demonstration. We will use a combination of Metropolitan Statistical Area (MSAs) and State boundaries to define geographic regions. Delineated by the Office of Management and Budget based on the decennial census, MSAs comprise one or more geographically contiguous counties linked to, and identified by, at least one core city. Based on the 2013 delineation, there were 381 MSAs in the United States. Many of these areas span more than one State, and in these cases, for analytic purposes we will divide MSAs into multiple areas along State boundaries. In each State, the counties that are not delineated into an MSA will be combined into one "rest-of-State" analytic area. Using these geographic grouping rules results in approximately 450 analytic areas in the United States. In demonstration States, we would remove from these analytic areas any counties in which the demonstration is offered. If the characteristics and population size for the potential comparison areas in the State are similar, then we will move forward with developing an in-State comparison group.

However, if a demonstration is statewide, or if we are unable to identify potential in-State comparison areas that are comparable to the demonstration counties and contain sufficient numbers of beneficiaries, we would consider an out-of-State comparison group or, potentially, a comparison group that includes both in-State and out-of-State areas.

b. Out-of-State Comparison Groups

If the demonstration is statewide, or if the excluded areas of a State are not representative of the demonstration areas, we will construct a comparison group from out-of-State areas or possibly a combination of in-State and out-of-State areas. We will limit the pool of potential comparison areas to States without a Financial Alignment Demonstration. We will compare demonstration and comparison areas on a range of measures, including spending per Medicare-Medicaid enrollee by each program, the shares of LTSS delivered in facility-based and community settings, and the extent of Medicare and Medicaid managed care penetration. Using statistical analysis, described below, we will select the individual comparison MSAs that most closely match the values found in the demonstration area. RTI and CMS may also consider other factors when selecting comparison States, such as timeliness of MSIS and/or encounter data submission.

In general, we expect to draw out-of-State comparison groups using analytic areas from multiple comparison States. The number of areas to be included in the comparison group will be determined by the closeness of the match with the demonstration State (or demonstration areas within the State) and the size of the Medicare-Medicaid enrollee population in the comparison area. The goal will be to identify a comparison group at least as large as the eligible population in the demonstration State. This will ensure a sufficient sample to support sensitivity analyses around the choice of comparison groups (discussed below). The first annual report will document decision rules for choosing the comparison area.

c. Statistical Distance Analysis Methodology

To identify comparison areas, the RTI team will conduct a statistical distance analysis to assess the similarity of a demonstration State with each of its potential comparison areas. The process entails the following three steps:

Step 1. **Identify characteristics that will be used to compare demonstration and comparison areas**

The first step is to identify characteristics reflecting State-level policies prior to the demonstration hypothesized to affect the outcomes of interest. The characteristics under consideration include the following combination of Medicare and Medicaid variables:

- Medicare spending per Medicare-Medicaid enrollee
- Medicaid spending per Medicare-Medicaid enrollee
- Nursing facility users per Medicaid beneficiary age 65 and over

- Home and community-based services (HCBS) users per Medicaid beneficiary age 65 and over
- Personal Care users per Medicaid beneficiary age 65 and over
- Medicare Advantage penetration
- Medicaid managed care penetration per full-benefit Medicare-Medicaid enrollee

The three LTSS variables capture how States differ in the settings in which they provide these services. Variation in LTSS State policy is most easily visible in the population using the most LTSS (i.e., those aged 65 and over). The relative importance of facility-based care observed in that population is expected to affect such use in the population under age 65 as well.

Step 2. **Compute distance scores for each demonstration area and potential comparison MSA**

To measure the similarity between the demonstration population and a potential comparison area, we will compute the squared Euclidean distance between their characteristics. This score is calculated by (1) converting the value of each State/ area characteristic to a standard score (mean = 0, standard deviation = 1), (2) subtracting the comparison area's standard score from the demonstration score, (3) squaring the score difference, and (4) summing the differences across all characteristics.

These scores will be computed for every demonstration State and each potential comparison area. The smaller the distance score, the more similar the States are with respect to the selected State-level characteristics. The lowest possible score is zero for two areas with identical values on all characteristics. The final analysis will be run with the most recent data available when the final list of demonstration States is confirmed.

Step 3. **Select comparison areas**

Distance scores will be sorted by magnitude to identify the comparison areas with the smallest scores. The number of areas to be selected will depend on their combined population of Medicare-Medicaid enrollees. Any given area may serve as a comparison for more than one demonstration.

In addition to using distance scores to identify the most appropriate comparison States, we will need to consider other factors, including the timeliness of a potential comparison State's Medicaid Statistical Information System (MSIS) submissions. Therefore, the cluster analysis results will be a starting point for CMS to consider in making a final decision on the most appropriate comparison States (or areas of States) for States that require an out-of-State comparison group.

Once the comparison areas have been selected, all Medicare-Medicaid enrollees who meet the eligibility criteria for the demonstration in those States or areas would serve as potential members of the comparison group. For demonstration States that identify the eligible population using claims-based criteria, such as age and diagnosis, the comparison group will be identified from administrative data sources based on those criteria.

For MFFS demonstration States that include individuals in the eligible population based on other criteria for which no reliable administrative data are available to the evaluation team (e.g., including high Body Mass Index or smoking status as criteria for Medicaid health home as well as demonstration eligibility), we will consider two comparison group strategies. If the subset of beneficiaries eligible for the demonstration based on nonadministratively verifiable information is relatively small, we will exclude this subset of participants from both the demonstration and comparison groups. If, however, the subset of such participants is substantial, the demonstration group members will be identified as all Medicare-Medicaid enrollees who are potentially eligible for the demonstration, so that it is possible to identify similar Medicare-Medicaid enrollees in a comparison group. For this reason, States are strongly encouraged to adopt demonstration eligibility criteria that are reflected in administrative data, and are available for identifying a comparison population.

5.2.3.2 Selecting Individuals for the Comparison Group

Regardless of whether the comparison group is to be in-State or out-of-State, the final step is to identify individuals within the comparison geographic areas and create analysis weights for them. We will accomplish this by estimating propensity scores and weighting comparison-group beneficiaries to look like the eligible population in the demonstration State.

We plan to use all available comparison-group members in a propensity analysis rather than attempting to construct one-to-one matches. In the context of this evaluation, the propensity score is an estimate of the probability that a beneficiary is in the demonstration group conditional on a set of observed characteristics. To compute the propensity scores, we will first identify beneficiary-level and MSA-level characteristics to serve as covariates in the propensity-score model. Example characteristics include the following:

- Beneficiary characteristics such as age, sex, MSIS eligibility information on socioeconomic status, prior Medicare and Medicaid expenditures, LTSS/HCBS, hierarchical condition category (HCC) risk scores, and end stage renal disease (ESRD) status, among others. These data will be obtained from unlinked Medicare and Medicaid files and will include encounter data or per member per month (PMPM) payments where appropriate and available.
- MSA-level characteristics from Census Bureau databases, and the Area Resource File (ARF) such as health care providers/100,000 population, morbidity/mortality, and urbanicity.
- State-level policy factors, such as the proportion of long-term services and supports spending that is for HCBS (rather than for facility-based care), Medicaid nursing facility eligibility criteria, and implementation of Health Home State Plan Amendment (HH SPA) (except for within-State evaluations, for which all beneficiaries are drawn from the same State).

Next, we will combine the beneficiaries eligible for the State's demonstration with the eligible comparison pool beneficiaries from the within-State counties or out-of-State

comparison States or areas. We will estimate a logistic model by regressing group status (demonstration vs. comparison pool) on the set of individual and area characteristics to determine the propensity scores for demonstration and comparison group beneficiaries. To ensure that the comparison group is similar to the demonstration group, we will remove from the comparison group any beneficiaries with a propensity score lower than the lowest score found in the demonstration group. This is known as being "outside common support."

The final step in the process is to create propensity-score weights for the comparison group. The weights for the demonstration group beneficiaries are all set to 1.0 because they have been assigned to the demonstration group. For beneficiaries in the comparison group, the propensity-score weight is PS/(1-PS), where PS is the individual's estimated propensity score. Comparison-group weights will be normalized to reflect the actual size of the comparison group sample. By weighting comparison group members by their predicted propensity score, the demonstration and comparison group samples will be more balanced, and the distribution of characteristics of the two groups will be similar to each other. A common practice is identifying a comparison group at least as large as the demonstration group.

The comparison areas will be determined within the first year of implementation of each demonstration, in order to use the timeliest data available. The comparison group members will be determined retrospectively at the end of each demonstration year, allowing us to include information on individuals newly eligible or ineligible for the demonstration during that year. The groups will be refreshed annually to incorporate individuals who become eligible for the demonstration over time. This schedule will accommodate any phased-in enrollments by a State.

5.2.3.3 Additional Considerations in Identifying Comparison Groups

In addition to the challenges inherent in selecting appropriate comparison groups for each demonstration, we list four additional issues that need to be considered in identifying comparison groups.

1. Identifying similar demonstration and comparison groups requires detailed Medicaid utilization and expenditure data from the predemonstration period for both the demonstration and comparison groups. Although we intend to collect at least 2 years (8 quarters) of baseline data to improve the precision of estimated demonstration effects, MSIS data quality and lags may limit the availability of the required data.
2. Some Medicare-Medicaid enrollees may not be suitable to be comparison group members if they are participating in other similar CMS programs or demonstrations operating in the comparison area. Accurate, timely data to identify whether individuals or counties must be excluded from the comparison group will be important to correctly identify suitable comparison group members.
3. Specific demonstration design criteria used by States should also be available for comparison group identification. State demonstrations that are limited to certain populations or service use require comparison groups with similar attributes.

Identification of comparison group members is subject to having the means to do so through claims, encounter data, or State-provided enrollment and eligibility information. Having data on these characteristics for both the demonstration and potential comparison groups is needed to select the comparison group. However, if a demonstration identifies its eligible population using criteria that cannot be defined using administrative data, such as having an excessive Body Mass Index or being a tobacco user, it will not be possible to identify individuals in comparison regions or States with those same characteristics.

4. Sample size is an important consideration in that it affects the precision of the estimates that can be obtained from the evaluation. Medicare-Medicaid enrollees typically have very high annual expenditures, and standard deviations that may be twice as high as the annual average. This may make it difficult to detect small demonstration group expenditure differences given that confidence intervals will be potentially wider than for non–Medicare-Medicaid enrollee populations. In addition, small sample sizes are likely to be a particular problem for some subgroup analyses, such as Medicare-Medicaid enrollees with rarer diagnoses, and for cases in which the expected demonstration effect is small.

5.3 Relationship of the State Demonstrations to Other CMS Demonstrations/Activities

The Financial Alignment Demonstrations will be implemented during a period when several other CMS demonstrations and initiatives are occurring. Prior to and during implementation, the evaluation will monitor what other CMS initiatives are occurring in the demonstration and comparison areas as well as whether there is overlap of implementation timelines. These initiatives may include the Multi-payer Advanced Primary Care Practice Demonstration (MAPCP), accountable care organizations (ACOs), the Comprehensive Primary Care Initiative (CPCi), and HH SPA activities. The RTI team will use the CMS Master Data Management (MDM) system as appropriate to determine demonstration and comparison group beneficiaries who are part of other CMS initiatives. As part of the evaluation framework in the last year of the evaluation, the RTI team will work to identify demonstration effects beyond those resulting from other demonstrations and programs, as appropriate.

5.4 Addressing Phase-In of New Cohorts and Rolling Enrollment of New Medicare-Medicaid Enrollees

In many demonstrations, enrollment occurs at the beginning of the demonstration. In this demonstration, States are planning two additional types of enrollment that the evaluation will need to take into account: (1) phase-in of new cohorts during a limited window of time after demonstration start, and (2) rolling enrollment of individuals who are new Medicare-Medicaid enrollees (e.g., Medicaid beneficiaries who age in to Medicare or Medicare beneficiaries who become eligible for Medicaid due to disability or low income). New cohorts may have the same or different characteristics from those of the initial cohort. For example, the new cohort may differ in that it has different eligibility or risk characteristics or is from an area that has much less managed care than earlier cohorts. We understand that these

new cohorts might enroll all at once or their enrollment may be staggered. Such cohorts will enter the evaluation at the beginning of a month to simplify pulling claims data. In the initial ramp-up period, Medicare-Medicaid Plans (MMPs) participating in some capitated model demonstrations may have limits on their new enrollments each month; we will also need to adjust for such a situation. In addition, some individuals will leave the demonstration each month through death, moving, or electing other insurance coverage.

In ***Table 14***, we have identified issues related to the planned phase-in of beneficiaries in some States, as well as rolling enrollment or exit from the demonstration and identified solutions.

Table 14
Issues and solutions for phase-in and rolling enrollment

Issues	Solutions
Differences in the observed number of months of demonstration exposure for each beneficiary	We will control for the number of months of exposure to a demonstration for both individual cohorts phased in over time, as well as for individuals who enter the demonstration later or leave for any reason before it ends, by weighting observations by the proportion of possible exposure time.
Comparison group selection	If the RTI team can replicate the eligibility criteria for each cohort of phased-in enrollment, then the same criteria can be implemented for identifying the comparison group. Examples of such criteria include geography, eligibility or risk characteristics, or random selection methodology.[1] Each demonstration cohort will have a comparison group for the same time period. Eligibility definitions for each State should be defined monthly.
Potential for small sample sizes	If the cohorts that are phased in are small, then analyses may not be able to detect small effects between the demonstration and comparison groups. On the other hand, if the demonstration effect on a particular cohort (e.g., decreased costs for persons with HIV/AIDS) is substantially greater than the average effect across cohorts, we may be more likely to detect it statistically in that cohort, where it is not diluted by the effects for other cohorts.

[1] It will be very important for cohorts to be identifiable using data that are available in comparison States as well as in the demonstration State.

5.5 Data Sources

The RTI team will obtain all Medicare and Medicaid data (eligibility, claims, and MMP and other encounter data) from CMS. To identify beneficiaries eligible for the demonstration, the States will need to provide data to RTI or another contractor.

5.5.1 Demonstration Evaluation (Finder) File of Medicare-Medicaid Enrollees with Eligibility, Enrollment, and Identifying Information

On a monthly or quarterly basis, States will submit a file that includes data elements needed for RTI to correctly identify Medicare-Medicaid enrollees for linking to Medicare and

Medicaid data, and information about whether the enrollees were eligible for and/or enrolled in a Financial Alignment Demonstration (***Table 15***). The file will list all of the Medicare-Medicaid eligible population for a State's demonstration, with additional variables in the file indicating monthly participation in the demonstration. For example, the file will have monthly participation variables with a 1/0 entry in each monthly variable. Eligible individuals who were not enrolled in the demonstration in a given month will still be part of the evaluation under the intent-to-treat research design. RTI will provide further information to States on the file format, method of transfer, and timeline for submitting these data.

Table 15
State demonstration evaluation (finder) file data fields

Data field	Length	Format	Valid value	Description
Medicare Beneficiary Claim Account Number (Health Insurance Claim Number [HICN])	11	CHAR	Alphanumeric	The HICN. Any Railroad Retirement Board (RRB) numbers should be converted to the HICN number prior to submission to the MDM.
MSIS number	20	CHAR	Alphanumeric	MSIS identification number.
Social security number (SSN)	9	CHAR	Numeric	Individual's SSN.
Sex	1	CHAR	Alphanumeric	Sex of beneficiary (1=male or 2=female).
Person first name	30	CHAR	Alphanumeric	The first name or given name of the beneficiary.
Person last name	40	CHAR	Alphanumeric	The last name or surname of the beneficiary.
Person birth date	8	CHAR	CCYYMMDD	The date of birth (DOB) of the beneficiary.
Person ZIP code	9	CHAR	Numeric	9-digit ZIP code.
Monthly facility status	1	CHAR	Alphanumeric	Each monthly flag variable would be coded 1 if in nursing facility, 0 if not.
Monthly HCBS waiver status	1	CHAR	Alphanumeric	Each monthly flag variable would be coded 1 if enrolled in HCBS waiver, 0 if not.
Eligibility identification flag	1	CHAR	Numeric	Coded 0 if identified as not eligible for the demonstration, 1 if identified as eligible from administrative data, 2 if identified as eligible from non-administrative data.
Monthly enrollment indicator	1	CHAR	Numeric	Each monthly enrollment flag variable would be coded 1 if enrolled, and zero if not. Quarterly demonstration evaluation (finder) files would have 3 such data fields; annual demonstration evaluation (finder) files would have 12 such data fields.

HCBS = home and community-based services; MDM = Master Data Management; MSIS = Medicaid Statistical Information System.

5.5.2 Medicare Fee-for-Service Data

The evaluation team will use Medicare enrollment data (i.e., the denominator file), claims from TAP files, and Nursing Home Minimum Data Set (MDS) data to create SAS or STATA data sets for programmers to analyze for the predemonstration period and during the demonstration, where appropriate. The Medicare data files will be linked using the beneficiary identifier to create a beneficiary-level file with summary variables on Medicare utilization and payment by service type (e.g., inpatient, skilled nursing, home health). By using the Medicare TAP files, monitoring results are expected to be available approximately 9 months following the beginning of any performance quarter. For example, for the period October through December 2013, Medicare cost and utilization results could be made available in July 2014. This timeline presumes that the State demonstration evaluation (finder) files will be obtained soon after the end of each performance quarter. Delays in obtaining the demonstration evaluation (finder) files for demonstration-eligible beneficiaries will delay analysis of claims and encounter data and our ability to report that information in quarterly reports. RTI has acquired all necessary data use agreements to access the needed Medicare fee-for-service data from CMS.

5.5.3 Medicaid Data

The RTI team has been working with CMS to review and assess each State's Medicaid data quality and submission timeliness. This assessment has been needed to understand what data will be available to satisfy two evaluation components: (1) data available on a rapid-cycle basis for the initial 6-month report and subsequent quarterly reports that will update CMS and demonstration States on the progress of each State demonstration, and (2) data available within the time necessary to conduct descriptive analyses for the annual reports and for the overall final evaluation of the demonstration.

The current time lag for Medicaid claims data is considerable and varies by State. For example, as of August 2013, about one-third of the potential demonstration and comparison group States had not submitted MSIS claims data for fiscal year 2012, Quarter 2, meaning that they were more than a year behind. For the rapid-cycle component of the evaluation to be successful, and comparison group selection and cost-savings analyses to be timely, MMPs participating in the demonstration will have to submit claims and encounter data to CMS in a timely manner. Ideally, the evaluation should be able to access finalized quarterly MSIS data within 4 to 6 months after the end of a quarter, as required in the MOUs between CMS and demonstration States.

Several potential demonstration States are moving to the t-MSIS system, so they will submit monthly t-MSIS data within 1 to 2 months of each claim month. Although the t-MSIS system will improve the timeliness of data submission during the demonstration period, the evaluation will also need to obtain data from these States for the full 2-year predemonstration period. If the t-MSIS demonstration States were behind on MSIS

submissions before they began submitting t-MSIS format files, obtaining complete predemonstration data could be difficult. RTI will incorporate Medicaid data in each annual report as they become available and will also include Medicaid data in the final evaluation report. The RTI team will obtain these data as they become available and will create a single analytic file that will be separable by year, to measure annual impact for each State during our analyses for the final evaluation report.

Medicaid claims data will be used to construct service use patterns, particularly for services not covered by Medicare—notably, facility-based long-term care, HCBS waiver services, and behavioral health services in the predemonstration period and during the demonstration, where appropriate. As with other medical claim files, depending upon the particular type of claim, the MSIS contains, among other things, information on the date of service (or beginning and end dates for facility-based services), principal diagnosis and up to nine other diagnosis codes, and principal procedure code.

To understand the services that enrollees are receiving, the evaluation will need to classify services consistently over time and, where possible, across States. In States using a MFFS model, claims data will be used in both the pre- and postperiods. In States moving from a Medicaid fee-for-service model to a capitated model, the claims would be used to construct predemonstration measures and for individuals in the eligible population who are not enrolled in Medicare-Medicaid plans whereas encounter data will be used to construct "dummy" claims in the postdemonstration period. For those States that already have Medicare-Medicaid enrollees in capitated Medicaid plans, encounter data for nondemonstration Medicaid plans and captured in the MSIS will be used for both the pre- and postperiods.

5.5.4 Medicare and Medicaid Encounter Data

Our goal is to construct high-quality claims and encounter analytic files using data obtained from CMS. Essential to the evaluation is information on patient diagnosis, service intensity (brief vs. comprehensive visits), type of visit (preventive vs. treatment), ancillary services, and facility charges. Encounter data will be needed from three sources: (1) Medicare Advantage plans, (2) nondemonstration-related Medicaid managed care plans, and (3) the MMPs in capitated model demonstration States. Because the evaluation uses an intent-to-treat design, data from both Medicare Advantage plans and nondemonstration Medicaid managed care plans will be needed for both capitated and MFFS model demonstrations, for the comparison States. These data will be needed for both the predemonstration and demonstration periods.

CMS provided guidance to Medicare Advantage plans to begin submitting encounter data for 2012 to CMS. Our initial review of part of this documentation indicates that Medicare Advantage plans will submit encounter data for all types of services to a single encounter data front-end system contractor, Palmetto GBA. The quality of the data is not yet known.

Given this relatively new reporting requirement, the RTI team may not have 2 complete years of predemonstration Medicare Advantage encounter data for constructing the baseline in some demonstration States and comparison States. Also, although Medicare Advantage plans will be required to submit only adjudicated encounter data monthly, plans have up to 12 months from the date of service to submit such data, so the RTI team may face the same data lag in acquiring Medicare Advantage encounter data directly from CMS as it faces in obtaining validated MSIS data for Medicaid claims. Any validation process implemented by CMS for Medicare Advantage data will also have implications for timely access to encounter data.

Medicaid managed care data for nondemonstration plans available through MSIS or t-MSIS will also be needed for both the predemonstration and demonstration periods for any demonstration or comparison group members enrolled in such plans. In the early days of Medicaid managed care, encounter data reported to CMS through the MSIS system tended to be incomplete because payment did not hinge on documenting services provided. However, as capitated models have become more common, Medicaid programs rely more heavily on this type of data to set payment rates, and many States have increased their requirements for participating plans to submit high-quality encounter data files. Reporting requirements for Medicaid managed care plans (not part of the demonstration) are not currently known, and it is expected that what is available through MSIS would vary by State. The RTI team will work with CMS to understand these requirements.

CMS will be collecting encounter data directly from the MMPs. The processes currently being established for MMP encounter data collection hold the promise of uniformity in format for such data and the means for the RTI team to have a single point of contact (with CMS) for obtaining encounter data from these plans. Also, CMS's plans to include flags indicating whether encounters represent Medicare or Medicaid-covered services will be valuable in evaluating utilization patterns in capitated model States. During 2013, the RTI team is monitoring the development of these requirements.

In addition to the Medicare and Medicaid encounter data, RTI will also obtain CMS data on prescription drug PMPM payments for beneficiaries from the monthly plan payment files at CMS and potentially Part D reconciliation costs directly from the CMS payment group to support analysis of Part D costs.

Data comparability across all of these data types will need to be carefully considered. For example, a Medicare-Medicaid enrollee may opt out of a capitated model demonstration but remain in a nondemonstration Medicaid managed care plan, and encounter data from both sources will need to be made comparable using an algorithm in order to assess utilization. Data comparability across demonstration and comparison States will also need to be carefully considered.

5.5.5 Data Preparation and Linking

We will obtain Medicare TAP data from CMS on a monthly basis. After initially checking the data for problems, we will clean the data quarterly for use in monitoring. Most of the intensive cleaning activities for the project should be completed through this effort unless variable formats change and new analytic measures need to be created (e.g., changing from ICD-9 to ICD-10 diagnosis codes). We will update these data each quarter for a 6-month run-out period.

We will obtain MSIS and t-MSIS data from CMS on a quarterly basis, as they become available. As discussed earlier, there are lags in obtaining Medicaid data, which will affect how soon we will be able to analyze Medicaid data and include in reports.

We will obtain encounter data for MMPs and Medicare Advantage from CMS as well, on a schedule to be confirmed once the processes for collecting and validating these data are finalized.

After each demonstration has ended, we will link Medicare and Medicaid data at the beneficiary level to do the impact evaluation analyses after each States' data are submitted to CMS, thereby linking all demonstration years of Medicare and Medicaid claim and encounter data at once. We would then conduct analyses for the individual State final reports as well as the aggregate analyses using claims and encounter data with sufficient run-outs to provide results as complete and unbiased as possible.

RTI and its subcontractors, the Urban Institute and Actuarial Research Corporation, will be working directly with data, so they will obtain approval for acquisition and analysis of these data from their respective Institutional Review Boards. All contractors will maintain these data in secure environments, as is the custom when working with CMS data containing data governed by the Health Insurance Portability and Accountability Act (HIPAA).

The activities to identify demonstration and comparison groups and to collect and utilize claims and encounter data may be revised if modifications are made to the demonstrations or if data sources are not available as anticipated. If modifications to this evaluation plan are required, they will be documented in the annual and final evaluation reports as appropriate.

Section 6. Analysis Overview

The quantitative components of the evaluation will consist of (1) monitoring individual States to track quarterly changes in selected beneficiary experience, quality, utilization, access to care, and cost measures for the demonstration group using pre- and postperiod data over the course of the demonstration using the State Data Reporting System (SDRS); (2) conducting individual-State descriptive analyses on quality, utilization, access to care, and cost measures for both the demonstration and comparison groups for annual reports; and (3) evaluating the impact of individual State demonstrations using multivariate regression with a comparison group as well as aggregate analyses for the demonstration as a whole for the final evaluation report. The approach to each of these analyses is outlined below. Additional detail regarding quality measures is provided in ***Section 8***.

Our ability to conduct the analyses described in this section on a timely basis will be affected by potential lag time in receiving Medicare-Medicaid Plan (MMP) encounter and Medicaid data (as discussed in ***Section 5***). We will also need to account for changes that may occur in States unrelated to the Financial Alignment Demonstrations, such as the effects of other demonstrations or other State-specific policy changes.

6.1 Monitoring Analyses

We will analyze available Medicare and Medicaid data each quarter to calculate means, counts, and proportions on selected quality, utilization, access to care, and cost measures to include in reports generated by the SDRS for CMS and the States. We will also analyze available Nursing Home Minimum Data Set (MDS) data to calculate facility admission rates. Generally, these measures will be for data elements that can be obtained across all States, or across all States having the same financial alignment arrangement (e.g., capitation). We will present the current values for each predemonstration and demonstration period quarter for each outcome, where available, for comparison. We will discuss with CMS whether other comparisons might be useful. See ***Section 4*** for additional detail about monitoring measures and the SDRS.

6.2 Individual State Descriptive Analyses

The RTI team will conduct individual-State descriptive analyses annually. These analyses will focus on estimates for a broad range of beneficiary experience, utilization, access to care, cost, and quality measures, as well as changes in these measures across years or subgroups of interest within each year. The results of these analyses will be presented in an annual report for each State. Analyses will also be presented for a comparison group. For these annual analyses, the RTI team will develop separate (unlinked) MDS, Medicare, and

Medicaid beneficiary-level analytic files to measure beneficiary experience, quality, utilization, access to care, and costs.

These analyses will be conducted for the 2-year baseline period and for each demonstration year. For managed fee-for-service (MFFS) demonstrations, we anticipate that analyses in the first annual report will focus on Medicare fee-for-service and MDS data. Medicaid data files will be incorporated into the analyses as they become available for each State. For capitated model demonstrations, the analyses in the first year will depend on the timeliness of MMP data, as well as Medicaid claims, and Medicaid managed care encounter data where relevant.

Consistent with the intent-to-treat approach, all individuals eligible to participate in the demonstration will be included, regardless of whether they opt out initially or disenroll, or whether they actually engage with the care model. The start date will vary by State and, therefore, may not represent calendar years. We will also use these data, but with additional claims run-out, for the final evaluation report, when we will link the Medicare and Medicaid data, including data for the comparison group for each State.

Ideally, individual State analyses will have predemonstration and demonstration year data available on quality, utilization, and costs. Capitated model States may have had beneficiaries enrolled in some type of Medicare or Medicaid managed care plan in the predemonstration period, which may limit data availability in the predemonstration period. RTI will consider the data availability issues for each State given its demonstration model(s). RTI will attempt to align the different types of data as closely as possible. For example, for estimating expenditures, we will use the per-member per-month (PMPM) total capitation payment as the total cost to the program of providing care under a capitated arrangement, and the sum of the paid claims as the total cost under MFFS models. However, it may not be possible to develop comparable utilization and cost measures for all States, which may limit analyses. For example, if States have different kinds of home and community-based services (HCBS) or behavioral health measures, finding a common metric may not be possible.

The annual reports will contain results of descriptive statistics (means, counts, and proportions; and tests of means, counts, and proportions across years or subgroups) and not multivariate results. For example, we will examine total costs (Medicare and Medicaid separately), rates of primary care and specialist care use, rates of avoidable hospitalization and inappropriate readmissions, counts of hospital and nursing facility admissions and length of stay, rates of HCBS use, and mortality. Providing results comparing beneficiary subgroups by age groups, subpopulations, and other important characteristics will inform CMS and States about improvements needed over time.

We are not planning to conduct multivariate time trend analyses each year, because data lags will lead to data analyses without sufficient claims run-out, causing us to over- or

underestimate results. We think it is important to allow for sufficient claims run-out before employing multivariate analyses. We also plan to incorporate the comparison group in each multivariate model in the final year of analysis to control for the changes over time that are not a direct result of the demonstrations. Finally, if we conduct multivariate modeling on incomplete data each year, we run the risk of finding different results after all the data have been received. Dummy variables for each demonstration year in the analysis will identify annual effects. The one exception to running multivariate analyses after all the data are in will be to estimate savings for each capitated model State. Savings for capitated models will be measured twice during the evaluation using a regression-based approach. Savings for MFFS will be done annually using an actuarial cost-savings analysis. The MFFS savings calculation methodology is described in MFFS State Memoranda of Understanding (MOUs) and Final Agreements. Comparison groups will be refreshed annually for all States for the descriptive analyses in annual reports and all savings calculations.

6.3 Impact Analysis for Final Report: Effects of Demonstrations Within States

We will assess the overall impact of the demonstration on quality, utilization, and cost measures using a difference-in-differences methodology with a comparison group for the final evaluation report for each State. Under the difference-in-differences methodology, pre- and postdemonstration changes in the dependent variables of those eligible for the demonstration will be compared with the pre- and postexperience of a comparison group. Whereas the comparison group will be determined for each individual State, the difference-in-differences methodology will be consistent across States. In addition, this methodological framework can be easily applied to each of the quality, utilization, access to care, and cost measures that will be tracked within States over time. These analyses will use linked Medicare and Medicaid claims and encounter data for the predemonstration and demonstration period. We will construct comparable measures, to the extent possible, for beneficiaries in the comparison group. We will identify measures for these analyses after the demonstration has concluded, after seeing the quality of the data in each State.

Data for 2 predemonstration years and each demonstration year for both the demonstration and comparison groups will be pooled, assuming we are able to obtain 2 comparable years of predemonstration data. A restricted difference-in-differences equation will be estimated as follows:

$$\text{Dependent variable}_i = \beta_0 + \beta_1 \text{PostYear} + \beta_2 \text{Demonstration} + \beta_3 \text{PostYear} * \text{Demonstration} + \beta_4 \text{Demographics} + \beta_{5\text{-}j} \text{Market} + \varepsilon \quad (1)$$

where separate models will be estimated for each dependent variable. *PostYear* is an indicator of whether the observation is from the pre- or postdemonstration period, *Demonstration* is an indicator of whether the beneficiary was in the demonstration group,

and *PostYear* * *Demonstration* is an interaction term. *Demographics* and *Market* represent vectors of beneficiary and market characteristics, respectively, and will most likely be similar to the ones used in propensity score analyses.[2]

Under this specification, the coefficient β_0 reflects the comparison group predemonstration period mean adjusted for demographic and market effects, β_1 reflects the average difference between postperiod and predemonstration period in the comparison group, β_2 reflects the difference in the demonstration group and comparison group at predemonstration, and β_3 is the overall average demonstration effect during the demonstration period. This last term is the difference-in-differences estimator and the primary policy variable of interest. In addition to demographic characteristics and geographic location, the model would also control for characteristics of the health care market, such as differences in provider supply.

In addition to estimating the model described in Equation 1, a less restrictive model will be estimated to produce overall and year-by-year effects of the demonstration within individual States. The specification of the unrestricted model is as follows:

$$\text{Dependent variable} = \beta_0 + \beta_{1\text{-}k}\text{PostYear}_{1\text{-}3} + \beta_2\text{Demonstration} + \beta_{3\text{-}k}\text{PostYear}_{1\text{-}3} * \text{Demonstration} + \beta_4\,\text{Demographics} + \beta_{5\text{-}j}\,\text{Market} + \varepsilon \quad (2)$$

This equation differs from the previous one in that separate difference-in-differences coefficients are estimated for each year. Under this specification, the coefficients $\beta_{3\text{-}k}$ would reflect the impact of the demonstration in each respective year, whereas the previous equation reflects the impact of the entire demonstration period. This specification will also allow testing of whether changes in dependent variables occur in the first year of the demonstration only, continuously over time, or in some other pattern. A Hausman test will be used to determine whether the more restrictive or less restrictive model is more appropriate for modeling demonstration effects. Depending on the outcome of interest, we will estimate the equations using logistic regression, Generalized Linear Models (GLM, a more general version of ordinary least squares [OLS] for cost models, which is easier to interpret than logged OLS models), or count models such as negative binomial or Poisson regressions (e.g., for the number of readmissions). For the analysis of utilization and costs, a two-part model will be used to estimate the probability of any utilization using Logit, and the level of use conditional on having some utilization using GLM. For example, only some beneficiaries will have inpatient expenditures or long-term care services. The two-part model allows one to account for the beneficiaries with no utilization for the outcome in question using Logit, and only the beneficiaries with use using GLM. In addition to

[2] Comparable models would be estimated for dichotomous variables, and a Norton correction would be used to obtain appropriate estimates and standard errors for interaction terms in models using nonlinear estimation.

interpreting regression coefficients, we will also use regression results to calculate the marginal effects of demonstration impact.

6.3.1 Sensitivity Analysis

We will look for opportunities to conduct sensitivity analyses of the impact of State demonstrations where data allow us to do so. The validity of the difference-in-differences approach depends in large part on the assumption that changes over time in the comparison group are a reasonable counterfactual for what would have happened to the demonstration group. As discussed, there are challenges to identifying an appropriate comparison group for this demonstration. Particular concerns include other demonstrations or initiatives occurring in potential comparison States, such as the Health Home State Plan Amendments (SPAs). In addition, accountable care organizations (ACOs) are being developed in demonstration and potential comparison States. Finally, delivery systems prior to the demonstration should be similar for demonstration and comparison groups, a requirement that may be hard to meet after States with other potential confounders are excluded. These issues may result in potential comparison groups or comparison States that differ from the demonstration groups and demonstration States on important dimensions and why it will be important to conduct sensitivity analyses.

Testing the sensitivity of impact estimates to a variety of modeling decisions is an important part of generating these estimates. The choice of comparison groups is one such decision. As part of our efforts to check the consistency of the impact estimates for a few States, we will compare the findings from the core models to estimates based on assumptions, such as different combination of States for out-of-State comparison groups and different propensity-score models. Consistency in the estimates across models will give us more confidence in the reliability of the impact estimates.

6.3.2 Additional Analyses

In addition to the claims and encounter data analyses that look at utilization and costs associated with medical and behavioral health supports, and long-term services and supports (LTSS), we will use the Nursing Home MDS to analyze additional changes in patterns of facility-based LTSS quality and use. We will evaluate admission rates, acuity upon admission, and selected quality measures for both short-stay (i.e., skilled nursing facility users) and long-stay facility residents.

The addition of meaningful assessment data in the MDS discharge assessments since implementation of the MDS 3.0 in October 2010 provides an opportunity to evaluate the characteristics (such as functional status) and discharge destinations of beneficiaries leaving nursing facilities, to the extent that these assessments are actually completed. RTI recently analyzed the completeness of discharge data submitted, finding that functional status upon discharge is virtually complete (less than 2 percent missing). However, it appears that many

residents who have left nursing facilities have no discharge assessment in place, limiting the generalizability of any discharge analyses.

The RTI team will also conduct an analysis of encounter data coding intensity to assess for any upcoding by providers or MMPs. Given that capitated payments under the Financial Alignment Demonstration may be affected by coding intensity, CMS has asked the RTI team to assess the data for changes in coding patterns during the demonstration. These analyses will examine the extent to which changes in coding intensity observed in demonstration States compare with nondemonstration States or a predemonstration period.

Because enrollment is voluntary, RTI will compare the characteristics of those who enroll with those who are eligible but do not enroll in the care model and conduct analyses to further explore demonstration effects on demonstration enrollees, acknowledging that selection bias must be taken into account in interpreting the results.

6.4 Aggregate Analyses

After the final multivariate analyses have been performed for individual State evaluations, the RTI team will conduct aggregate analyses to inform CMS and the States about changes resulting from various State demonstration design features on quality, utilization, and cost outcomes. CMS is interested in this type of analysis in order to look across States that have implemented similar design features (e.g., MFFS or capitated payment model, demonstration design, contract vs. noncontract states, Health Home SPA States). The timing and final methodology for this analysis will be contingent on when individual States begin the demonstration and the design features used, both of which were unknown at the time of this report because implementation had not yet begun for the majority of demonstrations. The goal of this type of analysis is to disentangle the relative effectiveness of the demonstration design choices, whereas the individual State evaluations will estimate the impact of a chosen *set* of demonstration design choices relative to the status quo in the State.

Several research questions will be addressed by this part of the evaluation—for example:

- Which demonstration model (MFFS or capitated) has achieved greater savings?
- Are there differences in key outcomes (e.g., quality, utilization, expenditure types) that can be attributed to the type of financial alignment model used?
- Do the effects achieved by alternative integrated care models occur equally fast? Or does one model (MFFS vs. capitated) achieve gains more quickly than the other?
- Does the approach to enrollment (e.g., passive enrollment) affect access to care and costs?
- How does the relative degree of care management intensity and diversity across services affect outcomes?
- Do these effects vary across subgroups of beneficiaries?

RTI will carefully consider which States to include in various meta-analyses in order to provide thoughtful conclusions. Some States will be moving beneficiaries from MFFS to capitated approaches, some from capitated approaches to better-integrated capitated approaches. The populations eligible for the demonstration also differ across States. We will assess States on these dimensions and include States with similar approaches or populations in appropriate meta-analyses

Currently, we plan to design the quantitative portion of this analysis as a series of meta-analyses of the demonstration effects estimated in each of the individual State evaluation analyses. A meta-analysis is an analytic technique designed to explain differences in State demonstration design choices and other contextual variables that influence the effectiveness of interventions. The meta-analysis regressions will be estimated on various quality, utilization, and cost measures. The analyses will rely on impact estimates obtained from State-specific regressions on demonstration outcomes for which comparison groups will already have been identified for each State.

This analysis will differ from individual State analyses in several ways. First, these analyses will be conducted at the State level rather than the beneficiary level. As a result, these analyses will have fewer independent variables in regression analyses than regressions in individual State analyses. Second, because the analyses use previously estimated regression coefficients from each individual State, the use of comparison group information is implicit in that it is contained in these previously estimated regression coefficients from individual State regressions. Thus, the quality of the meta-analyses is dependent on the prior choice of appropriate comparison groups in each individual State. Prior sensitivity analyses will help to support the validity/appropriateness of the State-specific impact estimates, which will strengthen the meta-analyses.

The purpose of the aggregate analyses is twofold: first, to summarize the effects of integrated care models on changes in quality, utilization, and cost of care; and second, to study the sources of variation in those State-specific estimates with the goal of understanding how State policy choices might affect the key outcomes. This latter purpose will take advantage of data on State policy characteristics, such as the proportion of a State's long-term care budget that is represented by home and community-based services (HCBS) spending, or the relative acuity of a State's nursing facility population, or similar factors of interest that will help differentiate which types of integrated care models (e.g., MFFS or capitated) better improve quality, shift utilization toward nonacute services, and reduce costs.

The data elements necessary for these analyses would be largely developed in the other parts of the evaluation. The key dependent variables would be the impact estimates from individual State analyses on quality, utilization, and cost outcomes described in ***Sections 7*** to ***9*** below. These effects will be derived from the individual State difference-in-differences

analyses described above. These effects, $\hat{\beta}_{jstg}$, will be estimated by outcome (j) and State (s), and optionally by time period (t), and beneficiary subgroup (g).

For example, preceding this aggregate analysis task, the RTI team will have already estimated the effect of each State's demonstration on inpatient expenditures using one OLS beneficiary-level regression for each State. Each State's estimated coefficient representing the effect of the State's demonstration on inpatient expenditures would be saved to a new, single, State-level variable that represents the change in inpatient expenditures produced by each of the State demonstrations. This new variable containing an observation from each State will become a dependent variable in a State-level regression of the effects of various State policy characteristics on the *change* in inpatient expenditures resulting from the demonstrations. If separate regression estimates had been generated for three subgroups of beneficiaries (e.g., if we stratified estimates by age group) in each State, there would be multiple estimates. To test whether the length of a demonstration is related to its effectiveness, we may have also previously estimated separate effects for each year. Thus over 3 years, 3 age groups, and 20 States (as an example), we would have a total of 180 observations (3 x 3 x 20) with which to test hypotheses on effect of managed care arrangements. If data are available, meta-analyses could also be conducted within financial alignment models (e.g., capitated and MFFS States separately) to test the impact of demonstration design and policy variables.

The explanatory variables in these analyses will include key elements of the demonstration design and policy variables, D_s and control variables, X_s to account for differences across demonstrations in contextual factors (demographic, market, regulatory, fiscal, policy differences) that might also influence the estimates. Examples of contextual factors may include differences across States in demographic or characteristics and provider supply. These variables will be obtained from a variety of sources, including the analyses of the quantitative data collected for the individual-State analyses and qualitative data collected in State-specific focus groups and interviews with State officials coded in a consistent manner so as to be useable in the meta-analysis.

Models will be of the following form:

$$\beta_{jstg} = \Pi D_s + \Gamma X_s + u_{jstg} \quad (3)$$

where D_S is the set of demonstration design and policy variables of interest included in the regression, X_S is the set of contextual variables, and *u* is the error term. The contextual variables will reflect differences in the policy environment across States (such as a State's Medicaid eligibility levels prior to the demonstration). Because State decisions on Medicaid policy are not random, omitted variables that affect both State policy choices and the outcomes of interest may bias the estimates of the effects of the demonstration. Addressing these policy issues will require including in the regression models time-varying State-level

control variables that are associated with the State's policy decisions, such as regulatory policies and measures of the fiscal and policy environment to capture time-varying differences (beyond the demonstration) across the States. This is another area (beyond the demonstration policy changes and changes in other factors), where the evaluation will need detailed State information in order to address the potential bias in the quasi-experimental estimates.

The estimates of principal interest to the evaluation are the Π coefficients. These models will be estimated accounting for varying degrees of precision in State-specific estimates. In particular, States where impact estimates are very precisely estimated because of large sample sizes will be given more weight in the meta-analysis than States where estimates are based on many fewer individuals. In practice, we will control for these differences in precision by weighting the estimated variances of the effects, $s^2{}_{\beta_{jstg}}$, which will be obtained as a byproduct of the individual State analyses.

The activities for quantitative monitoring, descriptive, and impact analyses may be revised if modifications are made to the demonstrations or if data sources are not available as anticipated. The aggregate analyses methods will be finalized when more information is known on the participating States and their demonstration designs. If modifications to this evaluation plan are required, they will be documented in the annual and final evaluation reports as appropriate.

This page intentionally left blank

Section 7. Utilization and Access to Care

CMS expects State demonstrations to improve outcomes for Medicare-Medicaid enrollees. The impact of these demonstrations should result in changes in service use, in annual utilization patterns, and in specific patterns of care. Of particular interest is the impact across subpopulations of Medicare-Medicaid enrollees with particular risk factors or comorbidities. CMS is also interested in whether any observed impact is short term only, or if the demonstrations have effects that lead to continued improvement in outcomes and cost savings over the course of the demonstration. Research questions regarding utilization to be analyzed include the following:

- What is the impact of the State demonstrations on utilization patterns during the course of the demonstration?
- What is the impact on hospital and nursing facility admission rates, potentially avoidable hospitalization utilization rates by setting, and on long-term services and supports (LTSS) utilization rates? What is the impact of the demonstration on hospital and nursing facility length of stay?
- Do demonstrations change the balance between home and community-based services (HCBS) and nursing facility use, the types of Medicare-Medicaid enrollees who use these services, and utilization rates by type of HCBS such as personal care? Do Medicare-Medicaid enrollees receive more HCBS as a result of the demonstrations?
- Is any impact short term (e.g., lasting only for 1 year before returning to predemonstration level, increasing over time, reaching a plateau after a year or 2)?
- Does the observed impact vary by health condition or other beneficiary characteristics?
- Will case management or care coordination lead to lower hospital admission rates or, if admitted, shorter lengths of stay and shorter nursing facility and home health care episodes?
- Are demonstration group members using fewer inpatient services and more ambulatory services?
- Is the impact greater for more medically complex (multiple chronic condition), high-cost (top 10 percent) enrollees?

In addition, State demonstrations are expected to improve access to services, which should be evident through changes in utilization patterns of certain services. Research questions pertaining to access to care are as follows:

- Access to medical care: do demonstration participants experience increases in the mean number of primary care visits and increased visit rates by specialty type?
- Access to LTSS: does acuity on admission to nursing facilities increase? Do discharge rates back to the community from nursing facilities increase? Is there an increase in the proportion of HCBS users self-directing care?

- Access to behavioral health services: does the mental health outpatient utilization rate increase? Does the outpatient substance use disorder service utilization rate increase?

Rapid-cycle monitoring analyses, within-State annual descriptive analyses, and postdemonstration impact analyses will be used to address the research questions outlined above.

7.1 Ongoing Monitoring

To monitor States' progress during the demonstration, we will calculate high-level measures for each State to identify changes in utilization over time. Theoretically, rates of facility use such as long-term nursing facility use should decrease and rates of HCBS use should increase. In addition, primary care service use and access to mental health services should increase. Various inpatient and emergency room measures that RTI plans to include in quarterly reports are described in more detail in our section on quality measures. We will also identify a range of key utilization and access to care measures to include in quarterly reports to CMS and the States, annual, and final evaluation reports. For each utilization type, these measures will usually be expressed as visits per 1,000 eligible beneficiaries and users as a percentage of the demonstration-eligible population.

7.2 Within-State Descriptive Analyses

For annual reports, we will measure predemonstration (the 2-year period before demonstration start) and annual utilization rates of Medicare- and Medicaid-covered services for each State, using unlinked data to identify the effects of State demonstrations on the type and level of service use, ranging along a continuum from facility-based care to care provided at home (***Table 16***). Both payers reimburse for similar services, such as inpatient care and home health care, but each payer also reimburses for services not reimbursed by the other (e.g., Medicare short nursing facility stays, Medicaid HCBS, and behavioral health services), so both Medicare and Medicaid data will be used for this analysis.

We will calculate average utilization rates at predemonstration and at the beginning, middle, and end of each demonstration. Use rates for each State will be stratified by hierarchical condition categories (HCC) scores, health status measures, or similar measures. We will adjust for hospitalizations in the prior year using categorical HCC scores. Chi-square and t-tests will be used to test for significant differences in use across years and between subpopulations within a State.

Almost all analyses will be conducted at the beneficiary level. Annual aggregate reports that summarize across States will be limited by when States begin the demonstration and by differences in data availability.

Table 16
Service categories for reporting utilization measures

Service type	Medicare only	Medicaid only	Medicare and Medicaid
Inpatient	—	—	X
Emergency room	—	—	X
Skilled nursing facility	X	—	—
Nursing facility (long-term stay)	—	X	—
Other facility-based[1]	—	—	X
Outpatient[2]	—	—	X
Outpatient behavioral health (mental and substance use disorder)	—	X	—
Home health	—	—	X
HCBS (PAS, waiver services)	—	X	—
Dental	—	—	X

— = not applicable. HCBS = home and community-based services; PAS = personal assistance services.

[1] Includes long-term care hospital, rehabilitation hospital, State mental health facility stays.

[2] Includes visits to physician offices, hospital outpatient departments, rehabilitation agencies.

7.3 Impact Analyses

Multivariate difference–in-differences analyses to evaluate the impact of individual State demonstrations relative to their selected comparison groups will be conducted in the final year of the evaluation after sufficient claims run-out and encounter data have been received. Dependent variables in the multivariate analyses will include rates and lengths of short- and long-term nursing facility stays, number of primary care provider (PCP) visits, number of specialty physician visits, and rates and number of months of personal assistance services and HCBS waiver services. Any inpatient analyses other than rates of overall inpatient use will be discussed in the section on quality measures (***Section 8***).

To understand whether demonstration effects are short term (one time only) or longer term (over the course of the demonstration), we will include dummy variables for each year of the demonstration to test the effects at the beginning, middle, and end of each demonstration. Differences in the slopes of the regression coefficients will be tested to determine whether the effects last only for 1 year or are ongoing.

One key strategy for reducing costs without compromising quality of care is to improve care coordination by reducing fragmentation and redundancies in services. We will develop analyses to address this issue, such as analyzing patterns of primary versus specialty care. We hypothesize that PCPs and internists will provide an increasingly higher proportion of all

physician visits in the demonstration group relative to the comparison group over time, unless non-visit compensation is provided to physicians of comparison group members. RTI will identify PCPs for each beneficiary in fee for service States (and in capitation States, if possible using encounter data) in the demonstration and comparison populations using Unique Physician Identification Numbers (UPIN) on the Part B physician claims. Differences in the relative use of PCP/internists versus specialist physicians will be examined to measure the relative probability of demonstration and comparison beneficiaries' use of primary and specialty services and changes in the relative probability over time. We will explore other measures in addition to the UPIN code for identifying changes in primary and specialist care, taking into consideration that specialists may be providing primary care for individuals with chronic conditions. Measures for assessing fragmentation of care for LTSS and behavioral health services will also be explored after reviewing candidate measures.

7.4 Analytic Challenges

Potential problems with encounter data quality, lack of care coordination and case management data, and incomplete data regarding behavioral health services could affect our ability to conduct aspects of the utilization and access to care analyses.

The activities for analyzing utilization and access to care may be revised if modifications are made to the demonstrations or if data sources are not available as anticipated. If modifications to this evaluation plan are required, they will be documented in the annual and final evaluation reports as appropriate.

Section 8. Quality of Care

Across all States, we will examine a set of quality measures for monitoring and evaluation purposes. Some of these measures will also inform Medicare-Medicaid Plan (MMP) quality withhold payments for capitated demonstrations, or managed fee for service (MFFS) demonstration State performance payments. These measures may be supplemented by additional evaluation measures appropriate to individual State demonstrations. The measures discussed in this section are largely utilization-based measures reflecting quality of care. We discuss other aspects of quality, such as quality of life, satisfaction, and access to care in ***Section 3, Beneficiary Experience,*** and in ***Section 7, Utilization and Access to Care***. There are several data sources for quality measures: claims, encounter, and Nursing Home Minimum Data Set (MDS) data, which will be obtained and analyzed by the RTI team; and information provided by States, CMS, or other entities. The latter may include Healthcare Effectiveness Data and Information Set (HEDIS) measures collected as part of MMP performance, other data that States require their MMPs to report, or any results from beneficiary surveys collected by a State, CMS, or other entities.

Table 17 provides the list of quality measures, common to all States that RTI has identified for the evaluation. RTI will calculate these measures using data provided by CMS.

Many of the measures in ***Table 17*** are established HEDIS measures that MMPs will be required to report to CMS As specified in the Memorandum of Understanding (MOU); as such, extant reporting could be a source of data for our analyses. The National Committee for Quality Assurance (NCQA) definitions are established and standardized. Under the demonstration, MMPs will be required by CMS to report data specific to demonstration enrollees from HEDIS as well as Health Outcome Survey (HOS) and Consumer Assessment of Healthcare Providers and Systems (CAHPS) surveys. But, no such requirement exists in comparison States, and beneficiaries who do not participate in the demonstration will not have such data.

Table 17
Evaluation quality measures

Measure concept	State model (capitated or MFFS)
RTI team calculations based on data obtained from CMS	
30-day all-cause risk-standardized readmission rate	Capitated, MFFS
Influenza immunization	Capitated, MFFS
Pneumococcal vaccination for beneficiaries 65 and older	Capitated, MFFS
Ambulatory care sensitive condition admissions—overall composite (AHRQ PQI #90)	Capitated, MFFS
Ambulatory care sensitive condition admissions—chronic composite (AHRQ PQI #92)	Capitated, MFFS
Preventable ED visits	Capitated, MFFS
ED visits, excluding those resulting in inpatient admission or death	Capitated, MFFS
Admissions with primary diagnosis of a severe and persistent mental illness or substance use disorder	Capitated, MFFS
Follow-up after hospitalization for mental illness	Capitated, MFFS
Screening for clinical depression and follow-up	Capitated, MFFS
Cardiac rehabilitation following hospitalization for cardiac event	Capitated, MFFS
Percent of high-risk long-stay NF residents with pressure ulcers	Capitated, MFFS
Screening for fall risk	Capitated, MFFS
Initiation and engagement of alcohol and other drug dependence treatment	Capitated, MFFS
HEDIS data obtained from CMS	
Adult BMI assessment	Capitated
Annual monitoring for patients on persistent medications	Capitated
Antidepressant medication management	Capitated
Breast cancer screening	Capitated
Comprehensive diabetes care—selected components	Capitated
Controlling high blood pressure	Capitated

BH = behavioral health; BMI = body mass index; ED = emergency department; LTSS = long-term services and supports; MFFS = managed fee for service; NF = nursing facility.

8.1 Measure Development

The scope of work for this evaluation requires development of quality measures that serve several purposes: ongoing monitoring on a rapid-cycle basis, within-State descriptive? analyses, and meta-analysis and support of MFFS State performance payment calculations or MMP quality withholds.

8.1.1 Rapid-Cycle Monitoring

Because Medicaid claims and MMP encounter data will likely lag behind Medicare claims data, and linking records across data systems will not occur until after the demonstrations are complete, we have sought to develop a meaningful set of quality measures that do not require linked data for quarterly and annual reporting. Measures that require claims analysis will be produced by RTI with unlinked data, and thus, Medicare-based measures will likely be available first for all MFFS States, followed later by measures based on MMP encounter data and Medicaid claims. As States vary in the timeliness of their Medical Statistical Information System (MSIS) submissions, the lag time for Medicaid claims-based and Medicaid managed care encounter-based measures may also vary by State. The timeliness of MMP encounter data is not yet known. The measures used for quarterly monitoring will be based only on data from demonstration States and will not be presented in relation to a comparison group because comparison groups will not be identified until after the end of each demonstration year. This limitation may also affect the extent to which risk adjustment is possible at this stage, although some stratification by subgroups may be included. These measures will, however, be standardized across States to the extent possible and be useful for monitoring trends over time within a State and across the demonstration.

8.1.2 Within-State Descriptive Analyses

Rapid-cycle monitoring data will be limited to monitoring trends in the demonstrations as demonstration State data are available, but measures used for State-specific analyses will be developed and reported based on a different set of criteria. First, as described in earlier sections, these evaluations will be based on demonstration and comparison group analyses, meaning that comparable data must be available in nondemonstration States and so will be limited to claims/encounter-based measures. Second, to best capture the impact of integrated models, some quality measures will require linking Medicare, Medicaid, and encounter data, which will be done only once, near the end of the evaluation to acquire adequate Medicaid claims and encounter data. Measures that do not rely on linked data can be reported more frequently. Third, because States have been given the freedom to develop unique approaches to integration and may target specific groups or services, it will be necessary, and desirable, to develop some measures unique to individual States and aligned with the demonstration goals developed by the States. We will develop State-specific measures within the first 6 months of each demonstration

8.2 Methods

8.2.1 Research Design

The research design for the analysis of quality measures will be consistent with that described in the prior section, and will depend on the research question and the quality measure analysis activity (monitoring or evaluation).

8.2.2 Measures and Data Requirements

Table 18 displays in detail the set of evaluation measures and the types of data required for each. We will also use these measures as dependent variables in multivariate regression analyses in the final evaluation report to identify factors contributing to quality outcomes.

The measure set currently lacks measures of beneficiary experience of care such as those commonly collected via CAHPS. We will explore the potential inclusion of CAHPS measures in the MFFS model State evaluations after further discussions with CMS about its plans to field a version of the CAHPS in MFFS States. However, the response rates to CAHPS surveys are generally low, and it would be important to survey both those enrolled and those in the eligible population who are not enrolled in the demonstration. However, we will be requesting from CMS CAHPS results collected by States, CMS, or other entities as described in ***Sections 3*** and ***4***, and include information from these reports in the annual and final evaluation reports.

8.2.2.1 Control Variables

In addition to the development of quality indicators, we will also use the data sources above to develop variables to control for observable differences between individual beneficiaries, both within the demonstration group and between the demonstration and comparison groups. At minimum, these variables would include demographic information (age, sex, race) available from Medicare and Medicaid enrollment or eligibility files. Controlling for comorbidity, or health more generally, is often done through the use of indices built from the occurrence of diagnosis codes in claims data. Examples of indices used in the analysis of Medicare data or in calculating risk-adjusted payments to managed care plans include the hierarchal condition categories (HCC) model and the Charlson-Deyo index. In Medicaid analyses, a commonly used index is the Chronic Illness and Disability Payment System (CDPS). There are options for the types of claims used in these adjustments (hospital only or hospital and physician), and we will work with CMS to make final decisions on the choice of risk adjuster for these analyses.

We will also explore the development of measures from publicly available sources, such as the Area Resource File (ARF), that define the health care environment in each beneficiary's residential area. These may reflect variation in the supply of available providers or general economic conditions that may apply to health care markets.

Table 18
Quality measures for evaluation: Detailed definitions, use, and specifications

Measure concept (specific measure)	Data sources and responsibility for data collection	Domain (prevention, care coordination, beneficiary experience)	Will evaluation produce impact estimates?*	Definition (link to documentation if available)	Numerator/denominator description
All-cause readmission 30-day all-cause risk-standardized readmission rate	Claims/ encounter RTI will acquire and analyze	Care coordination	Yes	Risk-adjusted percentage of demonstration-eligible Medicare-Medicaid enrollees who were readmitted to a hospital within 30 days following discharge from hospital for the index admission. (https://www.cms.gov/sharedsavingsprogram/Downloads/ACO_QualityMeasures.pdf)	Numerator: Risk-adjusted readmissions at a non-Federal, short-stay, acute-care or critical access hospital, within 30 days of discharge from the index admission included in the denominator, and excluding planned readmissions. Denominator: All hospitalizations not related to medical treatment of cancer, primary psychiatric disease, or rehabilitation care, fitting of prostheses, and adjustment devices for beneficiaries at non-Federal, short-stay acute-care or critical access hospitals, where the beneficiary was continuously enrolled in Medicare and Medicaid for at least 1 month after discharge, was not discharged to another acute care hospital, was not discharged against medical advice, and was alive upon discharge and for 30 days postdischarge.
Immunizations Influenza immunization	Claims/ encounter RTI will acquire and analyze	Prevention	Yes	Percentage of patients aged 6 months and older seen for a visit between October 1 and March 31 of the 1-year measurement period who received an influenza immunization OR who reported previous receipt of an influenza immunization. (https://www.cms.gov/sharedsavingsprogram/Downloads/ACO_QualityMeasures.pdf)	Numerator: Demonstration-eligible Medicare-Medicaid enrollees who have received an influenza immunization OR who reported previous receipt of influenza immunization. Denominator: Demonstration-eligible Medicare-Medicaid enrollees seen for a visit between October 1 and March 31 (flu season), with some exclusions allowed.

(continued)

Table 18 (continued)
Quality measures for evaluation: Detailed definitions, use, and specifications

Measure concept (specific measure)	Data sources and responsibility for data collection	Domain (prevention, care coordination, beneficiary experience)	Will evaluation produce impact estimates?*	Definition (link to documentation if available)	Numerator/denominator description
Immunizations (continued) Pneumococcal vaccination for patients 65 years and older	Claims/ encounter RTI will acquire and analyze	Prevention	Yes	This measure has been developed for those aged 65 and over, but will be revised for State's eligible population aged 21–64. As originally developed, the measures definition is percentage of patients aged 65 years and older who have ever received a pneumococcal vaccine. Because CDC recommends pneumococcal vaccine for individuals under 65 with chronic conditions, we will revise specifications accordingly.	The specifications for this measure reflect the current definitions that will be revised for the entire eligible population without regard to age. Numerator: Demonstration-eligible Medicare-Medicaid enrollees 65 years and older who have ever received a pneumococcal vaccination. Denominator: All demonstration-eligible Medicare-Medicaid enrollees 65 years or older, excluding those with documented reason for not having one.
Ambulatory care-sensitive condition admission Ambulatory care sensitive condition admissions—overall composite (AHRQ PQI #90)	Claims/ encounter RTI will acquire and analyze	Prevention, care coordination	Yes	Combination using 12 individual ACSC diagnoses for chronic and acute conditions. For technical specifications of each diagnosis, see http://www.qualityindicators.ahrq.gov/Modules/PQI_TechSpec.aspx	Numerator: Total number of acute-care hospitalizations for 12 ambulatory care-sensitive conditions among demonstration-eligible Medicare-Medicaid enrollees, aged 18 or older. Conditions include diabetes—short-term complications; diabetes—long-term complications; COPD; hypertension; CHF; dehydration; bacterial pneumonia; UTI; angina without procedure; uncontrolled diabetes; adult asthma; lower-extremity amputations among diabetics. Denominator: Demonstration-eligible Medicare-Medicaid enrollees, aged 18 or older.

(continued)

Table 18 (continued)
Quality measures for evaluation: Detailed definitions, use, and specifications

Measure concept (specific measure)	Data sources and responsibility for data collection	Domain (prevention, care coordination, beneficiary experience)	Will evaluation produce impact estimates?*	Definition (link to documentation if available)	Numerator/denominator description
Ambulatory care sensitive condition admissions—chronic composite (AHRQ PQI #92)	Claims/ encounter RTI will acquire and analyze	Prevention, care coordination	Yes	Combination using nine individual ACSC diagnoses for chronic diseases. For technical specifications of each diagnosis, see http://www.qualityindicators.ahrq.gov/Modules/PQI_TechSpec.aspx	Numerator: Total number of acute-care hospitalizations for nine ambulatory care-sensitive chronic conditions among demonstration-eligible Medicare-Medicaid enrollees, aged 18 or older. Conditions include diabetes—short-term complications; diabetes—long-term complications; COPD; hypertension; CHF; angina without procedure; uncontrolled diabetes; adult asthma; lower extremity amputations among diabetics. Denominator: Demonstration-eligible Medicare-Medicaid enrollees, aged 18 or older.
Admissions with primary diagnosis of a severe and persistent mental illness or substance use disorder	Claims/ encounter RTI will acquire and analyze	Prevention, care coordination	Yes	Percentage of demonstration-eligible Medicare-Medicaid enrollees with a primary diagnosis of a severe and persistent mental illness or substance use disorder who are hospitalized.	Numerator: Total number of acute-care hospitalizations among demonstration-eligible Medicare-Medicaid enrollees, aged 18 or older with a primary diagnosis of a severe and persistent mental illness or substance use who are hospitalized. Denominator: Demonstration-eligible Medicare-Medicaid enrollees, aged 18 or older.
Avoidable emergency department visits Preventable/ avoidable and primary care treatable ED visits	Claims/ encounter RTI will acquire and analyze	Prevention, care coordination	Yes	Based on lists of diagnoses developed by researchers at the New York University Center for Health and Public Service Research, this measure calculates the rate of ED use for conditions that are either preventable/ avoidable, or treatable in a primary care setting. (http://wagner.nyu.edu//chpsr/index.html?p=61)	Numerator: Total number of ED visits with principal diagnoses defined in the NYU algorithm among demonstration-eligible Medicare-Medicaid enrollees. Denominator: Demonstration-eligible Medicare-Medicaid enrollees.

(continued)

Table 18 (continued)
Quality measures for evaluation: Detailed definitions, use, and specifications

Measure concept (specific measure)	Data sources and responsibility for data collection	Domain (prevention, care coordination, beneficiary experience)	Will evaluation produce impact estimates?*	Definition (link to documentation if available)	Numerator/denominator description
Emergency department visits ED visits excluding those that result in death or hospital admission	Claims/ encounter RTI will acquire and analyze	Prevention, care coordination	Yes	Percentage of demonstration-eligible Medicare-Medicaid enrollees with an emergency department visit.	Numerator: Total number of ED visits among demonstration-eligible Medicare-Medicaid enrollees excluding those that result in death or hospital admission. Denominator: Demonstration-eligible Medicare-Medicaid enrollees.
Follow-up after mental health hospitalization Follow-up after hospitalization for mental illness	Claims/ encounter RTI will acquire and analyze	Care coordination	Yes	Percentage of discharges for demonstration-eligible Medicare-Medicaid enrollees who were hospitalized for selected mental health disorders and who had an outpatient visit, an intensive outpatient encounter or partial hospitalization with a mental health practitioner. Two rates are reported: (1) percentage of members who received follow-up within 30 days of discharge; (2) percentage of members who received follow-up within 7 days of discharge. (http://www.qualityforum.org/QPS/)	Numerator: Rate 1: An outpatient visit, intensive outpatient encounter or partial hospitalization with a mental health practitioner within 30 days after discharge. Include outpatient visits, intensive outpatient encounters or partial hospitalizations that occur on the date of discharge; Rate 2: An outpatient visit, intensive outpatient encounter or partial hospitalization with a mental health practitioner within 7 days after discharge. Include outpatient visits, intensive outpatient encounters or partial hospitalizations that occur on the date of discharge. Denominator: Demonstration-eligible Medicare-Medicaid enrollees who were discharged alive from an acute inpatient setting (including acute-care psychiatric facilities) in the measurement year. The denominator for this measure is based on discharges, not members. Include all discharges for members who have more than one discharge in measurement year.

(continued)

Table 18 (continued)
Quality measures for evaluation: Detailed definitions, use, and specifications

Measure concept (specific measure)	Data sources and responsibility for data collection	Domain (prevention, care coordination, beneficiary experience)	Will evaluation produce impact estimates?*	Definition (link to documentation if available)	Numerator/denominator description
Fall prevention Screening for Fall Risk	Claims/ encounter RTI will acquire and analyze	Prevention, care coordination	Yes	Percentage of demonstration-eligible Medicare-Medicaid enrollees aged 65 years and older who were screened for future fall risk at least once within 12 months	Numerator: Demonstration-eligible Medicare-Medicaid enrollees who were screened for future fall risk at least once within 12 months. Denominator: All demonstration-eligible Medicare-Medicaid enrollees 65 years or older.
Cardiac rehabilitation Cardiac rehabilitation following hospitalization for AMI, angina CABG, PCI, CVA	Claims/ encounter RTI will acquire and analyze	Care coordination	Yes	Percentage of demonstration-eligible patients evaluated in outpatient setting who within past 12 months have experienced AMI, CABG surgery, PCI, CVA, or cardiac transplantation, or who have CVA and have not already participated in an early outpatient cardiac rehabilitation/secondary prevention (CR) program for the qualifying event/diagnosis who were referred to a CR program.	Numerator: Number of demonstration-eligible Medicare-Medicaid enrollees in an outpatient practice who have had a qualifying event/diagnosis in previous 12 months who have been referred to outpatient cardiac rehabilitation/secondary prevention program. Denominator: Number of demonstration-eligible Medicare-Medicaid enrollees in an outpatient clinical practice who have had a qualifying cardiovascular event in previous 12 months, who do not meet any of exclusion criteria, and who have not participated in out-patient cardiac rehabilitation program since cardiovascular event.
Pressure ulcers Percent of high-risk residents with pressure ulcers (long stay)	MDS RTI will acquire and analyze	Prevention, care coordination	Yes	Percentage of all demonstration-eligible long-stay residents in a nursing facility with an annual, quarterly, significant change, or significant correction MDS assessment during selected quarter (3-month period) who were identified as high-risk and who have one or more Stage 2-4 pressure ulcer(s).	Numerators: Number of demonstration-eligible long-stay nursing facility residents who have been assessed with annual, quarterly, significant change, or significant correction MDS 3.0 assessments during selected time window and who are defined as high-risk with one or more Stage 2-4 pressure ulcer(s). Denominators: All demonstration-eligible long-stay residents who received an annual, quarterly, or significant change or significant correction assessment during target quarter and who did not meet exclusion criteria.

(continued)

Table 18 (continued)
Quality measures for evaluation: Detailed definitions, use, and specifications

Measure concept (specific measure)	Data sources and responsibility for data collection	Domain (prevention, care coordination, beneficiary experience)	Will evaluation produce impact estimates?*	Definition (link to documentation if available)	Numerator/denominator description
Treatment of alcohol and substance use disorders Initiation and Engagement of Alcohol and Other Drug Dependence Treatment	Claims/ encounter RTI will acquire and analyze	Care coordination	Yes	The percentage of demonstration-eligible Medicare-Medicaid enrollees with a new episode of alcohol or other drug (AOD) dependence who received the following: a. Initiation of AOD Treatment. The percentage who initiate treatment through an inpatient AOD admission, outpatient visit, intensive outpatient encounter or partial hospitalization within 14 days of the diagnosis. b. Engagement of AOD Treatment. The percentage who initiated treatment and who had two or more additional services with a diagnosis of AOD within 30 days of the initiation visit. (http://www.qualityforum.org/QPS/)	Numerator: Among demonstration-eligible Medicare-Medicaid enrollees (a) Initiation: AOD treatment through an inpatient admission, outpatient visit, intensive outpatient encounter or partial hospitalization within 14 days of diagnosis; (b) Engagement: AOD treatment and two or more inpatient admissions, outpatient visits, intensive outpatient encounters or partial hospitalizations with any AOD diagnosis within 30 days after the date of the Initiation encounter (inclusive). Multiple engagement visits may occur on the same day, but they must be with different providers in order to be counted. Do not count engagement encounters that include detoxification codes (including inpatient detoxification) Denominator: Demonstration-eligible Medicare-Medicaid enrollees age 13 years and older who were diagnosed with a new episode of alcohol and drug dependency during the intake period of January 1–November 15 of the measurement year. EXCLUSIONS: Exclude those who had a claim/encounter with a diagnosis of AOD during the 60 days before the IESD. For an inpatient IESD, use the admission date to determine the Negative Diagnosis History. For an ED visit that results in an inpatient stay, use the ED date of service.

(continued)

Table 18 (continued)
Quality measures for evaluation: Detailed definitions, use, and specifications

Measure concept (specific measure)	Data sources and responsibility for data collection	Domain (prevention, care coordination, beneficiary experience)	Will evaluation produce impact estimates?*	Definition (link to documentation if available)	Numerator/denominator description
Depression screening and follow-up Screening for clinical depression and follow-up	Claims/ encounter RTI will acquire and analyze	Prevention, care coordination	Yes	Percentage of demonstration-eligible Medicare-Medicaid enrollees screened for clinical depression using an age- appropriate standardized tool AND follow-up plan documented. (http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/2014_eCQM_EP_June2013.zip)	Numerator: Demonstration-eligible Medicare-Medicaid enrollees whose screening for clinical depression using an age-appropriate standardized tool AND follow-up plan is documented. Denominator: All demonstration-eligible Medicare-Medicaid enrollees 18 years and older with certain exceptions (see source for the list).
Blood pressure control Controlling high blood pressure	Medical records (HEDIS EOC035)	Prevention, care coordination	No	Percentage of demonstration participants who had diagnosis of hypertension (HTN) and whose blood pressure (BP) was adequately controlled (<140/90mm Hg) during the measurement year. (http://www.qualityforum.org/QPS)	Numerator: Number of patients in denominator whose most recent, representative BP is adequately controlled during measurement year. For a member's BP to be controlled, both systolic and diastolic BP must be <140/90mm Hg. Denominator: Demonstration participants with hypertension. A patient is considered hypertensive if there is at least one outpatient encounter with a diagnosis of HTN during first 6 months of measurement year.
Weight screening and follow-up Adult BMI Assessment	Medical records (HEDIS EOC110)	Prevention	No	Percentage of demonstration participants aged 18 to 74 who had an outpatient visit and who had their BMI documented during measurement year or year prior to measurement.	Numerator: BMI documented during measurement year, or year prior. Denominator: Demonstration-eligible Medicare-Medicaid enrollees aged 18 to 74 who had outpatient visit.

(continued)

Table 18 (continued)
Quality measures for evaluation: Detailed definitions, use, and specifications

Measure concept (specific measure)	Data sources and responsibility for data collection	Domain (prevention, care coordination, beneficiary experience)	Will evaluation produce impact estimates?*	Definition (link to documentation if available)	Numerator/denominator description
Breast cancer screening	Medical records (HEDIS 0003)	Prevention	No	Percentage of women aged 40 to 69 and participating in demonstration who had a mammogram to screen for breast cancer.	Numerator: Number of women aged 40 to 69 receiving mammogram in measurement year. Denominator: Number of women aged 40 to 69 enrolled in demonstration.
Antidepressant medication management	Medical records (HEDIS EOC030)	Care coordination	No	Percentage of members aged 18 or older who were diagnosed with a new episode of major depression and treated with antidepressant medication, and who remained on an antidepressant medication treatment.	Numerator: Two rates are reported: (1) Effective Acute Phase Treatment—newly diagnosed and treated demonstration participants who remain on antidepressant medication for at least 84 days; (2) Effective Continuation Phase Treatment—newly diagnosed and treated demonstration participants who remained on antidepressant medication for at least 180 days. Denominator: Newly diagnosed and treated demonstration participants aged 18 or older.
Diabetes care Comprehensive Diabetes Care: selected components—HbA1c control, LDL-C control, retinal eye exam	Medical records (HEDIS EOC020)	Prevention/ care coordination	No	Percentage of demonstration participants aged 18 to 75 with diabetes (type 1 and type 2) who had each of the following: HbA1c control, LDL-C control, and retinal eye exam	Numerator: Number of these who had HbA1c control or LDL-C control, or retinal eye exam in the measurement year. Denominator: demonstration participants aged 18 to 75 with type 1 or type 2 diabetes.

(continued)

Table 18 (continued)
Quality measures for evaluation: Detailed definitions, use, and specifications

Measure concept (specific measure)	Data sources and responsibility for data collection	Domain (prevention, care coordination, beneficiary experience)	Will evaluation produce impact estimates?*	Definition (link to documentation if available)	Numerator/denominator description
Medication management Annual monitoring for patients on persistent medications	Medical records (HEDIS EOC075)	Care coordination	No	Percentage who received at least 180 treatment days of ambulatory medication therapy for a select therapeutic agent during measurement year and at least one therapeutic monitoring event for therapeutic agent in measurement year. Agents measured: (1) ACE inhibitors or ARB, (2) digoxin, (3) diuretics, and (4) anticonvulsants.	Numerator: Number with at least 180 days of treatment *and* a monitoring event in measurement year. Combined rate is sum of 4 numerators divided by sum of 4 denominators. Denominator: Demonstration participants with at least 180 days of treatment in year for a particular agent.

ACE = angiotensin converting enzyme; ACSC = ambulatory care-sensitive condition; AHRQ = Agency for Healthcare Research and Quality; AMI = acute myocardial infarction; ARB = angiotensin receptor blocker; BMI = body mass index; BP = blood pressure; CABG = coronary artery bypass graft; CDC = Centers for Disease Control and Prevention; CHF = chronic heart failure; COPD = chronic obstructive pulmonary disease; CR = cardiac rehabilitation; CVA = cerebrovascular accident; ED = emergency department; HbA1c = hemoglobin A1c test; HEDIS = Healthcare Effectiveness Data and Information Set; HTN = hypertension; IESD = Index Episode Start Date; LDL-C = low-density lipoprotein cholesterol; MDS = minimum data set; NYU = New York University; PCI = percutaneous coronary intervention; PQI = prevention quality indicators; UTI = urinary tract infection.

NOTE: This table does not indicate which measures may be included in payment determinations (i.e., required for savings sharing or managed care organization quality withholds).

* Impact estimates will be produced only for measures where data can also be obtained for the comparison group. Measures for which data are not expected to be available in the comparison group will be tracked only within the demonstration to measure changes over time.

8.2.3 Analytic Methods

The choice of analytic methods for quality measures will follow the same template as described above in ***Section 6***. These methods will differ depending on whether the outcome is binary or continuous. One possible refinement here is that some measures may be conditional on others—for example, the use of smoking cessation programs among those screened for tobacco use. In most cases, this will simply result in the choice of a restricted sample (those screened) on which to estimate an outcome (those using program).

8.2.4 Analytic Challenges and Proposed Solutions

The quality measures presented in ***Tables 17*** and ***18*** present some expected constraints for the evaluation due to data availability. They are consistent with the measures included in the recommendations of the National Quality Forum for measuring quality of care among Medicare-Medicaid enrollees.

The activities for analyzing utilization and access to care may be revised if modifications are made to the demonstrations or if data sources are not available as anticipated. If modifications to this evaluation plan are required, they will be documented in the annual and final evaluation reports as appropriate.

Section 9. Cost

CMS is particularly interested in learning whether State demonstrations achieve cost savings while improving or maintaining quality. The evaluation will use a multivariate, difference-in-differences regression analysis to determine the impact of the demonstration on cost in both capitated and managed fee-for-service (MFFS) models. It will examine how costs are associated with the variety of services that beneficiaries receive, including medical, behavioral health, and long-term services and supports (LTSS).

As discussed in ***Section 5***, the determination of whether cost savings were achieved can be made using two different approaches that may yield somewhat different results. The approach to be used in the evaluation, discussed in this section, uses the same descriptive and regression-based techniques as outlined earlier in this report for the analyses of quality and utilization of care. The second approach uses actuarial methods to calculate changes in costs, and will be performed annually by RTI in order to inform performance payment calculations and any resulting payments to MFFS model States.[3] The actuarial approach can be calculated on a faster timeline, accommodating the goal of estimating savings as soon as possible after the end of the demonstration performance year to inform the amount of any performance payments to MFFS States. This section addresses only the approach to cost-savings calculations that will be used in the evaluation—not the actuarial approach. RTI will use the regression-based approach to calculate the impact of the demonstration on costs for the evaluation because it provides information about how various factors relate to costs.

This section of the report outlines analyses to estimate the demonstrations' impact on Medicare and Medicaid costs using regression-based techniques usually employed in traditional demonstration evaluation activities to learn what factors contribute to cost savings or increases. For example, CMS will learn whether certain types of demonstration approaches save more money than other types of approaches, or whether costs are lower in the demonstration group compared with comparison group for certain subgroups. CMS is also interested in which types of services (e.g., inpatient, home and community-based services [HCBS]) contribute the most to cost differences between the demonstration and comparison groups as State demonstrations promote changes in utilization patterns through care management. These changes in utilization patterns may result in decreased spending for some services and increased spending for other types of services; leading to overall differences in spending for the demonstration and comparison groups. These analyses are important for deconstructing demonstration effects so that they can be replicated if desired.

[3] Note that RTI will not determine MFFS States' eligibility for performance payments, nor calculate the amount of those performance payments. RTI will perform the annual savings calculations, and the amount of any savings will be used as an input in the performance payment calculations to be conducted by CMS or other entities.

Research questions regarding cost analyses include the following:

- Do the demonstrations reduce costs?
- If so, how were the demonstrations able to reduce the costs of Medicare-Medicaid enrollees compared with the comparison group?
- How do the demonstrations differentially affect expenditures for beneficiaries at risk for having high costs?

9.1 Ongoing Monitoring

As in the utilization analyses, we will identify high-level cost measures that can be calculated for all States to monitor changes over time. For MFFS demonstration States, we will provide per-capita or per-user costs for key services (e.g., inpatient, outpatient, long-term nursing facility, mental health) from claims to understand how costs change quarterly. For capitated demonstration States, costs are the per-member per-month (PMPM) rates paid, combined with the costs for those who opt out or disenroll and receive their services under the traditional fee-for-service approach or who choose to enroll in other types of managed care organizations (e.g., Medicare Advantage plans or Medicaid managed care organizations). We will include Part D PMPM and any PMPM reconciliation data provided by CMS in the final assessment of cost impact to ensure that all data are available. Accounting for all of these types of costs is important because of the cost implications of possible selection bias.

9.2 Within-State Descriptive Analyses

We will measure predemonstration and annual spending on Medicare-Medicaid enrollees for both Medicare and Medicaid. In MFFS model States, we will look at spending by service. In capitated model States, we will look at total Medicare and Medicaid costs based on PMPMs (because contribution to blended capitation payments plus Part D spending, rather than individual services, will drive the analysis of spending by payer). For the first annual report, only Medicare costs may be available for most States. Even if Medicaid costs are available for a few States, availability of these costs may lag behind the Medicare costs by several quarters. Still, we will report Medicaid costs when available. For MFFS States, we will also assess costs for the service types shown in ***Table 19***. For capitated model States, RTI anticipates that service-level spending will not be available in the encounter data reported by MMPs, so the utilization analysis described in Section 7 will be used to understand the impact of the demonstration by type of service. Other factors, such as changes in coding intensity, could also play a role in demonstration costs, and we will consider such factors in our analysis.

We will present descriptive cost analyses in quarterly and annual reports. We will also present costs for various subgroups of interest, such as demographic groups, LTSS users, beneficiaries with intellectual and developmental disabilities, end stage renal disease (ESRD), and those with other chronic conditions or health status, such as diabetes, as

desired in annual reports. We will also test for differences across years of the demonstration. Some of these characteristics will also be used in multivariate analyses as control variables, and understanding their distributions in descriptive analyses will allow us to select the most important ones for impact analyses.

Table 19
Service categories and associated data sources for reporting cost in MFFS States

Service type	Encounter data (Medicare Advantage)	Medicaid only (FFS)	Medicare and Medicaid (FFS)
Inpatient	—	—	X
Emergency room	—	—	X
Nursing facility (short rehabilitation stay)	—	—	X
Nursing facility (long-term stay)	—	X	—
Other facility-based[1]	—	—	X
Outpatient[2]	—	—	X
Outpatient behavioral health (mental and substance use disorder)	—	X	—
Home health	—	—	X
HCBS (State Plan PAS, waiver services)	—	X	—
Dental	—	X	X
Prescription drug PMPM	—	—	X
Managed care PMPM	X	—	—

— = not applicable; FFS = fee for service; HCBS = home and community-based services; PAS = personal assistance services; PMPM= per member per month capitation payments.

[1] Includes long-term care hospital, rehabilitation hospital, State mental health facility stays.

[2] Includes visits to physician offices, hospital outpatient departments, rehabilitation agencies.

9.3 Cost Impact Analyses

Cost impact analyses using comparison groups, like the utilization analyses, will be conducted as part of the final evaluation and included in the final evaluation report.

As part of these analyses, we will test for differences between demonstration and comparison groups using both descriptive analyses and regression methods. These descriptive analyses will use many of the measures developed for the within-State analyses for various subgroups, including comparison group data to test for differences in costs between the demonstration and comparison groups.

In regression analyses, the dependent variable will be total costs for various service types for beneficiaries in MFFS States. We are unsure whether PMPM costs (Medicare Parts A and B, Medicare Part D, and Medicaid) in the capitated model States will vary sufficiently across age and sex to provide enough variation for regression analysis.

In addition to cost analyses for all Medicare-Medicaid enrollees eligible for the demonstration, MFFS demonstration States should be expected to reduce total costs for high-cost beneficiaries. Demonstration and comparison group beneficiaries will be stratified to identify the groups of beneficiaries that have traditionally been the most expensive service users in the demonstration State. High-cost beneficiaries may include those with multiple comorbidities, severe and persistent mental illness, LTSS-users, or prior inpatient and/or skilled nursing facility stays. High-cost beneficiaries will be assigned an indicator, and we will consider whether regression to the mean plays a part in any reduction of costs for these beneficiaries and how to control for it in these analyses. RTI also will conduct cost analyses exploring demonstration effects on demonstration enrollees, acknowledging that selection bias must be taken into account in interpreting the results.

Using descriptive analyses, we will compare the distribution of annual service costs for the potentially high-cost and non-high-cost groups in the demonstration and comparison groups to understand the percentage of spending attributable to potentially high-cost beneficiaries overall and by service type. Using logistic regression, we will predict the probability of being a high-cost beneficiary, controlling for beneficiary characteristics, months of participation, and mix of services.

We will also evaluate cost savings for capitated model demonstrations. We will estimate cost savings accruing to the Medicare and Medicaid programs separately in capitated model demonstrations. We will estimate cost savings twice during the demonstration using a regression-based approach and a comparison group. To determine annual total costs (overall and by payer) for these analyses, we will aggregate the Medicare and Medicaid PMPM payments paid to the MMPs, Medicare Advantage plans, and Medicaid managed care organizations; and the FFS costs for the eligible population that is not enrolled in the demonstration. If possible, we will include Part D PMPM and any PMPM reconciliation data provided by CMS. The details of this methodology are currently under development. The methodology will be reviewed and approved by the CMS Office of the Actuary.

9.4 Analytic Challenges

As discussed, the availability of complete and accurate priced encounter data is critical for service-specific cost-related information, and the timeliness of that data will determine whether it can be reported quarterly and for use in annual reports and final impact analyses. Again, we are unsure whether PMPM costs in the capitated demonstration States will vary across age and sex such that they provide sufficient variation for regression analysis. Other considerations include the timing of Medicaid data, reflecting the impact of other policy changes that may affect the cost trajectory for demonstration or comparison group beneficiaries, and determining whether and how to adjust for Medicare or Medicaid supplemental payments not included in claims data.

The activities for analyzing costs may be revised if modifications are made to the demonstrations or if data sources are not available as anticipated. If modifications to this evaluation plan are required, they will be documented in the annual and final evaluation reports as appropriate.

This page intentionally left blank

Section 10. Subpopulations and Health Disparities

Many State demonstrations will either target or place particular emphasis on enrolling specific subpopulations of Medicare-Medicaid enrollees. Several States have proposed changes to the service delivery system and care coordination intended to improve the quality of care for particular subpopulations. Examples include State demonstration designs focused on Medicare-Medicaid enrollees under age 65, or on expanded access to and improved integration of behavioral health services. Beyond these defined target groups or specific service delivery changes, it is important to understand whether the demonstrations have differential effects on subpopulations as defined by disability type, or demographic or clinical characteristics, such as cognitive status, clinical complexity, and residence (community-residing or in a residential setting). We identify potential subpopulations of interest in ***Section 10.2***, and will work with CMS to refine this list for analyses.

Exacerbating the challenges of their experience as Medicare-Medicaid enrollees, subpopulations, including beneficiaries with certain types of disabilities, clinical diagnoses, or racial and ethnic groups, often have diminished access to comprehensive coordinated medical and long-term services and supports (Agency for Healthcare Research and Quality, 2011). Because of the complexity of the services that they need, coordination of care may be particularly challenging and the quality of the care they receive may suffer as a result. A recent RTI analysis found that Medicare-Medicaid enrollees receiving Medicaid home and community-based (HCBS) waivers had higher hospitalization rates than those in nursing facilities (Walsh et al., 2010). Moreover, some of these subpopulations, such as people with multiple chronic conditions or disabilities, account for a disproportionate amount of health care costs (e.g., Anderson et al., 2011). Thus, the ability of integrated care systems to improve quality of care and reduce expenditures may depend on their effectiveness for these subpopulations.

10.1 Research Questions

The overarching research questions for subpopulations are the following: Does the demonstration have an impact on the quality of care, service utilization patterns, and the beneficiary experience for subpopulations and the costs incurred for their services, and do these effects differ from those on the overall population of Medicare-Medicaid enrollees? Thus, the research questions are basically the same as for other populations, although the mechanisms being demonstrated may be different. The analysis of subpopulations is the application of the research questions and measures developed in the other sections of this research design to subpopulations.

To answer the overarching questions, four specific research issues will be addressed and reflected in the qualitative protocol development and the quantitative analyses:

- How do the demonstrations, as implemented by the different States, address the unique needs of the subpopulations? Are there special initiatives designed to meet the needs of these populations (e.g., special care coordination efforts, new services for people with severe and persistent mental illness, or nursing facility diversion programs)? Do the demonstration States successfully implement what they proposed? Do the models that focus on subpopulations work better than those that are designed for more general populations?
- Do the demonstrations reduce expenditures and improve beneficiary experience, quality of care, and health outcomes for subpopulations? What is the effect on service use?
- Do the demonstrations reduce or eliminate undesirable disparities (e.g., between African Americans and whites) in access to care, beneficiary experience, health care utilization, expenditures, quality of care, and health outcomes?
- To the extent that the demonstrations have positive outcomes for subpopulations, what features of the demonstration account for these outcomes?

10.2 Subpopulation Selection

In addition to the demonstration eligible populations as a whole, possible subpopulation groups of interest include the following:

- Racial and ethnic groups
- People living in rural or inner-city areas
- Younger people with disabilities
- People age 65 and older
- People with severe and persistent mental illnesses (SPMI)
- People with developmental disabilities
- People with end stage renal disease (ESRD)
- People with multiple chronic illnesses
- Users of long-term services and supports (LTSS)
- High-cost beneficiaries

The RTI team will work with CMS to identify high-priority, policy-relevant populations to analyze for each State. The evaluation will not focus on all subpopulations for every State; we will need to identify certain groups for in-depth analysis in each State, and potentially analyze selected outcomes for a variety of subpopulations of interest across all States. In addition to policy importance, the choice of subpopulations will depend on the following:

- **Whether States target their demonstrations to particular populations.** For example, the Massachusetts demonstration will be designed specifically for people aged 21 to 64 with disabilities and will provide enhanced services to people with SPMI.

- **The size of subpopulations participating in the demonstration and how they are distributed across States.** For example, it is likely that large numbers of African Americans will be participating in the demonstration in New York and South Carolina, but not in Vermont or Colorado.
- **The ability of the data sets to identify the subpopulations in both the demonstration and comparison groups.** In some cases, data identifying the subpopulations will be part of existing data sets (e.g., racial/ethnic categories). In some cases, the data are available, but their accuracy or completeness is questionable (e.g., people with a diagnosis of Alzheimer's disease). In other cases, the exact data are not available, but a substitute may be possible. For example, HCBS waiver participation or receipt of personal care services indicates frailty, but does not identify all individuals with functional impairments residing in the community.

Sections 10.2.1 and ***10.2.2*** discuss two examples of how the evaluation will be targeted to beneficiaries with behavioral health conditions and individuals residing in nursing facilities.

10.2.1 Populations with Behavioral Health Conditions

People with behavioral health conditions are a subpopulation of particular interest among Medicare-Medicaid enrollees. People with behavioral illnesses have either an SPMI or substance use disorder (or both). These behavioral health conditions may be a primary disabling condition or co-occur with other chronic conditions. Using selected diagnosis codes from the International Classification of Diseases (ICD-9; and ICD-10 in the future), the evaluation team will identify individuals with behavioral illnesses using claims or encounter data focusing on SPMI and substance use disorders. Following the evaluation of preliminary data, the evaluation team will finalize the list of diagnosis codes to include in this analysis.

Similar to the analyses for each State's demonstration group, these subanalyses will also evaluate the impact of the demonstrations on quality, utilization, and access to care for medical, LTSS, and behavioral health services, and will also examine qualitative data gathered through interviews, focus groups, and surveys. Examples of the range of measures that will be examined for beneficiaries with behavioral health conditions include outpatient behavioral health services; HCBS services; new long-term nursing facility admissions for beneficiaries with SPMI; access to a full range of scheduled and urgent medical care, behavioral health services, and LTSS; beneficiary reports of improved quality of life as a result of access to the full range of services; beneficiary choice of medical, behavioral, and LTSS services and providers; beneficiary reports on satisfaction with their life; care coordination assessment processes that integrate/address health, behavioral health, and LTSS; hospitalizations for beneficiaries with SPMI; outpatient visits after hospitalization for mental illness; and initiation and engagement of alcohol and other drug dependence treatment. Results of descriptive analyses will be presented in annual reports. The final evaluation reports will include multivariate analyses.

10.2.2 Nursing Facility Residents

By aligning the Medicare and Medicaid incentives, the demonstrations have an opportunity to improve quality of care in nursing facilities, reduce potentially avoidable hospitalizations of nursing facility residents, and, through rebalancing efforts, prevent, delay or shorten facility stays. Conversely, if demonstration providers seek to achieve savings by negotiating lower-cost contracts with nursing facilities, lower quality of care could result. The evaluation will analyze nursing facility admission rates, acute-care utilization (e.g., physician visits, hospitalizations, emergency room use) and cost patterns for individuals receiving short-term skilled nursing facility care and for long-stay nursing facility residents. In addition, we will use the Nursing Home Minimum Data Set (MDS) to evaluate the level of impairment or acuity of new nursing facility entrants to evaluate the extent to which the demonstrations are succeeding in maintaining frail individuals in the community, and calculate and monitor selected nursing facility quality measures. We will monitor trends in nursing facility admissions and quality within the demonstration States (or regions within a State) and analyze demonstration impact in comparison with facilities in comparison States or regions, using multivariate techniques.

10.3 Methodology

To address these research questions, the RTI team will conduct within-State and aggregate analyses and include the results in individual State and aggregate reports.

10.3.1 Qualitative Analyses

The qualitative analyses will consist of interviews and focus groups with beneficiaries, as part of the overall approach to monitoring demonstration implementation and the beneficiary experience. The interviews with State officials, health plans with large concentrations of subpopulations, managed fee-for-service (MFFS) initiatives, and others; and the focus groups with Medicare-Medicaid enrollees will ask questions regarding any special requirements for health plans or other initiatives to specifically address the medical, LTSS, behavioral health, and other needs of subpopulations. Medicare-Medicaid Plans (MMPs) participating in capitated model demonstrations and MFFS initiatives serving minority populations and rural populations will be targeted for inclusion in interviews and focus groups. Depending on the subpopulation being analyzed, questions will include topics such as whether the MMPs or care coordination entities:

- refer beneficiaries to community services, such as Supplemental Nutrition Assistance Program and senior centers;
- have established protocols for the treatment of common medical and nonmedical problems among subpopulations; and
- have procedures to address the needs of people with low English proficiency.

We will also ask those we interview and focus group participants what features of the demonstration they believe are most effective.

10.3.2 Quantitative Analyses

The quantitative analyses will be subanalyses of those conducted for all beneficiaries, stratified by the relevant subpopulations, and quality, utilization, and access to care analyses specific to a subpopulation's needs. These analyses will assess whether, for defined subpopulations, the demonstrations improve outcomes.

The activities for examining subpopulations may be revised if modifications are made to the demonstrations or if data sources are not available as anticipated. If modifications to this evaluation plan are required, they will be documented in the annual and final evaluation reports as appropriate.

This page intentionally left blank

Section 11. Next Steps

This *Aggregate Evaluation Plan* presents the analytic design for the evaluation, including plans for qualitative and quantitative analysis, and discusses challenges that we have identified to date, such as issues related to data availability. Information about the cost-savings analyses for payment purposes in managed fee-for-service (MFFS) models has been developed and reported in a separate memo. RTI will continue to work with CMS to further refine the evaluation design. Although this document will not be revised to address all changes that may occur, the annual and final evaluation reports will note areas where the evaluation as executed differs from this evaluation design plan. Some individual State-specific evaluation plans have been finalized, whereas others are in progress and will be completed as the demonstration designs are finalized for each State.

The results of our analyses will be presented in a series of deliverables, including quarterly reports to CMS and States, annual reports, and a final evaluation report for each State as well as a final aggregate evaluation report. ***Table 20*** below highlights major remaining deliverables over the course of the evaluation as well as notes on the expected availability of data that will be incorporated into these deliverables. RTI will work with CMS to meet evaluation timelines and to incorporate data into the deliverables as they become available.

Table 20
Deliverable timeline for monitoring and evaluation activities

Deliverable	Timeline	Data issues/data included
State-Specific Evaluation Design Plans	Summer 2013 through 2014, on a rolling basis	The State-specific evaluation design plans detail the application of the overall research design for each State given the characteristics of each State's demonstration.
State-Specific Initial Reports	Reporting on the first 6 months of demonstration implementation in each State	Mainly based on qualitative data collected through interviews or other State reporting, these reports will provide information about early implementation experience. These reports will be available to CMS and each individual State.
Quarterly Reports to CMS and States	Quarterly, beginning the quarter after the State-specific initial 6-month report	The goal of these reports is to include preliminary information on enrollment, disenrollment, quality, utilization, and cost measures for ongoing monitoring in each State. These reports initially will include data reported by States in the SDRS, predemonstration Medicare fee-for-service data, predemonstration Medicaid data as available, and Medicare fee-for-service data for MFFS States. Because of potential lags in obtaining Medicaid data, we will incorporate measures using Medicaid data as they become available. We will use claims data for MFFS model States, and encounter data for capitated model States. Both claims and encounter data will be obtained from CMS. The timeliness and completeness of the encounter data are not yet known.

(continued)

Table 20 (continued)
Deliverable timeline for monitoring and evaluation activities

Deliverable	Timeline	Data issues/data included
Annual State-Specific and Aggregate Reports	Annually, for each of the demonstration performance years	These documents will summarize the material contained in quarterly reports to CMS and States, allowing additional time for data run-out to update preliminary quarterly report results and providing context for the analysis. Annual reports will provide a descriptive analysis of quality, utilization, and cost measures for the demonstrations. These reports will also include qualitative information collected during site visits, focus groups, and telephone interviews. RTI will develop separate (unlinked) MMP encounter, Medicare, and Medicaid beneficiary-level analytic files for the annual analysis. Predemonstration Medicaid data will be incorporated into the first annual report, and Medicaid fee-for-service and encounter data will be incorporated into the analysis as they become available. RTI will continue to investigate estimated timing for Medicaid data and encounter data. Consistent with the intent-to-treat approach, all individuals eligible to participate in the demonstration will be included in the annual analysis, regardless of whether they opt out of the demonstration or disenroll. Savings will be calculated at least twice during the demonstration for capitated model States using a regression-based methodology: once during the demonstration (included in the second annual report) and once after the end of the demonstration, for the final evaluation report. Savings will also be calculated annually for MFFS States using an actuarial methodology, for performance payment purposes.
Final State-Specific Evaluation Reports and Final Aggregate Evaluation Report	After the demonstration period has ended	The final State-specific reports and the final aggregate evaluation reports will contain analyses based on linked Medicare and Medicaid data files to provide beneficiary-level information on total Medicare and Medicaid utilization and spending. These data will be analyzed using a multivariate difference-in-differences method using both the intervention group and comparison group beneficiaries to provide a comprehensive understanding of the effects of the demonstration interventions on quality, utilization, and cost. The final report will also include cost-savings calculations for capitated and MFFS or alternative financial alignment models, as well as for any alternative demonstration designs. These reports will also include qualitative information collected during site visits, focus groups, and telephone interviews.

MFFS = managed fee for service; MMP = Medicare-Medicaid Plan; SDRS = State Data Reporting System; TAP files = monthly Medicare claims files.

Section 12. References

Agency for Healthcare Research and Quality: National Healthcare Disparities Report. AHRQ Publication No. 11-0005. Rockville, MD. Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services, 2011.

Anderson, W. L., Armour, B. S., Finkelstein, E. A., et al.: Estimates of State-level health-care expenditures associated with disability. Public Health Reports, 125(1): 44–51. 2011.

Lind, A., and Gore, S.: From the Beneficiary Perspective: Core Elements to Guide Integrated Care for Dual Eligibles. Hamilton, NJ. Center for Health Care Strategies, 2010.

Walsh, E. G., Freiman, M., Haber, S., et al.: Cost Drivers for Dually Eligible Beneficiaries: Potentially Avoidable Hospitalizations from Nursing Facility, Skilled Nursing Facility, and Home and Community-Based Services Waiver Programs. Waltham, MA. RTI International, 2010. https://www.cms.gov/reports/downloads/costdriverstask2.pdf.

This page intentionally left blank

Appendix A:
Site Visit Interview Protocol: State Demonstration Staff

Site visit interviews will be conducted with individuals in demonstration States to obtain their perspectives on progress to date, internal and external changes in demonstration design that have taken place or are under consideration, underlying motives for why specific decisions have been made, and the major challenges faced. Protocols and discussion guides will be tailored to each type of informant interviewed, including State demonstration staff; Medicaid officials; officials directing State programs for long-term services and supports (LTSS), behavioral supports, and aging and disability; managers of State primary care transformation initiatives where they exist; stakeholders representing segments of the demonstration's eligible populations; providers; and Medicare-Medicaid Plans. The evaluation staff will use information collected in the State Data Reporting System to refine and target the issues so that a fuller understanding of the context for demonstration implementation can be understood.

The aggregate evaluation design will synthesize the results of these interviews across the State demonstrations, with attention given to major themes that may emerge across the demonstration States or within specific models or for specific subpopulations. Factors to be examined through these interviews will include the following:

1. The most significant changes being made to State delivery systems, payment mechanisms, and administrative processes under the demonstration. A sense of which changes are most critical and important to achieving desired outcomes and which have been most challenging.
2. Changes in staffing patterns required to implement and monitor the demonstration.
3. Administrative investments made to date because of the demonstration and investments planned for the future.
4. Aspects of State demonstrations that are being implemented according to plan (policy and timetable) and which ones have been revised.
5. Factors external to the demonstrations that have positively or negatively influenced their design and implementation. Aspects of State policy environments that are enabling the demonstrations and those that impede implementation.
6. Involvement in any new CMS initiatives in the demonstration areas that might affect the demonstration. If applicable, what are they? What areas are involved? What is the timeline for these initiatives and do they overlap with the demonstration? What impact might these initiatives have on the demonstration?
7. Methods for achieving better integration of Medicare and Medicaid, especially the integration of primary, acute, behavioral, and long-term services and supports.
8. Ways in which States have had to adjust State policies and practices because of the inclusion of Medicare in this demonstration.

9. Factors influencing State decisions to phase in elements of their demonstration.
10. Changes required to care coordination structures; resistance to re-structuring care coordination, and if so, ways that States found to overcome those obstacles.
11. Measures used to define success; the most important outcomes that States expect to achieve; any early results.
12. Use of Medicare, Medicaid, and any other data to drive demonstration design, implementation, and improvements or modifications.
13. Beyond better integration of care for Medicare-Medicaid enrollees, other types of system changes that States are anticipating, such as rebalancing LTSS, transforming delivery of primary care. Strategies for achieving these goals.
14. Acceptable and realistic cost-savings targets that States are anticipating.
15. Recommendations that States have for other States seeking to reform the way care overall, as well as LTSS in particular, is provided to Medicare-Medicaid enrollees.

Appendix B:
Interview Outline for Beneficiary Experience

GOAL: To identify trends and themes concerning beneficiary experience during the planning/design phase and early implementation of demonstrations across States.

Participants: Representatives of major beneficiary groups whose constituents are served by the demonstration (e.g., representatives of independent living centers, Area Agencies on Aging, United Cerebral Palsy, Alzheimer's Associations, affiliates of the United Cerebral Palsy, National Alliance on Mental Illness, Autism Association, Brain Injury Associations, Patients' Rights Advocates, Public Guardians)

Format: One-on-one interviews of approximately 1 hour duration

A thematic analysis will be conducted based on the following issues addressed during interviews early in the implementation of State demonstrations.

1. When and how beneficiaries first learned about their State demonstration.
2. What beneficiaries understand are the goals of the demonstration.
3. Additional goals that beneficiaries identified for the demonstration.
4. The role, if any, that beneficiaries served in the development of the demonstrations.
5. The positive impact that beneficiaries expect the demonstrations to have on beneficiary experience, including health outcomes and quality of life.
6. Concerns that beneficiaries have about the potential for negative impact on beneficiary experience.
7. Opportunities that beneficiaries have had to provide ongoing guidance to managers for the demonstration.
8. Reflections and examples from beneficiaries on each of the following:
   - Beneficiaries understood their options at the time of **enrollment** into the demonstration.
   - Beneficiaries understand their **rights** under the demonstration, including the right to disenroll.
   - **Information about the demonstration** and services is easy to understand.
   - The ability to have a full **assessment** of medical, long-term care services and supports and behavioral health needs.
   - The development of a **plan of care** that reflects the beneficiary's unique strengths, challenges, goals and preferences.
   - The ability to **self-direct**.
   - The opportunity for beneficiaries to actively participate in and **share decisionmaking** about their service options.
   - **Quality of life**, including social and emotional supports and days free from pain.

- Knowledge and **communication** across providers and service settings about the beneficiary and his/her service history and needs.
- The availability of assistance to help **navigate** across services and settings.
- Options to **involve caregivers and family** in planning and service delivery.
- Access to **qualified providers**, especially for those with complex needs.
- Access to **multilingual** and **culturally competent** providers for LTSS and other types of services, including providers who understand the culture of beneficiaries with different types of disabilities (e.g., physical disabilities, severe and persistent mental illness [SPMI], intellectual or developmental disabilities [I/DD]).

Appendix C:
Preliminary Focus Group Outline

GOAL: To identify trends and themes related to the impact of the demonstrations on beneficiary experience and, where possible, to identify factors and features that contribute to that experience.

Participants: Four focus groups per state with 8–10 beneficiaries, family members, and caregivers per focus group.

The evaluation will synthesize and identify themes in each of the following areas:

1. **Beneficiaries' understanding of the demonstration and their options**
   - How beneficiaries first learned about the demonstrations
   - Ease with which beneficiaries were able to understand the demonstrations
   - Extent to which beneficiaries understood how the demonstration differed from how they formerly accessed services
   - Beneficiaries' understanding of their enrollment choices
2. **Why beneficiaries decided to enroll**
   - The main reason why beneficiaries enrolled
   - Best features of the demonstration
   - Features of the demonstration that caused beneficiaries concern
   - The availability of a neutral person to help beneficiaries in making enrollment decisions
3. **Early enrollment experience**
   - Beneficiaries understanding about their choices for
     - Selecting a PCP
     - Self-directing their own services
   - Adequacy of information to help make these choices
   - Availability of information about beneficiary rights under the demonstration, including what to do if something went wrong or if a beneficiary wanted to dis-enroll
4. **Assessment of need and plan of care**
   - The level of understanding that primary care providers have about beneficiary needs, including medical, long-term care services and supports and mental health needs
   - Opportunities that beneficiaries have to talk with their primary care providers and others about their own goals, needs, and preferences

5. **Person-centered care**
    - The degree to which beneficiaries feel they are actively included in decisions and developing their own care plans (medical, LTSS, and BH)
    - The degree to which beneficiaries feel empowered by the services they receive
6. **Beneficiary assessment of demonstration outcomes**
    - Beneficiary assessment of their health outcomes
        - Beneficiary report of quality of life, including social and emotional supports and days free from pain
7. **Care coordination**
    - The availability of a knowledgeable person that can be called upon to help beneficiaries in getting the services they need
    - The ease with which services can be accessed
    - Knowledge of primary care providers about the care received from specialists or other providers
8. **Opportunities for improvement.**
    - Areas that beneficiaries think have improved since before the demonstrations
    - Areas that beneficiaries think have gotten worse since before the demonstration