



HHS Public Access

Author manuscript

Qual Life Res. Author manuscript; available in PMC 2015 October 08.

Published in final edited form as:

Qual Life Res. 2013 August ; 22(6): 1201–1211. doi:10.1007/s11136-012-0246-z.

Monitoring population health for Healthy People 2020: evaluation of the NIH PROMIS® Global Health, CDC Healthy Days, and satisfaction with life instruments

John P. Barile,

Department of Psychology, University of Hawai'i at Mānoa, 2530 Dole Street, Sakamaki Hall
C404, Honolulu, HI 96822-2294, USA

Bryce B. Reeve,

Lineberger Comprehensive Cancer Center and Department of Health Policy and Management,
Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill,
NC, USA

Ashley Wilder Smith,

Outcomes Research Branch, Applied Research Program, Division of Cancer Control and
Population Sciences, National Cancer Institute, Bethesda, MD, USA

Matthew M. Zack,

Division of Population Health, NCCDPHP, US Centers for Disease Control and Prevention, 4770
Buford Highway, NE, Mailstop K-51, Atlanta, GA 30331-3717, USA

Sandra A. Mitchell,

Outcomes Research Branch, Applied Research Program, Division of Cancer Control and
Population Sciences, National Cancer Institute, Bethesda, MD, USA

Rosemarie Kobau,

Division of Population Health, NCCDPHP, US Centers for Disease Control and Prevention, 4770
Buford Highway, NE, Mailstop K-51, Atlanta, GA 30331-3717, USA

David F. Cella,

Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine,
Chicago, IL, USA

Cecily Luncheon, and

Division of Population Health, NCCDPHP, US Centers for Disease Control and Prevention, 4770
Buford Highway, NE, Mailstop K-51, Atlanta, GA 30331-3717, USA

William W. Thompson

Division of Population Health, NCCDPHP, US Centers for Disease Control and Prevention, 4770
Buford Highway, NE, Mailstop K-51, Atlanta, GA 30331-3717, USA

Barile@Hawaii.edu.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Conflict of interest The authors do not report any conflicts of interest with presenting these findings.

Abstract

Purpose—Healthy People 2020 identified health-related quality of life and well-being (WB) as indicators of population health for the next decade. This study examined the measurement properties of the NIH PROMIS® Global Health Scale, the CDC Healthy Days items, and associations with the Satisfaction with Life Scale.

Methods—A total of 4,184 adults completed the Porter Novelli's HealthStyles mailed survey. Physical and mental health (9 items from PROMIS Global Scale and 3 items from CDC Healthy days measure), and 4 WB factor items were tested for measurement equivalence using multiple-group confirmatory factor analysis.

Results—The CDC items accounted for similar variance as the PROMIS items on physical and mental health factors; both factors were moderately correlated with WB. Measurement invariance was supported across gender and age; the magnitude of some factor loadings differed between those with and without a chronic medical condition.

Conclusions—The PROMIS, CDC, and WB items all performed well. The PROMIS items captured a broad range of functioning across the entire continuum of physical and mental health, while the CDC items appear appropriate for assessing burden of disease for chronic conditions and are brief and easily interpretable. All three measures under study appear to be appropriate measures for monitoring several aspects of the Healthy People 2020 goals and objectives.

Keywords

Health-related quality of life; Well-being; Measurement invariance; Structural equation modeling; Population health; Healthy People 2020

Introduction

Healthy People 2020, a national agenda to guide disease prevention and health promotion activities in the United States over the next decade, has identified several cross-cutting health measures to monitor progress in improving population health [1]. Three indices that reflect health-related quality of life (HRQOL) and well-being (WB), specifically the Centers for Disease Control and Prevention (CDC) Healthy Days Measures [2], the National Institutes of Health (NIH) Patient-Reported Outcome Measurement Information System® (PROMIS®) Global Health Measure [3], and the Satisfaction with Life Scale (SWLS) [4] comprise one of the measure sets for Health People 2020 [1].

Each of the three measures has particular strengths related to their underlying content and measurement approach, feasibility for surveillance, applicability in different population subgroups, and relevance for clinical studies and public health program development. The Centers for Disease Control and Prevention (CDC) Healthy Days Measures [5] and the National Institutes of Health (NIH) Patient-Reported Outcome Measurement Information System® (PROMIS®) Global Health Measure [3] are also currently being used on large national and state surveys, including the National Health Interview Survey, Behavioral Risk Factor Surveillance System, and the National Health and Nutrition Examination Survey. However, prior research has not examined the associations among the measures in a

representative sample of the US population, including adults with and without chronic disease—a focus of particular salience for population health assessment and for the ongoing Healthy People 2020 initiative.

Quality of life is generally considered the state of physical, mental, and social well-being [6]. Specifically, health-related quality of life is measured by the aspects of overall quality of life that can be clearly shown to affect physical and mental health [4]. The measurement of well-being includes individuals' global satisfaction with life (e.g., Satisfaction with Life Scale [4]), which serves as a foundational measure for Healthy People 2020 [1]. The purpose of this study was to evaluate the measurement properties of these three scales, including testing the extent to which each of these scales measures similar constructs, using a large, nationally representative sample of community-dwelling adults, and to assess the strengths and limitations of each measure for monitoring population HRQOL and WB.

Within the context of an aging population and the increasing burden of chronic diseases, CDC developed a set of brief questions for public health surveillance to assess the impact of impaired HRQOL by asking about the number of unhealthy days experienced in the past 30 days [2]. These Healthy Days questions were developed as an integrated set of single-item indicators of perceived physical or mental health associated with the leading causes of death and disability tracked on the Behavioral Risk Factor Surveillance System (BRFSS) [7, 8]. These questions were designed to identify population health burden of disease to guide public health program planning [8, 9]. Because HRQOL limitations are measured by an individual's number of unhealthy days, legislators and policy makers can easily interpret study results [10]. The Healthy Days questions have been evaluated [2, 5, 10–13], used in state and national surveillance systems, and established as brief indicators that can guide public health programs and planning [5, 14]. National and state trends (1993–2009) by select demographic subgroups are available online at <http://www.cdc.gov/hrqol>.

In 2004, the NIH funded the Patient-Reported Outcomes Measurement Information System® (PROMIS®) initiative (www.nihpromis.org). PROMIS uses modern measurement theory to reliably and validly measure several patient-reported outcomes (PRO) including general health, fatigue, pain, physical function, and negative affect. PROMIS measures offer researchers efficient and precise HRQOL assessment tools for use in clinical research and outcome evaluation in both chronically ill and healthy populations [15]. The core of PROMIS is its multiple item banks that include anywhere from 30 to over 100 questions per PRO domain. These banks have undergone qualitative appraisal by patients and experts, as well as extensive quantitative evaluation in both clinical and healthy samples of adults. The PROMIS Global Health Scale, a 10-item measure, captures physical, mental, and social aspects of HRQOL [3]. This scale was included in the 2010 National Health Interview Survey (NHIS) with plans for inclusion on the 2015 and 2020 NHIS.

Besides HRQOL, many national and international efforts are underway to measure positive aspects of health—sometimes referred to as WB—at the population level [1, 14, 16]. WB generally includes the presence of positive emotions and moods (e.g., contentment and happiness); the absence of negative emotions (e.g., depression and anxiety); and satisfaction with life, fulfillment, and positive functioning [17–20]. The Satisfaction with Life Scale

(SWLS) is one of the most established and extensively used WB instruments [4] that is being evaluated as a WB surveillance measure for the BRFSS [21].

While studies have examined the reliability and validity of the CDC Healthy Days measure, the NIH PROMIS Global Health Scale, and the SWLS in isolation, less is known regarding the relationships between these measures. Furthermore, no studies have simultaneously examined the measurement properties of these instruments using large, representative, community-dwelling populations by gender, age, and among those with and without chronic conditions. In this study, we examine the measurement equivalence of the nine PROMIS Global Health Scale and three CDC Healthy Days items, and their associations with SWLS in an adult sample with a large percentage over 65 years of age; a general self-rated health item appears on both the PROMIS and the CDC Healthy Days measures. We extend previous work in three ways: (1) determining whether the previously identified factor structures for the PROMIS Global Health Scale and SWLS fit data from another sample of the US general population, (2) assessing whether the CDC Healthy Days items reflect latent constructs comparable to those that underlie the PROMIS physical and mental health subscales, and (3) determining whether the factor structures of the PROMIS subscales combined with the relevant CDC Healthy Days indicators are equivalent across age (<64 vs. 65 years), gender, and the presence/absence of comorbidities. This information will provide data to support the validity of these measures and will aid public health officials in selecting and interpreting measures to monitor population health at the national, state, and local levels. It may also provide context for comparing HRQOL and WB estimates derived from different national and state surveys. We tested three hypotheses:

1. The PROMIS Global Health items and CDC Healthy Days items load on the same latent HRQOL factors, reflecting measures of physical and mental health.
2. These physical health and mental health latent factors, as well as a latent WB factor will maintain the aspects of configural and metric invariance across age groups, gender, and the occurrence of chronic medical conditions.
3. The PROMIS and CDC physical health and mental health factors will be significantly and positively associated with the WB factor, with the WB factor being more strongly associated with the mental health factor.

Materials and methods

Sample and setting

The CDC and the National Cancer Institute (NCI) sponsored this study. Relevant data were licensed from Porter Novelli. In 2010, all three measures (NIH PROMIS Global Health Scale, CDC Healthy Days, and the SWLS) were included on their HealthStyles Survey. Porter Novelli is a private firm that has designed and administered the HealthStyles survey, conducted annually since 1995, with input from public health agencies, such as the CDC and the Substance Abuse and Mental Health Services Administration. The CDC contracts with Porter Novelli to administer the mailed panel survey in an effort to gather data on how adult Americans think, feel, and act about their health [22]. The sampling design includes stratification by region, household income, population density, age, and household size. In

late summer 2010, 4,184 of 6,255 mailed surveys to adults 18 years of age or older were completed and returned (a 66.9 % response rate). Survey respondents were offered five dollars cash and entry into a lottery (one first place prize of \$1,000 and 20 s place prizes of \$50) as compensation for their time. No personal identifiers were included in the licensed data. Survey data were weighted to match US population estimates based on demographic factors taken from the Current Population Survey.

Measures

PROMIS® Global Health Scale—This scale includes ten items that tap into physical, mental, and social aspects of health. In a previous psychometric study [3], PROMIS investigators identified two overall factors—physical health and mental health. However, because one item (in general, please rate how well you carry out your usual social activities and roles) did not load on either factor, PROMIS does not use it to estimate their general population physical and mental health summary T-scores. To be consistent, we removed this item from the current analysis. The remaining nine PROMIS items all use a Likert-type response scale and appear in the “Appendix.”

CDC Healthy Days items—The CDC Healthy Days measure (<http://www.cdc.gov/hrqol/methods.htm>) includes four core items and ten supplemental items. The four core items included on the HealthyStyles survey measured general self-rated health, physically unhealthy days, mentally unhealthy days, and activity limitation days. The general self-rated health item is the same item as the one included on the PROMIS Global Health Scale (See Appendix). The activity limitation days item was not included in the current analyses because the item was not included as a measure for HP2010 (the other three CDC items were included). The CDC items have all demonstrated content, construct, and criterion validity [2, 5, 10–12].

Satisfaction with Life Scale—WB was measured by four items from the SWLS [4]. The SWLS has shown acceptable test–retest reliability, is sensitive to life events, and serves as a criterion measure for testing new WB scales [23–28]. The SWLS asks participants to report how much they agreed, from strongly disagree to strongly agree (a 5-point Likert-type scale), with four statements (Appendix). The fifth item from the SWLS (If I could live my life over, I would change almost nothing) was not a part of the HealthStyles survey because it was difficult to interpret in a content validity exercise and because its omission does not significantly reduce the reliability of the 4-item scale but does reduce respondent burden on lengthy surveys [29]. The SWLS factor structure has been replicated in diverse populations and has acceptable convergent and discriminant validity [30].

Demographic characteristics and chronic medical conditions—Survey respondents reported their age, gender, and the presence of one or more chronic health conditions. Chronic health condition status was determined by asking respondents: During the past year, have you had (or do you currently have) any of these health conditions? The queried conditions included diabetes, arthritis, cancer (other than skin cancer), chronic pain (or sciatica), atrial fibrillation, multiple sclerosis, emphysema/COPD, insomnia/sleep disorder, depression, anxiety, heart disease, congestive heart failure, epilepsy/seizure

disorder, irritable bowel syndrome, bipolar disorder, eating disorder, schizophrenia, or substance abuse. Respondents reporting any of the chronic conditions in the last year were coded as having at least one chronic medical condition, but otherwise coded as not having any chronic medical conditions. Age was reported categorically and recoded into a single dichotomous variable, <65 and ≥65 years.

Data analysis

We used confirmatory factor analyses (CFAs) to determine whether the CDC Healthy Days and the PROMIS Global Health measures shared significant variance reflecting the domains of physical and mental health. We also used CFA to test for measurement equivalence across demographic groups, to identify latent mean differences, and to measure the associations of physical and mental health with a latent measure of WB. All analyses were conducted using Mplus 6.0 [31] with weighted least square estimation using a diagonal weight matrix and a mean and variance-adjusted chi-square test. To account for the ordinal measurement scale of the items (except the CDC indicators, which were treated as continuous and therefore estimated using linear techniques), we employed a nonlinear factor model where probabilities of respondents' choosing a particular response option were nonlinear functions of the latent factor using a probit link function. In this model, factor loadings reflect the model-implied change in the latent response variable for a unit change in the latent factor [32]. We examined fit of the models using a variety of fit indices including the confirmatory fit index (CFI; good model fit ≥.95)[33, 34], the Tucker-Lewis index (TLI, good model fit ≥.95), and the root mean square error of approximation (RMSEA; good model fit ≤.05)[35, 36]. We also included chi-square fit statistics realizing that they would likely reject the null model statistically because they are sensitive to large sample sizes. Figure 1 represents the proposed latent factor model graphically.

We then examined configural and metric measurement invariance, as well as latent mean difference testing, by estimating a series of nested multi-group CFA models, comparing models that allowed different factor loadings between groups with models that constrained these loadings to be equal between groups. Configural invariance is substantiated when the pattern of factor loadings, when freely estimated between groups, is similar and the overall model fits well for both groups [37, 38]. Metric invariance is substantiated when the factor loadings between two groups are statistically equivalent, suggesting that the factors have the same meaning for both groups [37, 38]. We evaluated both the factor loadings and changes in the fit statistics to determine whether they measured similar constructs in both groups and permitted the examination of latent mean differences [36]. Configural and metric measurement invariance were tested by gender, age (<64 vs. ≥65 years), and chronic condition status (none vs. one or more chronic conditions).

Results

The HealthStyles sample included adults of all ages and slightly more females (52 %) than males. They were predominantly white (69 %), and more often married (55 %) than not, and 43 % reported annual incomes of \$60,000 a year (Table 1). Two-thirds of the respondents reported no physically or mentally unhealthy days in the past 30 days (Table 2). Most

respondents reported very good or good health based on the PROMIS physical health and mental health measures, except for physical function: Two-thirds of the respondents reported being completely able to carry out their everyday physical activities. Most respondents slightly agreed or were neutral with respect to the SWLS items.

Table 3 provides the factor loadings for the latent physical health, mental health, and WB factors using two different model specifications. The physical and mental health factors were estimated without (Model 1) and with (Model 2) the inclusion of the CDC physically and mentally Healthy Days indicators, respectively. Different model specifications permitted the examination of the amount of variance the CDC Healthy Days items shared with the PROMIS items.

Model 1 is a single group, three-factor latent measurement CFA model that incorporated the 4-item WB factor (SWLS), a 5-item physical health factor (PROMIS Global Health Scale physical health items), and a 4-item mental health factor (PROMIS Global Health Scale mental health items). Based on prior findings [3], this CFA model included a correlation between the residual error variances for the items, *in general, would you say your health is* and *in general, how would you rate your physical health*, and the items *in general, how would you rate your mental health, including your mood...* and *how often have you been bothered by emotional problems...*. We also correlated the residual error variances for the WB items *in most ways my life is close to my ideal* and *the conditions of my life are excellent* because it significantly improved model fit. The overall CFA model fits well, $\chi^2(59) = 1,243.21, p < .001, CFI = .98, TLI = .97, RMSEA = .07$. Additionally, the physical health factor was strongly correlated with the mental health factor ($r = .85$) and moderately correlated with the WB factor ($r = .46$), and the mental health factor was strongly correlated with the WB factor ($r = .70$) (Table 3, Model 1).

This CFA model was then modified to include the CDC physically and mentally Healthy Days items to load onto the physical health and mental health factors, respectively, (Table 3, Model 2) to determine whether the CDC items measured a similar construct as those measured by the PROMIS HRQOL items. The CDC physically Healthy Days item ($\lambda = .47$) and the mentally Healthy Days ($\lambda = .51$) item loaded moderately well on the physical health and mental health factors, respectively, and the model still exhibited acceptable overall model fit, $\chi^2(84) = 1,752.66, p < .001, CFI = .97, RMSEA = .07, TLI = .96$. Additionally, the inclusion of the CDC items did not alter the correlations between the factors (physical health and mental health, $r = .85$; physical health and WB, $r = .46$; and mental health and WB, $r = .70$).

To further determine the level of shared variance, the CDC indicators had with the PROMIS HRQOL items, two additional CFAs were conducted using negative binomial regression techniques to model the CDC indicators and to account for their skewed distributions (e.g., 69 % of the participants reported zero physically unhealthy days and 73 % report zero mentally unhealthy days). These models suggested that the inclusion of the CDC indicators had virtually no impact on the measurement of physical and mental health (as evidenced by little change in the PROMIS factor loadings), further suggesting that the CDC indicators do not alter the meaning of the physical and mental health factors.

We tested configural and metric invariance across gender (Table 4). Models in which the factor loadings were constrained across groups fit as well [$\chi^2(216) = 1,282.65, p < .001$, CFI = .98, TLI = .98, RMSEA = .05] as when these loadings were free to vary across groups [$\chi^2(204) = 1,653.26, p < .001$, CFI = .97, TLI = .97, RMSEA = .06], suggesting configural invariance. Furthermore, changes in the mean and variance-adjusted chi-square between the two groups (Mplus DIFFTEST option) were not statistically significant, suggesting metric invariance [adjusted $\chi^2(12) = 13.69, p = .32$]. This finding suggests that the factor structure for the physical and mental health HRQOL domains, as measured by the PROMIS items, the CDC Healthy Days items, and the WB factor is invariant across men and women. Examination of the two-tailed latent mean differences between the men and the women in the constrained model revealed that women reported significantly worse physical ($\alpha = -.08, p = .04$) and mental health ($\alpha = -.10, p = .04$) than men but equivalent WB ($\alpha = .05, p = .25$).

We also tested configural and metric invariance across age groups. Those ≥ 65 years old had factor loadings like those < 65 years, except for the PROMIS pain item and the mentally Healthy Days item, which loaded more strongly for those < 65 ; factor loadings constrained across groups, $\chi^2(216) = 1,900.81, p < .001$, CFI = .97, TLI = .97, RMSEA = .06; factor loadings free across groups, $\chi^2(204) = 2,069.71, p < .001$, CFI = .97, TLI = .97, RMSEA = .07; adjusted $\chi^2(12) = 124.42, p < .001$ (Table 4). While the constrained and unconstrained models differed significantly using a chi-square statistic (as was expected due to the large sample size), the approximate fit statistics differed very little between these groups (CFI = .003, TLI = .005, RMSEA = .005), suggesting that the models were essentially equivalent [39, 40]. Those ≥ 65 years old reported significantly worse physical health ($\alpha = -.20, p < .001$) but significantly better mental health ($\alpha = .15, p < .001$) and WB ($\alpha = .32, p < .001$) than those < 65 .

Those with at least one reported chronic medical condition in the last year had factor loadings comparable to those without such conditions for most of the PROMIS HRQOL indicators and all of the WB indicators, but not for the CDC Healthy Days items and the PROMIS fatigue item (Table 4). Comparing models with and without the factor loadings constrained resulted in well-fitting fit indices but significantly different chi-square values and moderate differences in fit indices; factor loadings constrained across groups, $\chi^2(216) = 2,140.48, p < .001$, CFI = .96, TLI = .96, RMSEA = .07; factor loadings free across groups, $\chi^2(204) = 1,678.76, p < .001$, CFI = .97, TLI = .97, RMSEA = .06; adjusted $\chi^2(12) = 420.15, p < .001$, CFI = .010, TLI = .008, RMSEA = .006. The strength of the factor loadings for the PROMIS fatigue item and the CDC healthy days items was weaker in the group with no chronic medical conditions, compared to the group with one or more medical conditions in the past year. This suggests that these indicators shared a greater proportion of variance with their corresponding latent variables when individuals reported a medical condition, compared to those who reported no medical conditions in the past year. Finally, those who reported at least one chronic medical condition in the past year had significantly worse physical health ($\alpha = -.91, p < .001$), mental health ($\alpha = -.63, p < .001$), and WB ($\alpha = -.26, p < .001$) than those who did not.

Discussion

This is the first study to simultaneously examine the factor structure of the PROMIS® Global Health Scale, CDC Healthy Days items, and the SWLS in a sample of community-dwelling adults. For population health surveillance, both the PROMIS Global Health Scale and the CDC Healthy Days items represent measures of physical and mental health with established psychometric properties but until now, little was known regarding how much these two measures overlapped, and the extent to which they demonstrated measurement invariance across subgroups, or were associated with a measure of well-being.

Several key findings emerged from this study. (1) This study confirmed the factor structure found in previous studies measuring physical and mental health [3, 13] (hypothesis 1). (2) The PROMIS Global Health Scale and CDC items measured similar physical and mental health constructs across gender and age but showed some differences in factor loadings depending on reports of a chronic medical condition (hypothesis 2). (3) The modified four-item SWLS measure of well-being fits well and upheld measurement equivalence across gender, age, and whether or not individuals reported a chronic medical condition (hypothesis 2), and was moderately correlated with the physical and mental health factors (hypothesis 3).

The CDC items and PROMIS Global Health Scale measured similar constructs, and the factor loadings were invariant across gender. There were some age-based and chronic disease-based differences in the factor loadings for some of the items that were evaluated. The CDC mentally unhealthy days item had a lower loading on the mental health factor for those aged 65 and older while the physically unhealthy days item had a lower loading among those with no chronic conditions. In both cases, reduced variability in these outcomes may account for differences in the magnitude of factor loadings. Among those with no chronic medical conditions, 86 % reported no physically unhealthy days and 87 % reported no mentally unhealthy days in the last 30 days. Reliance on only the CDC physically and mentally unhealthy days items may not capture some of the variability in HRQOL among healthy populations due to floor effects similar to those observed with other health status measures such as the SF-36 [41] but the CDC items appear well suited to estimate the burden of disease among the general population and will have discriminatory power in subpopulations of older adults and those with one or more chronic conditions. In some cases, researchers may want to consider incorporating additional supplementary CDC Healthy Days items, such as those monitoring sleep and vitality that traditionally result in more variability (data from other surveys available upon request) or by using count-derived distributions, such as Poisson or negative binomial, when factor scores are estimated.

In contrast, the PROMIS items did not show the same floor effects except for physical functioning and fatigue items, which also resulted in lower factor loadings for those without a medical condition. Differences in the magnitude of the factor loadings may reflect restricted variability in fatigue and physical functioning among those without a chronic health condition or may suggest that the physical functioning and fatigue item better reflects the construct of physical health for individuals with at least one comorbidity. However, the PROMIS physical health factor aggregates across five PROMIS items that are sensitive to differences in physical health, both in general and chronically ill populations. For the

PROMIS items, there was also a lower loading for the pain item on the physical health factor for those aged 65 or older. The reduced variability of both the PROMIS items and the CDC items for those over 65 years of age may also reflect a response shift, that is, changes in interpretation, priorities and values, accommodation, or the respondent's internal frame of Ref. [42–44]. Further research is needed to understand differences in factor loadings in subgroups based on age, comorbid conditions, and functional status, and whether additional issues need to be considered when evaluating HRQOL within diverse populations.

This study also fit a WB factor using 4 of the original 5 items from the SWLS [4]. This 4-item factor fits well and maintained equivalence across gender, age, and presence of a chronic medical condition. WB was moderately associated with the physical health factor and strongly associated with the mental health factor. Finally, all three factors were correlated with one another and the differences in the latent means across age, gender, and chronic health condition status were all in expected directions.

When determining the most appropriate measure of HRQOL, one should consider not only the population of interest (e.g., age and chronic disease status) but also the response burden and associated costs. Converging evidence in the current study and from other investigations indicates that the CDC Healthy Days and PROMIS Global Health instruments have established psychometric properties as measures of HRQOL [3, 5, 11, 12, 15], although they differ in some ways. One notable difference between them is their reference period. Both CDC Healthy Days Items ask the respondents to recall their health during the previous 30 days; three of the PROMIS items (mental health, fatigue, and pain) use a 7-day reference period, and the rest refers only to one's current health without specifying a specific time period. These differences in the reference period for each of the indicators may partially account for the differences in the magnitude of the factor loadings observed in this study. Other differences between the scales are the: (1) relative brevity of the three CDC indicators versus the 9 PROMIS items, and (2) the previously mentioned differences in performance by subgroup, with the majority of the PROMIS items performing well within both healthy and ill populations, while the CDC items performed well as global measures of burden of disease. Researchers may also want to consider the interpretability of the metric for each set of items. The PROMIS items can be rescaled to correspond to a *T* score metric (mean of 50 and standard deviation of 10) for research purposes and for comparison with national norms [3], while the CDC items may provide a more intuitive metric for some clinicians and policy makers, because the CDC items quantify the number of physically or mentally unhealthy days.

This study has several strengths including being the first study to assess the PROMIS Global Health scale, the CDC Healthy Days items, and the SWLS items simultaneously using a large, nationally representative community-dwelling population. We tested the measurement equivalence of the three factors across gender, age, and chronic condition status to understand the appropriateness of the measures among different populations. This study also had some limitations. These included relying on a survey administered through the mail, which led to the recruitment of a relatively healthy, English-speaking sample that excluded those living in institutions. In addition, we did not investigate whether specific chronic conditions are more or less associated with the observed differences in the factor loadings

and latent means. The negative impact on HRQOL of one or more comorbidities is not necessarily equal across all chronic conditions or across all HRQOL domains [45]. Future studies are needed to probe the associations between patterns of comorbidity and impairments in HRQOL.

Because the PROMIS Global Health Scale and CDC Healthy Days items measure similar factors but are administered on different national surveys, future studies should consider developing crosswalks for interpreting differences or combining data, as well as determine the extent to which they are able to detect meaningful changes overtime. This would allow researchers to compare and contrast national and state level trends in HRQOL and WB from these surveys. This approach would advance the interpretation of population-based health status indicators toward achievement of Healthy People 2020 goals and objectives. It would also allow policy makers the ability to quickly assess and compare trends in the different surveys. Lastly, since the group with one or more chronic conditions was heterogeneous relative to the number and characteristics of chronic illness, additional research is warranted to explore the extent to which distinct patterns of multimorbidity are associated with differential effects on HRQOL and WB [46, 47]. Such analyses may further inform decisions about the best measures for monitoring population health in these subgroups.

Acknowledgments

Barile and Luncheon were supported in part by an appointment to the Research Participation Program for the Centers for Disease Control and Prevention (CDC) administered by the Oak Ridge Institute for Science and Education through an agreement between the US Department of Energy and CDC. Cella was supported by a grant from the National Institutes of Health to the PROMIS Statistical Center (U54AR057951-02). We would like to thank Adam Burns and Bill Pollard of Porter Novelli for reviewing drafts of this manuscript.

Appendix

See Table 5.

Table 5

Measures of well-being, physical health, and mental health

Item	Range of response categories	Domain
Satisfaction with Life Scale		
In most ways, my life is close to my ideal	Strongly disagree (1)–strongly agree (5)	Well-being
The conditions of my life are excellent	Strongly disagree (1)–strongly agree (5)	Well-being
I am satisfied with my life	Strongly disagree (1)–strongly agree (5)	Well-being
So far I have gotten the important things I want in life.	Strongly disagree (1)–strongly agree (5)	Well-being
<i>PROMIS® items</i>		
Physical health		
In general would you say your health is? ^a	Excellent (5)–poor (1)	General health
In general, how would you rate your physical health?	Excellent (5)–poor (1)	Physical health

Item	Range of response categories	Domain
To what extent are you able to carry out your everyday physical activities such as walking, climbing stairs, carrying groceries, or moving a chair?	Completely (5)–not at all (1)	Physical function
In the past 7 days, how would you rate your pain on average?	No pain (0)–worst imaginable pain (10)	Pain
In the past 7 days, how would you rate your fatigue on average?	None (1)–very severe (5)	Fatigue
Mental Health		
In general, how would you say your quality of life is?	Excellent (5)–poor (1)	Quality of life
In general, how would you rate your mental health, including your mood and your ability to think?	Excellent (5)–poor (1)	Mental health
In general, how would you rate satisfaction with your social activities and relationship?	Excellent (5)–poor (1)	Social discretionary
In the past 7 days, how often have you been bothered by emotional problems such as feeling anxious, depressed or irritable?	Always (5)–never (1)	Emotional problems
CDC Healthy Days Items		
Now thinking about your physical health, which includes physical illness and injuries, for how many days during the past 30 days was your physical health not good?	Number of days —	Physical health
Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?	Number of days —	Mental health

^aThe general health item appears on both the PROMIS and CDC measures

References

1. US Department of Health and Human Services. Foundation health measures. Healthy People 2020. 2011
2. Centers for Disease Control and Prevention. Measuring Healthy Days. U.S. Department of Health and Human Services; Atlanta, GA: 2000.
3. Hays RD, Bjorner JB, Revicki DA, Spritzer KL, Cella D. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. *Quality of Life Research*. 2009; 18(7):873–880. [PubMed: 19543809]
4. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. *Journal of Personality Assessment*. 1985; 49:71–75. [PubMed: 16367493]
5. Moriarty DG, Kobau R, Zack MM, Zahran HS. Tracking Healthy Days—A window on the health of older adults. *Preventing Chronic Disease*. 2005; 2(3):A16. [PubMed: 15963318]
6. Fayers, PM.; Machin, D. *Quality of life: The assessment, analysis and interpretation of patient-reported outcomes*. Wiley; Chichester: 2007.
7. National Center for Health Statistics. In Proceedings of the 1993 NCHS conference on the cognitive aspects of self-reported health status. 1994 Unpublished manuscript.
8. National Center for Chronic Disease Prevention and Health Promotion. Consultation on functional status surveillance for states and communities: Decatur, Georgia, June 4–5, 1992: Meeting report. US Dept. of Health & Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion; 1993.
9. Booske, BC.; Kindig, DA.; Remington, PL.; Kempf, AM.; Peppard, PE. How should we measure health-related quality of life in Wisconsin? University of Wisconsin Population Health Institute Brief Report (Vol. 1). University of Wisconsin Population Health Institute; 2006.

10. 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. [PubMed: 14498988]
11. Jiang Y, Hesser JE. Using item response theory to analyze the relationship between health-related quality of life and health risk factors. *Preventing Chronic Disease*. 2009; 6(1):1–10.
12. Mielenz T, Jackson E, Currey S, DeVellis R, Callahan LF. Psychometric properties of the Centers for Disease Control and Prevention Health-related quality of life (CDC HRQOL) items in adults with arthritis. *Health and Quality of Life Outcomes*. 2006; 4:66–84. [PubMed: 16995955]
13. Horner-Johnson W, Suzuki R, Krahn G, Andresen E, Drum C, The RRTC Expert Panel on Health. Structure of health-related quality of life among people with and without functional limitations. *Quality of Life Research*. 2010; 19(7):977–984. [PubMed: 20467819]
14. Parrish RG. Measuring population health outcomes. *Preventing Chronic Disease: Public Health Research, Practice, and Policy*. 2010; 7(4):1–11.
15. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology*. 2010; 63(11): 1179–1194. [PubMed: 20685078]
16. Gallup-Healthways Well-Being Index. 2008. www.well-being-index.com
17. Frey BS, Stutzer A. What can economists learn from happiness research? *Journal of Economic Literature*. 2002; 40(2):402–435.
18. Andrews, FM.; Withey, SB. *Social indicators of well-being: Americans' perceptions of life quality*. Plenum Press; New York: 1976.
19. Diener E. Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist*. 2000; 55(1):34–43. [PubMed: 11392863]
20. Ryff CD, Keyes CL. The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*. 1995; 69(4):719–727. [PubMed: 7473027]
21. Bann CM, Kobau R, Lewis MA, Zack MM, Luncheon C, Thompson WW. Development and psychometric evaluation of the public health surveillance well-being scale. *Quality of Life Research*. 2011; 21(6):1031–1043. [PubMed: 21947657]
22. Porter Novelli International. *HealthStyles survey*. Porter Novelli; Washington, DC: 2010.
23. Pavot W, Diener E, Colvin CR, Sandvik E. Further validation of the satisfaction with life scale: Evidence for the cross-method convergence of well-being measures. *Journal of Personality Assessment*. 1991; 57:149–161. [PubMed: 1920028]
24. Blais MR, Vallerand RJ, Pelletier LG, Brière NM. L'échelle de satisfaction de vie: Validation canadiennefrançaise du "Satisfaction with Life Scale". *Canadian Journal of Behavioural Science*. 1989; 21(2):210–223.
25. Magnus K, Diener E, Fujita F, Pavot W. Extra-version and neuroticism as predictors of objective life events: A longitudinal analysis. *Journal of Personality and Social Psychology*. 1993; 65(5): 1046–1053. [PubMed: 8246112]
26. Pavot, W. *The assessment of subjective well-being*. The Guilford Press; New York: 2008.
27. Pavot W, Diener E. The satisfaction with life scale and the emerging construct of life satisfaction. *The Journal of Positive Psychology*. 2008; 3(2):137–152.
28. Larsen, RJ.; Eid, M. *Ed Diener and the science of subjective well-being*. The Guilford Press; New York: 2008.
29. Kobau R, Sniezek J, Zack MM, Lucas RE, Burns A. Well-being assessment: An evaluation of well-being scales for public health and population estimates of well-being among US adults. *Applied Psychology: Health and Well-Being*. 2010; 2:272–297.
30. Diener, E. *The science of well-being: The collected works of Ed Diener*. Springer; New York: 2009.
31. Muthen, BO.; Muthen, LK. *Mplus statistical analysis with latent variables (Version 6.0)*. Muthen & Muthen; Los Angeles: 2010.

32. Flora D, Curran P. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*. 2004; 9(4):466–491. [PubMed: 15598100]
33. Bentler PM. Comparative fit indexes in structural models. *Psychological Bulletin*. 1990; 107(2): 238–246. [PubMed: 2320703]
34. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6:1–55.
35. Browne MW, Cudeck R. Alternative ways of assessing model fit. *Sociological Methods Research*. 1992; 21:230–258.
36. Kline, RB. Principles and practice of structural equation modeling. 2nd ed.. The Guilford Press; New York: 2005.
37. Gregorich SE. Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*. 2006; 44(11 Suppl 3):S78–S94. [PubMed: 17060839]
38. Vandenberg RJ. Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods*. 2002; 5(2):139–158.
39. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*. 2002; 9:233–255.
40. Little TD. Mean and covariance structures (MACS) Analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research*. 1997; 32:53–76.
41. Gandek B, Ware JE Jr, Aaronson NK, Alonso J, Apolone G, Bjorner J, et al. Tests of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: Results from the IQOLA Project. International quality of life assessment. *Journal of Clinical Epidemiology*. 1998; 51(11): 1149–1158. [PubMed: 9817132]
42. Albrecht GL, Devlieger PJ. The disability paradox: High quality of life against all odds. *Social Science and Medicine*. 1999; 48(8):977–988. [PubMed: 10390038]
43. Krahn G, Fujiura G, Drum C, Cardinal B, Nosek M, RRTC Expert Panel on Health Measurement. The dilemma of measuring perceived health status in the context of disability. *Disability and Health Journal*. 2009; 2(2):49–56. [PubMed: 21122743]
44. Schwartz CE, Andresen EM, Nosek MA, Krahn GL. Response shift theory: Important implications for measuring quality of life in people with disability. *Archives of Physical Medicine and Rehabilitation*. 2007; 88(4):529–536. [PubMed: 17398257]
45. Rothrock N, Hays R, Spritzer K, Yount S, Riley W, Cella D. Relative to the general US population, chronic diseases are associated with poorer health-related quality of life as measured by the Patient-Reported Outcomes Measurement Information System (PROMIS). *Journal of Clinical Epidemiology*. 2010
46. U.S. Department of Health and Human Services. Multiple chronic conditions—A strategic framework: Optimum health and quality of life for individuals with multiple chronic conditions. U.S. Department of Health and Human Services; Washington, DC: 2010.
47. Boyd CM, Fortin M. Future of multimorbidity research: How should understanding of multimorbidity inform health system design? *Public Health Reviews*. 2010; 32(2):451–474.

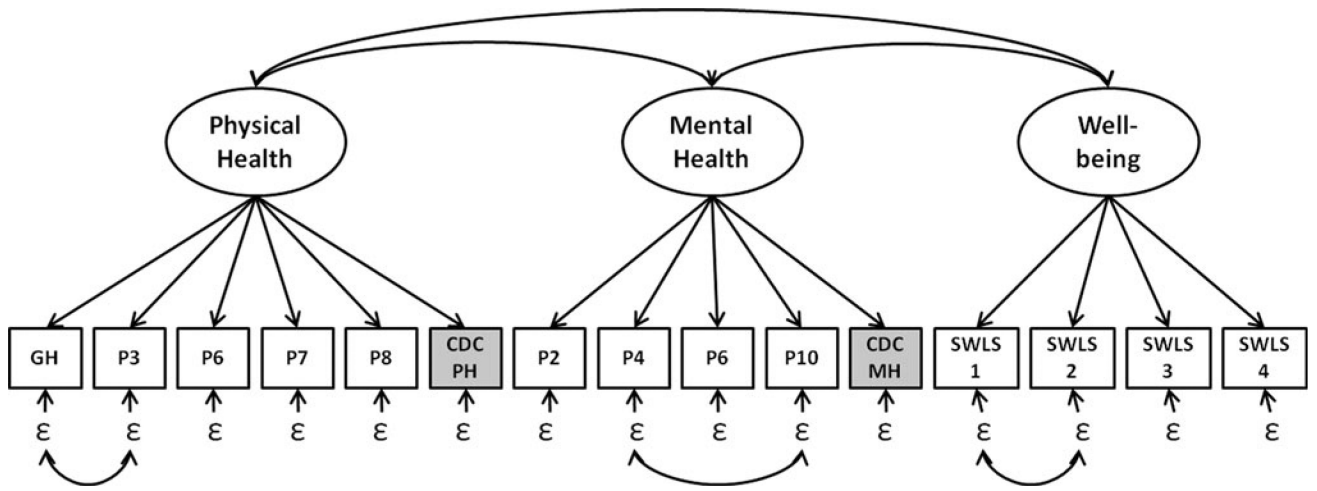


Fig. 1. The latent measurement model. The three large ovals represent the three latent constructs measured by PROMIS® Global Health, CDC Healthy Days, and SWLS items. GH represents the general health indicator, which appears on both the PROMIS and CDC measures, and all *P* indicators represent PROMIS indicators. The CDC unhealthy days items are *gray* because they are not included in the initial estimation of the model. The *curved double-headed* errors represent estimated correlations between latent factors and correlations between error variances

Table 1

Frequency of participants identifying with various demographics

	<u>Unweighted</u>		Weighted (%)
	Frequency	(%)	
Race or ethnicity			
White	2,842	68	69
Black	477	11	12
Latino	495	12	14
Other race or ethnicity	370	9	6
Female	2,181	52	52
Age of respondent			
18–24	60	1	13
25–34	414	10	18
35–44	707	17	18
45–54	1,269	30	20
55–64	806	19	15
65+	928	22	17
Income			
Under \$15,000	700	17	13
\$15,000–\$24,999	351	8	12
\$25,000–\$39,999	523	13	16
\$40,000–\$59,999	614	15	17
\$60,000 and over	1,996	48	43
Education			
Completed 8th grade or less	52	1	1
Completed 9–11th grade	182	4	4
Graduated high school or GED	916	22	21
Some college	1,537	37	42
College graduate	848	20	19
Completed Graduate School	620	15	13
Marital status			
Married	2,885	69	55
Widowed	231	6	6
Divorced	396	10	11
Separated	54	1	1
Never married	511	12	23
Domestic partnership	107	3	3
Employment			
Works for someone else full time	1,981	48	52
Self-employed	321	8	7
Works for someone else part-time only	364	9	10
Temporarily unemployed	128	3	3

	<u>Unweighted</u>		Weighted (%)
	Frequency	(%)	
Retired and not employed	782	19	15
Disabled, student, etc., and not employed	202	5	6
Full-time homemaker	387	9	7

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Means and distributions of the items included in the CDC, PROMIS, and satisfaction with life measures ($N = 4,184$)

	Mean	Frequency					
		0 days (%)	1–10 days (%)	11–20 days (%)	21–29 days (%)	30 days (%)	
CDC items							
# of Physically unhealthy days (0–30)	3.24	69	21	4	2	5	
# of Mentally unhealthy days (0–30)	2.51	73	20	3	1	3	
	Mean	Better health ^b		Worse health ^b			
		5 (%)	4 (%)	3 (%)	2 (%)	1 (%)	
PROMIS® physical health items							
PROMIS/CDC general health ^a	3.38	10	38	37	14	3	
Physical health	3.33	8	37	38	14	3	
Physical function	4.44	66	19	11	4	1	
Pain	3.85	27	45	16	12	1	
Fatigue	3.80	23	41	31	5	1	
PROMIS® mental health items							
Quality of life	3.62	16	43	31	9	2	
Mental health	3.71	19	43	29	8	1	
Social discretionary	3.44	14	37	33	13	4	
Emotional problems	3.57	23	41	31	5	2	
	Mean	Strongly agree				Strongly disagree	
		(%)	(%)	(%)	(%)	(%)	(%)
Satisfaction with Life Scale items							
In most ways, my life is close to my ideal	3.29	10	37	33	15	6	
The conditions of my life are excellent	3.32	11	36	33	14	6	
I am satisfied with my life	3.58	20	39	26	10	5	
So far I have gotten the important things I want in life	3.63	21	41	23	10	5	

^aThe general health item is part of the CDC Healthy Days measure as well as the PROMIS® Global Health scale.

^bThe headings for the PROMIS® item response categories have been altered from those that appeared in the survey in order to simplify the table. Please see the “Appendix” for the exact wording and response corresponding response categories for each item

Table 3Standardized loadings (CFA) for physical health, mental health, and well-being factors ($N = 4,184$)

Indicators	Model 1 λ	Model 2 λ
Physical health factor		
PROMIS/CDC general health	.83	.83
PROMIS physical health	.91	.90
PROMIS physical function	.63	.63
PROMIS pain	.62	.62
PROMIS fatigue	.63	.63
CDC physically Healthy Days	–	.54
Mental health factor		
PROMIS quality of life	.93	.93
PROMIS mental health	.78	.79
PROMIS social discretionary	.80	.80
PROMIS emotional problems	.52	.56
CDC mentally Healthy Days	–	.51
Well-being factor		
In most ways, my life is close to my ideal	.79	.79
The conditions of my life are excellent	.88	.88
I am satisfied with my life	.93	.93
So far I have gotten the important things I want in life	.80	.79

The CDC Healthy Days items were reverse scored. Model 1: $\chi^2 (59) = 1,243.21, p < .001, CFI = .98, RMSEA = .07, TLI = .97$. Model 2: $\chi^2 (84) = 1,752.66, p < .001, CFI = .97, RMSEA = .07, TLI = .96$. The physical health factor and the mental health factor at $r = .85$; the physical health factor and the well-being factor at $r = .46$; and the mental health factor and the well-being factor at $r = .70$ (Model 2)

Table 4

Standardized factor loadings from a multi-group confirmatory factor analyses of physical health, mental health, and well-being factors, by gender, age, and number of medical conditions

Indicators	Fully constrained model ($N = 4,184$) λ	Gender model		Age model		Medical conditions model	
		Men ($n = 1,990$) λ	Women ($n = 2,173$) λ	< 65 model ($n = 3,256$) λ	65 model ($n = 928$) λ	No conditions ($n = 1,723$) λ	At least 1 condition ($n = 2,461$) λ
Physical health factor							
PROMIS/CDC general health	.83	.81	.85	.82	.86	.81	.79
PROMIS physical health	.90	.89	.92	.89	.96	.89	.89
PROMIS physical function	.63	.63	.64	.65	.65	.48	.59
PROMIS pain	.62	.58	.66	.65	.48	.53	.55
PROMIS fatigue	.63	.64	.63	.62	.68	.43	.63
CDC physically Healthy Days	.54	.52	.53	.53	.50	.19	.58
Mental health factor							
PROMIS quality of life	.93	.92	.95	.93	.94	.91	.93
PROMIS mental health	.79	.80	.79	.79	.77	.78	.77
PROMIS social discretionary	.80	.79	.81	.80	.79	.79	.81
PROMIS emotional problems	.56	.57	.55	.57	.55	.49	.54
CDC mentally Healthy Days	.51	.49	.51	.51	.38	.31	.53
Well-being factor							
In most ways, my life is close...	.79	.78	.79	.78	.81	.82	.77
The conditions of my life are...	.88	.89	.87	.88	.89	.86	.88
I am satisfied with my life	.93	.94	.93	.94	.92	.93	.94
So far I have gotten the important...	.79	.81	.79	.80	.77	.80	.80

The CDC Healthy Days items were reverse scored. Model fit, gender: $\chi^2(204) = 1,653.26, p < .001, CFI = .97, RMSEA = .06$. Model fit, age: $\chi^2(204) = 2,069.71, p < .001, CFI = .97, RMSEA = .07$. Model, fit, medical conditions: $\chi^2(204) = 1,678.76, p < .001, CFI = .97, RMSEA = .06$