

Health-Related Quality of Life in High-Cost, High-Need Populations

Lindsey Jeanne Leininger, PhD; Marisa Tomaino, BA; and Ellen Meara, PhD

When states aim to improve care among Medicaid “superutilizers,”¹ federal alternative payment models incorporate refinements targeting high-cost, high-need (HCHN) populations,² and novel HCHN-focused quality improvement interventions are developed and tested by payers and health systems, the dominant measures of success are utilization and spending.³ Typical interventions augment standard care with enhanced care coordination and patient engagement strategies such as home visits, access to 24/7 nurse call lines, increased monitoring during transitions across care settings, and by connecting patients with social service providers.^{4,5} Associated impact evaluations rely heavily on administrative data on resource utilization as key outcome measures.^{6,7} Although elevated use of expensive health care services can be an independent marker of vulnerability in its own right,⁸ overreliance on utilization measures has several shortcomings. Common algorithms to identify HCHN patients for targeted intervention in health systems, due to reliance on health care utilization, can reinforce racial and other biases in access to care (eg, sex, geographic, economic).⁹ Furthermore, utilization measures ignore outcomes that are most valued by patients.

Patient-reported outcome measures offer one avenue to improve measurement of patient progress. These data capture patient perceptions of their underlying health status and functioning—fundamental dimensions to overall quality of life.¹⁰ Although patient-reported outcome measures have ample evidence of validity, reliability, and responsiveness in broad populations,^{11,12} one may worry that for HCHN patients, such measures will suffer from floor effects (ie, everyone is in poor health) or may not add value over more widely available utilization measures such as emergency department (ED) visits and hospitalizations. In this article, we ask whether a well-validated patient-reported outcome measure construct—health-related quality of life (HRQOL)—identifies meaningful differences within HCHN patients and whether and how HRQOL changes over time for this population. In addition to characterizing levels and trends in HRQOL among HCHN patients, we test the scientific appropriateness of adopting HRQOL for population and performance monitoring. Our overarching goal is to understand whether the

ABSTRACT

OBJECTIVES: To characterize patient-reported health and assess the psychometric performance of health-related quality of life (HRQOL) in high-cost, high-need (HCHN) populations.

STUDY DESIGN: Retrospective longitudinal study examining health care utilization, expenditures, and patient-reported health comparing a baseline (year 1) and follow-up year (year 2).

METHODS: The sample includes adults (n = 46,934) participating in the Medical Expenditure Panel Survey between 2011 and 2016. We estimated HRQOL for each sample member using the physical and mental health scales from the Medical Outcomes Study Short Form 12. We compared HRQOL stratified by HCHN, defined as patients whose baseline (year 1) demographics, utilization, and clinical characteristics predicted top decile health spending in year 2. Analyses assessed the validity, reliability, and responsiveness of physical and, separately, mental health scales.

RESULTS: Among HCHN adults, the physical health scale exhibited robust measure validity, reliability, and responsiveness across all age groups; the mental health scale did not. Mean physical health was 1.25 SDs lower in HCHN vs other patients (37.9 vs 51.0 on a 0-100 scale increasing in self-perceived health; pooled SD, 10.5). Regressions indicated that a 0.5-SD increase in year 1 physical health among HCHN adults predicted a 5-percentage-point (10%) decrease in the probability of top decile health spending in year 2. In contrast to health care spending, HRQOL did not exhibit reversion to the mean in HCHN patients.

CONCLUSIONS: Patient-reported health outcomes remain poor in HCHN populations, even after health care utilization recedes. HRQOL is a promising outcome measure for HCHN-focused payment and delivery interventions.

Am J Manag Care. 2023;29(7):362-368. doi:10.37765/ajmc.2023.89396

current utilization-centric focus on HCHN populations and initiatives that serve them could benefit from measuring and prioritizing patient-centric health outcome improvement.

METHODS

Data

We used the Medical Expenditure Panel Survey (MEPS) to characterize patterns of self-reported physical and mental health and to assess the performance of these measures among HCHN populations. The analytic sample of 46,934 adults 18 years and older drew from 5 two-year panels of the Agency for Healthcare Research and Quality's nationally representative MEPS spanning 2011 to 2016. The MEPS includes patient-reported health measures and health care utilization and cost outcomes among HCHN populations.¹³ The **eAppendix Figure** (eAppendices available at ajmc.com) depicts the timing of data capture by round for each study measure described below.

Measures

HRQOL. HRQOL reflects an individual's health and its impact on several domains, including physical, mental, emotional, and social functioning. A large literature validates the use of HRQOL across clinical and community-based settings (examples include Moriarty et al,¹⁴ Slabaugh et al,¹⁵ Zack,¹⁶ and Vilagut et al¹⁷).¹⁸ The widespread, standardized collection and dissemination of HRQOL data enable interested stakeholders to benchmark health across place, time, and population subgroups; this makes HRQOL an attractive measure in new settings and for new subgroups—such as HCHN populations—for whom no such benchmarks currently exist. In contrast to condition-specific patient-reported outcome measures, the global nature of HRQOL is potentially valuable in the context of HCHN populations, which are characterized by a high prevalence of multimorbidity.¹⁹ To date, HRQOL validation has not targeted HCHN populations.

The MEPS captures HRQOL using the Medical Outcomes Study Short Form 12 (SF-12, version 2), developed in the 1980s.²⁰ The SF-12 is constructed of 2 subscales, the Physical Health Composite Scale and Mental Health Composite Scale, referred to as “physical health” and “mental health” throughout this article. The physical and mental health composites use 12 identical items (**eAppendix Table 1**); however, the (proprietary) item weighting differs across them. Both composites are scaled to have a mean (SD) of 50 (10). Higher scores represent better health.²⁰ Numerous MEPS-specific studies have demonstrated the reliability and validity of both scales in cohorts defined by the presence of either specific symptoms (eg, a study of individuals with noncancer pain),²¹ or specific conditions (eg, a study of individuals with diabetes).²²

Other measures. We used the MEPS summary measures capturing annual total ED visits and overnight inpatient stays, as well as total annual health care expenditures. We adapted the method developed

TAKEAWAY POINTS

We provide the first psychometric assessment of health-related quality of life (HRQOL) among high-cost, high-need (HCHN) populations. This is important for the following reasons:

- ▶ Health care utilization and expenditures revert to the mean among HCHN populations.
- ▶ In stark contrast, patient-reported health outcomes do not; poor self-rated health persists.
- ▶ Outcome measurement for HCHN-focused interventions must move beyond narrowly focused health care utilization and expenditures to appropriately capture patient suffering.
- ▶ We demonstrate that HRQOL is a psychometrically sound candidate for patient-reported outcome measurement in HCHN-focused interventions.

by Fleishman and Cohen²³ to identify the following chronic conditions: anxiety, arthritis, asthma, cancer, cerebrovascular disease, diabetes, emphysema, heart disease, high blood pressure, high cholesterol, mood disorder (including depression), schizophrenia, and substance use disorder. **eAppendix A** provides further detail.

We used the following measures as covariates in adjusted analyses, or to stratify analyses as appropriate: age, sex, education, marital status, income, and geographic region. MEPS participants who were pregnant were excluded from the analytic sample. In keeping with related work,²¹ we incorporated 2 common measures of self-reported health status that are conceptually related to HRQOL for validity testing: self-reported general health and self-reported mental health.

Empirical Approach

HCHN cohort definition. Following related literature^{23,24} and common implementation practice,²⁵ we identified HCHN sample members using predicted spending. We estimated the probability that an individual is in the top expenditure decile next year (year 2) as a function of baseline (year 1) measures of health care utilization, expenditures, chronic conditions, and sociodemographics. We then selected sample members in the top decile of predicted year 2 spending for our analytic sample. **eAppendix B** includes methodological details and regression results.

Psychometric assessment. Following the CDC,²⁶ the scientific literature (eg, Aaronson et al²⁷), the National Quality Forum,²⁸ and CMS,²⁹ we designed empirical tests assessing the validity, reliability, and responsiveness of HRQOL among our HCHN cohort.

Validity. To assess validity, we asked the following questions:

1. Does HRQOL correlate with other well-established, conceptually related patient-reported outcome measures in a cross-section (*construct validity*)?
2. Is there appreciable regression to the mean in HRQOL, biasing longitudinal comparisons (*internal validity*)?
3. Does HRQOL predict future outcomes of interest (*predictive validity*)?

We assessed construct validity by estimating the cross-sectional association between year 1 values of mental health and physical health and the year 1 number of chronic conditions among HCHN

TABLE 1. Descriptive Statistics*

Characteristics	Full sample (N = 46,934)	Non-HCHN (n = 42,244)	HCHN (n = 4690)
Age in years, mean (SD)	47.6 (17.0)	45.5 (17.1)	64.2 (12.4)
Female	50.5%	49.5%	58.2%
Education			
Less than high school	14.1%	14.1%	14.1%
High school	45.7%	45.5%	47.4%
College	40.2%	40.4%	38.5%
Income category, year 1			
Negative or poor	12.3%	12.1%	14.0%
Near poor	4.6%	4.4%	5.9%
Low income	13.4%	13.2%	14.7%
Middle income	29.9%	30.4%	25.9%
High income	39.8%	39.9%	39.6%
Geographic region, year 1			
Northeast	17.6%	17.1%	21.0%
Midwest	21.8%	21.3%	25.8%
South	37.5%	37.6%	36.2%
West	23.1%	23.9%	17.1%

(continued)

sample members, adjusting for sociodemographic factors (age, sex, and education). This “sum-of-conditions” approach proxies the Charlson and Elixhauser indices, both validated comorbidity measures.³⁰ A negative correlation or gradient between the number of chronic conditions and HRQOL suggests that the latter demonstrates appropriate construct validity.

We estimated 1-year changes (year 2 minus year 1) to assess the extent of regression to the mean in HRQOL among HCHN sample members, a key internal validity check. Spending- and utilization-based measures exhibit considerable regression to the mean among HCHN patients.^{5,31} To our knowledge, however, no analogous evidence tests whether the pattern holds true for HCHN patients’ HRQOL. We compared the pattern of mental health and physical health scales over time with that of other patient-reported outcomes—specifically, single-item self-reported general and mental health—as well as with those exhibited by traditional resource-based measures capturing elevated overall expenditures and ED and inpatient use. Additionally, we compared the longitudinal HRQOL patterns of HCHN vs non-HCHN patients, formalized using a difference-in-differences framework (associated model specification detailed in [eAppendix C](#)).

We estimated the capacity of HRQOL to predict future expenditures and utilization—indicators of HCHN persistence—as our key predictive validity test. To our knowledge, no research exists indicating whether HRQOL independently predicts expenditures and utilization among HCHN populations. Using logit models, we estimated whether an individual is in the top decile of medical expenditures in year 2 as a function of year 1 measures of HRQOL scales of mental and physical health among HCHN patients,

controlling for year 1 sociodemographic characteristics (results in [eAppendix Table 2](#)). A positive coefficient exhibiting meaningful magnitude and statistical significance is indicative of predictive validity. We estimated similar models for year 2 top decile ED visits and for any inpatient hospitalization in year 2.

Reliability

Our primary reliability concern was ensuring HRQOL’s robustness across the variety of HCHN definitions employed by different stakeholders. The 2 main approaches to HCHN cohort definition are prospective assignment—the approach we used—which employs predictive modeling to identify those patients who are at elevated probability of HCHN status in the upcoming year, and rules-based assignment, employing cutoffs based on the prior year’s utilization and/or expenditures. The key related reliability concern for HRQOL is whether its values are comparable across the various HCHN cohort classifications and whether these values differ meaningfully relative to non-HCHN patients. We tested this by computing mental and physical health scales for our preferred cohort and 4 additional (rules-based) cohorts: top decile expenditures in year 1; 2 or more ED visits in year 1; any hospitalization in year 1; and 3 or more chronic conditions in year 1. We computed unadjusted means overall and stratified by age, comparing values within each of the 5 HCHN definitions and across the HCHN vs never-HCHN cohorts.

Responsiveness

We assessed the potential responsiveness of HRQOL by exploring whether HRQOL measures exhibit meaningful spread within the HCHN cohort. Spread can indicate unwarranted variation amenable to health care intervention.³² We were particularly mindful of the potential presence of floor effects among HCHN patients, given their high levels of morbidity. The extent of underlying spread also has implications for statistical power requirements, a key implementation consideration. We provided simple exploratory power calculations assuming effect sizes of 0.3, 0.5, and 1 SD. Values of 0.3 and 0.5 SD reflect “minimally important differences” in HRQOL documented in clinical studies of treatments targeting vulnerable populations (eg, Jayadevappa et al,³³ Norman et al³⁴) and in population-based studies (eg, Alghnam et al,³⁵ Baragaba et al,³⁶ Utah Department of Health³⁷). Finally, we estimated the marginal effect of increasing year 1 HRQOL by 0.3 and 0.5 SDs on the probability of year 2 top decile expenditure, using the associated predictive validity model discussed above. This model, although observational, demonstrates whether feasible improvements in mental and physical health might translate into future decreases in resource-intensive utilization.

RESULTS

Table 1 displays descriptive characteristics for the entire sample and separately across HCHN and non-HCHN sample members (subsequently referred to as HCHNs and non-HCHNs for brevity). The sociodemographic composition across the 2 strata were fairly

similar, with the notable exceptions of age and sex. The mean ages of HCHNs and non-HCHNs were 64.2 and 45.5 years, respectively. A higher percentage of HCHNs were women compared with non-HCHNs (58.2% vs 49.5%).

Our key measures of interest—HRQOL measures of mental and physical health—were lower in HCHNs than in non-HCHNs. However, the relative difference across the 2 measures was much larger for physical than for mental health. The mean physical health score of 37.9 among HCHNs was 1.25 SDs lower than the mean of 51.0 among non-HCHNs (using the full sample SD). In contrast, the mean mental health score for the 2 groups differed by only 3.4 points (48.6 vs 52.0 among HCHNs vs non-HCHNs), or a 0.35-SD difference (again, using the full sample SD).

Figure 1 and Figure 2 display cross-sectional evidence regarding reliability and concurrent validity. The panels comprising Figure 1 plot HRQOL distributions stratified by age and various definitions of HCHN status. The age stratification helped us rule out the possibility that the HCHN vs non-HCHN differences in HRQOL are driven solely by differences in age distribution. Computing mental and physical health across different HCHN definitions served as our primary reliability test, assessing the robustness of HRQOL's use across varying but related HCHN cohorts. Within each age group, the never-HCHN sample—defined as not meeting any of the HCHN definitions—exhibited considerably higher physical health scores relative to all HCHN populations (all differences statistically significant). Moreover, the distributions exhibited considerable spread within HCHNs, providing reassurance regarding floor and ceiling effect concerns. The mental health gradient across HCHN and never-HCHN populations, in contrast, differed markedly across age strata. In younger age strata, mental health was much lower among HCHN than never-HCHN adults. For older age strata, which comprised the majority of HCHN adults, there was very little gradient in mental health. Small sample sizes for the younger age strata indicate a need for caution when interpreting these different age patterns; notably, only 393 of the HCHN cohort members were younger than 45 years (age strata sample sizes displayed in eAppendix Table 3).

Figure 2 demonstrates the results from our concurrent validity analysis, which assessed the association between year 1 HRQOL and chronic condition count among HCHN adults, adjusted for age, sex, and education. The estimates indicated that both mental and physical health declined fairly monotonically as the condition count increases, demonstrating their validity as benchmarked against increasing levels of comorbidity. Here again we saw that the gradient is stronger for physical than for mental health. For example, moving from 3 chronic conditions (the mean among HCHN adults) to 5 or more chronic conditions (22.4% of the HCHN sample) was associated with an almost 5-point decrease in physical health, compared with an approximately 3.5-point decrease in mental health.

Table 2 incorporates longitudinal analyses. Patterns for utilization- and resource-based measures among HCHNs comparing years 1 and 2 demonstrated mean reversion patterns comparable with those

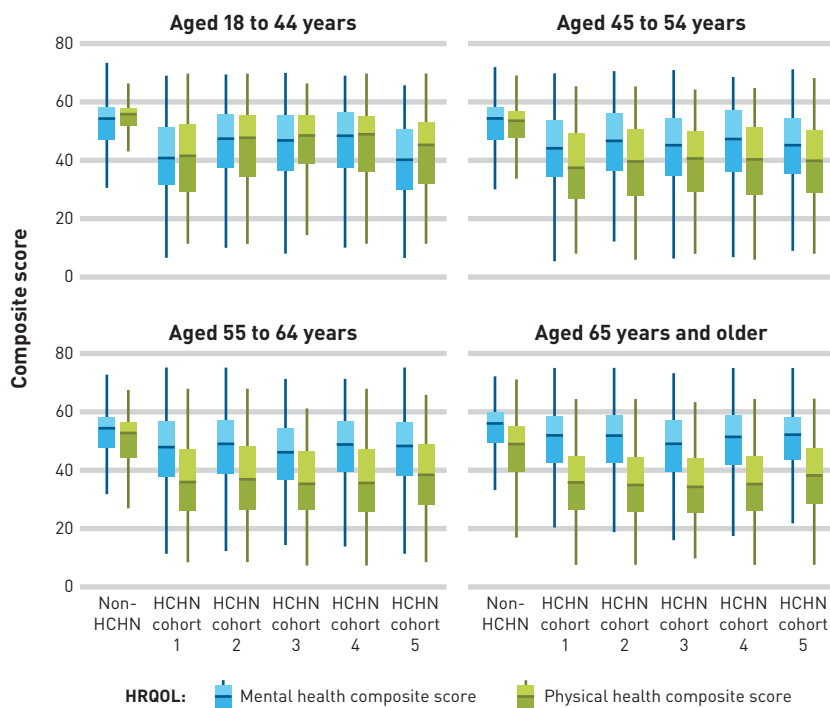
TABLE 1. (Continued) Descriptive Statistics*

Characteristics	Full sample (N=46,934)	Non-HCHN (n=42,244)	HCHN (n=4690)
Married	52.9%	53.0%	52.6%
Chronic conditions, year 1			
Anxiety	9.0%	7.6%	19.9%
Arthritis	7.3%	4.3%	30.8%
Asthma	4.9%	3.7%	14.3%
Cancer	5.8%	4.1%	19.5%
Cerebrovascular disease	1.2%	0.6%	6.4%
Diabetes	9.9%	6.6%	36.3%
Emphysema	4.1%	2.7%	15.4%
Heart disease	6.4%	3.6%	28.3%
High blood pressure	25.2%	20.4%	63.6%
High cholesterol	19.6%	15.5%	52.6%
Mood disorder	9.7%	7.4%	28.4%
Schizophrenia	0.3%	0.2%	1.5%
Substance use disorder	0.4%	0.3%	1.7%
No. of conditions, mean (SD)	1.0 (1.4)	0.8 (1.1)	3.2 (1.6)
Behavioral and physical conditions			
Behavioral health condition	16.3%	13.4%	39.8%
Physical health condition	42.1%	36.0%	91.3%
Physical self-reported health			
Poor/fair	12.8%	9.8%	36.4%
Good	28.0%	27.0%	35.8%
Very good	34.3%	36.1%	20.2%
Excellent	24.9%	27.1%	7.6%
Mental self-reported health			
Poor/fair	7.4%	6.0%	18.8%
Good	23.8%	22.7%	32.6%
Very good	31.4%	32.2%	25.3%
Excellent	37.3%	39.1%	23.3%
Physical health score, mean (SD)	49.6 (10.5)	51.0 (9.4)	37.9 (12.0)
Mental health score, mean (SD)	51.6 (9.5)	52.0 (9.3)	48.6 (11.1)
Total expenditures, mean (SD)	\$4888 (\$15,689)	\$2442 (\$12,955)	\$24,429 (\$23,920)

HCHN, high-cost, high-need.

*All differences between HCHN and non-HCHN subgroups are significant at the $P < .01$ level, except for education.

of the broader related literature.^{5,31} For example, among people in the top decile of predicted year 2 spending, 83.7% had observed top decile spending in year 1, but only 47.6% remained in the top decile of observed spending in year 2. In stark contrast, neither the mental nor physical health scores among the HCHN cohort exhibited qualitatively meaningful or statistically significant year-over-year regression to the mean. The measures of patient-reported fair or poor health exhibited similar stability, providing additional evidence that decreases in resource-intensive utilization do not translate to

FIGURE 1. Health-Related Quality of Life Stratified by Age and HCHN Status*

ED, emergency department; HCHN, high-cost, high need; HRQOL, health-related quality of life.

*Boxplot of year 1 physical and mental health composite scores for 5 HCHN cohorts and a non-HCHN cohort stratified by age. The 5 HCHN cohorts are not mutually exclusive and are defined as (1) being in the top decile of year 2 predicted expenditures, (2) being in the top decile of year 1 expenditures, (3) having 2 or more ED visits in year 1, (4) being hospitalized in year 1, and (5) having 3 or more chronic conditions in year 1. The non-HCHN cohort contains those who did not meet any of the 5 HCHN definitions. Cohort sizes are listed in eAppendix Table 3. All age groups exhibit higher physical health scores in the non-HCHN cohort. Mental health scores are much lower among the HCHN cohorts in younger age strata but are similar in older age strata.

comparable improvements in patient-reported outcome measures in an observational context. The differences-in-differences results, comparing year-over-year changes across HCHN and non-HCHN cohorts, demonstrated that although changes in expenditures and resource utilization among HCHN adults varied meaningfully from those among non-HCHN adults, changes in patient-reported outcome measures did not.

Estimates from the predictive validity regressions (Table 3) indicated that among the HCHN cohort, year 1 physical health exhibits robust, independent predictive capacity across all year 2 resource utilization measures (odds ratios in eAppendix Table 4). In contrast, year 1 mental health exhibited statistically significant predictive ability only for top decile expenditure membership. Moreover, the magnitude of the association between year 1 mental health and year 2 resource utilization measures was uniformly smaller than that of physical health. Table 3 displays predicted probabilities derived from the top decile expenditures regression. A 0.3-SD improvement in year 1 physical health was independently associated with a 3-percentage-point (6%) decrease in the probability

of year 2 top decile expenditure membership. A 0.5-SD improvement in year 1 physical health was associated with a 5-percentage-point (10%) decrease, and a 1-SD improvement in year 1 physical health was associated with a 9-percentage-point (19%) decrease.

Power calculations indicated that for both the mental and physical health scores, fewer than 20 HCHNs are required in each treatment arm to achieve 80% power for expected impacts of 1 SD (eAppendix Table 5). Sample sizes of approximately 65 individuals per group were needed to achieve 80% power for expected impacts of 0.5 SD, and approximately 145 per group were needed for expected impacts of 0.3 SD. Assuming that 10% of the commonly cited primary care panel size of 2500 patients qualifies as HCHN suggests that initiatives at the practice level—even those using within-practice treatment and control arms—may be suitably powered for expected impacts of 0.5 SD or higher.³⁸ Expected impacts on the order of 0.3 SD would likely require cross-practice treatment and comparison arms. Populations tracked by multipractice medical groups, even modestly sized integrated health systems, payers, and policy makers, easily exceed all of the minimum sample size requirements.

DISCUSSION

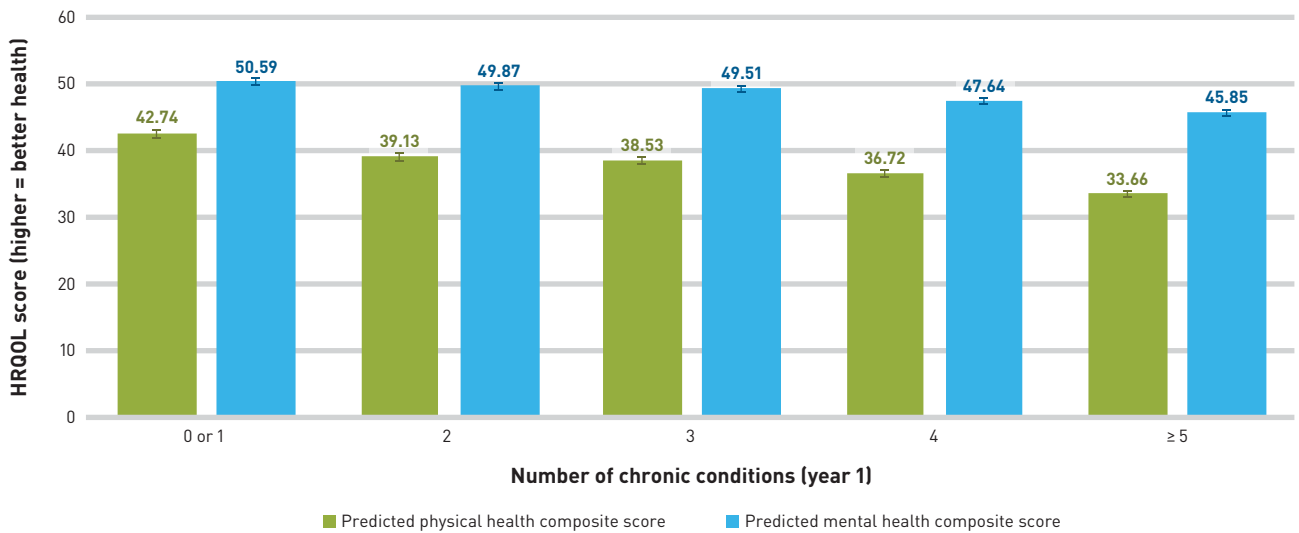
We found that HCHN adults experienced lower HRQOL relative to non-HCHN adults and that

this gradient was appreciably stronger for physical health compared with mental health. Together, our empirical tests suggested that physical health exhibited strong validity, reliability, and responsiveness as a measurement tool for HCHN adults, whereas mental health—which, unlike physical health, improves at older ages—was weaker along certain dimensions of validity and responsiveness. Unlike measures of resource utilization, HRQOL exhibited minimal regression to the mean among HCHN adults. Our findings imply that HRQOL represents an important, distinct construct from the utilization-based metrics currently used to hold providers or payers accountable for care of HCHN populations.

Limitations

Several limitations of the current study must be addressed prior to the use of HRQOL for higher-stakes HCHN measurement contexts, such as those involving meaningful financial or reputational impacts. First, our findings do not address the internal validity threat arising from likely differences in patient case mix, which would require risk adjustment. Moreover, tracing longer-term trajectories of the

FIGURE 2. Regression-Adjusted Association Between HRQOL and Chronic Condition Count^a



HCHN, high-cost, high-need; HRQOL, health-related quality of life.

^aThis bar chart shows year 1 physical and mental health composite scores adjusted for age, sex, and education by the number of chronic conditions in the HCHN cohort (n=4690). Whiskers represent SEs. Physical and mental health decline monotonically as the number of conditions increases.

TABLE 2. Year-Over-Year Changes in Outcome Measures: HCHN Cohort vs Non-HCHN Cohort^a

Outcome	HCHN cohort				Non-HCHN cohort				Difference-in-difference	
	Year 1	Year 2	Difference	P	Year 1	Year 2	Difference	P	Value	P
Physical health score	37.94	37.93	-0.01	<.97	51.00	50.95	-0.05	<.26	0.04	<.01
Mental health score	48.59	48.77	0.18	<.25	51.95	52.35	0.39	<.01	-0.21	<.01
Total expenditures	\$24,429	\$19,986	-\$4,443	<.01	\$2442	\$3721	\$1279	<.01	-\$5722	<.01
Top decile expenditures	0.84	0.48	-0.36	<.01	0.02	0.06	0.04	<.01	-0.41	<.01
Top decile ED visits	0.15	0.11	-0.04	<.01	0.02	0.03	0.01	<.01	-0.05	<.01
Any hospitalization	0.36	0.22	-0.14	<.01	0.02	0.05	0.02	<.01	-0.17	<.01
Fair/poor physical health	0.36	0.36	0.00	<.01	0.10	0.10	0.00	<.52	0.00	<.51
Fair/poor mental health	0.19	0.20	0.01	<.01	0.06	0.06	0.00	<.02	0.01	<.42

ED, emergency department; HCHN, high-cost, high-need.

^aThese estimates document the year-over-year change in health status and health care resource utilization, stratified by HCHN status. Health care expenditures and utilization exhibit considerable year-over-year mean reversion among the HCHN cohort, whereas health status measures do not. The difference in year-over-year trends across HCHN and non-HCHN cohorts is reflected in the difference-in-difference value, demonstrating the extent to which mean reversion differs across the 2 groups. Sample size for all models: 93,868 person-years.

interplay between HRQOL and resource utilization among HCHNs could improve predictive validity and yield clues on effective intervention design. Further work should explore trade-offs associated with using a shorter HRQOL instrument among HCHNs—for example, a single self-rated health item—to balance psychometric soundness vs measurement burden, as was done for non-HCHN populations in the seminal HRQOL development work.³⁹

Our estimates also revealed a need for further exploration of the divergent mental health patterns across different HCHN age strata. Statistical power limitations precluded our ability to produce the associated age-stratified validity assessments. Exploring the

potential for using mental health scores as measurement tools in interventions targeting younger ages is a potentially promising direction for future research. Additionally, further work on the interplay among sex, HCHN status, and HRQOL is important, as women seem to experience a higher prevalence of HCHN and have a different set of factors that drive HRQOL relative to men.⁴⁰

CONCLUSIONS

Our work contributes to the call for capturing the patient voice in measuring HCHN outcomes. We demonstrate that HCHN adults

TABLE 3. Predicted Probability of Membership in Year 2 Top Expenditure Decile by Values of Year 1 Health-Related Quality of Life^a

Value	HCHN cohort only (n = 4690)			
	Physical health		Mental health	
	Score	Phat	Score	Phat
Mean	37.94	0.48	48.59	0.48
Mean + 0.3 SD	41.93	0.45	52.30	0.47
Mean + 0.5 SD	43.92	0.43	54.15	0.46
Mean + 1 SD	49.91	0.39	59.72	0.45

HCHN, high-cost, high-need; phat, predicted probability.

^aThese estimates are predicted probabilities derived from a logistic regression of year 2 top decile expenditure membership on physical health score, mental health score, and a variety of other year 1 covariates (see eAppendix Table 4 for the full regression specification). The mental health score is held at its mean (48.59) for models varying the physical health score; the physical health score is held at its mean (37.94) for models varying the mental health score. All predicted probabilities are pairwise significant compared with the mean.

experienced significant health burdens that persist after their health care interactions recede. As such, an overemphasis on utilization when targeting care delivery efforts to HCHN adults excludes many who are still suffering and ignores the full burden of illness, which extends far beyond its cost. ■

Author Affiliations: Tuck School of Business, Dartmouth College (LJL), Hanover, NH; The Dartmouth Institute, Geisel School of Medicine, Dartmouth College (MT), Hanover, NH; Harvard School of Public Health, Harvard University (EM), Boston, MA.

Source of Funding: Agency for Healthcare Research and Quality's Comparative Health System Performance Initiative under grant No. U19HS024075.

Author Disclosures: The authors report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

Authorship Information: Concept and design (LJL, EM); acquisition of data (LJL, MT); analysis and interpretation of data (LJL, MT); drafting of the manuscript (LJL); critical revision of the manuscript for important intellectual content (LJL, MT); statistical analysis (LJL, MT); obtaining funding (EM); administrative, technical, or logistic support (LJL, MT); and supervision (LJL, EM).

Address Correspondence to: Lindsey Jeanne Leininger, PhD, Tuck School of Business, Dartmouth College, 100 Tuck Hall, Hanover, NH 03755. Email: lindsey.j.leininger@tuck.dartmouth.edu.

REFERENCES

- Medicaid Accelerated eXchange (MAX) series program: final report. New York State Department of Health. Updated December 2022. Accessed June 6, 2023. https://www.health.ny.gov/health_care/medicaid/redesign/dsrp/pops_workshops/max/docs/9-21_12-22_final_rpt.pdf
- Cotla CH, Lewis VA, Kao LS, O'Malley AJ, Chang CH, Fisher ES. Association between Medicare accountable care organization implementation and spending among clinically vulnerable beneficiaries. *JAMA Intern Med*. 2016;176(8):1167-1175. doi:10.1001/jamainternmed.2016.2827
- Das LT, Abramson EL, Kaushal R. High-need, high-cost patients offer solutions for improving their care and reducing costs. *NEJM Catalyst*. Published online February 5, 2019. <https://catalyst.nejm.org/high-need-high-cost-patients/solutions/>
- Long P, Abrams M, Milstein A, et al, eds. *Effective Care for High-Need Patients: Opportunities for Improving Outcomes, Value, and Health*. National Academy of Medicine; 2017. Accessed October 4, 2019. <https://ccn.loc.gov/2017041343>
- Johnson TL, Rinehart DJ, Durfee J, et al. For many patients who use large amounts of health care services, the need is intense yet temporary. *Health Aff (Millwood)*. 2015;34(8):1312-1319. doi:10.1377/hlthaff.2014.1186
- Finkelstein A, Zhou A, Taubman S, Doyle J. Health care hotspots — a randomized, controlled trial. *N Engl J Med*. 2020;382(2):152-162. doi:10.1056/NEJMs1906848
- Komaromy M, Bartlett J, Gonzales-van Horn SR, et al. A novel intervention for high-need, high-cost Medicaid patients: a study of ECHO care. *J Gen Intern Med*. 2020;35(1):21-27. doi:10.1007/s11606-019-05206-0
- Ryan J, Abrams MK, Doty MM, Shah T, Schneider EC. How high-need patients experience health care in the United States. *The Commonwealth Fund*. December 9, 2016. Accessed June 6, 2023. <https://bit.ly/44pWV3>
- Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-453. doi:10.1126/science.aax2342
- Health-related quality of life (HROQL): well-being concepts. CDC. Updated October 31, 2018. Accessed June 6, 2023. <https://www.cdc.gov/hrqol/wellbeing.htm>
- Cella D, Hahn EA, Jensen SE, et al. *Patient-Reported Outcomes in Performance Measurement*. RTI Press; 2015. doi:10.3768/rtipress.2015.bk.0014.1509
- Kallenberg GA. The origins and promise of PROMIS – Patient Reported Outcomes Measurement Information System. Workshop on Practice Based Research. December 5, 2015. Accessed October 4, 2019. https://medschool.ucsd.edu/som/fmph/research/centers/integrativehealth/TrainingEducation/Documents/Kallenberg_The_Promise_of_PROMIS_11-27-15.pdf
- Burns M, Mullaly J. Healthy-time measures of health outcomes and healthcare quality. National Bureau of Economic Research. August 2016. doi:10.3386/w22562
- Moriarty DG, Zack MM, Kobau R. The Centers for Disease Control and Prevention's healthy days measures – population tracking of perceived physical and mental health over time. *Health Qual Life Outcomes*. 2003;1:37. doi:10.1186/1477-7525-1-37
- Stabaugh SL, Shah M, Zack M, et al. Leveraging health-related quality of life in population health management: the case for healthy days. *Popul Health Manag*. 2017;20(1):13-22. doi:10.1089/pop.2015.0162
- Zack MM, CDC. Health-related quality of life – United States, 2006 and 2010. *MMWR Suppl*. 2013;62(3):105-111.
- Vilagut G, Forero CG, Pinto-Meza A, et al. The mental component of the Short-Form 12 Health Survey (SF-12) as a measure of depressive disorders in the general population: results with three alternative scoring methods. *Value Health*. 2013;16(4):564-573. doi:10.1016/j.jval.2013.01.006
- Search string "health-related quality of life." ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/results?cond=&term=health-related+quality+of+life&cntry=&state=&city=&dist=>
- Mann C. Targeting Medicaid super-utilizers to decrease costs and improve quality. Medicaid. 2013. Accessed October 30, 2019. <https://www.medicare.gov/federal-policy-guidance/downloads/cib-07-24-2013.pdf>
- Hays RD, Sherbourne CD, Mazel R. *User's Manual for the Medical Outcomes Study (MOS) Core Measures of Health-Related Quality of Life*. RAND Corporation; 1995.
- Hayes CJ, Bhandari NR, Kathe N, Payakachat N. Reliability and validity of the Medical Outcomes Study Short Form-12 version 2 (SF-12v2) in adults with non-cancer pain. *Healthcare (Basel)*. 2017;5(2):22. doi:10.3390/healthcare502022
- Kathe N, Hayes CJ, Bhandari NR, Payakachat N. Assessment of reliability and validity of SF-12v2 among a diabetic population. *Value Health*. 2018;21(4):432-440. doi:10.1016/j.jval.2017.09.007
- Fleishman JA, Cohen JW. Using information on clinical conditions to predict high-cost patients. *Health Serv Res*. 2010;45(2):532-552. doi:10.1111/j.1475-6773.2009.01080.x
- Wherry LR, Burns ME, Leininger LJ. Using self-reported health measures to predict high-need cases among Medicaid-eligible adults. *Health Serv Res*. 2014;49(S2):2147-2172. doi:10.1111/1475-6773.12222
- Jean-Baptiste D, O'Malley A, Shah T. Population segmentation and targeting of health care resources: findings from a literature review. *Mathematica*. December 1, 2017. Accessed October 4, 2019. <https://www.mathematica.org/our-publications-and-findings/publications/population-segmentation-and-targeting-of-health-care-resources-findings-from-a-literature-review>
- Measuring Healthy Days: Population Assessment of Health-Related Quality of Life*. CDC; 2000. Accessed October 4, 2019. <https://www.cdc.gov/hrqol/pdfs/mhd.pdf>
- Aronson N, Alonso J, Burnam A, et al. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res*. 2002;11(3):193-205. doi:10.1023/A:1015291021312
- Patient-reported outcomes in performance measurement. National Quality Forum. December 2012. Accessed October 8, 2019. https://www.qualityforum.org/Publications/2012/12/Patient-Reported_Outcomes_in_Performance_Measurement.aspx
- CMS Measures Management System Blueprint quickstart guide. CMS. December 2022. Accessed June 6, 2023. <https://mmshub.cms.gov/sites/default/files/Quickstart.pdf>
- Austin SR, Wong Y-N, Uzzo RG, Beck JR, Eggleston BL. Why summary comorbidity measures such as the Charlson Comorbidity Index and Elixhauser Score work. *Med Care*. 2015;53(9):e65-e72. doi:10.1097/MLR.0b013e318297429c
- Chakravarty S, Cantor JC. Informing the design and evaluation of superuser care management initiatives: accounting for regression-to-the-mean. *Med Care*. 2016;54(9):860-867. doi:10.1097/MLR.0000000000000568
- Wennberg JE. *Tracking Medicine*. Oxford University Press; 2010.
- Jayadevappa R, Cook R, Chhatre S. Minimal important difference to infer changes in health-related quality of life—a systematic review. *J Clin Epidemiol*. 2017;89:188-198. doi:10.1016/j.jclinepi.2017.06.009
- Norman GR, Sloan JA, Wyrych KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41(5):582-592. doi:10.1097/01.MLR.0000062554.74615.4C
- Alghnam S, Schneider EB, Castillo RC. Insurance status and health-related quality-of-life disparities after trauma: results from a nationally representative survey in the US. *Qual Life Res*. 2016;25(4):987-995. doi:10.1007/s11136-015-1126-0
- Baragaba B, Alghnam S, Bernacki EJ. Work-related injuries and health-related quality of life among US workers: a longitudinal study of a population-based sample. *J Occup Environ Med*. 2016;58(4):385-390. doi:10.1097/JOM.0000000000000667
- Health status in Utah: the Medical Outcomes Study SF-12. Utah Department of Health. 2001. Accessed June 6, 2023. https://ibis.health.utah.gov/ibisph-view/pdf/opha/publication/uhas/2001HSS_SF12.pdf
- Yarnall KSH, Poltak KI, Østbye T, Krause KM, Michener JL. Primary care: is there enough time for prevention? *Am J Public Health*. 2003;93(4):635-641. doi:10.2105/ajph.93.4.635
- Manning WG, Newhouse JP, Ware JE. The status of health in demand estimation; or, beyond excellent, good, fair, poor. In: Fuchs VR, ed. *Economic Aspects of Health*. University of Chicago Press; 1982:141-184.
- Włodarczyk D, Zietalewicz U. How gender-specific are predictors of post-MI HROQL? a longitudinal study. *Health Qual Life Outcomes*. 2020;18(1):202. doi:10.1186/s12955-020-01439-4

Visit ajmc.com/link/89396 to download PDF and eAppendix

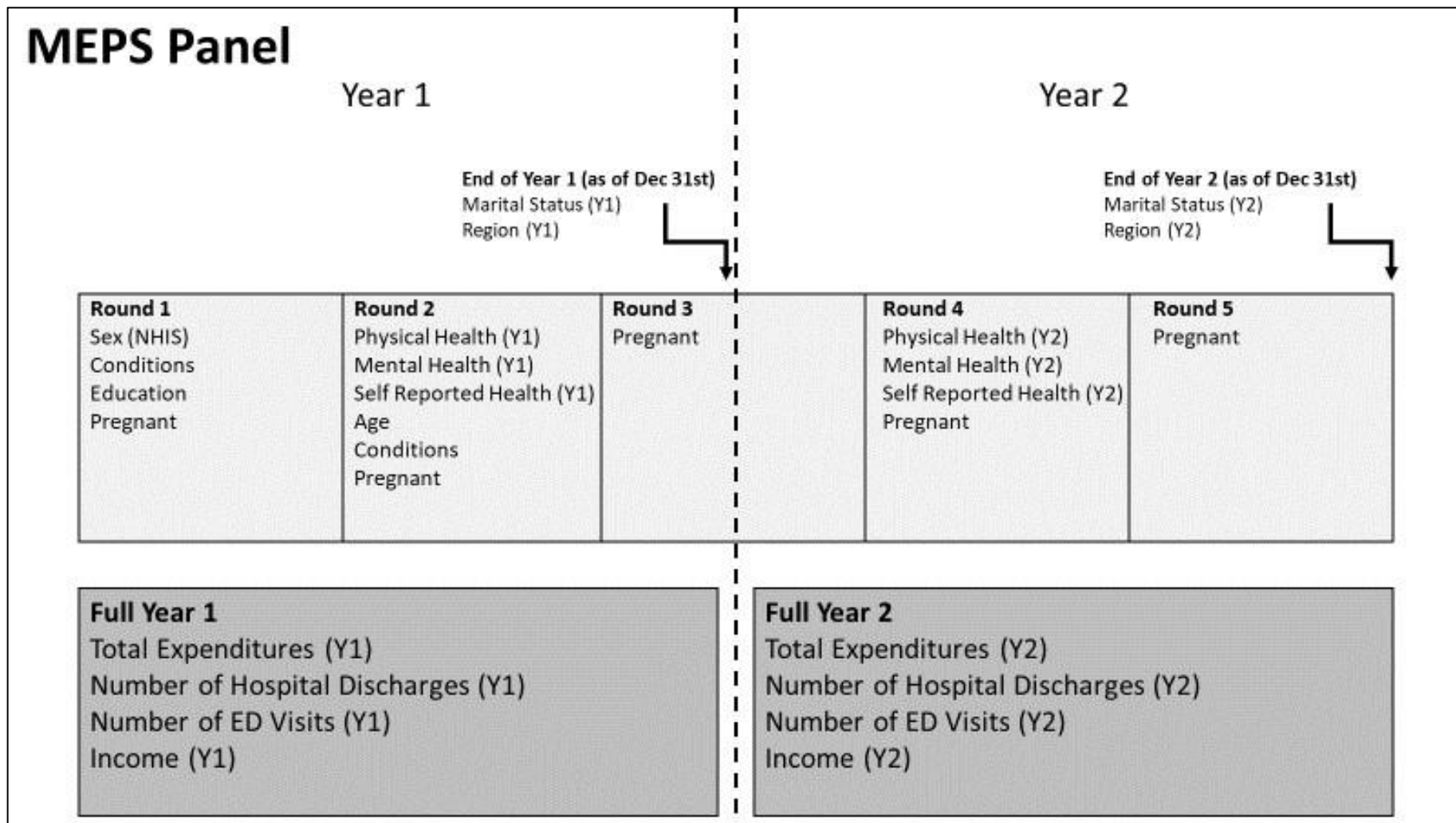
Health-Related Quality of Life in High-Cost, High-Need Populations

SUPPLEMENTAL APPENDIX

Table of Contents

eAppendix A. Chronic Conditions Measure Creation	5
eAppendix B. HCHN Cohort Identification	7
eAppendix C. Difference-in-Differences Model Specification	10
eAppendix D. Predictive Validity Model Specifications and Regression Results	11
eAppendix E. Power Calculations	15
References	17
eAppendix Figure. Measurement Timeline	2
eAppendix Table 1. Short Form 12 (SF-12v2) Items: MEPS Variable Names	3
eAppendix Table 2. Predicted Cohort	9
eAppendix Table 3. Sample Sizes for Age Strata	12
eAppendix Table 4. Association between Year 1 Mental Health and Physical Health and Year 2 Top Decile Expenditure Membership	13
eAppendix Table 5. Power Calculations	16

eAppendix Figure. Measurement Timeline



The five rounds of interviews occur over two and a half years, but data is collected for a two year period for households due to the overlapping panel design¹.

eAppendix Table 1. Short Form 12 (SF-12v2) Items: MEPS Variable Names

MEPS Variable Name	MEPS Question Label[†]	SF-36 Dimension	Dimension Description
ADGENH#	General health today	General Health (GH)	Respondents' perception of their current overall health status, how their health compares to others and expectations of future health
ADDAYA#	During a typical day, limitations in moderate activities	Physical Functioning (PF)	Respondents' reported level of difficulty in carrying out a range of physical tasks from low exertion to high exertion.
ADCLIM#	During a typical day, limitations in climbing several flights of stairs	Physical Functioning (PF)	Respondents' reported level of difficulty in carrying out a range of physical tasks from low exertion to high exertion.
ADPALS#	During past 4 weeks, as result of physical health, accomplished less than would like	Role-Physical (RP)	Limitations in performance of regular daily activities due to physical problems
ADPWLM#	During past 4 weeks, as result of physical health, limited in kind of work or other activities	Role-Physical (RP)	Limitations in performance of regular daily activities due to physical problems
ADMALS#	During past 4 weeks, as result of mental problems, accomplished less than you would like	Role-Emotional (RE)	Limitations in performance of regular daily activities due to emotional problems.
ADMWLM#	During past 4 weeks, as result of mental problems, did work or other activities less carefully than usual	Role-Emotional (RE)	Limitations in performance of regular daily activities due to emotional problems.
ADPAIN#	During past 4 weeks, pain interfered with normal work outside the home and housework	Bodily Pain (BP)	Respondents' experience of bodily pain and the extent to which pain interferes with their normal work

[†] https://meps.ahrq.gov/survey_comp/hc_survey/paper_quest/2016/2016_SAQ_ENG.pdf

ADCAPE#	During the past 4 weeks, felt calm and peaceful	Mental Health (MH)	The amount of time over the last four weeks a respondent felt happy, calm, down, depressed and nervous, respectively.
ADNRGY#	During the past 4 weeks, had a lot of energy	Vitality (VT)	Respondents' perceived levels of energy.
ADDOWN#	During the past 4 weeks, felt downhearted and depressed	Mental Health (MH)	The amount of time over the last four weeks a respondent felt happy, calm, down, depressed and nervous, respectively.
ADSOCA#	During the past 4 weeks, physical health or emotional problems interfered with social activities	Social Functioning (SF)	The extent to which physical or emotional problems impact on respondents' participation in social activities.

Dimension descriptions taken from Ware et al 1996².

SF-36 dimensions taken from Frieling et al 2013³.

eAppendix A. Chronic Conditions Measure Creation

All conditions are defined based on clinical classification codes present in rounds one and two. The MEPS conditions file records conditions mentioned in the priority conditions enumeration section, connected to a specific procedure or hospitalization reported by the respondent, or reported by the household-level respondent as a condition “bothering” the person during the reference period[‡]. These conditions are converted into ICD-9 codes by professional coders, and then aggregated into clinical classification codes (CCCODEX) that group clinically meaningful similar conditions. Because respondents are asked only for conditions that are affecting them currently, they do not catch all cases of chronic conditions.

Physical Conditions:

These definitions were taken from Fleishman and Cohen (2010)⁴ with minor modifications to asthma, emphysema, and cancer. We generally tried to be as inclusive as possible in condition definitions.

Condition	Clinical Classification Codes	Notes
Diabetes	49, 50	
High blood pressure	98, 99	Category 99 is “hypertension with complications and secondary hypertension”. While included in our definition, it never appears in the first two rounds of the conditions files.
Asthma [§]	128	
Arthritis	201, 202, 203	
Joint Disease	204	This was created to compare with the arthritis measure, but not included in our regression or the count of chronic conditions.
High Cholesterol	53	This clinical classification code is “disorders of lipid metabolism”, and includes ICD-9s 272.0-272.4.
Emphysema ^{**}	127	

[‡] https://www.meps.ahrq.gov/data_stats/download_data/pufs/h180/h180doc.pdf

[§] Fleishman and Cohen added three ICD-9 codes to this definition (491,492,496). Our definition only uses clinical classification aggregates and does not have these ICD-9s. The ICD-9s are included in our emphysema definition.

** Fleishman and Cohen do not give a definition of emphysema. Code 127 is “Chronic obstructive pulmonary disease and bronchiectasis”. Part of its ICD-9 codes are the additional codes (491,492,496) that Fleishman and Cohen put in their asthma measure.

Cancer ^{††}	11-45	Code 20, “Cancer; other respiratory and intrathoracic”, is included in our definition, but doesn’t have any observations in the time period we observe. Many other codes are only seen in panel 16 for our time frame.
Heart Disease	96, 97, 100, 103-108	
Cerebrovascular Disease	109-113	

Behavioral Health Conditions:

MEPS defines behavioral health conditions as those with a clinical classification code of 650 or higher, but we chose a subset of these conditions to include in our model. Substance use disorder is a combination of alcohol-related disorders and substance-related disorders.

Condition	Clinical Classification Codes
Mood Disorder	657
Anxiety	651
Schizophrenia	659
Substance Use Disorder	660, 661

^{††} Includes skin cancers (22,23), which Fleishman and Cohen excluded.

eAppendix B. HCHN Cohort Identification

We adapted the model specification published in Wherry, Burns, and Leininger (2014)⁵ to predict top decile expenditure membership in year 2 of the MEPS panel using year 1 covariates. Specifically, we estimated panel-specific logistic regressions to estimate Equation (1), with p indexing individuals:

$$(1) \text{TopDecileExpenditure_Year2}_i = \alpha + \mathbf{ResourceUtilization_Year1}_p \phi + \mathbf{Conditions_Year1}_p \gamma + \mathbf{Controls_Year1}_p \eta + \varepsilon_i$$

TopDecileExp_Year2 is a dummy variable reflecting whether or not the sample member was in the top expenditure decile in year 2. The vector ***ResourceUtilization_Year1_p*** represents the lagged dependent variable, specifically top decile expenditure membership in year 1, along with any inpatient hospitalization in year 1 and top decile ED membership in year 1. Dummies for each of the 13 chronic conditions documented in the section above as well as a dummy indicating the presence of behavioral and physical health comorbidity are captured by the vector ***Conditions_Year1_p***. Additional controls (vector ***Controls_Year1_p***) include sex (female = 1, male = 0); education (dummies for less than high school and high school (HS) graduate/some college; reference = college +); income (dummies for < 100% FPL, 100 to less than 125% FPL; 125% to less than 200% FPL; 200% FPL to less than 400% FPL; reference = 400% + FPL); and region (dummies for South, Midwest, and West; reference = Northeast). After estimating Equation (1) we compute \hat{p} – the predicted probability of a sample member belonging to the top decile of year 2 expenditures in her panel. We then selected the top decile in the \hat{p} distribution as our HCHN cohort. We perform this calculation separately for each of the 5 panels used in the analytic sample and append the observations. Table A2 displays the regression results for each panel. These coefficient magnitudes (displayed as odds ratios) are similar to the related literature, as is the overall predictive capacity of the model as assessed by the c-statistic (Wherry, Burns, and Leininger 2014⁵).

eAppendix Table 2. Predicted Cohort

Variable	Panel 16 (N = 10,596)		Panel 17 (N = 10,066)		Panel 18 (N = 8,934)		Panel 19 (N=8,329)		Panel 20 (N = 9,009)	
	OR	CI	OR	CI	OR	CI	OR	CI	OR	CI
Top decile total expenditure year 1	6.23	5.13, 7.57	6.11	4.97, 7.51	7.81	6.33, 9.63	6.71	5.38, 8.38	7.84	6.35, 9.69
Any hospitalization year 1	0.85	0.66, 1.09	0.81	0.63, 1.06	0.78	0.60, 1.03	0.68	0.51, 0.90	0.74	0.57, 0.97
Top decile ED visits year 1	1.86	1.37, 2.52	1.70	1.24, 2.33	1.58	1.17, 2.14	1.35	0.98, 1.86	1.66	1.23, 2.24
Female	1.21	1.04, 1.41	0.98	0.84, 1.15	1.03	0.87, 1.22	0.92	0.77, 1.09	0.96	0.81, 1.14
Age	1.03	1.02, 1.03	1.03	1.03, 1.04	1.02	1.02, 1.03	1.03	1.02, 1.03	1.03	1.02, 1.04
Married	0.95	0.81, 1.11	0.83	0.71, 0.97	0.92	0.78, 1.10	0.92	0.77, 1.10	0.87	0.74, 1.04
Education										
Less than high school	0.91	0.72, 1.16	1.08	0.87, 1.34	0.80	0.63, 1.03	0.91	0.70, 1.18	0.89	0.69, 1.15
High school	0.98	0.82, 1.16	0.89	0.74, 1.08	0.99	0.82, 1.21	1.07	0.88, 1.31	1.02	0.84, 1.23
Income Category Year 1										
Negative or Poor	0.73	0.57, 0.93	0.89	0.69, 1.14	1.03	0.79, 1.35	0.98	0.74, 1.29	0.97	0.74, 1.26
Near Poor	0.70	0.50, 1.00	0.94	0.68, 1.31	1.27	0.91, 1.80	0.92	0.63, 1.33	0.84	0.57, 1.23
Low Income	0.70	0.55, 0.89	0.83	0.65, 1.06	0.82	0.63, 1.07	1.09	0.83, 1.42	0.78	0.60, 1.02
Middle Income	0.74	0.60, 0.89	0.82	0.67, 1.01	0.86	0.69, 1.07	0.84	0.67, 1.06	0.93	0.75, 1.15
Geographic Region Year 1										
South	0.81	0.65, 1.00	0.81	0.66, 1.00	0.78	0.62, 0.98	0.65	0.51, 0.82	1.16	0.92, 1.47
Midwest	0.97	0.76, 1.22	0.96	0.75, 1.22	0.97	0.75, 1.24	0.73	0.56, 0.95	1.12	0.86, 1.46
West	0.91	0.72, 1.15	0.67	0.53, 0.85	0.68	0.53, 0.87	0.60	0.46, 0.77	0.97	0.75, 1.27
Chronic Conditions Year 1										
Anxiety	1.42	1.08, 1.87	1.27	0.93, 1.73	1.69	1.27, 2.25	1.74	1.30, 2.32	1.58	1.19, 2.10
Arthritis	1.90	1.55, 2.34	1.52	1.20, 1.92	1.45	1.15, 1.84	1.76	1.39, 2.22	1.73	1.38, 2.18
Asthma	1.46	1.11, 1.92	1.46	1.09, 1.97	1.58	1.19, 2.10	1.45	1.06, 1.99	1.39	1.03, 1.86
Cancer	1.11	0.8, 1.43	1.43	1.09, 1.89	1.31	0.98, 1.75	1.39	1.04, 1.86	1.52	1.16, 1.99
Cerebrovascular disease	1.33	0.88, 2.01	2.06	1.28, 3.32	1.14	0.69, 1.86	1.53	0.95, 2.48	0.98	0.62, 1.56
Diabetes	1.91	1.58, 2.32	2.13	1.75, 2.59	1.80	1.46, 2.22	1.73	1.40, 2.15	2.46	2.02, 3.00
Emphysema	1.50	1.13, 2.00	1.99	1.45, 2.72	1.19	0.86, 1.63	1.43	1.02, 2.00	1.32	0.96, 1.82
Heart disease	1.61	1.29, 2.02	1.51	1.18, 1.93	1.41	1.09, 1.82	2.05	1.59, 2.64	1.38	1.08, 1.76
High blood pressure	1.28	1.07, 1.54	1.30	1.09, 1.57	1.48	1.21, 1.80	1.21	0.98, 1.49	1.23	1.02, 1.49
High cholesterol	1.01	0.84, 1.21	1.01	0.84, 1.22	1.17	0.95, 1.43	1.23	1.00, 1.51	0.96	0.79, 1.17
Mood disorder	2.12	1.61, 2.77	2.49	1.85, 3.35	1.49	1.10, 2.01	2.24	1.68, 3.00	1.67	1.25, 2.24
Schizophrenia	2.87	1.20, 6.83	3.29	1.19, 9.13	3.73	1.72, 8.07	1.68	0.66, 4.25	2.36	0.94, 5.94
Substance Use Disorder	7.73	3.15, 18.99	1.22	0.39, 3.81	2.82	1.19, 6.70	3.22	1.40, 7.43	1.68	0.71, 4.01
Has behavioral and physical condition Year 1	0.93	0.68, 1.27	0.88	0.62, 1.23	0.95	0.68, 1.35	0.73	0.52, 1.02	0.80	0.58, 1.12

C- statistic	0.83			0.84		0.85	0.85		0.85	
--------------	------	--	--	------	--	------	------	--	------	--

eAppendix C. Difference-in-Differences Model Specification

To assess differences over time in self-reported health and utilization measures, we first create a person-year level analytic file with each sample member contributing two observations, corresponding with the MEPS panel duration of two years. We run the model specification detailed in equation (2) for the following eight outcomes: mental health, physical health, total expenditures, top decile expenditure membership, top decile ED membership, any hospitalization, fair or poor self-reported general health status, and fair or poor self-reported mental health status.

$$(2) Outcome_{!#} = \alpha + \beta_1 HCHN_{!} + \beta_2 POST_{\#} + \beta_3 (HCHN * POST)_{!#} + \varepsilon_{!#}$$

Note that p indexes people and y indexes year. In this specification, β_1 represents the difference in $Outcome_{py}$ across HCHN and non-HCHN populations (combining years), while β_2 represents the difference in $Outcome_{py}$ in year 2 of sample participation compared to year 1 (combining HCHN and non-HCHN cohorts). The difference in year-over-year trends across HCHN and non-HCHN cohorts is represented by β_3 -- often called the unadjusted “difference-in-differences” (DiD) result -- ascertaining the extent to which year-over-year regression-to-the mean differs across HCHNs and non-HCHNs. The associated results are displayed in Table 2 of the manuscript.

eAppendix D. Predictive Validity Model Specifications and Regression Results

Our predictive validity model specification is estimated on the person-level analytic sample and is restricted to the HCHN cohort. Equation (3) details the model specification, again with p indexing individuals:

$$(3) \textit{Utilization_Year2}_p = \alpha + \beta_1 \textit{MCS_Year1}_p + \beta_2 \textit{PCS_Year1}_p + \textit{Utilization_Year1}_p \delta + \textit{Conditions_Year1}_p \gamma + \textit{Controls_Year1}_p \gamma + \varepsilon$$

We fit a logistic regression for each of the following outcome measures, collectively represented by *Utilization_Year2*: top decile expenditure membership in year 2; top decile ED membership in year 2; and any hospital discharges in year 2. The vector ***Utilization_Year1_p*** represents lagged year 1 values of each of the three outcome measures; dummies for each of the 13 chronic conditions as well as dummy indicating the presence behavioral and physical health comorbidity are captured by the vector ***Conditions_Year1_p***. Additional controls (vector ***Controls_Year1_p***) include sex (female = 1, male = 0); education (dummies for less than high school and high school (HS) graduate/some college; reference = college +); income (dummies for < 100% FPL, 100 to less than 125% FPL; 125% to less than 200% FPL; 200% FPL to less than 400% FPL; reference = 400% + FPL); and region (dummies for South, Midwest, and West; reference = Northeast). The associated results (displayed as odds ratios) for the top decile expenditure membership are provided below in Table A3; the coefficients for the other two outcome measures exhibit similar patterns and are available upon request. As detailed in the manuscript, we facilitate the interpretation of the regression results by calculating predicted probabilities of top decile membership in year 2 at clinically meaningful levels of physical and mental health (Table 3).

Figure 1 in the manuscript displays physical and mental health scores stratified by age and HCHN status. The associated strata-specific sample sizes are displayed below in eAppendix Table 3.

eAppendix Table 3. Sample Sizes for Age Strata

	Not HCHN	HCHN 1	HCHN 2	HCHN 3	HCHN 4	HCHN 5
Ages 18-44	20256	393	896	741	654	655
Ages 45-54	7148	775	889	365	530	1201
Ages 55-64	5094	1281	1205	306	626	1960
Ages 65+	3884	2241	1700	390	1024	3081

eAppendix Table 4. Association between Year 1 Mental Health and Physical Health and Year 2 Top Decile Expenditure Membership
(n = 4,690 for all models)

Variable	(1)		(2)		(3)		(4)		(5)	
	OR	CI	OR	CI	OR	CI	OR	CI	OR	CI
Physical Health Composite Score Year 1	0.96*** 0.96, 0.97				0.97***	0.96, 0.98			0.97***	0.96, 0.98
Mental Health Composite Score Year 1			0.98***	0.98, 0.99			0.99**	0.98, 1.00	0.99**	0.98, 1.00
Top decile total expenditure year 1					4.27***	3.36, 5.43	4.71***	3.72, 5.98	4.25***	3.34, 5.40
Any hospitalization year 1					0.68***	0.57, 0.81	0.71***	0.60, 0.85	0.68***	0.57, 0.80
Top decile ED visits year 1					1.24	0.99, 1.54	1.29*	1.04, 1.60	1.22	0.98, 1.52
Female					0.87	0.72, 1.04	0.89	0.75, 1.07	0.87	0.73, 1.05
Age					1.01***	1.01, 1.02	1.02***	1.01, 1.02	1.01***	1.01, 1.02
Married					0.92	0.78, 1.09	0.96	0.81, 1.13	0.93	0.79, 1.11
Education										
Less than high school					0.90	0.70, 1.15	0.95	0.74, 1.24	0.87	0.68, 1.12
High school					0.93	0.78, 1.11	0.96	0.81, 1.14	0.92	0.77, 1.09
Income Category Year 1										
Negative or Poor					1.12	0.88, 1.43	1.28*	1.01, 1.62	1.08	0.84, 1.38
Near Poor					1.05	0.76, 1.44	1.20	0.88, 1.65	1.01	0.73, 1.39
Low Income					1.12	0.88, 1.44	1.26	1.00, 1.58	1.09	0.85, 1.40
Middle Income					0.93	0.76, 1.15	1.00	0.81, 1.23	0.92	0.75, 1.14
Geographic Region Year 1										
South					0.72*	0.57, 0.93	0.76*	0.60, 0.98	0.72**	0.56, 0.92
Midwest					0.68**	0.53, 0.88	0.72**	0.56, 0.92	0.68**	0.53, 0.87
West					0.70**	0.55, 0.90	0.73*	0.57, 0.93	0.69**	0.54, 0.89
Chronic Conditions Year 1										
Diabetes					1.14	0.90, 1.44	1.08	0.85, 1.37	1.09	0.86, 1.39
High blood pressure					1.23*	1.04, 1.47	1.48***	1.26, 1.75	1.23*	1.03, 1.46
Asthma					1.34*	1.05, 1.70	1.38**	1.09, 1.74	1.33*	1.04, 1.69
Arthritis					1.29*	1.05, 1.59	1.30*	1.06, 1.60	1.28*	1.04, 1.58
Mood disorder					1.15	0.85, 1.56	1.21	0.89, 1.64	1.12	0.82, 1.51
Anxiety					1.54***	1.28, 1.84	1.67***	1.40, 1.99	1.53***	1.28, 1.83
Schizophrenia					1.20	0.95, 1.52	1.33*	1.05, 1.67	1.19	0.94, 1.51
High cholesterol					1.23*	1.03, 1.48	1.31**	1.10, 1.57	1.23*	1.03, 1.47
Emphysema					1.15	0.96, 1.37	1.19*	1.00, 1.42	1.16	0.97, 1.38
Substance Use Disorder					0.84*	0.71, 1.00	0.87	0.73, 1.03	0.84	0.71, 1.00

Cancer		1.59*** 1.26, 2.02	1.50** 1.17, 1.92	1.48** 1.16, 1.89
Heart disease		3.49*** 1.95, 6.25	3.08*** 1.72, 5.52	3.45*** 1.94, 6.12
Cerebrovascular disease		1.95* 1.09, 3.47	1.85* 1.05, 3.23	1.92* 1.09, 3.38
Has behavioral and physical condition Year 1		1.10 0.86, 1.41	1.13 0.87, 1.45	1.10 0.85, 1.42

eAppendix E. Power Calculations

We compute a series of simple power calculations to provide exploratory sample size benchmarks for measure implementation. Specifically, we compute the required sample size for a two-tailed difference in means, setting power – the probability of correctly concluding that the means differ – at 0.80 and α – the probability of incorrectly concluding that the means differ when they do not – at 0.05. We use the *power* command in Stata for the computations, which are driven by the hypothesized effect size (difference in means) as well as the standard deviation. Of note is that the resulting estimates of required sample sizes for mental health and physical health (Table A4) are the same, as they exhibit very similar standard deviations among our HCHN cohort (11.13 and 11.98, respectively) and we choose the same effect size inputs for both measures (0.3 SD, 0.5 SD, and 1 SD). We offer a caveat that our calculations do not speak to sample size requirements for more complicated research designs, as such we offer them as potentially helpful implementation benchmarks as opposed to rigorous estimates underlying scientific trials.

eAppendix Table 5. Power Calculations

	Physical Health Year 1 (SD = 11.98)				Mental Health Year 1 (SD = 11.13)				
	Value	N		N per Group	Value	N	N per Group		
Mean + One Third SD	41.93	286	143	52.30	286	143	Mean + Half SD	43.92	128
64	54.15	128	64						
Mean + One SD		49.91		34	17	59.72		34	17

References

1. Agency for Healthcare Research and Quality. MEPS-HC Panel Design and Data Collection Process. https://meps.ahrq.gov/survey_comp/hc_data_collection.jsp. Accessed April 21, 2020.
2. Ware JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. *Med Care*. 1996;34(3):220-233. doi:10.1097/00005650-199603000-00003
3. Frieling MA, Davis WR, Chiang G. The SF-36v2 and SF-12v2 health surveys in New Zealand: norms, scoring coefficients and cross-country comparisons. *Aust N Z J Public Health*. 2013;37(1):24-31. doi:10.1111/1753-6405.12006
4. Fleishman JA, Cohen JW. Using Information on Clinical Conditions to Predict High-Cost Patients. *Health Serv Res*. 2010;45(2):532-552. doi:10.1111/j.1475-6773.2009.01080.x
5. Wherry LR, Burns ME, Leininger LJ. Using Self-Reported Health Measures to Predict High-Need Cases among Medicaid-Eligible Adults. *Health Serv Res*. 2014;49(S2):21472172. doi:10.1111/1475-6773.12222