LONG-TERM QUALITY ALLIANCE

Medical Utilization in Plans that Integrate Long-Term Services and Supports
Final Report

EXECUTIVE SUMMARY

The high and growing cost of health care in the U.S. is associated with the health care utilization of a small population with the most complex care needs. Only 5 percent of the population accounts for nearly half of all healthcare spending.[1] Within this subgroup, functional limitations are a major driver of spending. Medicare spends nearly three times as much per person on people who need long-term services and supports (LTSS) as on those who do not.[2]

High healthcare spending on the population with LTSS needs may be a consequence in part of inadequate funding in the U.S. for the services and supports people with functional limitations need to remain safe and independent in their home. Would greater funding for better quality LTSS help reduce avoidable medical events?

Health plans that accept a per-capita payment to provide integrated medical care, behavioral health care, and LTSS have all the necessary pieces to test the notion that the cost of meeting LTSS needs could be financed through reductions in avoidable high-cost medical care. However, while many researchers believe that financing and integrating LTSS with medical care could lower healthcare spending and improve outcomes, there is little evidence to support this belief. To spur investments in meeting LTSS needs, health plans and other risk holders need evidence of the return on investment (ROI) that is possible from LTSS integration.

Study Design

This project aims to compare the medical utilization of persons with serious functional limitations enrolled in plans that integrate medical care and LTSS with individuals who have a similar level of functional need who are enrolled in traditional fee-for-service Medicare. The hypothesis is that a population with complex care needs enrolled in an integrated health plan will have a lower rate of utilizing hospital and institutional care than a similarly complex population that is not enrolled in an integrated plan.

---

[1] https://meps.ahrq.gov/data_files/publications/st354/stat354.shtml

[2] People who need LTSS have functional limitations that prevent them from being able to perform activities of daily living (e.g., bathing, dressing, eating) without assistance from a family member or paid caregiver. For Medicare spending on people with and without LTSS need, see: http://www.ltqa.org/medicare-spending-on-older-adults-who-need-ltss/

Previous studies of medical utilization in integrated health plans have encountered difficulty obtaining data on a comparison population in traditional, non-integrated ('fee-for-service") health care. This study developed a new methodology that relies on the National Health and Aging Trends Survey (NHATS) as a data source. The NHATS was selected because it has sufficient information on individuals age 65 and older with functional limitations to enable the creation of a benchmark population that could be matched on the basis of functional need with similar populations enrolled in integrated health plans.

A subpopulation with a high-level of LTSS need[3] (the "study population") was drawn from the matched NHATS-Medicare Claims file for 2015 [traditional fee-for-service Medicare claims] and, similarly, from the participating integrated health plans. For this study, five participating plans were selected from a group of "exemplar" plans on which LTQA had previously developed extensive case studies. This project is intended to provide quantitative evidence of medical utilization in those plans to supplement the qualitative findings from the case studies on aspects of each plan's unique care model most linked to achieving health care savings and preferred outcomes.[4] The health plans that participated include three Senior Care Options (SCO) Fully-Integrated Dual-Eligible Special Needs Plans (FIDE-SNPs) in Massachusetts, a large multi-site Program of All-Inclusive Care for the Elderly (PACE) program in New York, and a Medicare-Medicaid Plan (MMP) in Texas.

The research team included staff from the Long-Term Quality Alliance (LTQA) and the Department of Health Policy and Management at Johns Hopkins Bloomberg School of Public Health. LTQA is an alliance of national organizations that represent stakeholders across the spectrum of LTSS. LTQA's mission is to advance the availability, affordability, and access to high-quality, person- and family-centered, integrated LTSS.

The research team created a model of each plan's study population in the NHATS study population and used the model to predict the expected medical utilization for a group that on average possessed comparable characteristics to that health plan's study population but were enrolled in traditional fee-for-service Medicare.

From each participating plan, the team collected data on the characteristics and the medical utilization of the study population. "Observed" utilization – the actual utilization experienced by enrollees selected for the study population – was compared to "Predicted" utilization, which was the utilization generated by the model for a hypothetical population with similar characteristics to that of plan enrollees but enrolled in fee-for-service Medicare.

[3] Defined for purposes of the study as individuals who received help with two or more of six standard activities of daily living (ADLs): eating, bathing, toilet use, transfer, indoor mobility, and dressing. This definition is similar to the criteria used by many state Medicaid programs and long-term care insurance programs to determine eligibility for long-term care benefits.
[4] http://www.ltqa.org/case-studies/

Results

Overall, medical utilization by the study populations enrolled in the integrated plans differs from the medical utilization predicted by the model for a similar population enrolled in traditional fee-for-service Medicare in the following ways:

- Hospitalization rates are lower than the rates predicted for a similar FFS population for four of the five integrated plans. These results are significant at the 95 percent confidence interval for one of the plans and nearly significant for two of the others. For these three plans, the observed hospitalization rates are between 40 and 70 percent lower than the predicted for a similar fee-for-service population.

- Emergency department (ED) visits are lower for four of the five integrated plans than for fee-for-service Medicare, but these differences are significant for only two of the four plans. For the two plans with significant results, the plans' rates are half or less than half of the predicted fee-for-service rates.

- Skilled Nursing Facility (SNF) utilization is lower for three of the five integrated plans than for fee-for-service Medicare, and these results are significant for all three. In the other two plans, there is not a significant difference either way. For the three plans that are significantly lower, the plan rates are a third or less of the predicted fee-for-service Medicare rates.

Conclusion

The results from this study of medical utilization in the five integrated plans are generally directionally positive (show lower utilization rates for integrated plans than the rates predicted for a traditional Medicare population) – but these differences are not all significant at a 95 percent confidence interval. The type of utilization (hospital, ED, or SNF) that showed the greatest difference between observed and predicted rates varied by plan. None of the plans had observed rates that were substantially lower on all three measures.

The clearest result is that the three SCO plans had consistently lower hospitalization rates than the fee-for-service population. SCO hospitalization rates were 40 to 70 percent lower than would be expected if this population was enrolled in traditional fee-for-service Medicare (one result was significant, and the other two were at the lower edge of the margin of error). It is important to note that the observed hospitalization rates for the high-need members that are in the study populations from these three plans are equivalent to the hospitalization rate for the average Medicare beneficiary (most of whom would not be considered "high need").

This result did not hold for the MMP and PACE plans. We believe that different results for these two plans are largely due to an underestimate of hospital admissions for these plans by the predictive

model, resulting from an inability to adjust the model's population to fully match what the researchers surmise are higher acuity populations in those two plans. For both plans, the study populations had a "nursing home level of care" need. This level of acuity would be greater than was true for the study populations from the NHATS and the three SCO plans.

The differences between observed rates for integrated plans and predicted rates on ED visits was mixed. Two of the plans (one SCO plan and the PACE plan) had a substantially (50 to 70 percent) lower rate of ED visits. Both results were significant. The other three plans had the same or slightly (10 to 20 percent) lower rate, but none of these results were significant.

The differences between observed and predicted rates for integrated plans on SNF admissions was similarly mixed. Three of the plans (one SCO, the MMP, and the PACE plan) had substantially lower-than-predicted rates on SNF admissions (70 percent or greater difference). All three results were significant. The other two SCO plans had SNF rates matching the rate predicted for their members had they been enrolled in traditional Medicare – there was no significant difference.

The variation in differences between observed and predicted medical utilization across plans suggest that differences in the care models of the individual plans and their approach to meeting the needs of their members may influence whether the plan has lower rates for ED visits, hospitalizations, or SNF admissions. All the plans had a substantially lower-than-predicted rate on at least one of the utilization measures – with a difference on that aspect of utilization that was substantial (in the range of 40 to 70 percent or more difference).

This study was undertaken to test a novel methodology for measuring differences between medical utilization in integrated plans compared to fee-for-service. Given limitations in the methodology, it is not possible to prove that there was a causal relationship or to calculate the medical savings that resulted from providing holistic services and supports through integrated plans. However, the directionality of the results is consistent and strong across the plans and provides a view of the potential in integrated LTSS to achieve substantial health care savings.

Introduction[5]

The U.S. spends a far greater share of its GDP on health care than any other country but falls short of many other developed nations on measures of the health status of its population.[6] A contributing factor to this phenomenon may be that the U.S. spends far less than other developed nations on social services. This pattern of comparatively high healthcare spending and low social services spending has particular significance for those Americans with functional limitations who require long-term services and supports (LTSS) for assistance with the basic activities of daily life and whose high use of medical care may be precipitated by the lack of adequate social services and supports.

Medicare spends nearly three times as much per person on people who need LTSS than those who do not need LTSS.[7] This is consistent with the fact that a small percentage of the population (5%) with the most complex care needs in a given year account for nearly half of all healthcare spending.[8]

High healthcare spending on the population that needs LTSS may reflect a tendency in the U.S. to respond with medical care to situations that have developed from the lack of the services and supports individuals with functional limitations need to remain safe and independent in their home. Medical treatment is often provided with little regard for the total circumstances and array of needs of people who have long-term complex conditions. When these people do have access to LTSS, it is rarely as part of a comprehensive program that addresses a full range of medical, behavioral health, and LTSS needs.

A more holistic approach to addressing complex conditions will require a transformation in the way we pay for care. New types of organizations have emerged in the last two decades that are taking financial risk for the overall care and wellbeing of people with complex care needs and functional limitations, and have the incentive to help these members remain in their homes and communities for as long as possible and avoid or defer the use of expensive and avoidable hospital and nursing home care. Through plans that accept a per-capita payment to manage all medical, behavioral health, and LTSS needs there is the potential to redirect resources high-cost medical care that is preventable to the social services that contribute to quality of life.

Despite a growing belief among researchers that financing and integrating LTSS with medical care could lower healthcare spending and improve outcomes, there is little evidence to support this belief. To spur investments in meeting LTSS needs and providing social services, health plans and other risk holders need evidence of the return on investment (ROI) that is possible from LTSS integration.

[5] This research was funded by a grant from the Laura and John Arnold Foundation. The views expressed are those of the authors and not necessarily those of the funder.
[6] https://www.commonwealthfund.org/publications/issue-briefs/2015/oct/us-health-care-global-perspective
[7] http://www.ltqa.org/medicare-spending-on-older-adults-who-need-ltss/
[8] https://meps.ahrq.gov/data_files/publications/st354/stat354.shtml

The need for evidence is stronger than ever with passage of the CHRONIC Care Act of 2018. This law gives Medicare Advantage (MA) plans new flexibility to offer certain non-medical services, such as personal care at home, to enrollees who qualify. The development of credible evidence of lower healthcare spending associated with these types of benefits would encourage MA plans to offer the new benefits and promote greater overall adoption of integrated models.

This project was designed to develop and disseminate foundational evidence of the effects of financial coverage and integrated LTSS on medical spending. This project sought to contribute evidence that might facilitate the ongoing shift in healthcare delivery and payment away from high-cost, unnecessary medical care towards social services that improve health and quality of life.[9]

Project Description

This project sought to compare the healthcare utilization of high-need individuals enrolled in programs that integrate medical care with LTSS to high-need individuals in traditional fee-for-service Medicare. The hypothesis was that a population with complex care needs and is enrolled in an integrated health plan that supports their LTSS needs and maintains their independence in their home and community has less use of expensive medical and institutional care than a similarly complex population that is not enrolled in an integrated plan.

The study used the National Health and Aging Trends Survey (NHATS) to develop high-need benchmark populations that were matched with similar populations enrolled in health plans we identified as being "exemplar," and whose integrated programs we have profiled.[10] The study performed regression analyses on these benchmark populations to examine the relationship between population characteristics and healthcare utilization. The analysis enabled us to predict the expected utilization for a group that on average possessed comparable characteristics to health plan enrollees but were enrolled in fee-for-service Medicare.

The research team included staff from the Long-Term Quality Alliance (LTQA) and the Department of Health Policy and Management at Johns Hopkins Bloomberg School of Public Health. LTQA is an alliance of national organizations that represent stakeholders across the spectrum of LTSS. LTQA's mission is to advance the availability, affordability, and access to high-quality, person- and family-centered, integrated LTSS. The health plans that participated include three Senior Care Options (SCO) Fully-Integrated Dual-Eligible Special Needs Plans (FIDE-SNPs) in Massachusetts, a large multi-site Program of All-Inclusive Care for the Elderly (PACE) program in New York, and a Medicare-Medicaid Plan (MMP) in Texas.

[9] G Peterson et al. (2014) "Impacts of Waiting Periods For HCBS On Consumers and Medicaid Long-Term Care Costs in Iowa." Report prepared for the Office of Disability, Aging and Long-Term Care Policy by RTI. Available at: https://aspe.hhs.gov/system/files/pdf/73211/IAWaitPd.pdf

[10] http://www.ltqa.org/case-studies/

Methodology

NHATS Data

The research team used the National Health and Aging Trends Study (NHATS) dataset linked to Medicare claims. NHATS collects annual data from a nationally-representative cohort of Medicare beneficiaries age 65 and older. This study used data from the 2015 wave of NHATS, which was fielded to approximately 8,000 older adults. NHATS builds and improves upon previous survey research to assess disability among older adults, collecting both self-reported and performance-based measures of physical and cognitive capacity, as well as the accommodations respondents make to enhance their capacity (for example by using assistive devices). The analysis of this study relied on self- and proxy reports of how NHATS study participants complete self-care, mobility, and household activities. The research team also examined Medicare claims for services received by study participants enrolled in fee-for-service Medicare for one year following their 2015 survey date, which allowed for analysis of Medicare utilization as it relates to individual characteristics measured in the survey.

Plan Data

Health plans were recruited for this study from LTQA's network of relationships. Only integrated plans were eligible to participate—plans that include both medical and LTSS benefits in a single, comprehensive package. Common types of integrated plans include PACE programs, FIDE-SNP Medicare Advantage plans, and Medicare-Medicaid plans. Initial candidates for inclusion were drawn from plans on which LTQA had already published case studies. Additional Massachusetts SCO plans were recruited to complete the study. Additional health plans expressed interest in participating but were not able to due to inadequate data capabilities.

First, the research team reviewed health plan assessment tools for variables of interest: activities of daily living (ADLs), instrumental activities of daily living (IADLs), chronic conditions, cognitive impairment, and demographic information. Then, the variables of interest from health plan assessments were cross-walked with those available in NHATS to ensure that variables were defined similarly in both sources.

Next, the research team worked with analysts at each health plan to develop detailed specifications for the data request, defining the study population, descriptive variables, and utilization variables. This iterative process resulted in a mutually agreed upon specifications memo and data request table shell for each health plan.

Health plans completed the table shells with population-level data and health care utilization for the designated study population. The results of this study are contingent on the completeness and accuracy of the data the plans shared. Health plan data was reviewed for consistency and logical errors by the research team and anomalies in the data were reconciled with the plans. Because individual-level data was not shared by the health plan, comprehensive data validation at the plan level was not possible.

However, the consistency of many of the population variables across plans provided additional validation of the aggregate data.

Research Plan

Regression models were constructed to predict the healthcare utilization of individuals with specific characteristics enrolled in each integrated health plan, relative to the benchmark population of individuals with LTSS needs enrolled in traditional fee-for-service Medicare. Then, the predicted utilization was statistically compared to actual utilization of integrated plan enrollees. Implementing this methodology proceeded in five steps.

1. Defining the study population: Assembling comparable integrated health plan and benchmark variables

The first step of the research process involved defining the study population. The research team used the available data to identify similar benchmark populations across the participating plans.

The research team restricted the study population from each plan and NHATS to individuals with high LTSS need, representing the population for which integrated care is expected to be the most impactful. In general, the research team defined the study population as individuals who received help with two or more of six standard activities of daily living (ADLs): eating, bathing, toilet use, transfer, indoor mobility, and dressing. This definition is similar to the criteria used by many state Medicaid programs and long-term care insurance programs to determine eligibility for long-term care benefits. We worked collaboratively with health plans to ensure that the definitions of ADLs and "receiving help" were reasonably consistent across plans as part of the data specification process.

The research team imposed additional eligibility criteria to match plan data to NHATS participants. The study population was limited to individuals age 65 and older, enrolled in Medicare, and who did not live in a nursing home at the beginning of the study period. To ensure data completeness, the study population for each plan was limited to enrollees who had a completed assessment during the study year (the year for which data was collected varied by plan), and who were enrolled at the beginning of the year and either (a) died before the end of the year or (b) remained enrolled for the entire year.

The study populations were subsets drawn from the NHATS national sample (representative of the population 65 and older enrolled in traditional Medicare) and from the plans' memberships, matched on their high level of functional limitation (LTSS need). To put the measured outcomes for these high-need populations in context, the research team also computed the observed rates for the entire NHATS population (the total population age 65 and older, without regard to level of LTSS need). As expected, given the average better health of the total population, these rates of medical utilization are much lower than the predicted and observed outcomes for the study populations. In instances where observed medical utilization in the plans for the high-need study population was actually similar to the medical utilization for the total population, it provided further evidence that the rates reported by the

plans was unusually low. (Table A-3 in the Appendix provides the observed rates for the total NHATS population for comparison with the results reported in Table 3 below).

2. Identifying key independent variables

The research team selected independent variables for inclusion in the regression model that were expected to be associated with utilization outcomes of interest. Next, variables were defined in the health plans and the NHATS to measure constructs consistently in both populations. This effort required a thorough understanding of the types of information collected by plans, as well as how information is collected.

Where plan data permitted, the following independent variables were collected: activities of daily living (ADLs), instrumental activities of daily living (IADLs), measures of cognitive impairment, chronic conditions, Medicaid eligibility, gender, age, race/ethnicity, educational level, primary language, living arrangements, and residential setting. Data availability varied across participating plans (see Table 1).

The study population from the NHATS (the NHATS benchmark population) and the study population from the plans were matched on the degree of functional need (as noted in the previous section). Since a high level of functional need is associated with a high level of medical utilization, it is expected that, all things being equal, there would be equally high levels of medical utilization in both the benchmark and the plan groups. At the same time, the NHATS population, from which the benchmark population for this study is drawn, is a representative sample of the national population 65 and older who are enrolled in traditional (fee-for-service) Medicare, while the plans' 65 and older members, from which the study populations are drawn, are all beneficiaries enrolled in Medicare Advantage who are dually eligible as Medicaid beneficiaries. While only 38 percent of the NHATS benchmark population was enrolled in Medicaid, all or nearly all members of the study plans were enrolled in Medicaid.

As a result, there are notable differences in the characteristics of the benchmark study population and the plans' study populations (see Table 2).[11] It is important to keep in mind, however, that the methodology does not simply compare medical utilization in these two groups. Rather, the methodology creates a model in the benchmark population of a specific plan population and then generates, through that model, a prediction of the rate of medical utilization for that specific plan population if it were enrolled in traditional (fee-for-service) Medicare. The modeling of the plan demographics in the NHATS study population is intended to control for the influence of the independent variables on the outcome, so that influence of the intervention (integrated care– as compared to fee-for-service Medicare) is apparent.

[11] The NHATS benchmark and plan study populations are comparable in terms of average age and numbers of ADLs, IADLs, and chronic conditions. However, the benchmark study population has a higher mortality rate than any of the study populations. Similarly, although only 9 percent of the NHATS benchmark study population was non-English speaking, much higher shares of plan enrollees were non-English speaking, ranging from 27 percent to 72 percent. Plan enrollees were also, on average, less educated, more likely to live alone, and less likely to live in independent or assisted living than the NHATS benchmark population.

Table 1: Plan Data Availability					
	SCO 1	**SCO 2**	**SCO 3**	**MMP**	**PACE**
Available Data	• ADLs • IADLs • Cognitive Impairment • Chronic Conditions • Age • Sex • Race • Ethnicity (Hispanic) • Education • Language • Residential Setting (i.e., assisted living) • Living Arrangement (i.e., alone, with others) • Medicaid Enrollment	• ADLs • IADLs • Cognitive Impairment • Age • Sex • Education • Language • Residential Setting • Living Arrangement • Medicaid Enrollment	• ADLs • IADLs • Cognitive Impairment • Chronic Conditions • Age • Sex • Education • Language • Residential Setting • Living Arrangement • Medicaid Enrollment	• ADLs • Cognitive Impairment • Chronic Conditions • Age • Sex • Race • Ethnicity (Hispanic) • Language • Medicaid Enrollment	• ADLs • IADLs • Cognitive Impairment • Chronic Conditions • Age • Sex • Race • Ethnicity (Hispanic) • Education • Language • Residential Setting • Living Arrangement • Medicaid Enrollment
Missing Data	None	• Chronic Conditions • Race • Ethnicity (Hispanic) • Education	• Race • Ethnicity (Hispanic)	• IADLs • Education • Residential Setting • Living Arrangement	None
Included Populations	• Study population including decedents • Study population excluding decedents • Decedents			• Study population excluding decedents	• Study population including decedents
Missing Populations	None			• Study population including decedents • Decedents	• Study population excluding decedents • Decedents
Data Year	1/1/2015-12/31/2015			1/1/2016-12/31/2016	1/1/2017-12/31/2017

Table 2: Descriptive Characteristics for NHATS Benchmark Population and Participating Plans						
	NHATS Study Population[12]	SCO 1	SCO 2	SCO 3	MMP	PACE
Average Age	83.7	76.8	77.0	78.0	79.1	82.0
Female	69%	71%	70%	73%	78%	62%
Enrolled in Medicaid	38%	100%	100%	100%	100%	99%
High School Degree or More Education	69%	43%	Missing	53%	Missing	52%
Non-English Speaking	9%	72%	64%	27%	25%	36%
Living Alone	29%	43%	42%	46%	Missing	37%
Living in Independent / Assisted Living	18%	3%	8%	11%	Missing	0%
Average # ADLs	3.7/6	3.6/6	3.8/6	4.5/6	4.6/6	4.3/6
Average # IADLs	2.7/4	3.3/4	4.0/4	3.6/4	3.9/4 (24% response)	3.7/4
Cognitive Impairment	51%	47%	75%	66%	67%	69%
Average # Chronic Conditions	3.3/9	2.8/9	Missing	3.9/9	3.0/9	1.8/9
Mortality Rate	19.1%	2.3%	4.3%	6.5%	Missing	Missing
Sample Size	493	>2,000	>2,000	>2,000	419	397

[12] The NHATS benchmark population described in this column was used to build models for each that predict utilization for a hypothetical population with similar characteristics enrolled in traditional fee-for-service Medicare. This approach adjusts for the observed differences between the benchmark and plan populations.

Two areas of difference between the benchmark and plan study populations were explored further. A separate sensitivity analysis to test the separate effect of Medicaid eligibility was not possible because the "n" of persons with Medicaid eligibility in the benchmark study population was too small to obtain meaningful results. Medicaid enrollment was a factor that was controlled for in the model, though, and should not contribute to differences in the predicted and actual outcomes. The modeling is intended to adjust for these population differences and to compare utilization rates for two similar populations.

A second factor we explored was mortality. Mortality in the NHATS benchmark study population was 19.1 percent, while mortality in the three SCO plans was substantially lower: between 2.3 and 6.5 percent. The research team was able to run a sensitivity analysis comparing outcomes of the model with and without decedents included. There were not notable differences between the two sets of results (Table 3 below includes decedents, while Table A2 in Appendix III excludes them).

3. Defining utilization outcomes

The primary outcome measure was average hospitalization rate, operationalized as the aggregate number of hospitalizations per person incurred by the enrolled population over 365 days. Hospitalizations in fee-for-service Medicare were defined on the basis of admission dates within 365 days of the first day of the month in which the NHATS interview occurred. The research team worked with each health plan to construct comparable measures from claims data. Secondary outcomes included average number of Emergency Department (ED) visits per person and average number of Medicare-reimbursed skilled nursing facility admissions per person (i.e., post-acute admissions).

When possible, distribution of each outcome was examined for the benchmark population and each health plan. In addition to mean values for each outcome from the plans, plans shared median utilization and maximum values. The lack of individual-level information from plans prevented examination of utilization distributions or the sensitivity of findings to outliers in the data.

4. Constructing regression models

For the benchmark study population (described above) drawn from the NHATS, a multivariate regression model was developed for each utilization outcome separately for each health plan. The research team examined the distribution of each outcome to identify the most appropriate regression analysis (e.g., whether a high number of zeros in the outcome variable need to be accounted for), as well as distributions of the independent variables to uncover potential challenges of outliers and influential points or missing data. Scatter plots were used to graphically examine the type of relationship (linear or non-linear) between each outcome variable and the independent variables.

The research team then constructed separate multivariate regression models for each health plan using all variables for which information was available and comparably measured by the health plan and NHATS.

5. Predicting health plan enrollee outcomes in a fee-for-service payment environment

a. Applying health plan summary data to model coefficients and testing for statistical significance

Using the regression models in the NHATS benchmark population, the research team predicted utilization by outcome based on characteristics of each health plan's enrolled population (i.e. average number of ADLs, average number of chronic conditions). To calculate the predicted utilization, the health plan value (average or proportions) for each independent variable was multiplied by regression coefficients generated by the NHATS benchmark regression analysis for the plan, producing predicted utilization (e.g., average number hospitalizations per person). Logistic regression was used to predict the percent of the population with any utilization, and negative binomial regression was used to predict average utilization for each outcome. All regressions incorporated complex survey design variables.

b. Contrasting predicted and actual utilization by health plan

The research team compared the predicted and actual outcomes by generating a 95 percent confidence interval around the mean for the NHATS estimate using delta method generated standard errors. This confidence interval was examined in relation to the point estimate of the actual value from the health plans. Because individual-level data from each health plan was not available, it was not possible to compute the joint distribution or covariance structure. Therefore, the approach is subject to the assumption that the covariance correlation term is zero, or that there is no difference in the covariance structure that would be observed for the health plan and the actual covariance structure that is observed for the NHATS comparison.[13]

c. Sensitivity testing

In the cases where health plans varied in the availability of measures, the research team examined the extent to which including particular measures of interest made a substantial difference to the predicted outcomes. Specifically, sensitivity analyses examined the extent to which excluding data for decedents, and including a measure of Ethnicity (i.e., Hispanic/non-Hispanic) affected model estimates in plans for which these data were available (see Appendices II and III). It was not possible to do a sensitivity analysis on the effect of the Medicaid eligibility variable on the outcomes due to the small number of Medicaid eligible persons in the NHATS benchmark study population. However, Medicaid eligibility was a variable included in the model and a factor controlled for in predicting the utilization rates.

[13] Further discussion of the implications of this assumption can be found in Appendix I

Results

Table 3 summarizes results from plan-specific regression models. For each participating plan, observed utilization refers to the actual utilization experienced by enrollees who met eligibility criteria during the time period of interest. Predicted utilization is the utilization generated by the model for a hypothetical population with similar characteristics to that of plan enrollees but enrolled in fee-for-service Medicare.

Due to the focus for this study on people with high levels of functional limitation, the sizes (the "n"s) of the populations for both the NHATS study benchmark and the models drawn from the benchmark are small. This is true also for the size of the high-functional-need populations that we looked at in the plans. As a result, at a 95 percent confidence interval, the predicted rates from the model have a wide range, making it more difficult to demonstrate statistically significant results. While almost all the results from the plans are differentiated in the same direction from the predicted FFS results, many are not statistically significant at a 95 percent confidence interval. The research team believes several of these results would be significant at a 90 percent level, but also feel that the directionality of these results is meaningful.

Table 3: NHATS Model Results for Plan Outcomes

	SCO 1	SCO 2	SCO 3	MMP	PACE
Observed Hospitalization Rate	0.35	0.47	0.25	0.80	0.65
Predicted Hospitalization Rate	0.72 (0.35-1.08)	0.83 (0.42-1.24)	0.79 (0.50-1.08)	0.68 (0.43-0.93)	0.71 (0.40-1.02)
Observed ED Visit Rate	1.20	0.84	1.49	1.30	0.42
Predicted ED Visit Rate	1.47 (0.91-2.04)	1.61 (1.00-2.22)	1.49 (1.06-1.93)	1.42 (1.03-1.81)	1.25 (0.83-1.67)
Observed # SNF Events	0.20	0.20	0.08	0.07	0.00
Predicted # SNF Events	0.18 (0.03-0.32)	0.25 (0.04-0.46)	0.29 (0.14-0.44)	0.23 (0.11-0.34)	0.25 (0.10-0.39)

Overall, (as shown in Table 3) medical utilization by the study populations enrolled in the integrated plans differs from the medical utilization predicted by the model for a similar population enrolled in traditional fee-for-service Medicare in the following ways:

- Hospitalization rates are lower than the rates predicted for a similar FFS population for four of the five integrated plans. These results are significant at the 95 percent confidence interval for one of the plans and nearly significant for two of the others. For the plan with significant results, the plans' hospitalization rate is less than one-third of the rate predicted for a similar fee-for-service population.

- Emergency department (ED) visits are lower for four of the five integrated plans than for fee-for-service Medicare, but these differences are significant for only two of the four plans. For the two plans with significant results, the plans' rates are half or less than half of the predicted fee-for-service rates.

- Skilled Nursing Facility (SNF) utilization is lower for three of the five integrated plans than for fee-for-service Medicare, and these results are significant for all three. In the other two plans, there is not a significant difference either way. For the three plans that are significantly lower, the plan rates are a third or less of the predicted fee-for-service Medicare rates.

Discussion

Three Distinct Types of Plans

This study included different types of health plans: three Fully-Integrated Dual-Eligible Special Needs Plans (FIDE-SNPs) one Medicare-Medicaid Plan (MMP), and one Program of All-Inclusive Care for the Elderly (PACE). Each of these three types of plans can be considered fully-integrated plans and care models along the continuum of integration.[14] Although the plans are similar in their degree of integration, there are also substantial differences between the three types.

The three FIDE-SNP plans all participate in the Senior Care Options (SCO) program and have all been providing integrated care in SCO for many years—two for 14 years and the third for 10 years. The enrolled population includes any dual-eligible senior who chooses to join, and these plans tend to serve a lot of "community well" members. The study population selected from these plans was limited to people who needed help with two or more ADLs to match the population that was modeled in the NHATS data.

The Medicare-Medicaid Plan (MMP) is a new model created as part of the Affordable Care Act (ACA) in 2010 and implemented by the Centers for Medicare & Medicaid Services (CMS) in 2013. This particular plan was launched in March 2015, and the study data from calendar year 2016 represents its first full year of its operation in this model. In this plan, we were limited to a subgroup of the enrolled population for which assessment data were available. This subgroup was defined by a nursing home level of care (NHLOC) need, which in this state is a more restrictive definition and a population with a higher level of acuity than the study population. NHATS did not have the necessary variables to adjust the benchmark projection for this.

The PACE program dates back to a Medicare-funded demonstration in the 1980s of integrated LTSS and medical care at On Lok Senior Health Services in Northern California and was made a permanent part of Medicare in 1997. The program in this study was launched in 2009 and it has nearly a decade of

[14] For more detail on the characteristics of a fully-integrated plan, please see LTQA's "Taxonomy of LTSS Integration" available at: http://www.ltqa.org/wp-content/themes/ltqaMain/custom/images//Taxomony-of-LTSS-Integration.pdf

operation. Eligibility for PACE is limited to people who meet the state threshold for nursing home level of care. Again, we assume that this NHLOC population has a higher level of acuity than we are able to replicate in the NHATS study population and the model.

Hospitalization Rates

Observed Rates

Hospitalization rates reported by the health plans for the study population ("observed hospitalization rates") are quite low for the three SCO plans, and are substantially higher for the MMP and the PACE plans.

The observed SCO hospitalization rates for the study population (that have substantial functional assistance needs) are between 0.25 and 0.47. The observed hospitalization rates for the study population in the MMP plan and the PACE plan are substantially higher, 0.80 and 0.65 respectively).

The research team believes the MMP and PACE plan populations included in this study (both of which meet the nursing home level of care (NHLOC) threshold in their respective states) would be equivalent to a subpopulation of the NHATS and SCO study populations with higher average levels of acuity than the study populations as a whole in either the NHATS or the SCO plans. The research team expects a population with this higher level of acuity to experience a higher hospitalization rate than would be found in the study population as a whole, although the research team is unable to gauge how much higher this rate would be.

Predicted Rates

The hospitalization rates predicted by the model for traditional Medicare populations matched to each of the five plan populations fall within a small range of 0.68 to 0.83. The similarity in predicted hospitalization rates across all five plans reflects the similarity in the level of chronic conditions and functional assistance need in the different plan populations.

Differences in Hospitalization Rates

The greatest difference between plan hospitalization rates and predicted rates for a similar FFS population is evident in the three SCO plans. Actual (observed) hospitalization rates for the three SCO plans are substantially lower (40 to 70 percent lower) than the predicted rates. The result is statistically significant for one of the SCOs and approximates statistical significance for the other two SCOs.

Speculation on Factors Contributing to the Results

As noted above, the research team believes two of the plans (MMP and PACE) have populations with a higher acuity level (i.e., NHLOC) for which the research team was unable to make full adjustment in the

model.[15] These two plans have the lowest predicted hospitalization rates, 0.68 and 0.71. The research team believes that if the model could be adjusted to account for this higher acuity, it would show higher predicted rates of ED, hospital, and SNF use for the MMP and the PACE plans.

The results for these two plans do not show a significant difference in hospitalization rates. The MMP had the highest observed rate of the five plans, a rate that is higher than the predicted rate, but within the statistical margin of error. The PACE plan has an observed rate that was within the margin of error for the predicted rate and is, therefore, not significantly lower.

Hospitalization rates in the three SCO plans, though, are consistently low and statistically significant for one and approximating statistical significance for the other two. All three of the SCO plans have been operating for more than 11 years and may have services and supports in place for members that have become over time quite effective in preventing avoidable hospitalizations.

Emergency Department Visits

Observed Rates

Observed rates for emergency department (ED) visits for the plans seem to cluster in two groups:

- Three of the plans (two SCOs and the MMP) have observed rates between 1.20 and 1.49.
- The other two plans (one SCO and the PACE plan) have significantly lower observed ED visit rates of 0.84 and 0.42.

Predicted Rates

Predicted rates for ED visits for the five plans fall within a narrow range of 1.25 to 1.61. All five have fairly similar predicted rates, again reflecting the similarity in the populations served by the plans.

Differences in ED Visit Rates

For three of the plans, observed ED visit rates were the same or slightly lower than the predicted rates. Of these three plans, observed rates for one was nearly 10 percent and for another 20 percent lower than predicted. These observed results were all within the margin of error for the predicted results, however, and were not statistically significant. Overall, for these three plans, there was no evidence of a clear effect on ED visit rates from plan activity.

The difference in ED visit rates for the other two plans with lower observed rates was substantial and statistically significant. The SCO plan ED visit rate was 50 percent lower than projected while the PACE plan rate was nearly 70 percent lower than projected. Both of these results were well below the margin of error.

[15] State variation in the definition of NHLOC made it impossible to include a factor in the study population for that designation.

Speculation on Factors Contributing to the Results

The substantial difference among the three SCO plans in ED visit rates raises interesting questions about the plan interventions and whether there are differences in the care models or in plan protocols that may affect the rate of ED visits. Answers to that question would require a different study methodology.

Interestingly, among the three SCO plans, the plan with the lower ED visit rate had a slightly higher hospitalization rate than the other two (although still a 40 percent reduction relative to predicted). The research team speculates that the lower than expected ED visits in the PACE plan (nearly 70 percent lower than traditional Medicare) may be more directly related to the program. As with the SCO, the PACE program has a very low ED visit rate coupled with a relatively high hospitalization rate, suggesting that the program may minimize unnecessary ED visits through the more immediate and preventive engagement of plan members with the interdisciplinary care team (IDT) and the plan's primary physician. Changes in conditions and other medical events that might otherwise trigger an emergency room visit may be effectively handled by the plan's physician and IDT, leading to fewer ED visits and, when circumstances warrant, more direct access to the hospital.

It is possible that a similar scenario, involving greater engagement of members living in the community with the plan, may explain the one SCO's combination of lower ED visits and slightly higher hospitalizations. Plan protocols and greater connectivity with plan members may link the plan with the member to enable the care team to anticipate and stabilize situations that would otherwise precipitate an ED visit, and thus result in more direct access to a hospital when needed, rather than through an ED visit leading to a hospitalization.

Were this true, there would be a noticeable difference between one of the SCOs and the other two SCOs in the protocols and degree of connectivity of the plan with plan members that has an observable impact on the way medical care is provided and the overall cost of medical care for members with the most complex care needs to the plan.

Skilled Nursing Facility Admissions

Observed Rates

Observed rates for skilled nursing facility (SNF) admissions vary widely across the plans, suggesting variation in the ways plans utilize SNF stays.

- Three of the plans (one SCO and the MMP and PACE plans) have very low SNF admission rates between 0.00 and 0.08.
- The other two SCOs have much higher observed rates of 0.20.

Predicted Rates

The model predicted similar SNF admission rates for all five plans of between 0.18 and 0.29.

Differences in SNF Admission Rates

Three of the plans had SNF admissions rates that were 70 percent or more lower than the model predicted for beneficiaries in traditional Medicare. In all three cases, the lower rates were statistically significant (below the margin of error). At the same time, the other two plans were either the same or slightly lower than the rates predicted for traditional Medicare beneficiaries – and in both cases, the difference was not statistically significant.

Speculation on Factors Contributing to the Results

The observed SNF admission rates for the plans in the study fell into one of two clusters:

- extremely low (0.0 to 0.08) and substantially lower than predicted for the traditional fee-for-service (FFS) Medicare population, and
- similar to the rate predicted for the FFS Medicare population.

The fact that the actual rates cluster in this way suggests the difference in observed rates results from two different ways that plans are using SNF admissions.

The observed SNF admissions measures in this study apply to only part of institutional use by plans—short-stay Medicare-reimbursed admissions—and do not include admissions or transfers to nursing facilities for longer custodial stays. Longer NF stays are covered under Medicaid, but not under Medicare, and were, therefore, not available for this study.

The research team expected to see a difference between SNF admission rates in traditional fee-for-service Medicare and in Medicare Advantage due to the fact that traditional (fee-for-service) Medicare admission to a SNF requires a prior hospital stay for three nights. Managed care plans do not have to meet the "three-day rule" and can admit members directly to a Medicare-covered SNF. In traditional fee-for-service Medicare, we would expect hospital admissions to be higher, since they are the only route to SNF admissions, and SNF admissions to be lower because of the need for a prior hospitalization. Lower hospital admissions and higher SNF admissions for MA plans would align with the difference in incentives between the two Medicare coverages.

Difference in these rates is expected based on the different incentives. The full picture of institutional use in the integrated plans is not evident in this study, because longer nursing home stays are not captured. Nevertheless, there are factors that may contribute to the observed rates:

1. High SNF admissions as an alternative to hospitalization: It is possible that some of the plans were using admissions to SNFs as an alternative to hospitalization for individuals with complex care needs, since direct admission to the SNF might be a more appropriate placement for an individual with substantial functional limitation and a need for a skilled level of care than admission to a hospital, particularly for individuals without a safe housing situation or adequate supports for in-home care.

2. Low SNF admissions due to more intensive home care: It is also possible that some models of care provide intensive support in the home for individuals with the most complex care needs, and, thus, experience very low SNF or long-stay nursing home admissions as a result. The PACE model is an alternative care model to support individuals with nursing home level of care need in their home and community. PACE also has features that make it attractive to individuals who are more likely to remain in their home and community and thus likely, with sufficient support, to transition to institutional care.

3. Low SNF admissions due to access to long-stay admissions: Integrated plans would also be able to admit members directly to a long-stay nursing facility, without the need for a short-term SNF admission – which in Medicare is definitionally "post-acute." Lower costs for Medicaid nursing facility stays may incent plans to avoid SNF admissions for members making a transition to institutional care.

Directions for Future Research

This project used a novel research approach to evaluate plans that integrate LTSS and was unique in its access to linked assessment and utilization data from plans. If it were possible to work with similar data at the individual level, a more sophisticated matching methodology would allow for a deeper understanding of plan impact. However, there are substantial barriers to third-party researchers accessing this level of data from proprietary health plans.

The findings of this analysis identify some areas where integrated plans have lower than expected utilization of medical services by their high-need Medicare beneficiaries. Some of these differences are substantial and warrant further investigation. It will be important to include a better understanding of the interaction of Medicare-covered and Medicaid-covered services for the populations with high levels of functional limitation.

It will be important to do this kind of analysis with individual records. In some states, individual records that link Medicare and Medicaid claims data are starting to become available to researchers. In addition, there has been a recent suggestion from CMS that encounter data reported to CMS by health plans may be made available to researchers. These individual records linked with detailed assessment data would provide a particularly valuable data resource for measuring the impact of integrated care on medical utilization and outcomes.

To fully understand the differences that we have observed, it is necessary to understand how care is delivered in the plan and the incentives and objectives the plans are pursuing. In preparation for this study, the research team prepared several qualitative case studies of integrated plans that included the plans that elected to participate in this study. These case studies helped the team to understand each plan's unique care model and the factors that could be influencing the outcomes that were observed in this study. Future research needs to focus more attention on the way the plans deliver care and how that relates to the medical utilization patterns that are experienced by the plans' members.

Future research should also seek to better understand how integrated care affects the type and amount of medical utilization for specific groups of individuals in the plan. Particularly, how medical utilization differs for enrollees with a more moderate level of need.

Conclusion

The results from this study of medical utilization in the five integrated plans are generally directionally positive (show lower utilization rates for integrated plans than the rates predicted for a traditional Medicare population) – but these differences are not all significant at a 95 percent confidence interval. The type of utilization (hospital, ED, or SNF) that showed the greatest difference between observed and predicted rates varied by plan. None of the plans had observed rates that were substantially lower on all three measures.

The clearest result is that the three SCO plans had consistently lower hospitalization rates than the fee-for-service population. SCO hospitalization rates were 40 to 70 percent lower than would be expected if this population was enrolled in traditional fee-for-service Medicare (one result was significant, and the other two were at the lower edge of the margin of error). It is important to note that the observed hospitalization rates for the high-need members that are in the study populations from these three plans are equivalent to the hospitalization rate for the average Medicare beneficiary (most of whom would not be considered "high need").

This result did not hold for the MMP and PACE plans. We believe that different results for these two plans are largely due to an underestimate of hospital admissions for these plans by the predictive model, resulting from an inability to adjust the model's population to fully match what the researchers surmise are higher acuity populations in those two plans. For both plans, the study populations had a "nursing home level of care" need. This level of acuity would be greater than was true for the study populations from the NHATS and the three SCO plans.

Another factor that may be operating with the three SCO plans, and is not with the MMP is the time factor. The three SCO plans have been in operation for 11 to 14 years, and have well-established care models in place with a relatively stable population. The data from the MMP is from its first full year of operation. It is not unusual for integrated plans to experience high levels of medical utilization in their

first few of operation, as they enroll populations that may have deferred medical care. Over time, the activity of a care manager and development of a better set of services and supports in the community and at home may help reduce the frequency of avoidable ER visits and hospitalizations

The differences between observed rates for integrated plans and predicted rates on ED visits was mixed. Two of the plans (one SCO plan and the PACE plan) had a substantially (50 to 70 percent) lower rate of ED visits. Both results were significant. The other three plans had the same or slightly (10 to 20 percent) lower rate, but none of these results were significant.

The differences between observed and predicted rates for integrated plans on SNF admissions was similarly mixed. Three of the plans (one SCO, the MMP, and the PACE plan) had substantially lower-than-predicted rates on SNF admissions (70 percent or greater difference). All three results were significant. The other two SCO plans had SNF rates matching the rate predicted for their members had they been enrolled in traditional Medicare – there was no significant difference.

The variation in differences between observed and predicted medical utilization across plans suggest that differences in the care models of the individual plans and their approach to meeting the needs of their members may influence whether the plan has lower rates for ED visits, hospitalizations, or SNF admissions. All of the plans had a substantially lower-than-predicted rate on at least one of the utilization measures – with a difference on that aspect of utilization that was substantial (in the range of 40 to 70 percent or more difference).

This study was undertaken to test a novel methodology for measuring differences between medical utilization in integrated plans compared to fee-for-service. Given limitations in the methodology, it is not possible to prove that there was a causal relationship or to calculate the medical savings that resulted from providing holistic services and supports through integrated plans. However, the directionality of the results is consistent and strong across the plans and provides a view of the potential in integrated LTSS to achieve substantial health care savings.

Appendix I: Methodology Limitations and Adjustments

The study sought to assess the effects of integrated plans on services utilization. Past evaluations have typically relied exclusively on Medicare claims data to characterize a benchmark population, whereas this study used Medicare claims linked to NHATS survey responses. The NHATs survey offers more comprehensive information on characteristics such as ADLs, IADLs, cognitive impairment, and demographic variables that are not available in claims data alone. However, the NHATS sample size is limited for individuals with the high level of need that this study focuses on (n=493), which increased the degree of uncertainty surrounding our results.

To examine the intervention (integrated care), the research team relied on plan assessment and encounter data. Collaborating with the health plans to obtain plan-level information posed substantial logistical barriers. Unlike the federal Medicare databases, health plan data is not cleansed or adjudicated for research use and requires substantial up-front administrative work before analysis can begin. A detailed discussion of challenges encountered in the research process follows.

Absence of Independent Data Validation

This study relied on plans sharing data. Although plan data submissions were reviewed for logical inconsistencies and completeness, it was not possible independently validate the accuracy of the data. Therefore, the results of this study are contingent on the accuracy of the data plans submitted.

Lack of Individual Level Data from Plans

The approach to this analysis was constrained by the inability to access individual-level information from the health plans. Access to individual-level information for both NHATS and plan enrollees would have enabled use of a different approach involving individual-level matching for specific characteristics.

As noted in the methodology section of the Report, “because individual-level data from each health plan was not available, it was not possible to compute the joint distribution or covariance structure. Therefore, the approach is subject to the assumption that the covariance correlation terms is zero, or that there is no difference in the covariance structure that would be observed for the health plan and the actual covariance structure that is observed for the NHATS comparison.”

The possibility cannot be ruled out, therefore, that differences in utilization between the plan population and the traditional Medicare population that could result from selection bias toward one group or the other are not fully controlled by the variables available to construct the model. Without the ability to do a covariance structure analysis, this possibility cannot be ruled out. While the NHATS and plan study populations have the same high acuity levels, it is possible that those who selected into integrated plans were more amenable to management or were lower utilizers of health care.

Limited Sample Size

The NHATS benchmark sample is relatively small due to the focus on a subpopulation with significant impairment, which is relatively rare in the older population at large, but more common among enrollees in integrated plans. The availability of a relatively small sample was most clearly seen in the ethnicity variable (i.e., Hispanic/non-Hispanic). Only 6 percent of the NHATS benchmark population is Hispanic, while participating integrated plans had enrolled populations that were between 37 and 83 percent Hispanic. The small prevalence of Hispanics in NHATS—fewer than 40 individuals—means that the dataset may not be able to make reliable projections about outcomes for this community. When the variable for Hispanic ethnicity was included in the plan models, there was a substantial effect on predicted utilization. Specifically, a larger Hispanic population was associated with lower predicted utilization across the three outcome variables. Due to this outsize impact, the research team decided not to include the Hispanic variable in the plan models. The results of models including this variable are available in Appendix II of this report.

Inability to Adjust for Geographic Variation

Integrated health plans operate with state or local markets, and the NHATS benchmark does not have a large enough sample to account for geographic variability in population-level characteristics and utilization. The literature has well documented the variation in healthcare spending and utilization across U.S. geographic regions.[16] This analysis did not account for the potential impact of geographic variation available on results.

Factors Influencing Model Results

The predictions of the study models rely on adequately matching the NHATS benchmark population to the sample population enrolled in each plan. Several factors affected the ability to make this match.

- Omitted Variables: If the project cannot account for a key characteristic defining the plan population in the study model, this may influence the dependent variable in the equation predicting medical utilization. If the likely direction of the missing variable's impact and the relative difference along that variable is known for the benchmark and study populations, then the research team can surmise from the separation how including that variable might change the model predictions.

- Varying Data Availability Across Plans: If a variable was available in the NHATs study population and in the data from some of the plans but not all of the plans, the inclusion or exclusion of that variable could affect predicted rates. In this case, the research team tested the sensitivity of the predicted results (including and excluding the variable) and provided the results of sensitivity tests in the appendix.

[16] M Gornick (1982) "Trends and regional variations in hospital use under Medicare," *Health Care Finance Review* 3:41-73. JR Knickman and AM Foltz (1984) "Regional differences in hospital utilization. How much can be traced to population differences?" *Med Care* 22:971-86. WP Welch et al. (1993) "Geographic variation in expenditures for physicians' services in the United States," *NEJM* 328:621-7.

- Large Differences in Benchmark and Study Population Characteristics: A variable in the model for which the "n" in the baseline is a small percentage of the NHATS study population (e.g., Hispanic ethnicity) while the plan's "n" is a large percentage of the plan's population, may produce a result that is anomalous. In this case, the research team tested the sensitivity of the predicted results - including and excluding the variable – to determine the magnitude of the variable's impact on the predicted results.

Factors Affecting the Results from the Plans

- Population Acuity: Observed rates will vary with the acuity of the plan population. The NHATS model is designed to adjust predicted utilization for variance that can be explained in terms of the acuity of the population.

- Differences in the Plans' Care Models
    - Differences between managed care organizations and traditional fee-for-service Medicare in the way beneficiaries access the healthcare system as a result of differing regulations or reimbursement incentives (e.g., the requirement in traditional Medicare of a qualifying three-day prior hospitalization before SNF admission is covered).
    - Goals for the plan that are influenced by the structure of payment, the risk that plans hold, the priorities of the state that are laid out in the managed care organization contract with the state. For example, the state may have incentives in place that discourage SNF use for program participants.
    - Unique aspects of an organization's care model that may influence how resources in the system are used.

Appendix II: Sensitivity Analysis Including Variable for Hispanic Ethnicity

Several of the participating plans were able to provide data on ethnicity for the study population, specifically the share of the population that identified as being Hispanic. In the primary analysis the research team excluded ethnicity because the plan populations varied substantially from one another as well as the nationally representative NHATs sample. While between 37 and 83 percent of plan enrollees identified as Hispanic, this is true of only 7 percent of the NHATS benchmark population. The small sample of Hispanics in NHATS — only 31 individuals — means that the dataset did not have high enough prevalence of Hispanic ethnicity to make projections about outcomes for this community.

Table A1: NHATS Model Results Including a Variable for Ethnicity			
	SCO 1	MMP	PACE
Share Hispanic	37%	83%	37%
Observed Hospitalization Rate	0.35	0.80	0.65
Predicted Hospitalization Rate	0.61 (0.30-0.92)	0.41 (0.18 – 0.65)	0.61 (0.36-0.87)
Observed ED Visit Rate	1.20	1.30	0.42
Predicted ED Visit Rate	1.39 (0.84-1.94)	1.10 (0.62 – 1.58)	1.19 (0.78-1.59)
Observed # SNF Events	0.20	0.07	0.00
Predicted # SNF Events	0.13 (0.02-0.24)	0.09 (0.00 – 0.24)	0.20 (0.05-0.34)
Sample Size	>2,000	419	397

When the study included a variable for Hispanic ethnicity in the regression, there was a dramatic effect on predicted utilization (Table A1), relative to the predicted utilization from regressions excluding this variable. Specifically, including ethnicity in the model lead to decreases in predicted utilization for the plans due to the fact that Hispanics have lower utilization that respondents who are white or other races in the NHATS benchmark population. This is a common health services research phenomenon that is referred to as "the Hispanic Paradox" and has been well documented.[17]

[17] See, for example: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3673509/

Appendix III: Sensitivity Analysis Excluding Decedents

Three of the participating plans were able to provide data for the study sample with and without individuals who died during the observation period. The primary analysis for this study included decedents where this data was available. There were substantial differences in mortality rates between the NHATS benchmark population (19.1%) and the plan populations (2.3% to 6.5%). The magnitude of this difference suggests there may be underlying variation between the benchmark and plan populations that is not captured in the model. However, there are not notable differences between results when decedents were excluded (Table A2 below) and results when decedents were included in the population (Table 3 in the main text).

Table A2: NHATS Model Results Excluding Decedents			
	SCO 1	SCO 2	SCO 3
Observed Hospitalization Rate	0.35	0.41	0.26
Predicted Hospitalization Rate	0.63 (0.24-1.02)	0.82 (0.32-1.32)	0.76 (0.44-1.09)
Observed ED Visit Rate	1.16	0.75	1.45
Predicted ED Visit Rate	1.42 (0.80-2.04)	1.58 (0.88-2.29)	1.48 (0.98-1.98)
Observed # SNF Events	0.17	0.18	0.08
Predicted # SNF Events	0.15 (0.02-0.28)	0.26 (0.04-0.47)	0.26 (0.11-0.41)
Sample Size	>2,000	>2,000	>2,000

Appendix IV: Observed Utilization in the Overall Medicare Population

As a further benchmark to understand the magnitude of the plan outcomes, the research team ran comparable measures for the entire NHATS population (which is a representative sample of the national population 65 and older enrolled in traditional Medicare (fee-for-service)). These rates of medical utilization are much lower for the entire population than would be expected for the predicted and observed outcomes for the high-need study populations. In instances where the outcomes for the study and total population are similar, however, it is further evidence of the magnitude of the rates reported by the plans. (Table A-3 below provides the observed rates for the total NHATS population for comparison with the results reported in Table 3 in the main text).

Table A3: Observed Utilization in the Overall Medicare Population[18]

	Overall Medicare (95% Confidence Interval)	NHATS Benchmark Population (95% Confidence Interval)
Hospitalization Rate	0.26 (0.24-0.29)	0.69 (0.56-0.82)
ED Visit Rate	0.54 (0.50-0.58)	1.29 (1.10-1.49)
# SNF Events	0.07 (0.06-0.08)	0.25 (0.19-0.31)

[18] Limited to traditional fee-for-service Medicare enrollees age 65 and older.

Acknowledgements

LTQA Project Team

G. Lawrence Atkins
Jennifer Windh
Anne Tumlinson

Johns Hopkins University Project Team

Jennifer Wolff
Judith Kaspar
Amber Willink
John Mulcahy

March 2019

LTQA appreciates the participation and suggestions from the members of our Technical Advisory Panel: Vincent Mor (Brown University), Randall Brown (Mathematica), Marc Cohen (UMass Boston), Tamara Kuznetski (University of Chicago), Howard Gleckman (Urban Institute), Paul Saucier (IBM), and Joshua Weiner (RTI-deceased). The methodology, its execution, and the analysis of the results are the product of the project team.

This Study was made possible through a grant from the Laura and John Arnold Foundation.