



Evaluating clinical and demographic influences on health perception: A translation of the SF-12 for use with NHANES

Elise C. Covert^{a,1}, Anna M. Baker^b, Owais Gilani^{a,*}

^a Department of Mathematics, Bucknell University, 1 Dent Drive, Lewisburg, PA, 17837, USA

^b Department of Psychology, Bucknell University, 1 Dent Drive, Lewisburg, PA, 17837, USA

ARTICLE INFO

Keywords:

Self-perceived health
Self-rated health
Measure
Short Form 12
National Health and Nutrition Examination Survey

ABSTRACT

Improving public health depends on an intricate understanding of the factors that influence how individuals perceive and self-report their personal health. Self-perceived health is an independent predictor of future health-related outcomes, but capturing self-perception of health is complex due to the intricate relationship between clinical and perceived health. A commonly used measure of self-perceived health is the Short Form 12 (SF-12), developed in the 1990s. In this study, we aim to evaluate clinical and demographic influences on self-perceived health among American adults using the National Health and Nutrition Examination Survey (NHANES). While NHANES captures information on a number of domains of health, including clinical assessments, it does not include SF-12 items necessary to measure self-perceived health. Therefore, to assess self-perceived health for our study, we constructed and validated a novel SF-12-equivalent measure for use with NHANES using analogous items from the 2015–2016 NHANES interview questionnaires. The developed measure reflects established knowledge of population health patterns and closely parallels the behavior of the original SF-12. An analysis of the clinical and demographic influences on this novel measure of health perception revealed that both clinical and demographic factors, such as depression status and race, influence how healthy individuals perceive themselves to be. Importantly, our analysis indicated that among American adults, while controlling for clinical and demographic covariates, an increase in low-density lipoprotein (i.e., “bad”) cholesterol level was associated with an improvement in self-perceived health. This study contributes significantly in two domains: it provides a novel measure of self-perceived health compatible for use with the widely used NHANES data (as well as details on how the process was developed), and it identifies a critical area in need of improved clinical education regarding the apparent confusion around cholesterol health.

1. Introduction

Individuals are increasingly taking an active role in their personal health: self-care, self-advocacy, disease prevention, and seeking health services. Improving patient care and health education depends on an intricate understanding of factors influencing how individuals perceive and self-report their personal health. Self-reports reflect a patient’s interpretation of various dimensions of health that are often invisible to an external clinical evaluator (Miilunpalo et al., 1997). Perceived health is an independent predictor of future health-related outcomes, including chronic disease incidence and mortality (Nayak et al., 2018; Shields & Shooshtari, 2001). Known determinants of self-rated health include

education level (Shields & Shooshtari, 2001), sex (Shields & Shooshtari, 2001), age (Loprinzi, 2015), race-ethnicity (Loprinzi, 2015), and socioeconomic status (SES) (Loprinzi, 2015).

The relationship between perceived and clinically measured health is complex. Perceived health is influenced by knowledge of objective health, but the clinical profile is also influenced by self-report of pain levels, symptom severity, and treatment adherence. Due to its abstract and subjective nature, however, measuring perceived health is a challenge. While the World Health Organization defines health as a state of complete physical, mental and social well-being, and not merely the absence of disease and infirmity (Karimi & Brazier, 2016), there is no guarantee that personal perceived health aligns with this professional

Abbreviations: SF-12, Short Form 12; NHANES, National Health and Nutrition Examination Survey.

* Corresponding author. Bucknell University, 1 Dent Drive, Lewisburg, PA, 17837, USA.

E-mail addresses: eccovert@umich.edu (E.C. Covert), amb060@bucknell.edu (A.M. Baker), owais.gilani@bucknell.edu (O. Gilani).

¹ Department of Biostatistics, School of Public Health, University of Michigan, 1215 Washington Heights, Ann Arbor, MI, 48109 (E. Covert present address).

<https://doi.org/10.1016/j.ssmph.2022.101081>

Received 5 October 2021; Received in revised form 2 January 2022; Accepted 22 March 2022

Available online 25 March 2022

2352-8273/© 2022 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

definition. The transition to patient-centered care in the 1980s called medical professionals to consider the patient experience in addition to traditional, objective external measures (Hunt et al., 1985). In response, medical researchers developed standard quantitative measures of perceived health, including the Short Form 12 (SF-12) (Ware et al., 1995). The SF-12 comprises 12 items targeting a variety of health constructs and strikes an effective balance between length, reliability, and validity (McDowell, 2006).

For health educators and practitioners, it is vitally important to understand how accurately self-reports of subjective health reflect objective clinical health. Inaccurate self-reports can misrepresent a patient's symptoms or treatment adherence, which in turn affects diagnoses and treatment plans and can lead to poor health outcomes in the future. Interpreting this interplay, however, requires the comparison of self-report data to clinical assessment data from the same individuals. National health surveys provide rich data on large sample sizes that can yield clinical insights into health and health behaviors of a population. To limit response burden, however, national health surveys often do not collect extensive perceived health measures. Instead, they typically administer a single-item global health rating. One such survey in the U. S. is the National Health and Nutrition Examination Survey (NHANES). Meanwhile, research has demonstrated that multi-item scales like the SF-12 are more valid than single-item measures like the global rating administered by NHANES (Ware & Sherbourne, 1992). There is currently no multi-item perceived health scale that is compatible with NHANES, despite the survey being utilized extensively in epidemiological and health sciences research and to inform national health standards.

Although NHANES does not include the SF-12, it does collect data on variables targeting similar constructs. Building an analogous measure from these variables would give NHANES researchers a tool to determine how perceived and clinical health are related, and the impact of demographic factors on this relationship. Understanding how perception of health is related to clinical health can inform clinical strategies to improve health outcomes. Thus, this study has two objectives: (i) to construct and validate a SF-12 equivalent perceived health measure using NHANES interview data; and (ii) to use this novel perceived health measure to identify patterns in how demographic and clinical health factors influence perceived health among American adults.

2. Materials and methods

2.1. Data

NHANES is a nationally-representative survey that produces vital statistics and monitors health trends. Importantly, the survey collects both self-reported and clinically assessed measures of health. NHANES employs a complex, multistage probability design including oversampling of minority subgroups of the U.S. population. All tables and figures presented account for NHANES sampling weights. Interview data were processed to account for questionnaire skip patterns that determine which respondents are eligible to answer various items. Failing to do so would inaccurately result in higher frequency of missing data. Additionally, "Refused" and "Don't know" responses were recoded to missing values and included in the non-response rate. Our sample cohort excluded respondents under age 20.

2.2. Measure construction & validation

Our first objective was to develop a SF-12 equivalent measure using NHANES interview responses. After data cleaning, we identified a NHANES analog for each of the original SF-12 items from among approximately 800 interview items (Centers for Disease Control and Prevention [CDC], 2018). This process was completed in three phases. First, we identified any items that were *exact or nearly exact matches* between the SF-12 and the NHANES questionnaire. Some items did have

exact matches, as with the "General health" question, which is worded identically in both sources. Second, for items without near-exact matches in the NHANES questionnaire, we chose analogs targeting *similar constructs*. For example, a SF-12 item asking how often the respondent felt calm and peaceful in recent weeks was matched to a NHANES item asking how often the respondent felt worried, nervous, or anxious.

Third, for items that did not have an exact or matched-construct analog in NHANES, we *constructed a new composite item*. Only one SF-12 item (regarding "Pain interference") fell in this category. We identified an interview item asking respondents to report conditions or health problems that caused them difficulty with common activities. Four response categories for this item indicated pain conditions: arthritis or rheumatism, back or neck problem, fractures or bone/joint injury, and other injury. If a respondent reported a condition in one or more of these categories, they were indicated to have pain interfering with their work. Otherwise, we assumed that a pain condition was not present.

Table 1 provides full wording of the original SF-12 questions and their NHANES analogs, plus an item shorthand to more concisely reference specific items. We mapped the response options for each NHANES analog to the response options for the corresponding SF-12 item (Fig. 1) so that the established SF-12 scoring procedure (Ware et al., 1995) could be implemented to construct our novel perceived health measure.

Per SF-12 guidelines, an individual's responses inform two summary scores – the physical component score (PCS) and the mental component score (MCS) – using a linear combination of regression weights (based on the selected response to the survey item) and constants (Ware et al., 1995). To implement this scoring method on NHANES data, we constructed a set of indicator variables and corresponding regression weights. For each item, the response category indicating the highest state of health served as the reference level. A set of 26 indicator variables captured all possible response choices to the NHANES analogs.

Next, we applied regression weights to the indicator variables. The SF-12 scoring manual provides physical and mental component regression weights derived from general U.S. population distributions for each scale item. For simplicity, we refer to these regression weights as "coefficients." In the construction of the original SF-12, Ware et al. (1995) calculated coefficients such that both component scores would have mean 50 and standard deviation 10 in the general U.S. population. Higher scores indicate better perceived health.

With slightly different response options, the SF-12 has a set of 35 indicator variables and 35 corresponding coefficients. From the original coefficient set, we produced a modified set of 26 coefficients – one to weight each of our NHANES indicators – by combining certain factor levels. Two methods of doing so produced two measure variants. In the first measure (unweighted measure or Measure UW), for each of the 12 items, we assumed that responses were evenly distributed across all possible levels and combined the relevant coefficients using a simple algebraic mean. In the second measure (weighted measure or Measure W), we took a weighted average of the coefficients based on the estimated prevalence of factor levels in the general U.S. population (Office of Public Health Assessment, 2004). Methodological details and coefficients are provided in the supplement (Table A1).

Of 5,719 respondents in the sample, scores were constructed for 3,367 (58.87%) who provided valid responses for all 12 NHANES analogs. For those with missing data, we implemented an imputation procedure recommended for use with SF-12 (Perneger & Burnand, 2005). However, the distribution of MCS scores for respondents with complete data starkly differed from the distribution for respondents whose data had been imputed. This difference hinted that imputation was ineffective and may bias the final results. Further details are provided in the supplement.

We validated our measure using three approaches. First, we plotted score distributions to rule out floor and ceiling effects (Brazier et al., 1992). Next, we compared our score distributions to those from a

Table 1Short Form 12 (SF-12) item texts and corresponding National Health and Nutrition Examination Survey (NHANES) 2015–2016 analog item texts.^a

Item Shorthand	SF-12 Item	NHANES Analog
General health	In general, would you say your health is excellent, very good, good, fair, or poor? (GH1)	Would you say your health in general is excellent, very good, good, fair, or poor? (HSD010)
Moderate activity	Does your health now limit you in moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf? (PF02)	By yourself and without using any special equipment, how much difficulty do you have doing chores around the house? (PFQ061F)
Mobility	Does your health now limit you in climbing several flights of stairs? (PF04)	Do you have serious difficulty walking or climbing stairs? (DLQ050)
Physical limitation	During the past 4 weeks, have you accomplished less than you would like (work or other regular daily activities) as a result of your physical health? (RP2)	Are you limited in any way in any activity because of a physical, mental, or emotional problem? (PFQ059)
Work limitation	During the past 4 weeks, were you limited in the kind of work or other activities as a result of your physical health? (RP3)	Are you limited in the kind or amount of work you can do because of a physical, mental, or emotional problem? (PFQ051)
Emotional limitation	During the past 4 weeks, have you accomplished less than you would like (work or other regular daily activities) as a result of any emotional problems (such as feeling depressed or anxious)? (RE2)	How difficult have these {emotional problems} made it for you to do your work, take care of things at home, or get along with people? (DPQ100)
Concentration & care	During the past 4 weeks, have you not done work or other activities as carefully as usual as a result of any emotional problems (such as feeling depressed or anxious)? (RE3)	Because of a physical, mental, or emotional condition, do you have serious difficulty concentrating, remembering, or making decisions? (DLQ040)
Pain interference	During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)? (BP2)	What condition or health problem causes you to have difficulty with {these activities}? (PFQ063A-E)
Anxiety	How much of the time during the past 4 weeks have you felt calm and peaceful? (MH3)	How often do you feel worried, nervous or anxious? (DLQ100)
Energy level	How much of the time during the past 4 weeks did you have a lot of energy? (VT2)	[Over the last 2 weeks how often have you been bothered by the following problems:] feeling tired or having little energy? (DPQ040)
Depression	How much of the time during the past 4 weeks have you felt downhearted and blue? (MH4)	[Over the last 2 weeks how often have you been bothered by the following problems:] feeling down, depressed, or hopeless? (DPQ020)
Social activity	During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)? (SF2)	By yourself and without using any special equipment, how much difficulty do you have participating in social activities? (PFQ061R)

^a Variable names from the SF-12 survey and the NHANES data release are included in parentheses after their respective item texts.

previous study that used the SF-12 to measure perceived health in the adult U.S. population (Ware et al., 1995). Similar distributions verify that our scores reasonably replicate the SF-12; we anticipated only slight variation, partially due to sample-to-sample differences, and partially due to the fact that we have used different items on our scale. Finally, we established construct validity by correlating our measure scores with known demographic determinants of health (Brazier et al., 1992), checking that the direction and strength of the correlations match those documented in the literature. We predicted a negative correlation between perceived health and age (Loprinzi, 2015), a negative association between perceived health and the recent use of medical services (Hunt & McEwen, 1980), and a positive correlation between perceived health and income-to-poverty ratio (Brazier et al., 1992; Shields & Shooshtari, 2001).

2.3. Regression analysis

After validation, we used the novel measure to analyze the influence of demographic and clinical covariates on health perception. The outcome of interest was total perceived health score (the sum of PCS and MCS). From the examination and laboratory components of NHANES, we selected eight clinical covariates: body mass index (BMI), systolic blood pressure, low-density lipoprotein (LDL) cholesterol, glycohemoglobin (%), creatinine concentration, cotinine concentration, Patient Health Questionnaire (PHQ-9) depression severity measure, and the Healthy Eating Index (HEI). We do not claim that these eight measurements provide a comprehensive picture of any one person's health status, but they do represent a range of vital components comprising the clinical health profile. For example, blood pressure, creatinine, and cotinine cover cardiovascular, renal, and pulmonary health, respectively. Further, BMI and HEI give us a sense of some important health behaviors (exercise and diet). Because all of the clinical covariates were quantitative, we standardized each to a mean of 0 and a standard deviation (SD) of 1. We also considered six demographic covariates: age, gender, race-ethnicity, income-to-poverty ratio, education level, and U.S. citizenship status.

We fit multiple linear regression (MLR) models to data on 1,239

NHANES respondents with constructed perceived health scores and complete clinical and demographic data. Analyses were run in R statistical software using the *survey* package. The eight clinical covariates remained in the model regardless of statistical significance because we aimed to examine demographic associations while controlling for a basket of clinical conditions.

To determine demographic covariates included in the final model, we employed a backwards elimination approach, using the Akaike Information Criterion to compare model fit. We fit MLR models to both measure variants (UW and W), and compared the estimated coefficients using 95% confidence intervals (CI).

3. Results

Table 2 presents selected demographic characteristics of the sample. The majority of respondents were non-Hispanic white, over half had at least some college education, and 90% were U.S. citizens. Mean age was about 48 years. Fig. 2 illustrates the distribution of responses to selected NHANES analogs. Measure UW scores were left-skewed with descriptive statistics close to the SF-12 targets; PCS had mean 49.09 with SD 10.04, and MCS had mean 51.29 with SD 9.59. Measure W results were nearly identical; PCS had mean 49.77 with SD 9.24, and MCS had mean 51.60 with SD 9.47.

3.1. Validation

All three validation checks confirmed that both measure variants are valid measures of perceived health. Despite the left-skewed score distributions, we did not observe a case where the majority of respondents scored either the minimum or maximum value, ruling out the presence of floor or ceiling effects.

Fig. 3 presents side-by-side box plots for the original SF-12, Measure UW, and Measure W scores for the general U.S. population. Additional box plots by age category are presented in the supplement (Figure B2). The plots confirm that our novel measures reasonably replicate the original SF-12 in the general population, as well as within varying age ranges. In a final validation check, we correlated total perceived health

Item Shorthand	NHANES Response Options	SF-12 Response Options
General health	Excellent Very good Good Fair Poor	Excellent Very good Good Fair Poor
Moderate activity	Do not do this activity No difficulty Some difficulty Much difficulty Unable to do	Not limited Limited a little Limited a lot
Mobility	Yes No	Limited a lot Limited a little Not limited
Physical limitation	Yes No	Yes No
Work limitation	Yes No	Yes No
Emotional limitation	Not at all difficult Somewhat difficult Very difficult Extremely difficult	Yes No
Concentration & care	Yes No	Yes No
Pain interference	Pain condition present No pain condition present	Extremely Quite a bit Moderately A little bit Not at all
Anxiety*	Daily Weekly Monthly A few times a year Never	None of the time A little of the time Some of the time A good bit of the time Most of the time All the time
Energy level*	Not at all Several days More than half the days Nearly every day	All the time Most of the time A good bit of the time Some of the time A little of the time None of the time
Depression	Not at all Several days More than half the days Nearly every day	All the time Most of the time A good bit of the time Some of the time A little of the time None of the time
Social activity	Do not do this activity No difficulty Some difficulty Much difficulty Unable to do	None of the time A little of the time Some of the time Most of the time All the time

* Note that NHANES and SF-12 items ask opposite questions. Responses are mapped accordingly.

Fig. 1. National Health and Nutrition Examination Survey (NHANES) 2015–2016 analog response options mapped to original Short Form 12 (SF-12) item response options.

with age, income-to-poverty ratio, and use of medical services over the past year. Each pair had a statistically significant correlation of moderate strength, with $0.19 \leq |r| \leq 0.28$. Measure UW presented negative correlations with age ($r = -0.21$) and health care use ($r = -0.27$) and a positive correlation with income-to-poverty ratio ($r = 0.28$). Similar results held for Measure W (data not shown).

3.2. Regression

The backwards elimination process removed education level, so the final model predicted total perceived health from eight clinical covariates and five demographic covariates. Pairwise correlations indicated no concerns of multicollinearity among clinical covariates. For the Measure UW model, four clinical and four demographic covariates had statistically significant associations with perceived health (Fig. 4). Controlling for demographics and other clinical covariates, a one-SD increase in PHQ-9 score (4.21 points) was associated with a 9.78 point decrease in perceived health (95% CI: $-10.55, -9.01$). On average, a one-SD increase in BMI (7.25 kg/m^2) or cotinine concentration (128.32 ng/mL) corresponded to a decrease of 0.98 points (95% CI: $-1.56, -0.41$) and 0.97 points (95% CI: $-1.67, -0.27$), respectively, in perceived health score. A one-standard-deviation increase in LDL (36.11 mg/dL), corresponded to a 0.71 point increase in perceived health (95% CI: $0.26, 1.16$), on average.

Controlling for the clinical profile and other demographic covariates, a one-year increase in age corresponded to a 0.16 point decrease in perceived health (95% CI: $-0.19, -0.13$), on average, while a one-unit increase in income-to-poverty ratio was associated with an increase of 0.87 points (95% CI: $0.43, 1.31$). Further, a respondent who identified their ethnicity as Hispanic, on average, had perceived health 1.97 points higher than someone who identified as White (95% CI: $0.08, 3.86$), and a respondent who was not a U.S. citizen had perceived health 2.22 points higher than a U.S. citizen with the same clinical and demographic profile (95% CI: $0.21, 4.23$). Measure W results were nearly identical (Fig. 4). Model summary tables and diagnostic plots are presented in the supplement.

4. Discussion

This study produced two variants of a novel perceived health measure that translates the SF-12 for use in the NHANES setting. All selected NHANES analogs presented sufficient variability in their responses (Fig. 2). Both measure variants followed distributions and correlational patterns consistent with knowledge of population health and reflective of the original SF-12. Sample means and standard deviations for both measure variations very nearly parallel the SF-12 targets of a population mean of 50 and a population standard deviation of 10 for each component. The distribution of each component is left-skewed, with more participants having scores indicating better perceived physical and

Table 2

Selected demographic characteristics of National Health and Nutrition Examination Survey (NHANES) respondents in the sample cohort, 2015–2016 ($n = 5,719$).

Variable	Mean (SD)
Age, years	47.94 (17.19)
Income-to-poverty ratio	3.00 (1.65)
Variable	Percentage of cohort
Gender	
Male	48.07
Female	51.93
Race-Ethnicity Mexican	
American	8.87
Other Hispanic	6.41
Non-Hispanic White	63.85
Non-Hispanic Black	11.35
Non-Hispanic Asian	5.81
Other or mixed race	3.71
U.S. Citizen	90.03
Education level	
Less than high school diploma	14.56
High school diploma or equivalent	20.66
Some college or associate's degree	32.50
College graduate or above	32.29

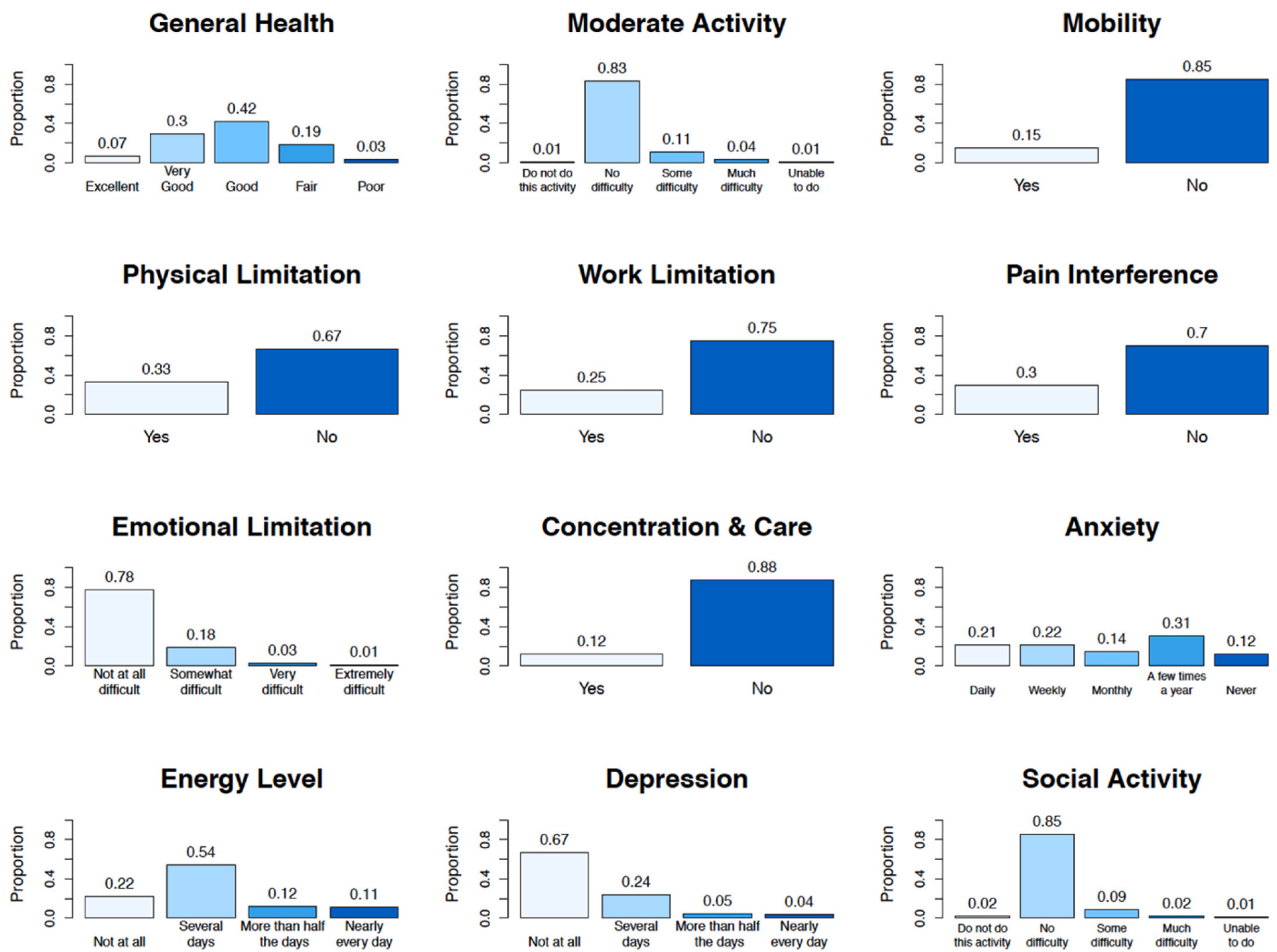


Fig. 2. Response distributions for National Health and Nutrition Examination Survey 2015–2016 analogs used in measure construction ($n = 3,367$).

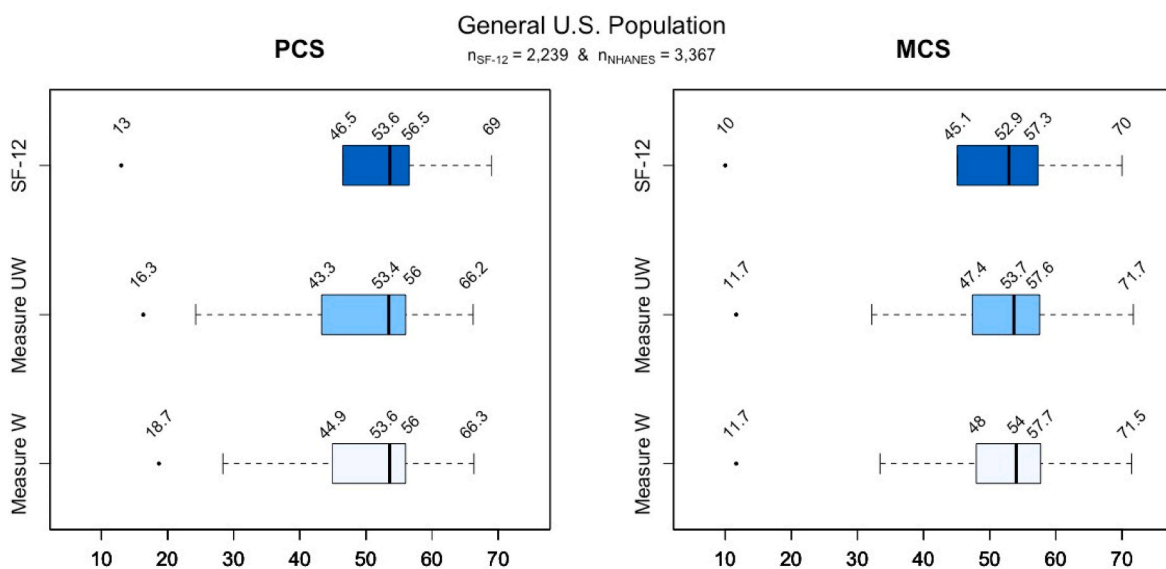


Fig. 3. Side-by-side box plots for Short Form 12, Measure UW, and Measure W score distributions for the general U.S. adult population. Abbreviations: UW, un-weighted; W, weighted; PCS, physical component score; MCS mental component score.

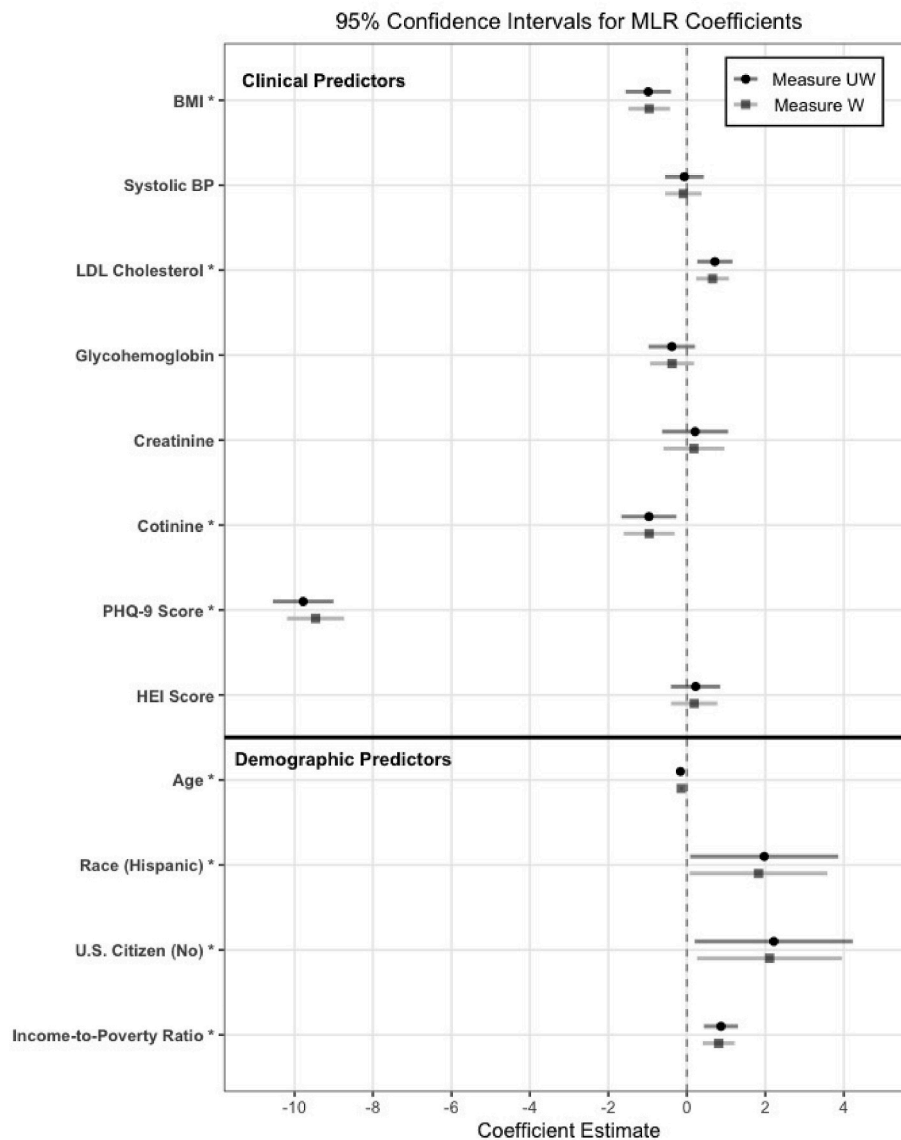


Fig. 4. Estimated regression coefficients with 95% confidence intervals for covariates in the final models for both measures. Estimates that reached statistical significance are marked with an asterisk. Abbreviations: UW, unweighted; W, weighted; PCS, physical component score; MCS mental component score.

mental health. The absence of floor and ceiling effects signals that the measures can distinguish between subjects with differing perceived health levels (Hunt & McEwen, 1980).

Further, the boxplots in Fig. 3 confirm that the numerical summaries of Measure UW, Measure W, and SF-12 score distributions are remarkably similar, giving reassurance that our measures reasonably replicate the SF-12 using NHANES questionnaire variables. As age increased, the median PCS score decreased, and the interquartile range widened (see supplemental Figure B1). This indicates that not only was increased age associated with wider variability in perceived physical health, but also correlated with lower average perceived physical health. These observations are consistent with general knowledge of how age and physical health interact. Minor discrepancies can be attributed to sample-to-sample differences, changes in population health since the SF-12 was developed in the 1990s, and the simple fact that the three measures are all slightly different.

Correlational patterns with demographic factors were consistent with documented knowledge of determinants of perceived health. Because many variables influence perceived health, it would be unreasonable to anticipate a strong linear relationship with any single

variable. Thus, we expected correlation coefficients of moderate strength, and each pair met this expectation. Regarding direction, we expected negative correlations with age (Loprinzi, 2015) and recent use of medical services (Hunt & McEwen, 1980). Correlations showed that as individuals got older, their perceived health did, in fact, worsen. Additionally, the more an individual accessed medical care in the past year, the worse their perceived health became. We anticipated a positive correlation with income-to-poverty ratio (Brazier et al., 1992; Shields & Shooshtari, 2001). The positive correlation coefficient confirmed that as SES improved, so did perceived health. Thus, we conclude that both variants are valid measures of perceived health for use in the NHANES setting.

Regression analysis demonstrated the utility of the novel measures in evaluating interactions between clinical health, demographic influences, and health perception. The strength and direction of the observed associations were the same for both measure variants (Fig. 4), confirming that the two act almost interchangeably in an analytic setting. Thus, despite its underlying simplifying assumptions, the more straightforward Measure UW may be preferable for use in practice.

Backwards elimination removed education level from the model. We

propose two mechanisms by which education level could feasibly affect perceived health. Education level is likely a proxy for SES. Because income-to-poverty ratio was also included in the model, including education level may not have supplied any additional information to predict perceived health. The second possible mechanism is that individuals who are more highly educated may be better equipped to make healthier decisions, access health services, and adhere to medical advice. In this instance, a higher education level manifests itself in improved clinical health. Thus, controlling for the clinical health profile in the model may have accounted for variance in perceived health attributable to education level. Either way, that education level was not a significant predictor does not mean that education level does not have an effect on health. Rather, it suggests that the effect of education on health perception is indirect and may be mediated through other factors like clinical health measurements or alternative measures of SES.

Four clinical covariates reached statistical significance: BMI, LDL cholesterol, cotinine, and PHQ-9 score. That BMI, cotinine, and PHQ-9 score should influence health perception was largely unsurprising. Whether visually evident (BMI), associated with conscious behaviors (smoking), or related to daily mood (depression), each of these three measures are often salient in our consideration of personal health. Because we standardized the clinical measurements, the large estimated coefficient for PHQ-9 suggests that depression score is by far the most influential predictor, likely because it accounts for much of the variation in MCS. Another potential explanation is that people with depression are more accurate in self-assessment, such as the assessment of one's own state of health (Bortolotti & Antrobus, 2015). If this depressive realism is present, then the strong negative effect of PHQ-9 on perceived health suggests that non-depressed people, on average, overestimate their good health.

It was surprising, however, that a rise in LDL cholesterol was correlated with an increase in perceived health; that is, all else equal, those with higher "bad" cholesterol perceived themselves as healthier. An association in the opposite direction would have been more consistent with the clinical reality. This important result implies that perhaps individuals are unaware of their cholesterol levels or are interpreting their health information incorrectly. Improved patient education about what cholesterol measurements mean and the associated health risks could help individuals more accurately perceive their level of health.

Four demographic covariates reached statistical significance: age, Hispanic ethnicity (compared to White), U.S. citizenship status, and income-to-poverty ratio. It is well-documented that increased age leads to poorer health and that those of higher SES experience better health (Barr, 2019), so the observed associations were consistent with the literature. The reasoning behind the significance of Hispanic ethnicity and citizenship status is less clear. A respondent who identifies as Hispanic will, on average, have perceived health nearly two points higher than someone who identifies as White when the two respondents have identical clinical profiles, ages, incomes relative to the poverty line, and citizenship status. This result may be explained by the Latino paradox, the reality that Latinos in the U.S. fare better in a variety of health outcomes than do other racial-ethnic groups, despite socioeconomic disadvantage (Campbell et al., 2012). Additionally, on average, a respondent who is not a U.S. citizen will have perceived health over two points higher than a citizen with the same clinical and demographic profile. Studies suggest that immigrants often display better health relative to native-born populations in industrialized nations like the U. S., despite disadvantages in access to health care and SES (Riosmena et al., 2017). This finding has been attributed to self-selection of healthier foreign-born individuals for migration and the protective effect of high social capital in immigrant communities. Critics suggest that data collection procedures may disproportionately overstate immigrant health, but if the immigrant health advantage does exist in reality – not just due to biased data – then it is interesting that this phenomenon is also evident in self-perception of health.

Study limitations included missing data and tenuous regression

assumptions. Details on methods to address missing data in both measure construction and regression are available in the supplement. The attempted imputation of interview responses, though validated for use with SF-12, was likely too simple for the NHANES setting. Health data, particularly from a multistage sample like NHANES, presents intricate challenges for imputation. We know that a respondent's health and their perception of it depend on many factors that imputation did not account for. Continuing research should focus on developing more nuanced imputation methods for NHANES because excluding individuals with incomplete data can induce selection bias. We excluded respondents from regression analysis for missing data, as well, mostly due to missing cholesterol measurements. Cholesterol measurements were only taken in morning examinations; however, participants were randomly assigned to testing slots. Thus, excluding respondents without cholesterol measurements reduced the sample size, but is not a source of bias. No other rates of missingness in clinical or demographic data were high enough to raise concern about additional selection bias.

Additionally, residual plots for both models demonstrated heteroskedasticity. The large variance in residuals for smaller fitted values of perceived health can be explained by the relatively few low perceived health scores present in the data, so the model has a lower level of certainty in predicting these values. Neither transformation of variables nor weighted least squares regression was successful in reducing the heteroskedasticity, so we should interpret regression output with added caution. Future studies might benefit from using a regression procedure that does not require assuming constant variance.

An important strength is that the developed measure construction framework is a replicable process that can be applied to future NHANES survey cycles, as well as retroactively on cycles dating back to the start of the continuous NHANES in 1999. Eleven of the 12 selected NHANES analogs have been asked in identical or minimally altered form on every survey cycle since 1999 (CDC, 2019). Only the "Mobility" item does not have an evident analog in NHANES prior to 2013. However, the physical functioning questionnaire could be used to select an alternative or construct a composite, as we did for the "Pain interference" item. The methodology could also be modified to create SF-12 equivalent measures for international health surveys. Further, perceived health measures that are compatible with national surveys like NHANES have the benefit of accessing large, population-representative samples. The large sample size should be noted as an additional strength of the current study.

5. Conclusions

In this study, we have successfully constructed a reproducible SF-12 equivalent perceived health measure compatible for use with NHANES, one of the United States' most robust sources of population health data. Creating a SF-12 analogous measure for use with NHANES expands research possibilities in a wide range of applied settings: population health, medical sociology, epidemiology, and more. Accurate perceived health measures like ours can be used to develop a more comprehensive understanding of how demographics, clinical health, and health perception influence one another. For example, we identified that an increase in "bad" cholesterol is paradoxically associated with improved perceived health, signaling a potentially important gap in health literacy. Insights such as this one have the potential to improve patient-provider communication, facilitate development of improved health education and promotion strategies, and possibly improve health outcomes by addressing gaps between perceived and clinical health, ultimately lowering health care costs.

Author statement

Elise C. Covert: Conceptualization, Methodology, Software, Formal analysis, Visualization, Writing – original draft.

Anna M. Baker: Methodology, Writing – review & editing,

Supervision.

Owais Gilani: Methodology, Formal analysis, Writing – review & editing, Supervision.

Financial disclosures

None.

Ethical statement

This study did not require ethical approval by an IRB. Publicly available NHANES data are deidentified to protect participant privacy.

Declaration of competing interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ssmph.2022.101081>.

References

- Barr, D. A. (2019). *Health disparities in the United States: Social class, race, ethnicity, and the social determinants of health*. JHU Press.
- Bortolotti, M., & Antrobus, M. (2015). Costs and benefits of realism and optimism. *Current Opinion in Psychiatry*, 28, 194–198.
- Brazier, J. E., Harper, R., Jones, N., et al. (1992). Validating the SF-36 health survey questionnaire: New outcome measure for primary care. *British Medical Journal*, 305, 160–164.
- Campbell, K., Garcia, D. M., Granillo, C. V., et al. (2012). Exploring the Latino paradox: How economic and citizenship status impact health. *Hispanic Journal of Behavioral Sciences*, 34, 187–207.
- Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey questionnaire. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention <https://www.cdc.gov/nchs/nhanes/search/datapage.aspx?Component=Questionnaire&CycleBeginYear=2015> [Accessed on 22 October 2019].
- Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey data. Hyattsville, MD: U. S. Department of Health and Human Services, Centers for Disease Control and Prevention <https://www.cdc.gov/nchs/nhanes/ContinuousNhanes/Default.aspx?BeginYear=2015> [Accessed on 3 October 2019].
- Hunt, S. M., & McEwen, J. (1980). The development of a subjective health indicator. *Sociology of Health & Illness*, 2, 231–246.
- Hunt, S. M., McEwen, J., & McKenna, S. P. (1985). Measuring health status: A new tool for clinicians and epidemiologists. *The Journal of the Royal College of General Practitioners*, 35, 185–188.
- Karimi, M., & Brazier, J. (2016). Health, health-related quality of life, and quality of life: What is the difference? *Pharmacoeconomics*, 34, 645–649.
- Loprinzi, P. D. (2015). Factors influencing the disconnect between self-perceived health status and actual health profile: Implications for improving self-awareness of health status. *Preventive Medicine*, 73, 37–39.
- McDowell, I. (2006). *Measuring health*. New York, NY: Oxford University Press.
- Miilunpalo, S., Vuori, I., Oja, P., et al. (1997). Self-rated health status as a health measure: The predictive value of self-reported health status on the use of physician services and on mortality in the working-age population. *Journal of Clinical Epidemiology*, 50, 517–528.
- Nayak, S., Hubbard, A., Sidney, S., et al. (2018). A recursive partitioning approach to investigating correlates of self-rated health: The CARDIA Study. *SSM - Population Health*, 4, 178–188.
- Office of Public Health Assessment. (2004). *Health status in Utah: The medical outcomes study SF-12 (2001 Utah health status survey report)*. Salt Lake City, UT: Utah Department of Health.
- Perneger, T. V., & Burnand, B. (2005). A simple imputation algorithm reduced missing data in SF-12 health surveys. *Journal of Clinical Epidemiology*, 58, 142–149.
- Riosmena, F., Kuhn, R., & Jochem, W. C. (2017). Explaining the immigrant health advantage: Self-selection and protection in health-related factors among five major national-origin immigrant groups in the United States. *Demography*, 54, 175–200.
- Shields, M., & Shooshtari, S. (2001). Determinants of self-perceived health. *Health Reports*, 13, 35–52.
- Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care*, 473–483.
- Ware, J., Kosinski, M., & Keller, S. (1995). *How to score the SF-12 physical and mental health summary scales*. Boston, MA: The Health Institute, New England Medical Center.