MATHEMATICA
Policy Research

REPORT

FINAL REPORT

Risk Adjustment of HCBS Composite Measures, Volume 2

September 16, 2015

Jessica Ross
Alex Bohl
Sheng Wang
Dejene Ayele

Submitted to:
U.S. Department of Health and Human Services
Centers for Medicare & Medicaid Services
7500 Security Blvd. Mailstop: C2-21-15
Baltimore, MD 21244
Project Officer: Effie George
Contract Number: HHSM-500-2010-00026I/HHSM-500-T0010

Submitted by:
Mathematica Policy Research
955 Massachusetts Avenue
Suite 801
Cambridge, MA 02139
Telephone: (617) 491-7900
Facsimile: (617) 491-8044
Project Director: Carol Irvin
Reference Number: 40137.C0F

This page left blank for double-sided copying.

ACKNOWLEDGMENTS

A number of people at Mathematica and other organizations made important contributions to this report. We would like to thank our government task lead, D.E.B. Potter, of the Office of Disability, Aging, and Long-Term Care Policy in the Office of the Assistant Secretary for Planning and Evaluation. We would also like to thank our project officer, Effie George, and Mike Smith, Director of the Division of Community Systems Transformation, Disabled & Elderly Health Programs Group, Center for Medicaid & CHIP Services at the Centers for Medicare & Medicaid Services (CMS). In addition, we would like to thank the members of the Agency for Healthcare Research and Quality (AHRQ) Quality Indicator team for their technical contributions. We also wish to acknowledge Eric Schone for his significant contributions to this report, which benefited greatly from his statistical expertise and overall guidance. In addition, Xiaojing Lin, Kerianne Hourihan, Justin Vigeant, Rachel Thompson, and Sarah Schoenfeldt all provided programming support for these calculations.

We also wish to thank the following participants in our HCBS Composite Measure Technical Expert Panel (TEP), who provided valuable guidance on the development of our risk-adjustment approach and methodology: Robert Applebaum, Ph.D., of Scripps Gerontology Center; Arlene Ash, Ph.D., of University of Massachusetts Medical School; Julie Bershadsky, Ph.D., of Human Services Research Institute; Peter Boling, M.D., of Virginia Commonwealth University School of Medicine; Alison Cuellar, Ph.D., M.B.A. of George Mason University; Lynda Flowers, J.D., M.S.N, R.N., of AARP; Sara Galantowicz, M.P.H., of Abt Associates; Teresa Johnson, M.B.A., of National Adult Day Services Association; Alice Lind, M.P.H., B.S.N., of Washington State Health Care Authority; Abby Marquand, M.P.H., of Paraprofessional Healthcare Institute; Thomas Meehan, M.D., M.P.H., of Qualidigm; Chris Murtaugh, Ph.D., M.P.A., of Visiting Nurse Service of New York; Cheryl Phillips, M.D., of LeadingAge; Jonathan Shaw, M.D., M.S., of Stanford University; and Michael Shwartz, Ph.D., M.B.A., of Boston University.

This page left blank for double-sided copying.

CONTENTS

TABLES

FIGURES

EXECUTIVE SUMMARY

This is the second of two reports describing the development of risk-adjustment models for two home- and community-based services (HCBS) composite measures that assess potentially avoidable hospitalizations due to acute or chronic ambulatory care sensitive conditions (ACSCs). This report (Volume 2) provides recommendations for how states and other stakeholders can best use the HCBS composites to assess the quality of care delivered to HCBS users and initial steps to help design quality improvement initiatives.

The Agency for Healthcare Research and Quality (AHRQ) began the development of the HCBS composite measures 10 years ago as directed by the Deficit Reduction Act of 2005. Through this process, AHRQ finalized a set of HCBS quality measures that included composite measures adapted from the Prevention Quality Indicators (PQIs) (Schultz et al. 2012), which report rates of potentially avoidable hospitalization for select acute or chronic ACSCs. These measures are intended to assess the quality of care delivered to Medicaid fee-for-service (FFS) beneficiaries using HCBS under a shared accountability framework: the measures profile the experience of the HCBS population and reflect care delivered by all community-based providers (not just HCBS providers). However, to fairly assess the quality of care provided to the HCBS population, the composites needed further methodological refinements to account for differences in age and health status across HCBS populations—achieved through statistical risk adjustment – and strategies for addressing challenges posed by small sample sizes.

To fulfill this need, the Centers for Medicare & Medicaid Services (CMS) and the Office of Disability, Aging, and Long-Term Care Policy (DALTCP) in the Office of the Assistant Secretary for Planning and Evaluation (ASPE), tasked Mathematica Policy Research (Mathematica) with building risk-adjustment models, and establishing recommendations for minimum denominator sizes, use of reliability-adjustment, and benchmarking approaches for the HCBS composites.[1] As a first step, Mathematica proposed recommendations for risk-adjustment, described in a previous report: *Risk Adjustment of HCBS Composite Measures, Volume 1* (Bohl et al. 2015b). Drawing on the guidance of our HCBS Composite Measure technical expert panel (TEP), in this report we then develop options to: (1) account for variation in the reliability of HCBS composite rates, (2) establish relevant HCBS composite benchmarks, (3) identify suitable methods to compare HCBS composites to benchmarks, and (4) report risk-adjusted HCBS composite rates for policy-relevant subgroups, such as persons who transition from institutional long-term care settings to HCBS.

The report evaluates potential methods for addressing these issues, including (1) implementation of reliability-adjustment or minimum case sizes, (2) comparing use of ranking, confidence intervals, and exceedance probabilities to identify statistically meaningful differences in results, and (3) consideration of different national and peer-group benchmarks.

[1] Mathematica is also tasked with the development of a risk-adjusted measure to assess potentially avoidable hospitalizations due to pressure ulcers in the HCBS user population. The final measure specifications and risk-adjustment models will be published in two volumes, which will be publicly available on CMS's MFP website (http://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Long-Term-Services-and-Supports/Balancing/Money-Follows-the-Person.html) by October 2015.

Based on the results of these analyses and feedback from the TEP, the following guidelines are recommended for reporting of the HCBS composite measures:

- For purposes of quality improvement, HCBS composite results should only be reported when denominators exceed a minimum sample size. A minimum of 1,200 HCBS person-years is recommended for reporting.
- Meaningful differences in performance should be determined using risk-adjusted rates with 95 percent confidence intervals.
- Benchmarks should not be pre-established; instead a flexible approach to benchmarking that allows stakeholders to identify their own peer groups is optimal.
- While providing risk-adjusted measure results is an important and necessary step for drawing equitable comparisons between states, provision of contextual information (e.g., managed care use, state HCBS spending, etc.) is equally important for interpretation of results.

These guidelines are used to report state-level HCBS composite results for HCBS users in 2009 and 2010 (Appendix D), and population-level results for Money Follows the Person (MFP) participants from 2008 to 2010, and those who recently transitioned from institutional care to HCBS outside of HCBS from 2008 to 2010. The report is also accompanied by detailed measure specifications and (SAS) programming code to calculate the risk-adjusted HCBS composites. The previous report (Volume 1) summarizes the development of risk-adjustment models, final model specifications, and risk-adjusted results for HCBS users in 2010 and 2009.

The goal of this work is to continue to develop quality measures that can be used to assess the care provided to Medicaid FFS beneficiaries receiving long-term services and supports in the community. This report, as well as other reports related to the effort to develop quality measures for the HCBS population, can be found at: http://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Long-Term-Services-and-Supports/Balancing/Money-Follows-the-Person.html.

I. INTRODUCTION

This report presents the recommended reporting framework for two composite measures that assess quality of care among Medicaid fee-for-service (FFS) beneficiaries using home- and community-based services (HCBS).[2] The development of these HCBS composite measures began 10 years ago when the Deficit Reduction Act of 2005 directed the Agency for Healthcare Research and Quality (AHRQ) to develop "program performance indicators, client function indicators, and measures of client satisfaction" for Medicaid beneficiaries receiving HCBS (U.S. Congress 2006). Subsequent work by AHRQ finalized a set of HCBS quality measures, including three composite measures adapted from the Prevention Quality Indicators (PQIs) (Schultz et al. 2012). These measures report the rate of potentially avoidable hospitalization as a result of either chronic or acute ambulatory care sensitive conditions (ACSCs), as shown in Table I.1. These HCBS composite measures monitor the occurrence of hospitalizations that should rarely occur when high quality outpatient care is provided, and as such, have been recognized by several expert panels as highly relevant to the HCBS community (Schultz et al. 2012; Davies et al. 2009).

Table I.1. Final AHRQ recommended measures

HCBS composites	Component indicators
ACSC Chronic Conditions Composite (PQI 92)	Diabetes, short-term complications (PQI 1) Diabetes, long-term complications (PQI 3) COPD (PQI 5) Hypertension (PQI 7) Heart Failure (PQI 8) Angina without procedure (PQI 13) Uncontrolled diabetes (PQI 14) Adult asthma (PQI 15) Lower extremity amputations among people with diabetes (PQI 16)
ACSC Acute Conditions Composite (PQI 91)	Dehydration (PQI 10) Bacterial pneumonia (PQI 11) Urinary tract infection (PQI 12)
ACSC Overall Composite (PQI 90)	All components from both the ACSC Chronic Conditions and ACSC Acute Conditions composites

Source: Adapted from Schultz, E., S. Davies, and K. McDonald. "Development of Quality Indicators for Home and Community-Based Services Population: Technical Report." June 2012.

Note: The individual PQIs are largely mutually exclusive, due to the utilization of the primary diagnosis field to identify qualifying numerator events. However, the PQI 16 numerator utilizes specific procedure codes in combination with a diabetes diagnosis in any diagnosis field. For this reason, the same discharge can qualify as both a PQI 16 event and a PQI 1, 3, or 14 event. The composites only flag discharges with at least one PQI component, meaning that such a discharge can contribute only once to the chronic or overall composite numerators.

ACSC = ambulatory care-sensitive condition; AHRQ = Agency for Healthcare Research and Quality; COPD = chronic obstructive pulmonary disease; PSI = Patient Safety Indicator; PQI = Prevention Quality Indicator.

[2] This report focuses on the development of the ACSC Acute and Chronic Composites developed by AHRQ; based on stakeholder feedback the ACSC Overall Composite is not included in this report.

These HCBS composites have the potential to inform states about the quality of care experienced by the HCBS user population. As state and federal governments set up performance-based payment programs, they are incorporating ACSCs as the basis of incentives to manage population health. However, before using the composites to compare the quality of care delivered to HCBS users in different states or programs, three types of methodological enhancements were needed. First, a method to account for differences in characteristics of the populations served, i.e., risk-adjustment; second, strategies to address the effect on the reliability of results due to variations in sample size – such as establishment of minimum denominator sizes or reliability-adjustment; and third, appropriate methods for comparison (incorporating statistical uncertainty, identifying relevant benchmarks and displaying the results of comparisons).

To address these needs, the Centers for Medicare & Medicaid Services (CMS) and the Office of Disability, Aging, and Long-Term Care Policy (DALTCP) in the Office of the Assistant Secretary for Planning and Evaluation (ASPE) directed Mathematica Policy Research (Mathematica) to develop a risk-adjustment methodology for these measures. As a first step, Mathematica proposed recommendations for risk-adjustment, described in a previous report: *Risk Adjustment of HCBS Composite Measures, Volume 1* (Bohl et al. 2015b). This methodology was vetted by CMS and ASPE, as well as measurement experts, clinicians and other stakeholders who participated in our HCBS Composite Measures technical expert panel (TEP). The HCBS Composite Measures TEP was convened twice, first to provide guidance on the proposed approach to risk-adjustment, and second, to review risk-adjustment results, and provide input on how to address small sample sizes, make appropriate comparisons, and display results.[3]

We also conducted research to test different options for addressing variations in reliability, incorporating statistical uncertainty in comparisons, choosing benchmarks and presenting results. The purpose of this report is to summarize the results of our testing, summarize the TEP's recommendations and present final recommendations on how to report and use the HCBS composites for the purpose of quality improvement. The remainder of this report describes:

- A summary of the data, methods, and measures used to calculate the HCBS composites
- Comparisons of strategies to address:
    1. reliability of estimates,
    2. setting benchmarks,
    3. establishing a statistical comparison framework, and
    4. displaying contextual information alongside the composites
- The HCBS Composite Measures TEP's recommendations

[3] The TEP did not recommend refinements to the statistical models or demographic or health conditions included in the risk-adjustment models; instead, the TEP focused on how to use and report the risk-adjusted HCBS composites to the target audiences: states and other HCBS stakeholders.

- Results for the following populations using the recommended approach:
    1. Medicaid beneficiaries using HCBS in 2010 (Appendix D)
    2. Medicaid beneficiaries using HCBS in 2009 (Appendix D)
    3. Medicaid beneficiaries transitioning to HCBS from institutions during 2008 to 2010 either through or outside of the Money Follows the Person (MFP) Demonstration
- Conclusions and technical resources

This page left blank for double-sided copying

II. DATA AND MEASURES

A. Analytic populations

The development population for this work utilized data for Medicaid beneficiaries using HCBS in 2010, which is the most recent year for which the required Medicare and Medicaid data are available for nearly all states. The 2010 HCBS user population includes persons enrolled in HCBS 1915(c) waiver plans or using HCBS state plan or 1915(c) waiver services at any point during 2010.[4] This population includes HCBS users who are enrolled only in Medicaid, as well as those eligible for both Medicare and Medicaid (referred to as Medicare–Medicaid eligible, or MME). The data are derived from Medicare and Medicaid administrative data, including the Medicaid Analytic eXtract (MAX) Person Summary (PS), Other Services/Therapies (OT), and Long-term Care (LT), and Inpatient (IP) files, Medicare Beneficiary Summary File (MBSF), and Medicare Part A (from the Medicare Provider Analysis and Review (MedPAR) files)[5], and B claims data available on the Chronic Conditions Data Warehouse (CCW).[6]

In alignment with AHRQ's recommended specifications, we imposed several important exclusions on these populations (Schultz et al. 2012). We excluded both Medicaid managed care and Medicare Advantage enrollees, because their claims are either unavailable or incomparable to those for beneficiaries enrolled in fee-for-service programs. The population is also limited to HCBS users who are age 18 or older as of January 1, 2010. Finally, we excluded people with a record of HCBS enrollment only (that is, no observed HCBS claims) and at least one month with an institutional claim for long-term care. This step removes individuals who are enrolled in HCBS 1915(c) waivers but are only receiving institutional long-term services and supports (LTSS) during the period of interest.

The same overall analytic approach was used to create the populations of 2009 HCBS users, MFP participants, and Medicaid beneficiaries transitioning to HCBS outside of MFP; however the MFP population was created utilizing MFP administrative files and the non-MFP transitioners were selected by identifying HCBS use that followed 181 or more days of institutional long-term care.[7]

[4] HCBS 1915(c) waivers include aged/disabled, aged only, disabled only, traumatic brain injury, HIV/AIDS, intellectually disabled/developmentally disabled, mental illness, technologically dependent, an unspecified waiver, or autism. HCBS 1915(c) or state plan services include personal care, at-home private duty nursing, adult day, home health of at least 90 days, residential care, at-home hospice, rehabilitation, case management, transportation, or durable medical equipment.

[5] For additional information on these data files see the Centers for Medicare & Medicaid Services (CMS) Research Data Assistance Center (ResDAC) at http://www.resdac.org/.

[6] For additional information see the Chronic Conditions Data Warehouse (CCW) at www.ccwdata.org/.

[7] Until March 23, 2010, Medicaid beneficiaries needed at least 181 days of institutionally-based long-term care to be eligible for the MFP program. After this date, the requirement decreased to at least 91 days of institutionally-based care. The non-MFP transitioner population includes individuals who utilized HCBS following either 181 days of institutional care (through March 31, 2010) or 91 days of institutional care (from April 1, 2010 onward) to increase comparability with the MFP population.

A more detailed description of the analytic population was described the previous report, *Risk Adjustment of HCBS Composite Measures, Volume 1* (Bohl et al. 2015b); in addition, technical details on the variables used to create the analytic files are included in the measure calculation package available at http://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Long-Term-Services-and-Supports/Balancing/Money-Follows-the-Person.html.

B. Measure definitions

1. Observed (Unadjusted) Composite Rate

The observed (unadjusted) composite rate for the time period of interest is calculated as the number of qualifying inpatient admissions divided by the number of months of HCBS use, i.e.,

$$\frac{\textit{Number of qualifying inpatient admissions during HCBS months}}{\textit{Total number of HCBS months}}$$
$$= \textit{Rate of all events during HCBS months.}$$

This rate will include qualifying inpatient admissions from HCBS users who are admitted to the hospital once, as well as admissions from those who are admitted to the hospital multiple times during the period of interest. For ease of discussion, we multiply rates by 12 to generate rates in person-years. In addition, we multiply rates by 100,000 to present the HCBS composites with units of ACSC events per 100,000 person-years.

The denominator is calculated by summing the total number of months during the period of interest when eligible Medicaid beneficiaries were either enrolled in or using HCBS 1915(c) waivers or state plan HCBS.

The numerator includes the total count of inpatient acute care hospital admissions with diagnosis or procedure codes meeting the criteria for any of the component measures (Table I.1). These specifications are taken from the AHRQ Prevention Quality Indicators (PQIs) software version 4.4. Admissions that meet the criteria for multiple component measures are counted only once in the composite numerator.[8] To better attribute events to the HCBS care experience, Mathematica imposed an additional restriction so that qualifying admissions are included in the numerator only if the admission date occurs during a month of HCBS use. In the event that an HCBS user is transferred between acute care settings, the second stay (the “transfer in”) is excluded from the analysis, to align with AHRQ’s specifications (Schultz et al. 2012).

2. Risk-Adjusted Composite

The risk-adjusted composite is calculated using indirect standardization, which takes results from a standard population to derive predicted results for a population of interest, given the presence of certain risk factors. A zero-inflated negative binomial (ZINB) model is used to

[8] The individual PQIs are largely mutually exclusive, due to the utilization of the primary diagnosis field to identify qualifying numerator events. However, the PQI 16 numerator utilizes specific procedure codes in combination with a diabetes diagnosis in any diagnosis field. For this reason, the same discharge can qualify as both a PQI 16 event and a PQI 1, 3, or 14 event. The composites only flag discharges with at least one PQI component, meaning that such a discharge can contribute only once to the chronic or overall composite numerators.

generate predictions. This model employs two components that correspond to two zero generating processes: (1) a logistic regression for a binary distribution that generates zeros, and (2) a negative binomial count model to predict the number of ACSC events, some of which may be zeros. When predictions from these models are combined, it produces an expected number of ACSC events per person.

3. Specification of the ZINB

Assume that $Y_{ACSC} = Y_{ACSC,1} \times Y_{ACSC,2}$, where the binary component $Y_{ACSC,1}$ and negative binomial component $Y_{ACSC,2}$ are independent and follow the following distribution.

Binary component $Y_{ACSC,1}$: $\text{Logit}(P(Y_{ACSC,1} = 1 \mid X_{aj}, Z_j)) = \beta_{a1,0} + \sum_{j=1}^{n} \beta_{a1,j} X_{aj} + \sum_{j=1}^{6} \gamma_{a1,j} Z_j$ and $P(Y_{ACSC,1} = 0 \mid X_{aj}, Z_j) = 1 - P(Y_{ACSC,1} = 1 \mid X_{aj}, Z_j)$

Negative binomial component $Y_{ACSC,2}$: Negative binomial distribution. In particular,

$$P(Y_{ACSC,2} = k) = \binom{k + r_a - 1}{k} p_a{}^k (1 - p_a)^r \text{ for k=0, 1, 2, …,}$$

where $E(Y_{ACSC,2}) = \frac{p_a r_a}{1 - p_a} = \beta_{a2,0} + \sum_{j=1}^{n} \beta_{a2,j} X_{aj} + \sum_{j=1}^{6} \gamma_{a2,j} Z_j$, r_a is the shape parameter and $p_a / (1 - p_a)$ is scale parameter for the negative binomial distribution that $Y_{ACSC,2}$ follows. Definitions of terms are as follows:

- Y_{ACSC} is the count of HCBS composite events per person
- ACSC denotes the type of composite (chronic or acute)
- The subindex α denotes the coefficients (β or γ) or set of risk factors (*X* or *Z*) used in acute or chronic models
- The subindex *j* denotes the index for the coefficients or set of risk factors used in acute or chronic models
- *k* denotes the positive integer values taken on by the ACSC composite

Additional details regarding the model development process and final coefficients are included in *Risk Adjustment of HCBS Composite Measures, Volume 1* (Bohl et al. 2015b)

The final risk-adjusted composite rates are produced by dividing the observed number of acute or chronic events by the model-predicted number of events, creating an observed-to-expected (O/E) ratio. Using this method, which is also used by AHRQ to produce PQI area-level results, we produced risk-adjusted HCBS composite rates for each state. The process for this calculation was:

1. For each state, sum the observed number of ACSC events separately across all MME HCBS users and all Medicaid-only HCBS users to yield the observed count of events for each population.
2. For each state, sum the predicted number of ACSC events separately across all MME HCBS users and all Medicaid-only HCBS users to yield the predicted count of events for each population.
3. For each state and population, divide the total number of observed and expected events calculated in steps 1 and 2 above.

Instead of transforming the O/E ratio into an indirectly-standardized rate, we can use the O/E ratio directly to assess state performance. An O/E ratio below 1.0 indicates that a state is performing better than average, and a ratio above 1.0 indicates worse-than-average performance. However, the point estimate alone is insufficient to understand and interpret a state's relative performance to a benchmark of interest (e.g., national MME rate, or peer-state); additional consideration must be given to the impact of small sample sizes and statistical uncertainty as discussed in the subsequent sections of this report.

III. POTENTIAL STRATEGIES TO IMPROVE RELIABILITY OF ESTIMATES

The HCBS composites provide information about the experience of HCBS users, but it is important to recognize and address their limitations for the purpose of assessing performance and guiding quality improvement. First, similar to all quality measures, the HCBS composite results should be interpreted as *estimates* of the average experience of an HCBS user in a given state. Statistical uncertainty (random error) around these estimates results from several sources:

- **Measurement error**: Because the numerator events included in the composite are captured using claims data, it is possible that some events are missed, while other events are wrongly attributed as numerator events.
- **Estimation variance**: The number of HCBS users and duration of HCBS use (months) varies markedly by state: whereas California had 390,239 users in 2010, Tennessee only had 234 (Bohl et al, 2015). As a result, Tennessee's rates are subject to random fluctuations to a greater extent than California's. When HCBS users are further separated into smaller policy-relevant subgroups, this challenge is further magnified.

When statistical uncertainty is large compared to variation in the results, the measure is unreliable, and comparisons do not provide useful information. Two methods that increase the reliability of estimates were explored by Mathematica:

- **Reliability adjustment**. Reliability adjustment is the process of removing statistical "noise" or random error from measure results in order to produce more accurate comparisons between entities of interest such as states. The general approach is to shrink risk-adjusted rates toward an overall mean, with the degree of shrinkage depending on the amount of variability in the data. In particular, imprecise risk-adjusted rates (that is, those with large standard errors, often due to small sample size) will be shrunken toward the overall mean to a greater degree. The gain in precision achieved by reliability adjustment can be substantial, yielding estimates that are more stable over time.
- **Establishing a minimum case size**. This process establishes a minimum threshold below which results are not reported, so that measure results with relatively higher random error or estimation variance cannot be used to draw incorrect conclusions about performance. States or subgroups of HCBS users with sample sizes below the threshold will not have results reported.

In the next sections, we outline potential approaches to implement these recommendations for the HCBS composite, which were discussed with the HCBS Composite Measures TEP.

A. Reliability adjustment of HCBS composites

Numerous approaches are available to reliability adjust the HCBS composites (Ash et al. 2012; Wang et al. 2014). Reliability adjustment is often performed using Bayesian shrinkage estimator models (also referred to as "smoothing" or "stabilizing" models), where information is borrowed from other sources, such as the published evidence, other observations from the same or similar dataset, or observations from a different time period.

To reliability adjust the HCBS composites, our preferred approach is the one currently used for AHRQ's Quality Indicators (QI). In this two-stage approach, the HCBS composites are risk-adjusted at the first stage and reliability-adjusted at the second stage (AHRQ, 2011). The advantage of this approach is that it is familiar to those using who use the QI measures, is relatively easy to implement, and only requires information from the analytic sample. The disadvantage of this approach is that it only incorporates information from the analytic sample (as opposed to historical or published information) and makes assumptions about the underlying distribution of the data that may not apply to the HCBS composites.

In the two-stage approach, the shrinkage target and the signal variance must be identified:

- The shrinkage target is known as the "prior" in Bayesian statistics. The prior represents the best estimate of the HCBS composite with limited information, which is often the mean HCBS composite across all states, or the mean of a subgroup of states or HCBS users.
- The signal variance is the variation in state-level HCBS composite rates that is due to differences in performance. In a two-stage model, the signal variance is used to calculate the reliability weight, which is based on the signal-to-noise ratio. Noise variance estimates are based on the number of HCBS users in the state and the expected number of events. States with low reliability weights have their composites pulled closely to the shrinkage target.

We tested several statistical models to model the total number of admissions: normal distribution, Poisson distribution, binominal distribution (Appendix B). We found the reliability adjusted rates are robust to different choices of distributions and the Poisson distribution slightly outperforms other two distributions. To understand how reliability adjustment affects the chronic composite results, we examined 2010 Medicaid-only HCBS users (Table III.1). Because the Medicaid-only population is only 25 percent of all FFS HCBS users in our sample, we anticipate that reliability adjustment will increase the stability of composite rates. After shrinking to the HCBS chronic composite average of Medicaid-only HCBS users, 42 of 47 state composite rates move by less than 5 percent (not shown). However, the maximum difference in state rates before and after adjustment was 578 percent. This state has a small Medicaid-only HCBS user population that has, on average, a much lower predicted risk of chronic events than other states. By reliability-adjusting, we reduce the influence of random error on this state's rate. The decrease in statistical uncertainty among all states is also illustrated by the reduction in the standard error of 38 percent, on average.

Table III.1 Effect of reliability adjustment on Medicaid-only chronic composite rate and standard error

Relative Difference	Chronic Composite	Standard Error
Mean	11.5%	-38.8%
Minimum	-29.8%	-79.2%
25th percentile	-0.4%	-26.6%
Median	-0.1%	-37.0%
75th percentile	0.5%	-45.1%
Maximum	578.7%	-9.1%

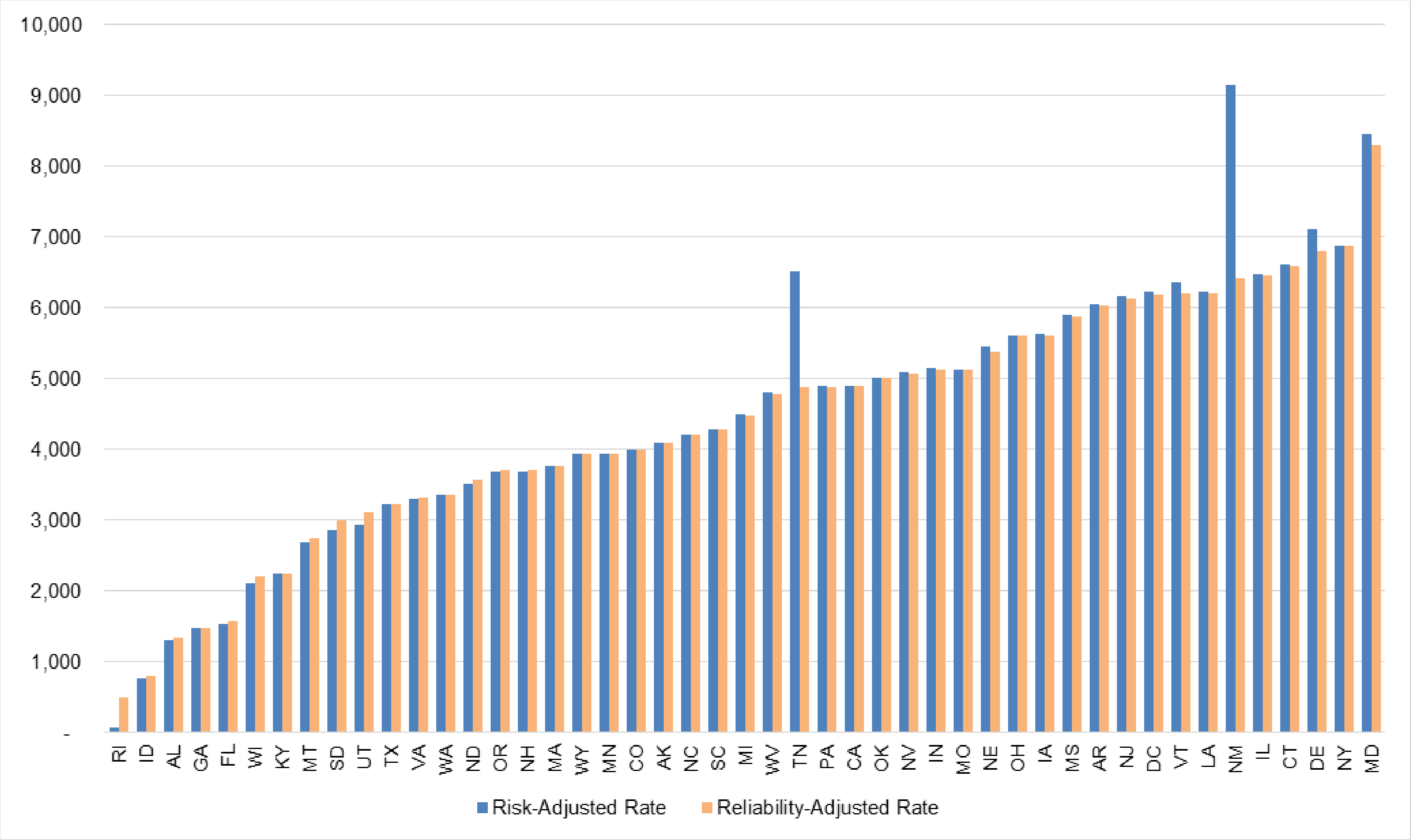
Source: Mathematica analysis of 2010 HCBS users. Data sources included the 2010 MAX PS, OT, and IP files, MedPAR file, MBSF, and CCW conditions.

Note: Reliability adjustment model assumes the following: (1) model follows a Poisson-Normal setup, (2) shrinking to the Medicaid-only mean in 2010 across all states. Relative difference is defined as the reliability adjusted estimate less the risk-adjusted estimate, divided by the risk-adjusted estimate.

Our results show that reliability adjustment, performed by Markov Chain Monte Carlo (MCMC) method under a Bayesian framework, confers the benefit of reducing the risk of drawing false conclusions due to low reliability of state estimates. For example, without reliability-adjustment, chronic composite results for the Medicaid-only populations in Tennessee and New Mexico appear much higher than most other states (Figure III.1). After reliability-adjustment, results for these states are no longer marked outliers in our sample.

Reliability adjustment is an important tool, particularly for scenarios when it is desirable or necessary to produce results for all entities of interest (e.g., states, hospitals, etc.). However, this comprehensive reporting is not the intended use for the HCBS composite measures, which are intended to provide states and stakeholders with information about their own performance that could help drive quality improvement. In addition, because almost all states have sufficient HCBS MME and Medicaid-only user populations to generate risk-adjusted rates that have a signal-to-noise ratio of at least 0.8, the potential benefits of reliability adjustment will be concentrated in a few states with fairly small numbers of HCBS users. Moreover, given that the states with small number of FFS HCBS users are likely to be highly specialized subpopulations (e.g., Tennessee FFS HCBS users are almost exclusively persons with intellectual disabilities), these states may not be comparable to other states. Instead, it may be preferential to impose a minimum case size.

Figure III.1 Risk-Adjusted and Reliability-Adjusted Chronic Composite Results for Medicaid-only 2010 HCBS users



Source: Mathematica analysis of 2010 HCBS users. Data sources included the 2010 MAX PS, OT, and IP files, MedPAR file, MBSF, and CCW conditions

B. Establishing minimum case sizes

The benefits of establishing a minimum case size are that (1) results are comparably easy to calculate and interpret, and (2) the intended end use of these composites (quality improvement) does not require all states to have a composite rate.

To establish a minimum case size, we used a power calculation to detect a 10 percent difference with 95 percent certainty. A 10 percent difference is a subjective choice, but we rationalize this decision by noting roughly an 80 percent difference in the inter-quartile range for most HCBS composites.

Using the standard equation for a power calculation for a difference in rates, i.e.

$$n = \frac{2\sigma^2 (Z_\beta + Z_{1-\alpha})^2}{(p_1 - p_2)^2}$$

Where $(p_1 - p_2)$ is the desired difference, σ is the standard deviation, $Z_{1-\alpha}$ is $100(1-\alpha)$ th percentile of the standard normal distribution with mean 0 and variance 1, $\alpha = 0.05$ is the level of statistical significance, and $\beta = 0.8$ is the desired power (Fleiss et al, 2003). We find that the minimum case size to calculate the HCBS composite is 1,200 HCBS users.

Using this approach, Tennessee should be excluded when calculating the HCBS composites for 2010 MME HCBS users, while Tennessee, New Mexico, Delaware and Wyoming should be excluded for 2010 Medicaid-only HCBS users. No states met this minimum case size standard for their MFP participants, while among non-MFP transitioner populations, only four states - California, Missouri, New York and Ohio – should have results reported for the MME transitioner population, and only New York should have results reported for the Medicaid-only transitioner population.[9]

[9] Risk-adjusted composite results could only be calculated for MFP and non-MFP participants that could be matched to CCW risk factor flags.

This page left blank for double-sided copying

IV. APPROPRIATE BENCHMARKS FOR THE HCBS COMPOSITES

The HCBS composites will be used by states and other stakeholders to guide quality improvement efforts, not establish payments or profile individual HCBS plans. In order for states to assess their baseline performance, meaningful benchmarks are needed. Because of the end use and the diversity of the HCBS user population, multiple benchmarks might be needed. In this section, we discuss potential benchmarks that were presented to the TEP for their input.

The following approaches to benchmarking are currently used in CMS programs, or in the published quality measurement literature:

National rate. The national rate is commonly used as a benchmark for provider-level quality measurement reporting, such as CMS' Hospital Inpatient Quality Reporting program. For the Medicaid-only population, the national Chronic HCBS composite risk-adjusted rate is 4,583.77 per 100,000 HCBS months, and 20 of 39 states with reportable results exceed this rate.

Achievement thresholds. Another approach is to set an achievement threshold, which implies that because a certain level of performance has been attained by at least some groups (e.g., providers, states, etc.), it should be attainable for all groups. For example, the AHRQ National Disparities Report utilizes the mean of the top five reporting states as an achievement threshold (AHRQ, 2013). The mean Chronic HCBS composite value for the Medicaid-only population among the top five states is 2,004.83 per 100,000. Twenty-five of 39 states with reportable results have rates at least twice this value (indicating worse outcomes); nine states have results three times or greater.

Penalty thresholds. As an incentive to spur poor-performers to improve, some payment programs use a set penalty threshold, such as cut-off for the worst performing quartile in the Hospital-Acquired Condition (HAC) Reduction Program. In the next section, we illustrate that 19 states have a non-zero probability of exceeding the 80th percentile value for the HCBS Chronic composite value among Medicaid-only HCBS beneficiaries.

Peer group rates. Peer-group benchmarks can also be established as the mean performance for comparable states, but defining a peer group is somewhat subjective. For example, peer groups could be comprised of states in a given region, those with similar HCBS program characteristics, or other policy-relevant characteristics. This approach was preferred by the members of our HCBS Composite Measure TEP due to the high degree of variation in the characteristics of state Medicaid programs and policies.

Population-specific benchmarks. Implementing population-specific benchmarks would impose a further level of stratification on the HCBS population, beyond MME status. For example, a national rate might be reported for persons with intellectual disabilities. Population-specific benchmarks may be desirable because they are more actionable from the state's perspective; however, dividing HCBS users into smaller subgroups decreases the reliability of estimates, as discussed previously.

This page left blank for double-sided copying

V. USING THE HCBS COMPOSITES IN STATISTICAL COMPARISONS

To gauge the best approach to statistical adjustment, we asked the TEP to address three common approaches to performance assessment: (1) ranking the entity's rate compared with that of its peers, (2) testing the significance of the difference between the entity's rate and a benchmark, where significance is measured using the confidence interval around the rate, and (3) expressing the likelihood that the entity is different from a benchmark (Shwartz 2014) (Table V.1). Each of these have a different benefits and lead to different interpretations, as discussed in detail in the *Proposed Methods for Developing and Testing Risk- and Reliability-Adjustment Models for HCBS Composite Measures* report (Bohl et al. 2015).

Table V.1. Summary of common methods for evaluating performance

Method	Description	Interpretation of lower rate[a]
Ranking	Ordering states based on their rates without making statistical inference	State A has the lowest rate, but this ranking may be due to chance
Performance categorization	Distinguishing which states are statistically different from a benchmark without reference to the magnitude of the difference	There less than a 5 percent chance of observing such a low rate for State A if its true quality is no different from average
Exceedance probability	Articulating the degree to which rates differ from a benchmark	State A has a 95 perfect probability of being lower than the benchmark

[a] This example is for interpreting results for a state with the lowest HCBS composite rate.

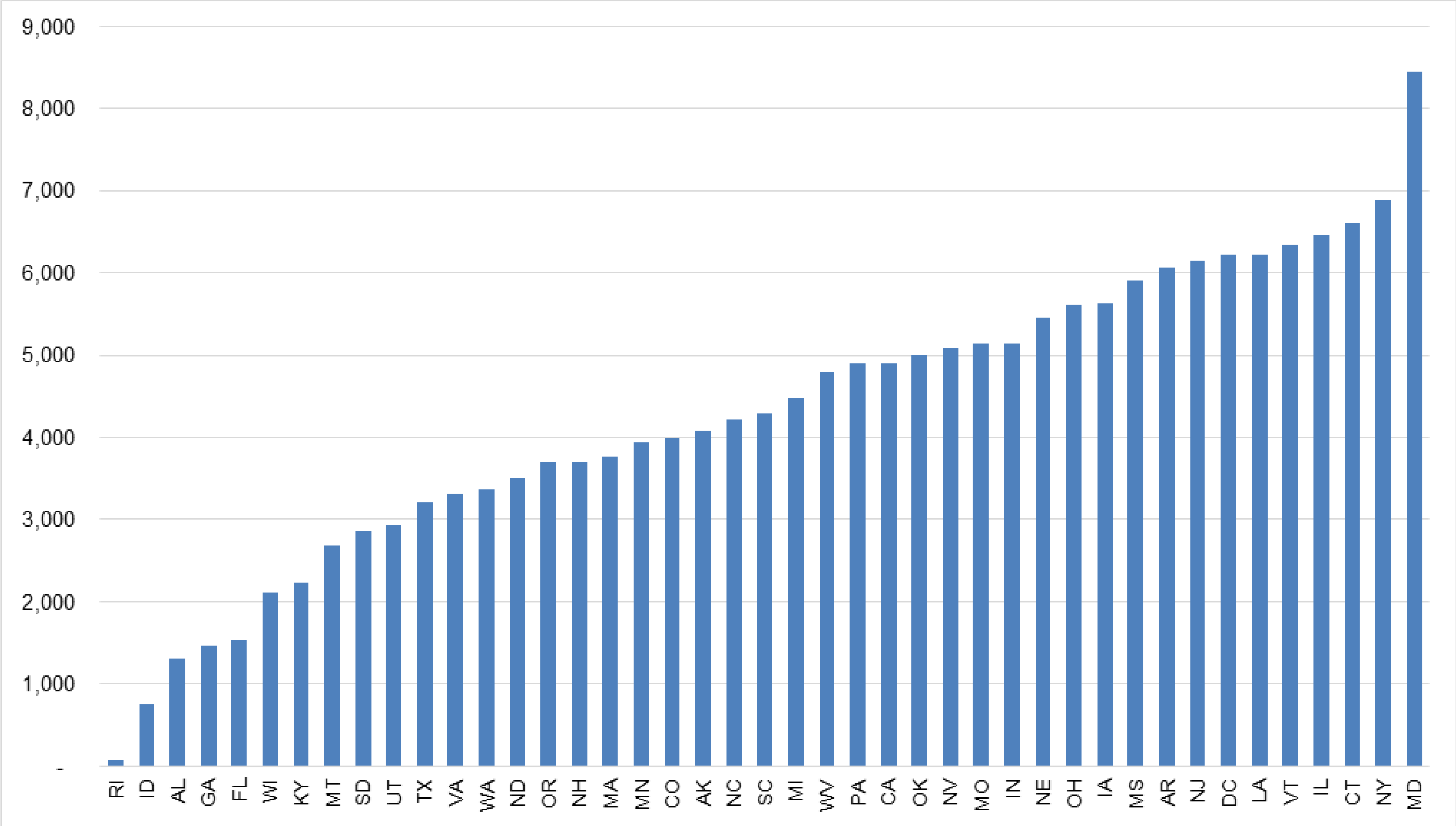
To compare the impact of using simple ranking, performance categorization, or exceedance probabilities to assess performance, we applied each approach to the risk-adjusted chronic composite results for the Medicaid-only population (Figures V.1 – 3). In the simple ranking approach, results are displayed from lowest to highest, but without any inference about whether Rhode Island truly has a lower rate than Maryland. In Figure V.2, we include 95 percent confidence intervals around risk-adjusted rates, which allows us to conclude that the likelihood of these states having the same performance is less than 5 percent.

In the Exceedance Probability example (Figure V.3), we compare the risk-adjusted chronic composite results for Medicaid-only HCBS users to a common benchmark for all states: the 80th percentile value of the HCBS composite among Medicaid-only users (which is considered a population-specific benchmark). While this approach focuses on states with higher rates relative to the benchmark (an undesirable outcome), it is easily adapted to identify states with rates lower than the benchmark. The identification of states with higher rates (worse performance) might identify the states where quality improvement is desired, while states with lower rates (better performance) might provide useful models for other states. Complete technical details for calculating Exceedance Probability estimates are described in Appendix C.

Figure V.3 illustrates how the information provided by the exceedance probability differs from that provided through a performance categorization approach (95 percent confidence intervals) that also uses the same national benchmark (the 80th percentile value). As shown by the orange bars, the performance categorization approach would identify only three states as being significantly higher than the benchmark: Illinois, Maryland, and New York. However, as indicated by the blue bars, the exceedance probability approach shows that Connecticut also has

a very high probability of exceeding the benchmark, and four other states (New Jersey, Vermont, DC, and Louisiana) have greater than 50 percent probability of exceeding the benchmark.
The exceedance probability provides more nuanced information than confidence intervals by articulating the degree to which rates differ from a benchmark, rather than just whether it's likely to differ from the benchmark with at least 95 percent confidence. Reporting a p-value from a hypothesis test would also provide similar information regarding the degree of confidence that two rates are statistically different.

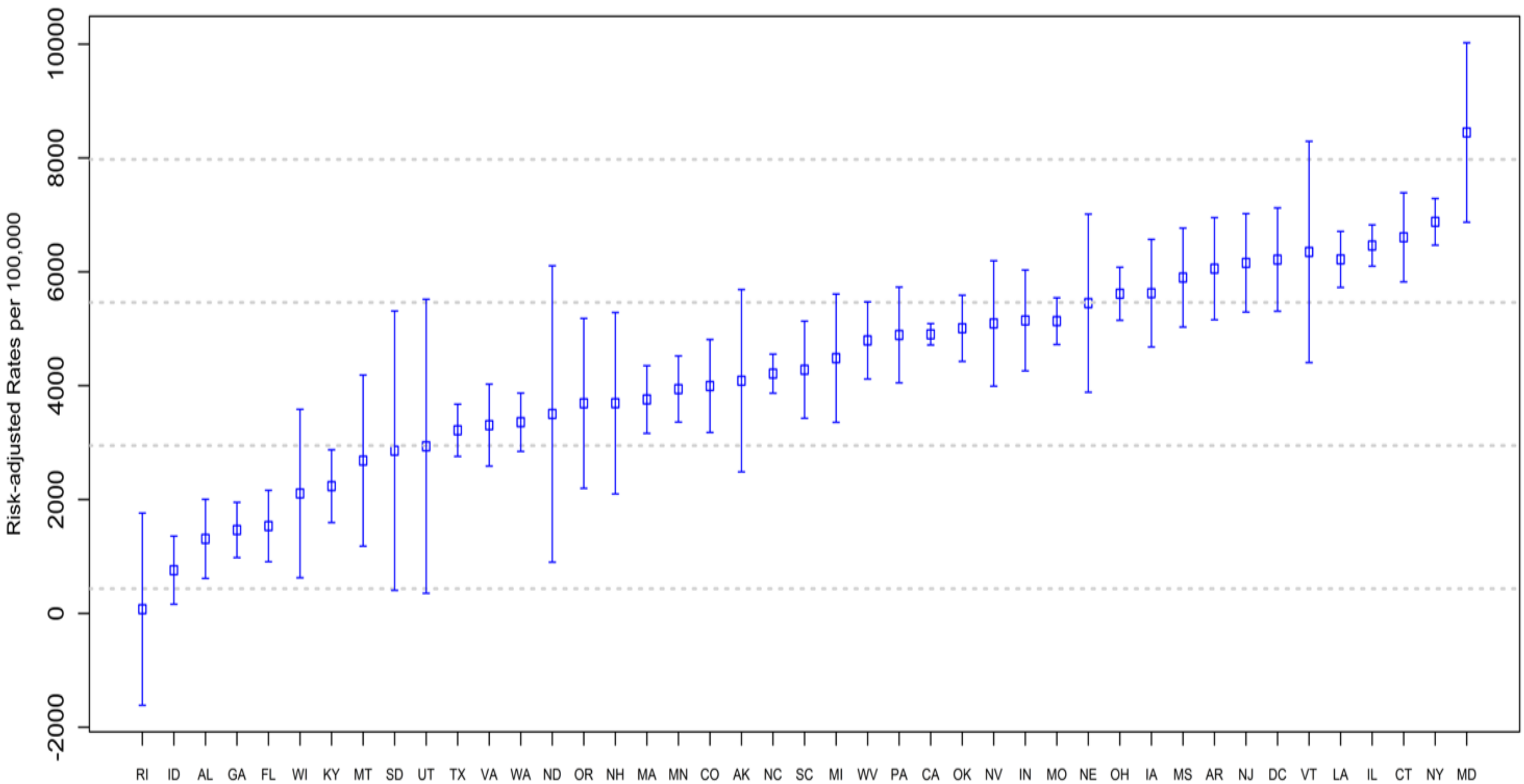
Figure V.1 Simple Ranking of Risk-Adjusted Chronic Composite Results for Medicaid-only 2010 HCBS users



Source: Mathematica analysis of 2010 HCBS users. Data sources included the 2010 MAX PS, OT, and IP files, MedPAR file, MBSF, and CCW conditions

Note: Results for Delaware, New Mexico, Tennessee, and Wyoming are omitted due to minimum sample size requirements.

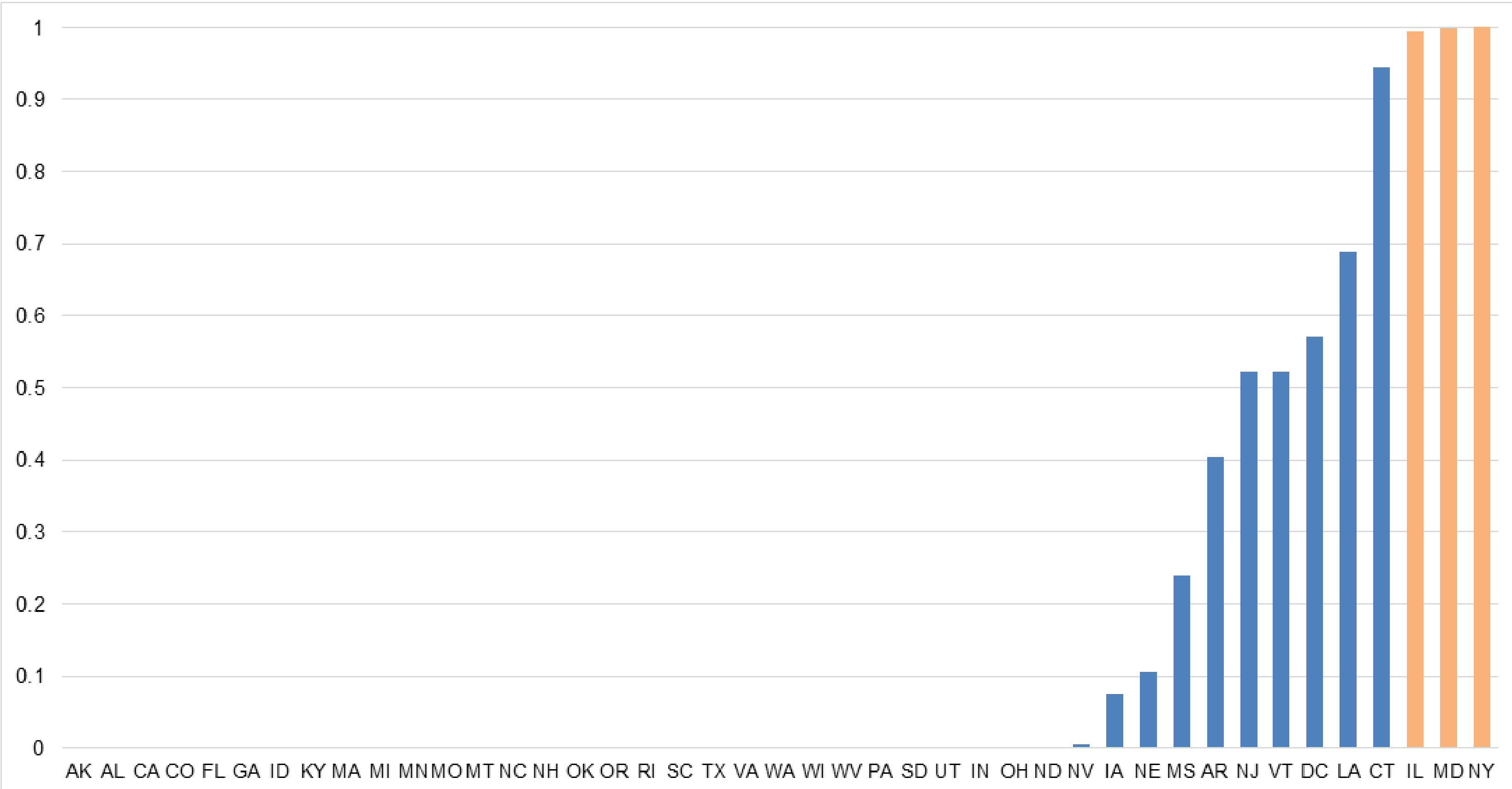
Figure V.2 Performance Categorization of Risk-Adjusted Chronic Composite Results for Medicaid-only 2010 HCBS users



Source: Mathematica analysis of 2010 HCBS users. Data sources included the 2010 MAX PS, OT, and IP files, MedPAR file, MBSF, and CCW conditions

Note: Results for Delaware, New Mexico, Tennessee, and Wyoming are omitted due to minimum sample size requirements.

Figure V.3. Probability that Risk-Adjusted Chronic Composite Result Exceeds Population-specific Benchmark for Medicaid-only 2010 HCBS users



Source: Mathematica analysis of 2010 HCBS users. Data sources included the 2010 MAX PS, OT, and IP files, MedPAR file, MBSF, and CCW conditions

Note: Blue or orange-shaded bars indicate states with at a non-zero probability of being greater than the 80th percentile value; orange-shaded bars have a 95 percent or greater probability of being greater than the 80th percentile value. Results for Delaware, New Mexico, Tennessee, and Wyoming are omitted due to minimum sample size requirements.

This page left blank for double-sided copying

VI. DISPLAY AND CONTEXTUAL INFORMATION

While risk-adjustment, statistical uncertainty, setting benchmarks, and performance frameworks play an important role in facilitating state-to-state comparisons, additional contextual information is crucial for interpreting these measure results. For example, these measures focus on the Medicaid fee-for-service population, and do not consider those enrolled in managed care. It is helpful to understand the proportion of HCBS users enrolled in FFS to determine the extent to which these measures reflect the complete HCBS care experience within a state. Such information would ideally be incorporated into a dynamic display, such as a website, that users could manipulate to create comparisons of interest. At a minimum, the following information should be displayed with composite results:

1. The number of FFS HCBS users on which the composite estimates are based
2. The proportion of all HCBS users in the state that are included in the estimates
3. An overall measure of risk for HCBS composite events, such as the expected rate
4. Statistics related to Medicaid LTSS policies in the state, such as per-person spending or the proportion of the LTSS population using HCBS
5. Statistical uncertainty estimates for the HCBS composite, such as the 95 percent confidence interval (performance category approach) or the percentile cutoffs from the posterior distribution (exceedance probability)
6. Benchmarks, with the default as the MME and Medicaid-only population specific means for the acute and chronic composite

Other sources of information also provide important insights into a state's long-term services and supports environment. For example, the AARP LTSS Scorecard (http://www.longtermscorecard.org/) provides state-level indicators of LTSS performance across five dimensions: 1) affordability and access, 2) choice of setting and provider, 3) quality of life and quality of care, 4) support for family caregivers, and 5) effective transitions. While these metrics would be difficult to incorporate into a single display, they may help inform quality improvement efforts.

This page left blank for double-sided copying

VII. HCBS COMPOSITE MEASURES TEP GUIDANCE

On May 18th, 2015, Mathematica Policy Research (Mathematica) convened the second meeting of the HCBS Composite Measures TEP to solicit input on the continued development of the HCBS composite measures. During this meeting, Mathematica presented information regarding the characteristics of the HCBS population, approach to model development and testing, recommended model structure and included risk factors, and risk-adjusted results for the 2010 HCBS population. The TEP was asked to provide input on the composite definitions, methods used for risk-adjustment testing and validation, final models, and face validity of state-level results (detailed in Volume 1). They provided the following feedback to Mathematica:

- The proposed HCBS measure definitions, risk-adjustment models, and methods of testing and validation were approved.
- The final HCBS composite measures should report results for the MME and Medicaid-only populations separately, and not include an aggregated measure that combines these two populations.
- Further development of the overall HCBS composite is not warranted. Results indicate that the important risk factors for the acute and chronic composites are markedly different, and states sometimes had discordant performance on the two measures. The combination of these two measures into an overall composite could potentially obscure these important findings.

In addition, the group discussed strategies for addressing uncertainty of rates, frameworks for comparison and appropriate benchmarks, and display and contextual information. The TEP's guidance can be summarized as follows:

- Strategies for addressing statistical uncertainty should focus on the development of minimum case sizes, rather than reliability-adjustment. While reliability-adjustment is useful and appropriate in some cases, minimum case sizes are better suited for quality improvement purposes.
- Frameworks for comparison should emphasize flexible approaches that allow states to determine their own peer groups, and provide appropriate contextual information to help facilitate appropriate comparisons. For example, TEP members states often have ideas regarding their "peer" states based on (a) similarity in HCBS program structure, (b) concentrations of Medicaid managed care, and (c) performance on other indicators of LTSS quality, such as those in the AARP Scorecard for LTSS (Reinhard et al, 2014).
- The TEP advised against the use of simple ranking (Figure V.1), and preferred performance categorization (Figure V.2) over exceedance probability (Figure V.3). Performance categorization, the TEP relayed, is a well-known approach that accounts for statistical uncertainty. Exceedance probability, on the other hand, is less familiar.
- Contextual information is also crucial for understanding these measures. Most importantly, because these measures were developed for the FFS Medicaid population, the TEP indicated it was critical to understand the proportion of HCBS users absent from these analyses due to enrollment in managed care. In addition, the TEP also suggested providing information on

sample size, case mix, and Medicaid LTSS policy information to place the composites in context and to elicit from users testable hypotheses and quality improvement ideas.

The complete summary of the meeting is included in Appendix A; the report concludes by describing Mathematica's final recommendations based on the TEP's input.

VIII. FINAL COMPOSITE RESULTS

The HCBS composite results for the 2009 and 2010 FFS HCBS user population are included in Appendix D of this report, available online at http://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Long-Term-Services-and-Supports/Balancing/Money-Follows-the-Person.html. Based on the analyses presented in the previous sections and advice from the HCBS Composite Measure TEP, results are reported using the following recommended guidelines:

- Risk-adjusted acute and chronic composite results are presented for each state, stratified by MME status. The overall composite is not calculated.
- States with a denominator less than 1,200 HCBS months do not have results reported.
- State-level results include information on the number of composite events (numerator), the number of HCBS months (denominator), expected, and risk-adjusted rate.
- Ninety-five percent confidence intervals surrounding the risk-adjusted rate are produced for purposes of identifying statistically significant differences.
- Results are intentionally displayed using a flexible approach to facilitate custom comparisons; national MME and Medicaid-only rates are provided for reference.
- Benchmarks based on the mean MME and Medicaid-only composite rates.
- Results are accompanied by contextual information on the number and percentage of HCBS users included in calculation and state Medicaid HCBS expenditures per person.

In addition to the 2009 and 2010 HCBS user populations, we produced HCBS composite risk-adjusted rates for MFP participants and non-MFP transitioners. These estimates are based on the transitioners for whom we could merge on risk factors from the 2009 and 2010 HCBS user population, which is a subset of the complete transitioner population. These results show that in the MME population, MFP participants had a higher risk-adjusted rate of acute events relative to the non-MFP transitioners, but the reverse is true for the Medicaid-only population (Table VII.1). The same pattern is repeated for the chronic composite measure: MFP participants who are MME show a higher risk-adjusted rate relative to non-MFP transitioners, but MFP participants enrolled in Medicaid-only have a lower risk-adjusted rate compared to non-MFP transitioners (Table VII.2). While the rates are risk-adjusted, it is possible that the non-MFP MME population is at lower risk for these events due to characteristics that aren’t captured in our data sources. For example, individuals who transition outside of MFP may be better able to access family and other social supports. Because few states met the minimum denominator cutoff for the measures, results are only reported at the national level for these populations, rather than by state.

Table VII.1 HCBS acute composite for transitioner subpopulations

Population	Numerator	Denominator	Expected Rate	Risk-Adjusted Rate
MFP MME	191	2,394	5,739.43	6,796.36
MFP Medicaid-only	50	1,274	3,442.96	2,671.40
Non-MFP MME	1,470	26,789	6,590.85	4,070.57
Non-MFP Medicaid-only	317	5,865	3,586.79	3,531.47

Source: Mathematica analysis of 2008-2010 MFP and non-MFP users. Data sources included the 2008-2010 MAX PS, OT, and IP files, MFP administrative files, MedPAR file, MBSF, and CCW conditions

Table VII.2 HCBS chronic composite for transitioner subpopulations

Population	Numerator	Denominator	Expected Rate	Risk-Adjusted Rate
MFP MME	166	2,394	8,267.49	6,052.14
MFP Medicaid-only	49	1,274	5,637.50	3,123.27
Non-MFP MME	1,324	26,789	7,965.67	4,477.20
Non-MFP Medicaid-only	281	5,865	6,170.63	3,554.50

Source: Mathematica analysis of 2008-2010 MFP and non-MFP users. Data sources included the 2008-2010 MAX PS, OT, and IP files, MFP administrative files, MedPAR file, MBSF, and CCW conditions

IX. CONCLUSIONS AND TECHNICAL RESOURCES

The measure specifications and recommendations included in this report and Volume 1 represent the fruition of 10 years of work, with contributions from multiple HHS agencies, federal contractors, TEPs and other forums. To gain a comprehensive picture of this measure development process, we point the reader to the initial environmental scan (AHRQ, 2007), preliminary measure development work (Schultz et al, 2010), the proposed methodology report for risk-adjustment (Bohl et al. 2015), and the final documentation of the statistical risk-adjustment models (Bohl et al. 2015b). In addition, we note that a similar process was used to develop risk-adjusted pressure ulcer measures (Ross et al, 2015). The acute and chronic composites along with the severe pressure ulcer measure represent the first state-level quality measures for the HCBS population that incorporate patient comorbidities and case-mix into risk-adjusted estimates.

With the TEP's guidance, we have developed risk-adjusted HCBS composites suitable for use in quality improvement initiatives, and recommendations for their application. Still, more work remains to give policymakers actionable information on the quality of care received by the HCBS user population. Most importantly, the next phase of measure development should focus on including information on managed care. Other topics of study include incorporating assessment data for risk-adjustment, strategies to define state peer-groups or HCBS user subgroups, or validating the composite measures to reassure the users that they are capturing meaningful variations in care quality.

This technical report includes state-level acute and chronic composite estimates for the 2009 and 2010 HCBS user populations, but we encourage users to re-estimate this information on more recent data. Because many states have expanded their Medicaid managed care programs, both overall and specific to LTSS, and because of other changes in the health care system, many changes in these results can be expected. To give the users the power to estimate their own rates, we have developed a measure calculation package, with step-by-step instructions on how to build analytic files and estimate rates. The measure calculation package is available at http://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Long-Term-Services-and-Supports/Balancing/Money-Follows-the-Person.html.

This page left blank for double-sided copying

REFERENCES

109th United States Congress. "Deficit Reduction Act of 2005." Washington, DC: Government Printing Office, 2006. Available at http://www.gpo.gov/fdsys/pkg/BILLS-109s1932enr/pdf/BILLS-109s1932enr.pdf. Accessed November 19, 2012.

Agency for Healthcare Research and Quality. "National Healthcare Disparities Report, 2013. Rockville, MD: May 2014. Available at: http://www.ahrq.gov/research/findings/nhqrdr/nhdr13/index.html

Agency for Healthcare Research and Quality. "Environmental Scan of Measures for Medicaid Title XIX Home and Community-Based Services Measure Scan as of July 5, 2007. Rockville, MD, Agency for Healthcare Research and Quality, June 2010. Available at http://archive.ahrq.gov/professionals/systems/long-term-care/resources/hcbs/hcbsreport/hcbssum.html

Agency for Healthcare Research and Quality. "Inpatient Quality Indicators Software Instruction, SAS, Version 4.3 (Prepared by Battelle, under Contract No. 290-04-0020)." Rockville, MD: August 2011. Available at http://www.qualityindicators.ahrq.gov/Downloads/Software/SAS/V43/SAS_Instruction_IQI_4.3.pdf

Ash, A.S., S.E. Fienberg, T.A. Louis, S.L.T. Normand, T.A. Stukel, and J. Utts, with the Committee of Presidents of Statistical Societies. "Statistical Issues in Assessing Hospital Performance." Baltimore, MD: Centers for Medicare & Medicaid Services, 2011, revised 2012.

Bohl, Alex, Mariel Finucane, Jessica Ross, Sheng Wang, and Dejene Ayele. "Proposed Methods for Developing and Testing Risk- and Reliability-Adjustment Models for HCBS Composite Measures." Final Report submitted to the Centers for Medicare & Medicaid Services. Cambridge, MA, February 12, 2015. Available at: http://www.medicaid.gov/medicaid-chip-program-information/by-topics/long-term-services-and-supports/balancing/money-follows-the-person.html. Accessed May 2015.

Bohl, Alex, Jessica Ross, and Dejene Ayele. "Risk Adjustment of HCBS Composite Measures, Volume 1." Final Report submitted to the Centers for Medicare & Medicaid Services. Cambridge, MA, July 30, 2015. Available at: http://www.medicaid.gov/medicaid-chip-program-information/by-topics/long-term-services-and-supports/balancing/money-follows-the-person.html. Accessed August 2015.

Clark, T.S., and D.A. Linzer. "Should I Use Fixed or Random Effects?" *Political Science Research and Methods*, 2014, pp. 1–10.

Davies, S.M., K.M. McDonald, E. Schmidt, E. Schultz, J. Geppert, and P.S. Romano. "Expanding Use of the Prevention Quality Indicators: Report of Clinical Expert Review Panel." Stanford, CA: Center for Primary Care and Outcomes Research, Stanford University, November 2009. Available at http://www.qualityindicators.ahrq.gov/Downloads/Modules/PQI/PQI_Summary_Report.pdf.

Gelman, A., and J. Hill. “Data Analysis Using Regression and Multilevel/Hierarchical Models.” New York: Cambridge University Press, 2007.

Gelman, A., X.L. Meng, and H. Stern. “Posterior Predictive Assessment of Model Fitness Via Realized Discrepancies.” *Statistica Sinica*, vol. 6, no. 4, 1996, pp. 733–760.

Mathematica Policy Research. "Medicaid Analytic Extract Eligibility Anomaly Tables, 2010." Report submitted to the Centers for Medicare & Medicaid Services. May 31, 2014. Available at http://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MAXGeneralInformation.html.

Paddock, S.M. "Statistical Benchmarks for Healthcare Provider Performance Assessment: A Comparison of Standard Approaches to a Hierarchical Bayesian Histogram-Based Method." *Health Services Research*, vol. 49, no. 3, 2014, pp. 1056–1073.

Reinhard, Susan, Enid Kassner, Ari Houser, Kathleen Ujvari, Robert Mollica, and Leslie Hendrickson. “Raising Expectations 2014: A State Scorecard on Long-Term Services and Supports for Older Adults, People with Physical Disabilities, and Family Caregivers.” Public Policy Institute, June 2014. Available at: http://www.aarp.org/home-family/caregiving/info-2014/raising-expectations-2014-AARP-ppi-health.html. Accessed August 2015.

Ross, Jessica, and Alex Bohl. “Development and Testing of Risk-Adjusted Composite Measures for Populations Using Home- and Community-Based Services.” Final report submitted to the Centers for Medicare & Medicaid Services. Cambridge, MA, September 27, 2013.

Ross, Jessica, Sam Simon, Carol Irvin, and Dean Miller. “Institutional Level of Care Among Money Follows the Person Participants.” The National Evaluation of the Money Follows the Person Demonstration Grant Program, Reports from the Field no. 10. Cambridge, MA: Mathematica Policy Research, October 2012.

Schultz, E., S. Davies, and K. McDonald. “Development of Quality Indicators for Home and Community-Based Services Population: Technical Report.” Stanford, CA: Center for Primary Care and Outcomes Research, Stanford University, June 2012. Available at: http://www.ahrq.gov/professionals/systems/long-term-care/resources/hcbs/index.html. Accessed May 2015.

Shwartz, M., E.A. Pekoz, J.F. Burgess, Jr., C.L. Christiansen, A.K. Rosen, and D. Berlowitz. “Probability Metric for Identifying High-Performing Facilities: An Application for Pay-for-Performance Programs.” *Medical Care*, vol. 52, no. 12, 2014, pp.1030–1036.

Wang, S., M. Finucane, X. Lin, F. Yoon, and D. Jones. “Reliability Adjustment of the AHRQ Quality Indicators Under an Empirical Bayes Framework.” Submitted to the U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality. Cambridge, MA: Mathematica Policy Research, December 2014.

APPENDIX A

HCBS COMPOSITE TEP MEETING 2 SUMMARY

This page left blank for double-sided copying

On May 18th, 2015, Mathematica Policy Research (Mathematica) convened the second of two Technical Expert Panels (TEPs) held on behalf of the Centers for Medicare & Medicaid Services (CMS), and the Office of the Assistant Secretary for Planning and Evaluation (ASPE) to solicit input on the development of risk- and reliability-adjustment models for three composite measures intended to measure quality of care among Medicaid beneficiaries using home- and community-based services (HCBS). The three composite measures, which act as indicators of potentially avoidable hospitalizations and are adapted from the AHRQ prevention quality indicators (PQIs), include (1) a chronic conditions composite, (2) an acute conditions composite, and (3) an overall composite that includes components from both the chronic conditions and acute conditions composites. During this TEP, Mathematica presented preliminary risk-adjustment results, and asked the TEP members to provide input on the following topics:

- Preliminary risk-adjustment results (reactions and feedback)
- Addressing low reliability of rates for small populations
- Establishing how to use HCBS composites for state-level comparisons and setting appropriate benchmarks
- Methods for displaying results and providing supporting information

The TEP participants included Robert Applebaum, Ph.D. of Scripps Gerontology Center, Arlene Ash, Ph.D. of University of Massachusetts Medical School, Julie Bershadsky, Ph.D. of Human Services Research Institute, Peter Boling, M.D. of Virginia Commonwealth University School of Medicine, Alison Cuellar, Ph.D., M.B.A. of George Mason University, Lynda Flowers, J.D., MSN, R.N. of AARP, Sara Galantowicz, M.P.H. of Abt Associates, Teresa Johnson, M.B.A. of National Adult Day Services Association, Alice Lind, M.P.H., BSN of Washington State Health Care Authority, Abby Marquand, M.P.H. of Paraprofessional Healthcare Institute, Thomas Meehan, M.D., M.P.H. of Qualidigm, Chris Murtaugh, Ph.D., M.P.A. of Visiting Nurse Service of New York, Cheryl Phillips, M.D. of LeadingAge, Jonathan Shaw, M.D., M.S. of Stanford University, and Michael Shwartz, Ph.D., M.B.A. of Boston University.[10]

The remainder of this memo summarizes the TEP's feedback and recommendations, and concludes with immediate next steps for this measure development effort.

A. Review of Risk Adjustment Results

Based on feedback and recommendations from the first meeting of the HCBS Composite Measures TEP Mathematica made the following decisions when developing risk-adjusted models for the three composite measures:

1. Directly model the chronic conditions composite and acute conditions composite then calculate the overall composite by combining the risk-adjusted acute and chronic composite rates for each state.

[10] Robert Applebaum, Julie Bershadsky, and Alice Lind could not attend the teleconference, but provided feedback via email and/or separate individual calls. Their input is also included in this summary.

2. Allow risk factors to vary between the acute and chronic conditions composites.
3. Use both clinical and statistical rationale to select risk factors for inclusion in the models.
4. Develop separate models for Medicare-Medicaid Eligible (MME) and Medicaid-only HCBS users.
5. Do not include past outcomes as a risk-adjustment factor.
6. Do not include duration of HCBS use directly in the models, but do evaluate model sensitivity by comparing results among continuous users versus short-term or sporadic users.

B. Summary of model development

Mathematica used the 2010 HCBS fee-for-service (FFS) user population as a model development sample, and employed a zero-inflated negative binomial (ZINB) model structure to model acute and chronic composites. The ZINB structure accounts for the over dispersion and high proportion of zeros in the development sample data, and also has a better model fit—measured by Akaike information criterion (AIC) and Bayesian information criterion (BIC)—than negative binomial, Poisson, and zero-inflated Poisson model structures. Testing revealed that removing non-significant risk factors has little impact on model fit as the ratio of observed-to-expected events (OE), and AIC and BIC are similar before and after removal. Short-term and sporadic HCBS users also have little effect on the model. Throughout validation the model showed consistent and intuitive results; however it is important to remember that characteristics of the development sample data influence results.

Mathematica asked the TEP for feedback on and reactions to the modeling approach. The TEP's main concerns were the implications of limiting the sample population to HCBS FFS users, leaving out HCBS users enrolled in managed care. This limitation greatly effects the proportion of Medicaid-only and MME events available to inform the model. Some important points were:

- Due to differences in state policies, the ratio of Medicaid-only to MME events varies widely by state. Some experts believed that the differences are so vast that simply adding Medicaid-only and MME events to get an overall composite may not be the most sensible option. Others felt this approach was acceptable, as long as the weights of Medicaid-only and MME events are shown alongside the composite measures in reporting.
- About 20 percent of HCBS users are Medicaid-only; thus reliability for Medicaid-only rates will be much lower than for MME rates due to sample size.
- When building a prior for reliability adjustment it will be important to take steps to ensure that Medicaid-only and MME inform themselves rather than each other.
- The proportion of Medicaid-only to MME users in the data is not necessarily representative of the actual population of HCBS users in a given state; it is more likely a representation of a state's FFS population. Many experts agreed that this point must be made clear in reporting and suggested including the following to provide context when reporting the composite measures:

- The proportion of a state's HCBS users enrolled in managed care (this would illuminate what percentage of the population of interest is not captured by the measure).
- State-specific policies that directly influence the HCBS FFS population.

- One expert recommended looking at states' OE ratios versus their median HCBS enrollment length, or stratifying results by length of enrollment to see if there is any correlation between the two.
- Another communicated concern with potential underreporting of chronic, disability, and mental health conditions in the Chronic Conditions Data Warehouse (CCW) and how this may affect our ability to compare results at the state-level. This expert recommended analyzing people who have just become Medicaid beneficiaries to see if there is a corresponding jump in their risk level.
- Reporting the expected number of events (E) with a state's OE ratio could help inform meaningful comparisons targeting states with similar case-mix risk profiles.

1. Summary of State-level Results

Overall, the results showed that most states (with the exception of MT, NM, SD, TN, UT, and WY) have higher chronic composite observed rates compared to acute composite observed rates. Risk adjustment only moderately affected the ranking of states, and its impact was slightly greater on the acute composite than on the chronic composite. However, risk adjustment did impact relative state performance. For example, the acute observed rate for TN is lower than the national average, but TN's acute OE ratio is above the national average. Similarly, the chronic observed rate for KY is higher than the national average, but the chronic OE ratio is below the national average.

Several TEP members expressed that some of Mathematica's results did not align with their initial expectations. Comments and concerns were as follows:

- Some TEP members wished to see state sample population sizes alongside reported rates to try to deduce if there is a volume-outcome relationship.
- One expert noted that the model appears to over-predict rates for states with low numbers of outcomes, and under-predict rates for states with high numbers of outcomes, which is a typical problem encountered in risk-adjustment. A suggestion for addressing this was to incorporate states' overall hospitalization rates into the model to adjust for the overall sickness of states' Medicaid populations, which may be the cause of the over/under predictions. Alternatively, the model could adjust for states' PQI rates to account for differences in overall wellness.
- Another expert noted that historically there has been a pattern of higher hospitalization use in states with more liberal Medicaid policies, and more poverty. Based on this information, he would expect OE ratios for OH and VA to be lower that they are in the risk-adjusted results, and the OE ratio for NY to be higher than it is in the risk-adjusted results.
- Several experts agreed that contextual factors could help explain why Mathematica's results sometimes differed from many of the panel members' intuitions. While one expert supported including contextual factors in the risk-adjustment model, the majority of the panel felt such

factors should not be controlled for, but should be reported. Other suggestions for adding context included reporting both adjusted and unadjusted rates, or reporting coefficients associated with state policies.

2. Impact of Reliability Adjustment

Due to sample sizes, results have low reliability for small states. Mathematica presented two options to account for uncertainty surrounding results for small states. One option is to establish a minimum denominator size for reporting, and the other is to use reliability-adjustment. The advantage of establishing a minimum denominator size is that it removes states with uncertain estimates; however, with this approach some states would not be available for comparison.
As a second and potentially complementary choice, reliability-adjustment would not remove any states, and could reduce the risk of false conclusions based on uncertain estimates by shrinking each state's rate toward a prior estimate of true performance. Some options for establishing a prior include: the national rate in the analytic time period; a sub-group specific rate, such as one based on the MME population or the ID/DD population; the past year's reliability-adjusted state rate; or a peer group rate, such as the rate of a state with a similar HCBS FFS population. The disadvantages of reliability adjustment are that it may remove between-state variation of interest, and small states will look similar to the prior.

On the day of the TEP, there was no clear consensus on what method to employ. A few arguments were presented in favor of accounting for uncertainty resulting from small sample sizes, but some experts cautioned against further adjusting results. However, in post-TEP follow-up discussions, a preference for minimum case (denominator) sizes began to emerge. The main points from this discussion were:

- Losing some states through imposing a minimum denominator size may not necessarily be a disadvantage since the goal of the measure is to spur actionable change. If a state's HCBS FFS service population is very small, changes based on outcomes for that population alone are not necessarily beneficial for the state's HCBS population as a whole.
- Experts in favor of reliability-adjustment supported either using a MME-specific prior, because about 80 percent of the sample population is MME, or building a prior based on peer grouping, which has a lot of intuitive appeal as it could be based on expected rates or policy similarities.
- The main argument against building a prior based on national MME rates was that the national rate is influenced most by large states and thus the prior would pull the rates of small states towards the rates of larger states. For instance, RI's results would be pulled towards CA's results.
- It was also stressed that the underlying causes for differences between adjusted and un-adjusted rates must be fully understood before employing a "statistical fix," and more importantly, before releasing results to states.
- Other experts reaffirmed that if certain states or populations of interest don't have adequate sample sizes to draw meaningful conclusions, it's better to state this, rather than reliability-adjust and end up with a result that can't be used for quality improvement.

C. Comparison Framework and Benchmarks

Experts were next asked to comment on methods for reporting the HCBS composite measures. Specifically, they were asked to consider several comparison frameworks, and different ways of selecting benchmarks. The three potential comparison frameworks are:

1. Ranking: ordering states based on their rates without making statistical inferences.
2. Performance categorization: distinguishing which states are statistically different from a benchmark (using each state's 95 percent confidence interval) without reference to the magnitude of difference.
3. Exceedance probability: articulating the degree to which states differ from a benchmark; for example, state X has a 40 percent chance of performing worse than the national average on the HCBS composite measure.

Possible benchmarks for comparison are the national rate, or a peer-group rate determined by region or characteristics of a state's HCBS population. Another option is to allow benchmarks to vary depending on policy objectives. For example, one benchmark could be an achievement threshold such as the mean of the top 5 performing states. Alternatively, results could express improvement over time using a benchmark based on past year performance; however there are limitations to this option as it is difficult to account for year-to-year variations that are not related to quality of care, and quality of Medicaid data varies from state-to-state and year-to-year.

Most TEP members agreed that one of the biggest challenges in reporting results will be the states' inclination to compare their HCBS composite measures to the PQIs results calculated for their overall populations, which would not be a meaningful comparison. For this reason it was suggested that a benchmark enabling comparisons within a state over time may be the best choice. In terms of a comparison framework, many experts expressed that exceedance probability is very difficult to communicate to non-technical users, and leaned towards the use of performance categories. Subsequent post-TEP conversations revealed more support for simple comparison of HCBS composite rates to peer groups or benchmarks deemed relevant by each state. Important takeaways from this conversation were:

- The end users of this measure are likely more comfortable with confidence intervals than with exceedance probabilities. Thus, results can be presented much more concisely with less technical explanation using performance categories.
- Performance categories offer a lot of visual information that users would not necessarily get from exceedance probabilities. Looking at performance categories allows states to compare themselves to the national average, and also highlights to which states they most closely compare.
- One negative aspect of performance categories is that they can give a false picture when comparing states with rates near category borders. In particular, two states could have rates that differ only slightly but put them into different categories which gives the impression of a significant difference between them.
- One well received suggestion for improving the understandability of exceedance probability was to include "complement" probabilities in the presentation. In other words, panel

members agreed that showing the probability of being under or over the national average (as opposed to just showing the probability of being over the national average) would make exceedance probabilities more understandable and useful to states.

D. Display and Use

Finally, the TEP discussed communication and display of results. The panel was asked to keep in mind that the goal of the HCBS composite measures is to provide actionable information to states and stakeholders. There are three rates to consider for reporting: observed, risk-adjusted (which illustrates performance relative to case mix), and reliability-adjusted (which communicates the degree to which reported rates may be due to chance and is especially useful in analyzing small populations). Potential methods for communicating rates are a web-based application such as MONAHRQ that would allow data manipulation, or a periodic report with key findings similar to the AARP Scorecard. Some ideas of contextual information that could be reported alongside state rates are: variations in Medicaid eligibility standards, variations in HCBS level of care criteria or waiver caps, differences in available benefits for specific populations, and state Medicaid HCBS expenditures per person.

Throughout the day's discussion TEP members mentioned including the following contextual information alongside reported measures:

- Proportion of MME and Medicaid-only events in the measure numerator when events are aggregated together for both populations
- Percentage of HCBS users enrolled in managed care
- Medicaid policies that would influence the HCBS FFS population
- Coefficients of risk factors related to state policies
- Expected number of numerator events
- Sample population size
- Usage of nursing homes

Experts reiterated the importance of underscoring the fact that the HCBS composite measures are not necessarily a measure of the states' HCBS programs when reporting results. The majority of TEP members seemed to believe that a web-based application would be the most useful method for reporting results; reasons included:

- Web-based reporting would be extremely useful to states when they are writing their own reports because it could allow them to extract and export the information in which they are most interested.
- Many states have pre-conceived notions of other states with which they should compare; however this is sometimes based on outdated or incorrect information. A web-based application could help guide meaningful comparisons between states through a list of default comparison states, or by flagging states that are well suited to comparison.

- One expert did mention that a disadvantage to web-based reporting is that additional contextual information could potentially be lost. This is of particular concern in instances where someone using the system is looking for a simple yes/no answer to a nuanced question.

E. Next Steps

Based on the TEP's feedback, Mathematica, CMS and ASPE anticipate taking the following next steps in the HCBS measure development process:

- To be responsive to concerns about aggregating information for the MME and Medicaid-only populations, the final proposed HCBS composite measures will only report results for each population separately.
- Similarly, our analyses and the TEP's desire to have actionable information suggests further development of the overall HCBS composite is not warranted. Results indicate that the important risk factors for the acute and chronic composites are markedly different, and states sometimes had discordant performance on the two measures. The combination of these two measures into an overall composite could potentially obscure these important findings.
- Strategies for addressing statistical uncertainty will primarily focus on the development of minimum case sizes, rather than reliability-adjustment. While reliability-adjustment will still be discussed in the final HCBS measure report, further development in this area will be minimal.
- Frameworks for comparison will emphasize flexible approaches that allow states to determine their own peer groups, and provide appropriate contextual information to help facilitate appropriate comparisons.
- Mathematica will continue to develop options for display, keeping in mind the TEP's feedback regarding the typical end user for this information.

This page left blank for double-sided copying

APPENDIX B

RELIABILITY ADJUSTMENT METHODS

This page left blank for double-sided copying

To assess the stability of the HCBS composite rates, Mathematica tested a two-part reliability-adjustment model for the HCBS composites, similar to the approach used by the AHRQ Quality Indicators.11 Part one is the person-level risk-adjustment model already detailed in the HCBS Composite Volume 1 report (Bohl et al. 2015b). Part two reliability adjusts the HCBS composites by specifying a likelihood function (for observed composite rates) and prior (representing the mean of state rates for a given population). With this setup, we can estimate reliability-adjusted rates from the posterior distribution. In this section, we given a plain-language description of the reliability adjustment model, and evaluate model fit by reviewing posterior predictive checks and Markov chain Monte Carlo convergence statistics.

Stage 1: Fit the patient-level risk adjustment model

Reliability adjustment begins with the risk-adjustment model outlined in the HCBS composite report Volume 1 (Bohl et al. 2015b). To model the risk of HCBS composite events, we fit a zero-inflated negative binomial (ZINB) model. The outcome is the count of HCBS composite events, and risk factors include age, gender, chronic conditions, physical disabilities, mental health conditions, and substance use disorders. Models were fit separately for the acute and chronic composite, and models were stratified by whether HCBS users were only enrolled in Medicaid or were Medicare-Medicaid eligible (MME).

The risk-adjustment model is then used to construct risk-adjusted rate and variance estimates. Risk-adjusted rates are estimated by indirect standardization as follows:

1. For each state, sum the observed number of pressure ulcers across MME or Medicaid-only HCBS users.
2. For each state, sum the predicted number of pressure ulcer events across MME or Medicaid-only HCBS users. Predicted events are estimated from the risk-adjustment model.
3. For each state, divide the total number of observed and expected events calculated in steps 1 and 2 above. This is the observed-to-expected (O/E) ratio.
4. Multiply each state's O/E ratio by the national observed composite rate for a given composite (acute or chronic) and subgroup (Medicaid-only or MME).

Stage 2: State level (reliability adjustment)

Reliability adjustment is performed after risk adjustment. At Stage 2, the unit of analysis is the state. We estimate reliability-adjusted rates through the posterior distribution, modeled as a function of the likelihood and the prior. The likelihood function specifies the distribution of each state's observed rate, and the prior specifies the distribution of true state performance. In our analysis, we investigated three specifications of the likelihood and prior using different parametric distributions: a normal likelihood and normal prior (referred to as normal-normal), Poisson-normal, and Binomial-normal. In each case, the distribution of risk-adjusted rates from

[11] Quality Indicator Empirical Methods (Revised by Truven Health Analytics, Stanford University (prime contractor), under Contract No. HHSA290201200003I). Rockville, MD: Agency for Healthcare Research and Quality. November 2014. Available at: http://www.qualityindicators.ahrq.gov/Downloads/Resources/Publications/2015/Empirical_Methods_2015.pdf Accessed August 2015.

Part 1 is used to specify the prior in Part 2. In addition, in the case of the normal-normal model, the noise variance of the risk-adjusted rate is used to model the variance of the likelihood function.

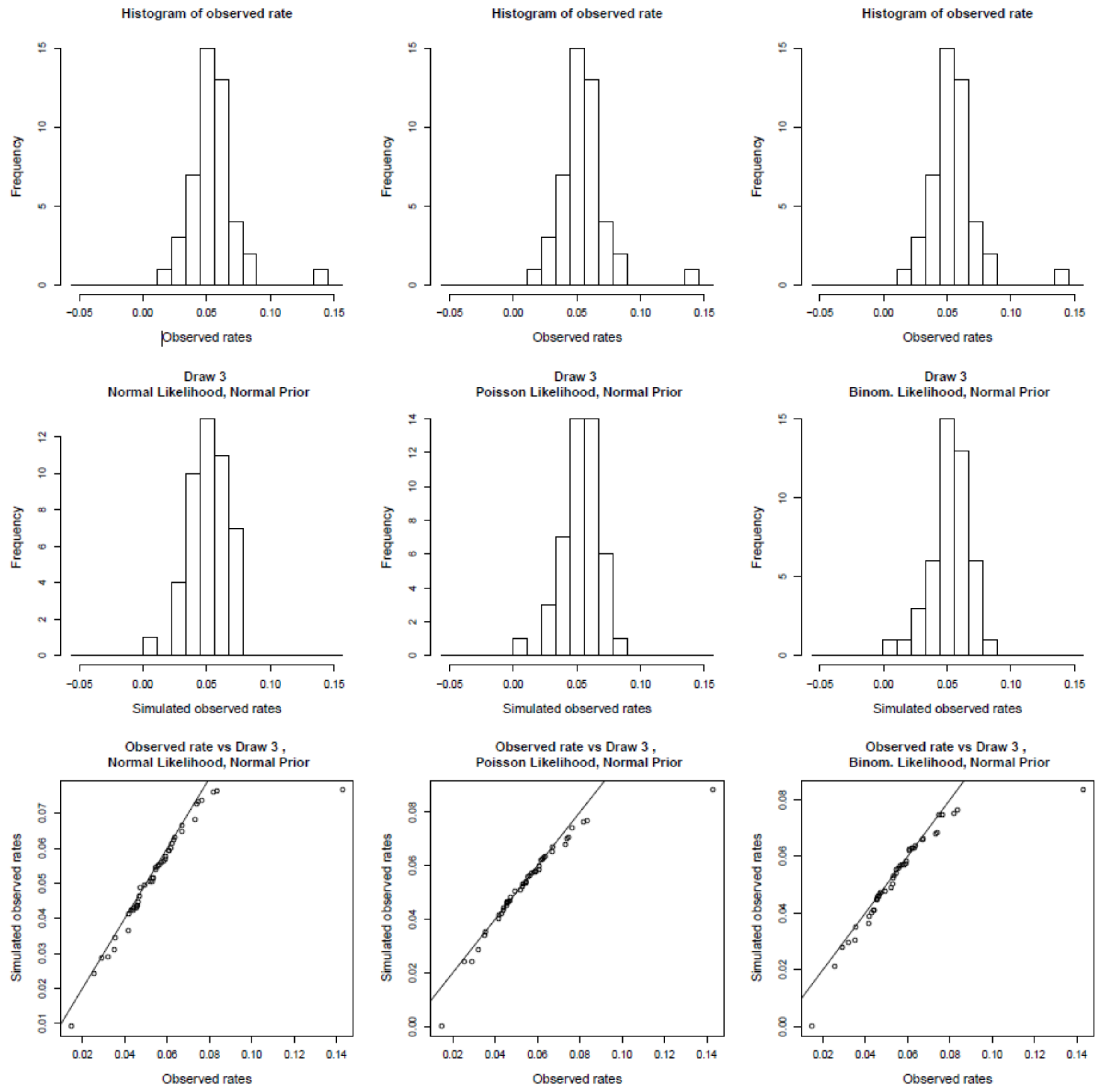
The reliability-adjusted rates comes from the posterior distribution from the Bayesian two-part model. To estimate the prior distribution, we used MCMC simulation methods. By using random draws from the simulation, we estimated the posterior distribution for each state. The mean of each state’s posterior distribution is the reliability-adjusted rate.

To validate the model, we compared the estimated state distribution to the observed state rate distribution using posterior predictive draws. In Bayesian analysis, it is imperative to ensure that the likelihood reflects the data-generating mechanism and that posterior inference is not being unduly driven by prior assumptions (Gelman and Hill 2007). We did rigorous and comprehensive model checking to ensure the validity of any estimates based on a Bayesian framework. We first generated state-level risk-adjusted rates under the assumed model and compare the actual state-level risk-adjusted rates to the simulated rates. Systematic differences between data and replications provide an indication of poor model fit. In our study, we find that three likelihood functions produced very similar patterns in the comparisons. Figure B.1 shows the posterior predictive check plots for the acute composite for MME population. The first row is the histogram for observed rates for three likelihood functions. The second row is one set of simulation results. The last row is a quantile-quantile plot comparison of simulated rates and observed rates. We can see that except for one outlier on the top right of each quantile-quantile plot, the simulated risk-adjusted rates have a similar distribution compared to the observed risk-adjusted rates for all three likelihood functions.

We also tracked the convergence of the MCMC method. For the same set of analyses, Figure B.2 (for the normal-normal case) shows the Gelman-Rubin statistics for the five poorest chains on the top left corner as well as the MCMC chains for those poor runs. We can see that the Gelman-Rubin statistics are very close to 1 and even the poorest five runs have very stable distributions, which means that the MCMC chains have converged for all parameters.

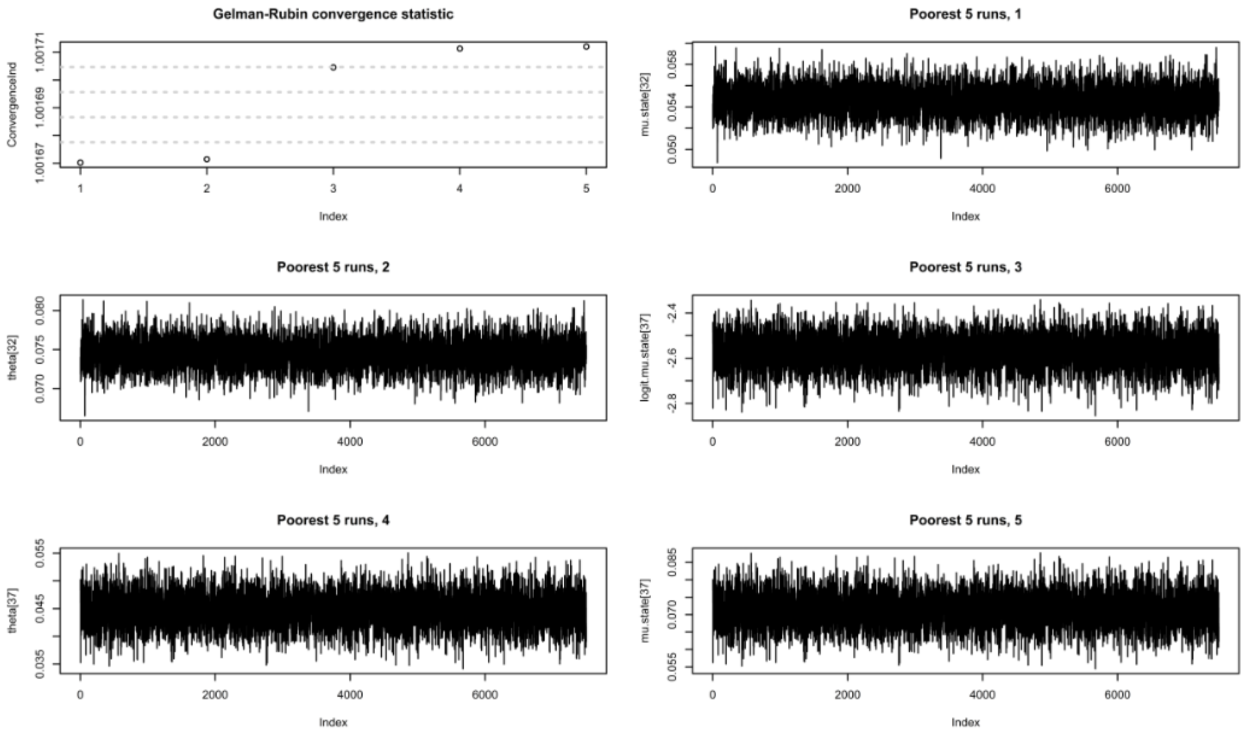
We conclude that it is feasible to calculate valid reliability-adjusted HCBS composites using the two-stage approach. The choice of the likelihood and prior combination had minimal impact on results, but the Poisson-normal had the best fit according to posterior predictive checks. We recommend that the reliability-adjusted HCBS composites (a) when it is feasible for end users or (b) when groups smaller that MME or Medicaid-only HCBS users in a state are the subpopulation or interest.

Figure B.1 Posterior Predictive Checks for the Acute Composite for 2010 MME HCBS Users



Source: Acute composite for 2010 HCBS users that are Medicare-Medicaid eligible

Figure B.2 Convergence of the MCMC method for the Acute Composite among 2010 MME HCBS users



Source: Acute composite for 2010 HCBS users that are Medicare-Medicaid eligible

APPENDIX C

EXCEEDANCE PROBABILITY METHODS

This page left blank for double-sided copying

Exceedance probability

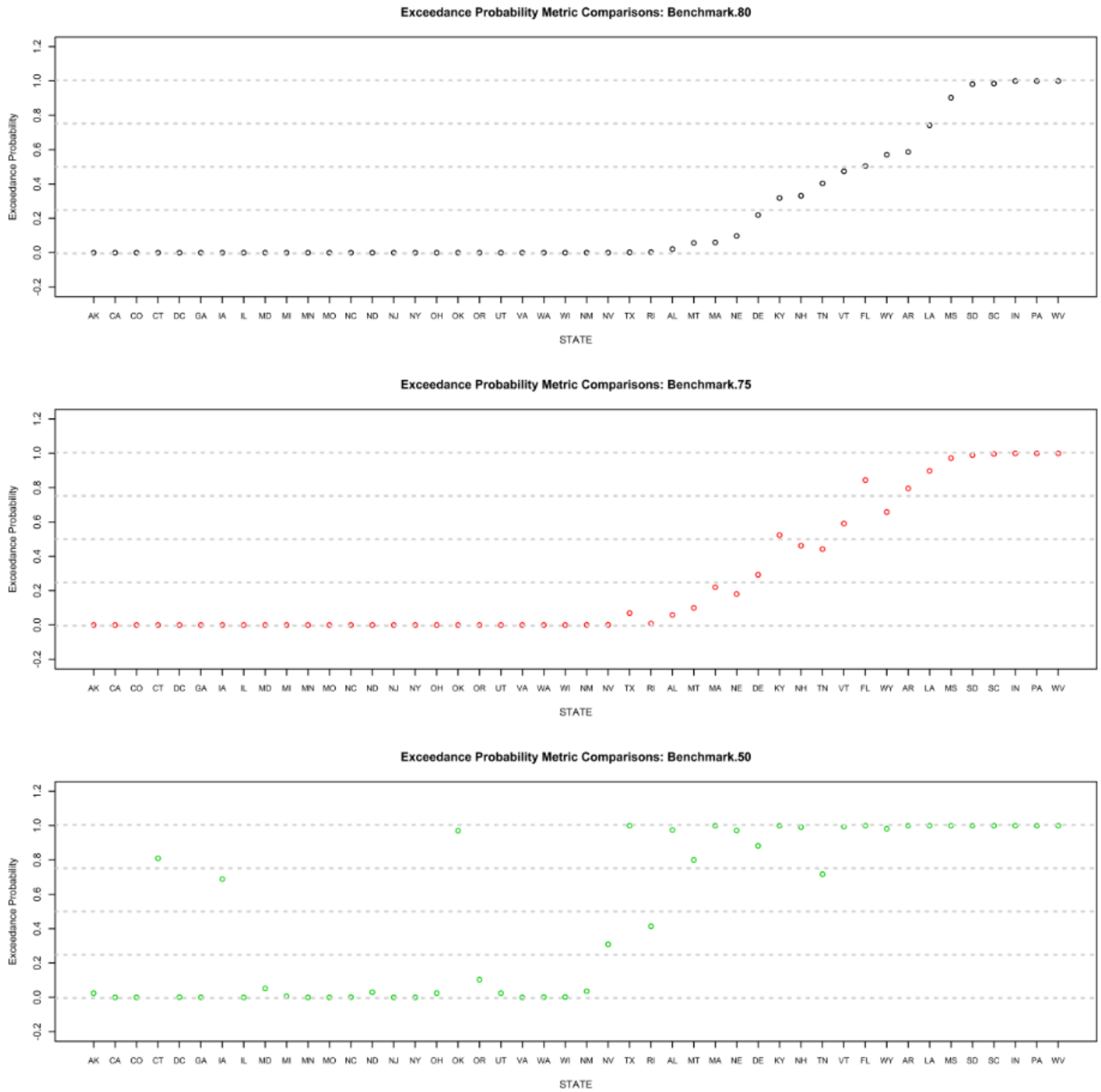
To measure the uncertainty of reliability-adjusted HCBS composites, we tested the feasibility of calculating the exceedance probability. The exceedance probability is estimated using each state's posterior distribution compared to a performance benchmark (Shwartz et al. 2014). Each state's posterior distribution is derived using the two-stage reliability adjustment described in Appendix B. To understand whether the exceedance probability provided different results compared to the performance categorization approach, we examined how exceedance probability varied across states using different benchmarks.

We specify a plain-language description of the exceedance probability through an example. Let us assume a benchmark that is the national unweighted average of acute composite events for HCBS users only eligible for Medicaid; furthermore, let us assume that we are interested in identifying the probability that each state's reliability-adjusted acute composite rate is greater than the benchmark. For each state, we estimate the exceedance probability as the proportion of posterior draws that are greater than this benchmark.

Figure C.1 shows an example for the acute composite for MME population under the normal-normal case. The three charts in the plot (from top to bottom) used 80%, 75%, and 50% percentile of reliability adjusted rates as the cut-off points. Within each chart, we plotted the exceedance probabilities (as on the y-axis) against states (as on the x-axis), sorting by the exceedance probability from lowest (good performance) to largest (bad performance). The plots show variability in the exceedance probabilities between 0 and 1, meaning that it's difficult to classify states into categories.

Our results find that the exceedance probability provides information that is not contained in the performance categorization approach. The exceedance probability varied by state and by HCBS composite. We recommend that those users calculating reliability-adjusted rates use the exceedance probability as a measure of uncertainty using a benchmark that is more relevant to their HCBS population.

Figure C.1. Exceedance probability for Acute Composite among 2010 MME HCBS users, for benchmarks at the 80th, 75th, and 50th percentile of rates



Source: Acute composite for 2010 HCBS users that are Medicare-Medicaid eligible

This page left blank for double-sided copying