



# HHS Public Access

Author manuscript

*Disabil Health J.* Author manuscript; available in PMC 2017 October 01.

Published in final edited form as:

*Disabil Health J.* 2016 October ; 9(4): 567–574. doi:10.1016/j.dhjo.2016.04.008.

## Measurement characteristics for two health-related quality of life measures in older adults: The SF-36 and the CDC Healthy Days items

John P. Barile, Ph.D.<sup>a,\*</sup>, Willi Horner-Johnson, Ph.D.<sup>b</sup>, Gloria Krahn, Ph.D.<sup>c</sup>, Matthew Zack, M.D.<sup>d</sup>, David Miranda, Ph.D.<sup>e</sup>, Kimberly DeMichele, Ph.D.<sup>e</sup>, Derek Ford, Ph.D.<sup>f</sup>, and William W. Thompson, Ph.D.<sup>g</sup>

<sup>a</sup>University of Hawai'i at Manoa, Department of Psychology, 2530 Dole Street, Sakamaki Hall C404, Honolulu, HI 96822-2294, United States

<sup>b</sup>Oregon Health & Science University, Institute on Development & Disability, 3181 S.W. Sam Jackson Park Rd., Portland, OR 97239-3098, United States

<sup>c</sup>Oregon State University, College of Public Health and Human Sciences, Hallie Ford Center 255, Corvallis, OR 97331, United States

<sup>d</sup>US Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, 1600 Clifton Road, Atlanta, GA 30329-402, United States

<sup>e</sup>Centers for Medicare and Medicaid Services, Division of Consumer Assessment & Plan Performance, 7500 Security Boulevard, Baltimore, MD 21244, United States

<sup>f</sup>US Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, 1600 Clifton Road, Atlanta, GA 30329-402, United States

<sup>g</sup>US Centers for Disease Control and Prevention, National Center on Birth Defects and Developmental Disabilities, 1600 Clifton Road, Atlanta, GA 30329-402, United States

### Abstract

**Background**—The Short Form Health Survey (SF-36) and the Centers for Disease Control and Prevention (CDC) Healthy Days items are well known measures of health-related quality of life. The validity of the SF-36 for older adults and those with disabilities has been questioned.

**Objective**—Assess the extent to which the SF-36 and the Centers for Disease Control and Prevention (CDC) Healthy Days items measure the same aspects of health; whether the SF-36 and the CDC unhealthy days items are invariant across gender, functional status, or the presence of chronic health conditions of older adults; and whether each of the SF-36's eight subscales is independently associated with the CDC Healthy Days items.

\*Corresponding author. Tel.: +1 808 956 6271; fax: +1 808 956 4700. Barile@Hawaii.edu (J.P. Barile).

Disclaimer: 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.

Disclosures: This study did not receive any financial support. The authors do not report any conflicts of interest.

**Methods**—We analyzed data from 66,269 adult Medicare advantage members age 65 and older. We used confirmatory factor analyses and regression modeling to test associations between the CDC Healthy Days items and subscales of the SF-36.

**Results**—The CDC Healthy Days items were associated with the SF-36 global measures of physical and mental health. The CDC physically unhealthy days item was associated with the SF-36 subscales for bodily pain, physical role limitations, and general health, while the CDC mentally unhealthy days item was associated with the SF-36 subscales for mental health, emotional role limitations, vitality and social functioning. The SF-36 physical functioning subscale was not independently associated with either of the CDC Healthy Days items.

**Conclusions**—The CDC Healthy Days items measure similar domains as the SF-36 but appear to assess HRQOL without regard to limitations in functioning.

## Keywords

Health-related quality of life; Measurement; SF-36; Healthy Days; Functional limitations

---

Health-related quality of life (HRQOL) is an important indicator of broad physical and mental health and is used to monitor both progress in clinical settings (e.g. rehabilitation) and population health (e.g., through the Healthy People 2020 [HP2020] initiative).<sup>1</sup> In 2010, HP2020 created a new Health-Related Quality of Life and Well-Being Workgroup of experts to develop 10-year national health objectives using scientifically valid measures of both HRQOL and Well-Being. One of the challenges for HP2020, as well as for health professionals, is to understand how different HRQOL measures are related and whether they measure the same constructs.

In 2012, the Centers for Disease Control and Prevention (CDC) and the National Cancer Institute (NCI) assessed the relationship between the CDC's core Healthy Days measure and the NIH Patient-Reported Outcomes Measurement Information System's (PROMIS) Global Health measure because HP2020 used both measures to assess population health.<sup>2</sup> Both measures were administered on the 2010 Health Styles survey. In the survey, the CDC Physically Unhealthy Days item loaded well onto the PROMIS physical health factor while the CDC Mentally Unhealthy Days item loaded well on to the PROMIS mental health factor. Another prominent HRQOL measure, the Short Form Health Survey (SF-36), is also widely used to assess physical and mental health. As with the PROMIS measure, questions have been raised regarding whether the CDC Healthy Days items and the SF-36 measure similar constructs.<sup>3</sup>

The measurement structure of the SF-36 has been assessed across a wide range of samples collected in many countries.<sup>4</sup> De Vet, Ader, Terwee, and Pouter reviewed the measurement studies to date and reported that among 29 different studies, only six studies carried out confirmatory factor analyses and only three of these studies reported associations between factors.<sup>5</sup> They suggested that although the developers of the SF-36 attempted to achieve high standards with respect to the psychometric properties of the SF-36 instrument, the quality of the factor analytic studies suggests the methods and hypothesis testing could be improved upon. In an early measurement study carried out in ten different countries, Keller, Ware,

Bentler, et al, estimated eight first-order factors, three second-order factors, and one overall third-order factor.<sup>6</sup> The three second-order factors included the original physical functioning and mental functioning factors as well as a third factor that they referred to as general wellbeing. The general wellbeing factor was fit with loadings on the first-order general health factor and the vitality factor. The second-order factor correlations ranged from 0.77 to 0.95.

Our study examined the factor structure of the SF-36 and its congruence with the CDC Healthy Days items in a large sample of older adults. We also investigated potential differences by gender, by the presence of chronic conditions, and by limitations in functioning, as defined by the International Classification of Functioning, Disability, and Health (ICF).<sup>7</sup>

## HRQOL and disability

One concern about the SF-36 centers on the relationship between HRQOL and functioning. Because disability has historically been interpreted as equivalent to poor health, this interpretation has been reflected in measures of health-related quality of life, including the SF-36.<sup>8</sup> While some researchers maintain that functioning is an essential component of HRQOL, others argue that functioning and HRQOL are distinct concepts that should be measured separately.<sup>8-18</sup> The ICF supports this latter view, as it depicts a dynamic interplay between functioning, activities, social participation, environments, and health.<sup>7</sup> This interplay acknowledges the possibility that a person with limitations in functioning can enjoy good health and quality of life. For example, a wheelchair athlete with excellent physical and mental health could in fact be fitter than a non-athlete without limitations in functioning. However, on the SF-36, the wheelchair athlete will score lower than the non-athlete because items on the Physical Functioning scale ask about the ability to walk or climb stairs.

Horner-Johnson and colleagues tested the SF-36 for differential item functioning (DIF) in a sample of adults with and without disabilities who reported no major health problems.<sup>19</sup> DIF is a statistical methodology that estimates the probability of endorsing a single item within a scale differently across populations. As expected, people with spinal cord injuries showed negative DIF on all items of the Physical Functioning scale. People who were blind also showed negative DIF on items pertaining to walking long distances. Because people who were blind performed well on other Physical Functioning items, the authors concluded that the distance walking items reflected issues with navigating in the environment rather than poor health. People who were blind, deaf, or hard-of hearing showed negative DIF on items about ability to work or accomplish tasks. Rather than reflecting health problems, these findings could again be related to environmental barriers (e.g., communication challenges) experienced by people with sensory impairments. The key concern is that items assessing functioning used to assess HRQOL may result in artificially low scores for people with disabilities. Importantly, the extent of this reduction is not always clear and likely differs by disability type.<sup>8,20</sup> A similar DIF analysis of the CDC Healthy Days items showed very little evidence of negative DIF for people with disabilities.<sup>21</sup> That finding suggests the CDC items

may be more appropriate than the SF-36 when researchers or clinicians wish to measure HRQOL separate from limitations in functioning.

## Current investigation

The CDC collects HRQOL data on several national and state surveys to track population-based trends in HRQOL. The indices include measures of the number of physically unhealthy days, mentally unhealthy days, and activity limitation days that adults experienced during the previous 30 days. Despite widespread use of these measures, few studies have examined their associations with the SF-36, one of the most frequently used HRQOL measures in the literature. The first study to examine the measurement characteristics of the CDC Healthy Days items and the SF-36 simultaneously used a relatively small sample of adults ( $N=401$ ) from a single state and inferred from correlation analyses that the two measures assessed the same constructs.<sup>3</sup> A similar study with a sample of veterans with spinal cord injuries also found high correlations between SF-36 subscales and CDC Healthy Days items, although the SF-36 received a low acceptability rating because questions about walking were deemed inappropriate by some individuals.<sup>20</sup> We are unaware of more recent or larger studies or any that used more sophisticated measurement models to assess the similarities and differences between these two HRQOL instruments.<sup>2,22,23</sup> In particular, the relationship between these measures among older adults has yet to be examined. The purpose of the present study was to better understand how closely the CDC Healthy Days measures correspond to the subscales of the SF-36 in older adults and whether any differences between these measures are consistent across gender, limitations in activities of daily living (ADL), and the presence of multiple chronic conditions.

Based on previous literature, we formulated the following research questions:

1. To what extent do latent factors derived from the SF-36 and CDC unhealthy days items measure the same aspects of health?
2. Does the extent to which the SF-36 and the CDC unhealthy days items measure the same constructs depend upon the gender, functional status, or the presence of chronic health conditions of older adults?
3. What are the associations between the CDC physically unhealthy and mentally unhealthy days items and the eight subscales of the SF-36?

## Methods

### Participants

The study utilized data obtained from 66,269 older adults (≥ 65 years of age) who completed the Medicare Health Outcomes Survey (MHOS) in 2004 (response rate = 65.3%). Use of these data has been approved by the U.S. Centers for Medicare and Medicaid Services (CMS). The MHOS is a patient-reported outcomes survey required by CMS and administered to a random sample of Medicare beneficiaries from participating Medicare Advantage Organizations. Medicare Advantage (MA) plans include seniors 65 years or older as well as individuals younger than 65 years with disabilities (not included in this study)

enrolled in Medicare Part A and Part B who also chose to enroll in the optional MA plan. Individuals that enroll in optional MA plans, compared to traditional Medicare, tend to have a higher annual income, live in urban environments, are more likely to be Hispanic/Latino, report slightly better health, and utilize fewer services.<sup>24,25</sup> The purpose of this survey is to monitor health plan performance by assessing the health of participating individuals. Most respondents were women, white, aged 65–74 years, and had annual household incomes below \$30,000 (Table 1).

## Measures

**Medical outcomes study short Form-36 health survey (SF-36)**—The SF-36 ([http://www.rand.org/health/surveys\\_tools/mos/mos\\_core\\_36item.html](http://www.rand.org/health/surveys_tools/mos/mos_core_36item.html)) is a 36-item general health questionnaire designed to provide physical and mental health summary scores based on eight subscales.<sup>26</sup> The measurement characteristics of the SF-36 have been studied extensively to assess its reliability and validity.<sup>6,15,27</sup>

**CDC Healthy Days**—The CDC Healthy Days measures (<http://www.cdc.gov/hrqol/methods.htm>) include four core items and ten supplemental items. The CDC items have all demonstrated content, construct, and criterion validity.<sup>28–32</sup> The MHOS used the four core items that measure general self-rated health, physically unhealthy days, mentally unhealthy days, and activity limitation days. The general self-rated health item (*Would you say that in general your health is: excellent, very good, good, fair, or poor?*) is identical to the one included on the SF-36. For this study, we focused specifically on the physically unhealthy days item (*Now thinking about your physical health, which includes physical illness and injury, for how many days during the past 30 days was your physical health not good?*) and the mentally unhealthy days item (*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?*).

## Activities of daily living and chronic diseases

Activities of Daily Living (ADL) were assessed using a six-item ADL limitations measure that inventoried whether participants had difficulty bathing, dressing, eating, getting in or out of chairs, walking, and using the toilet (based on Katz et al.'s index of ADL).<sup>33</sup> Respondents selected 1 (I am unable to do this activity), 2 (Yes, I have difficulty), or 3 (No, I do not have difficulty).<sup>33</sup> If respondents stated they had difficulty (either 1 or 2) with one or more activities, they were considered to have a limitation.

Multiple chronic conditions were assessed by asking participants whether they had ever been diagnosed with each of 12 chronic medical conditions. Self-reported conditions in the MHOS correspond closely with information available in medical records.<sup>34</sup> Conditions were coded as 0 (never) or 1 (diagnosed) and totaled to create a summary score representing the number of medical conditions experienced by each participant.

## Analyses

We tested associations between the CDC items and the SF-36 subscales using confirmatory factor analyses (CFA) and multivariate regression. We considered two latent models for the

SF-36. Model 1 (the Two Factor Model), based on Ware and Gandek,<sup>4</sup> comprised two latent factors (physical health and mental health), each estimated using five subscales. Two of the subscales, general health and vitality, were included as indicators of both factors. Model 2 (the Three Factor Model), supported by Keller et al,<sup>6</sup> consisted of three latent factors: physical health, mental health, and general wellbeing.<sup>6</sup> After testing both models, we conducted two additional CFAs to determine whether the CDC Healthy Days items loaded onto the SF-36 physical health and mental health factors. For each model, we then specified multiple-group CFA models to examine measurement equivalence across gender, ADL limitations, and number of chronic medical conditions by constraining factor loadings to be equal for both groups (e.g., women vs. men; people with vs. without ADL limitations) and by comparing these results to a model that did not constrain the factor loadings to be equal.<sup>35–37</sup> Finally, we conducted multivariate negative binomial regressions to identify which SF-36 subscales were most closely associated with the CDC Healthy Days items. Missing data on SF-36 and CDC Healthy Days items were addressed using full-information maximum likelihood.

## Results

Survey respondents reported an average of 4.38 physically unhealthy days (Standard Deviation (SD) = 8.8) and 2.1 mentally unhealthy days in the past 30 days (SD = 5.8) (Table 2). SF-36 subscale means ranged from 58.80 (Vitality) to 82.75 (Social Functioning); means for subscales related to physical health were lower than those related to mental health.

The Two Factor Model for the SF-36 had a comparative fit index (CFI) = 0.95 and a standardized root mean residual (SRMR) = 0.03, indicating that the model met minimal fit criteria, but the root mean square error of approximation (RMSEA) = 0.10, and  $\chi^2(17) = 11,873.00$  ( $p < 0.001$ ), suggested a poor fitting model.<sup>38</sup> Despite this limited model fit, we were interested in determining whether the CDC Healthy Days items loaded onto the two latent factors derived from the SF-36. The CDC physically unhealthy days item measures a latent physical health factor similar to that of the relevant SF-36 subscales ( $\lambda = 0.64$ ), and the CDC mentally unhealthy days item measures a latent mental health factor similar to that of the relevant SF-36 subscales ( $\lambda = 0.60$ ) (Table 3). The overall fit for this model was poor (see Table 3), but similar to statistics from recently published models that estimated the SF-36 using latent factors.<sup>6,39</sup> The correlation between the physical and mental health factors was high ( $r = 0.83$ ).

The Three Factor Model resulted in a better fit,  $\chi^2(17) = 7428.35$  ( $p < 0.001$ ), CFI = 0.97, RMSEA = 0.08, SRMR = 0.03. Although model fit was less good with the CDC items added, CDC item loadings on the relevant factors were strong (Table 4). Like the Two Factor Model, the correlation between the physical and mental health factors was high ( $r = 0.79$ ), as were the correlations between the physical factor and the wellbeing factor ( $r = 0.92$ ) and between the mental health factor and the wellbeing factor ( $r = 0.84$ ).

In the multi-group CFA analyses, loadings for all items for the physical and the mental health factors were similar in both models for each between-group comparison (Tables 3 and 4). With respect to the latent factor means (data not shown), women reported significantly



better physical (the Two Factor Model:  $\alpha = 2.21, p < 0.001$ ) and mental health (the Two Factor Model:  $\alpha = 1.12, p < 0.001$ ) than men and better physical health (the Three Factor Model:  $\alpha = 2.40, p < 0.01$ ), mental health (the Three Factor Model:  $\alpha = 0.16, p < 0.001$ ), and wellbeing (the Three Factor Model:  $\alpha = 0.10, p < 0.001$ ). Compared to those without ADL limitations, those with ADL limitations had significantly worse physical health (the Two Factor Model:  $\alpha = 2.15, p < 0.001$ ; the Three Factor Model:  $\alpha = 2.33, p < 0.001$ ), mental health (the Two Factor Model:  $\alpha = 1.41, p < 0.001$ ; the Three Factor Model:  $\alpha = 1.53, p < 0.001$ ), and wellbeing (the Three Factor Model:  $\alpha = 1.53, p < 0.001$ ). Similarly, people with two or more chronic medical conditions had significantly worse HRQOL than those with no or one chronic medical condition (the Two Factor Model: physical health,  $\alpha = 0.94, p < 0.001$ , mental health,  $\alpha = 0.55, p < 0.001$ ; the Three Factor Model: physical health,  $\alpha = 0.91, p < 0.001$ , mental health,  $\alpha = 0.55, p < 0.001$ , and wellbeing,  $\alpha = 0.84, p < 0.001$ ).

The CDC mentally unhealthy days item was closely associated with the four SF-36 subscales that make up its mental health component. The CDC physically unhealthy days item was closely associated with all of the SF-36 physical health subscales except the physical functioning subscale (Table 5).

## Discussion

Regardless of the particular factor structure used for the SF-36, the CDC items load well on the appropriate factors, indicating that the CDC items measure similar physical and mental health constructs as the SF-36. The CDC physically unhealthy days item is strongly associated with SF-36 measures of bodily pain, physical role limitations, and general health, while the CDC mentally unhealthy days measure is strongly associated with SF-36 measures of mental health, emotional role limitations, and social functioning. The exception was the SF-36 physical functioning subscale, which was not associated with either CDC item.

We speculate that the exception regarding physical functioning may have occurred because our sample of older adults did not consider physical functioning as particularly important when evaluating their overall physical health. Rather, the capacity to fulfill broader physical roles (with or without assistance) and to be generally healthy and pain free may be much more salient than the ability to perform specific physical functions. This finding sheds important light on the ongoing debate about the relevance of measuring functional abilities when assessing HRQOL. Given that physical functioning may not be as pertinent for older adults, the CDC items may be a more efficient way of assessing physical and mental health status for this population. This study thus supports other recent research that drew similar conclusion regarding use of the CDC items for individuals with limitations in functioning.<sup>21</sup> The CDC items appear to assess overall physical and mental health in a way similar to the SF-36, but more concisely and without including questions about ambulation and other physical functions that may or may not be directly tied to HRQOL.

Understanding the associations between different HRQOL instruments is becoming a higher priority for government agencies so that they can interpret the potential benefits of different health policies from studies that use either of these instruments. The SF-36 has been used to

assess the efficacy of interventions in clinical trials and in comparative studies and to develop normative data in representative samples of different countries.

The CDC Healthy Days items have been proposed as measures of “General Health Status,” a Foundation Health Measure in Healthy People 2020 (<http://www.healthypeople.gov/2020/about/GenHealthAbout.aspx#physically>). Our findings reinforce the utility of these items for monitoring population health and support the continued use of the CDC Healthy Days items on the Behavioral Risk Factor Surveillance System and the CMS MHOS Survey.

Furthermore, the CDC items warrant consideration for inclusion on other surveys that need only a few items to assess global physical and mental HRQOL constructs in a manner that is independent of physical functioning.

### Study limitations

The CDC Healthy Days items have limitations previously described, including digit preferences, ceiling effects, and floor effects.<sup>28</sup> Additionally, these measures focus on perceptions of recent health, within the past 30 days, so that adaptation to chronic conditions including response shifts may make them less sensitive to change. By focusing on “not good” or unhealthy days to improve recall over the past 30 days, they also do not cover positive aspects of health that other types of items can better assess. Because the current investigation included only adults at least 65 years old, most of whom were white woman, these findings may not generalize to other populations, including Medicare beneficiaries who do not choose the MA option.

### Conclusions

This study used a large sample of mostly white, older adults to determine the level of association between two prominent measures of health-related quality of life. Using latent modeling, this study found that the SF-36 and the CDC Healthy Days measures assess similar physical and mental HRQOL constructs, except that the Healthy Days measures do not correlate with SF-36 indices measuring physical functioning. These findings have important implications for monitoring individual and population health, particularly for older adults and those with limitations in physical functioning. For example, these findings support the consideration of the CDC Healthy Days measures in epidemiological and prevention research studies that need a brief set of items to efficiently assess global physical and mental health independent of function.

### References

1. U.S. Department of Health and Human Services. Health-related Quality of Life and Well-being. Healthy People 2020-Improving the Health of Americans. 2014. <http://www.healthypeople.gov/2020/topicsobjectives2020/overview.aspx?topicid=19>
2. Barile JP, Reeve BB, Smith AW, et al. Monitoring population health for healthy people 2020: evaluation of the NIH PROMIS® global health, CDC Healthy Days, and satisfaction with life instruments. *Qual Life Res.* 2013; 22:1201–1211. [PubMed: 23404737]
3. Newschaffer, CJ. Validation of Behavioral Risk Factor Surveillance System (BRFSS) HRQOL Measures in a Statewide Sample. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion; 1998.



4. Ware JE Jr, Kosinski M, Gandek B, et al. The factor structure of the SF-36 Health Survey in 10 countries: results from the IQOLA Project. *J Clin Epidemiol.* 1998; 51:1159–1165. [PubMed: 9817133]
5. De Vet HC, Adèr HJ, Terwee CB, Pouwer F. Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36. *Qual Life Res.* 2005; 14:1203–1218. [PubMed: 16047498]
6. Keller SD, Ware JE Jr, Bentler PM, et al. Use of structural equation modeling to test the construct validity of the SF-36 health survey in ten countries: results from the IQOLA project. *J Clin Epidemiol.* 1998; 51:1179–1188. [PubMed: 9817136]
7. World Health Organization. *International Classification of Functioning, Disability and Health: ICF.* World Health Organization; Geneva: 2001.
8. Krahn G, Fujiura G, Drum CE, et al. The dilemma of measuring perceived health status in the context of disability. *Disabil Health J.* 2009; 2:49–56. [PubMed: 21122743]
9. Johnson RJ, Wolinsky FD. The structure of health status among older adults: disease, disability, functional limitation, and perceived health. *J Health Soc Behav.* 1993; 34:105–121. [PubMed: 8277124]
10. Hays RD, Hahn H, Marshall G. Use of the SF-36 and other health-related quality of life measures to assess persons with disabilities. *Arch Phys Med Rehabil.* 2002; 83:s4–s9. [PubMed: 12474166]
11. Schwartz CE, Andresen EM, Nosek MA, Krahn GL. Response shift theory: important implications for measuring quality of life in people with disability. *Arch Phys Med Rehabil.* 2007; 88:529–536. [PubMed: 17398257]
12. Hall T, Krahn GL, Horner-Johnson W, Lamb G. Examining functional content in widely used health-related quality of life scales. *Rehabil Psychol.* 2011; 56:94–99. [PubMed: 21574727]
13. Helmes E. Function and disability or quality of life? Issues illustrated by the Osteoporosis Functional Disability Questionnaire (OFDQ). *Qual Life Res.* 2000; 9:755–761.
14. Huguet N, Kaplan MS, Feeny D. Socioeconomic status and health-related quality of life among elderly people: results from the Joint Canada/United States Survey of Health. *Soc Sci Med.* 2008; 66:803–810. [PubMed: 18155337]
15. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care.* 1993;247–263. [PubMed: 8450681]
16. Ware JE, Kosinski M. Interpreting SF&-36 summary health measures: a response. *Qual Life Res.* 2001; 10:405–413. [PubMed: 11763203]
17. Andresen EM, Meyers AR. Health-related quality of life outcomes measures. *Arch Phys Med Rehabil.* 2000; 81:S30–S45. [PubMed: 11128902]
18. Krahn GL, Fox MH. Health Disparities of Adults with Intellectual Disabilities: what do we know? What do we do? *J Appl Res Intellect Disabil.* 2014; 27:431–446. [PubMed: 23913632]
19. Horner-Johnson W, Krahn GL, Suzuki R, et al. Differential performance of SF-36 items in healthy adults with and without functional limitations. *Arch Phys Med Rehabil.* 2010; 91:570–575. [PubMed: 20382289]
20. Andresen EM, Fouts BS, Romeis JC, Brownson CA. Performance of health-related quality-of-life instruments in a spinal cord injured population. *Arch Phys Med Rehabil.* 1999; 80:877–884. [PubMed: 10453762]
21. Horner-Johnson W, Krahn GL, Suzuki R, Roid G. Performance of nine health-related quality of life surveillance items among healthy adults with and without functional limitations. (unpublished manuscript).
22. 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. *Qual Life Res.* 2009; 18:873–880. [PubMed: 19543809]
23. Reeve BB, Hays RD, Bjorner JB, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care.* 2007; 45:S22–S31. [PubMed: 17443115]
24. Cubanski, J., Neuman, T., Stollo, MK., Damico, A., Gonzales, R. Examining Sources of Coverage Among Medicare Beneficiaries: Supplemental Insurance, Medicare Advantage, and Prescription

- Drug Coverage: Findings from the Medicare Current Beneficiary Survey, 2006. Menlo Park, CA: Henry J. Kaiser Family Foundation; 2008.
25. Landon Bruce E, Zaslavsky AM, Saunders RC, Pawlson LG, Newhouse JP, Ayanian JZ. Analysis of Medicare Advantage HMOs compared with traditional Medicare shows lower use of many services during 2003–09. *Health Aff.* 2012; 31:2609–2617.
  26. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Med Care.* 1992; 30:473–483. [PubMed: 1593914]
  27. Brazier J, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *Br Med J.* 1992; 305:160–164. [PubMed: 1285753]
  28. 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]
  29. Moriarty DG, Kobau R, Zack MM, Zahran HS. Tracking Healthy Days – a window on the health of older adults. *Prev Chronic Dis.* 2005; 2:A16.
  30. Centers for Disease Control and Prevention. *Measuring Healthy Days.* Atlanta, GA: 2000.
  31. Jiang Y, Hesser JE. Using item response theory to analyze the relationship between health-related quality of life and health risk factors. *Prev Chronic Dis.* 2009; 6:1–10.
  32. 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 Qual Life Outcomes.* 2006; 4:66–84. [PubMed: 16995955]
  33. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. *JAMA.* 1963; 185:914–919. [PubMed: 14044222]
  34. Miller DR, Rogers WH, Kazis LE, Spiro A III, Ren XS, Haffer SC. Patients’ self-report of diseases in the Medicare Health Outcomes Survey based on comparisons with linked survey and medical data from the Veterans Health Administration. *J Ambul Care Manage.* 2008; 31:161–177. [PubMed: 18360178]
  35. Vandenberg RJ. Toward a further understanding of and improvement in measurement invariance methods and procedures. *Org Res Methods.* 2002; 5:139–158.
  36. Meredith W. Measurement invariance, factor analysis and factorial invariance. *Psychometrika.* 1993; 58:525–543.
  37. Widaman, KF., Reise, SP. *Exploring the Measurement Invariance of Psychological Instruments: Applications in the Substance Use Domain.* 1997.
  38. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling.* 1999; 6:1–55.
  39. Reed PJ. Medical outcomes study short form 36: testing and cross-validating a second-order factorial structure for health system employees. *Health Serv Res.* 1998; 33:1361–1380. [PubMed: 9865224]

**Table 1**

## Demographic characteristics

Characteristic	Number	Percent
Age group		
65–69	14,735	22%
70–74	19,505	29%
75–79	15,958	24%
80 or older	16,070	24%
Annual household income		
Less than \$5000	1632	3%
\$5000–\$9999	4889	10%
\$10,000–\$19,999	16,676	32%
\$20,000–\$29,999	11,858	23%
\$30,000–\$39,999	6790	13%
\$40,000–\$49,999	3909	7%
\$50,000–\$79,999	4019	8%
\$80,000–\$99,999	963	2%
\$100,000 or more	994	2%
Education		
Less than high school diploma	17,179	27%
High school diploma or GED	24,916	39%
Some college	13,388	21%
College degree	9078	14%
Race		
White	58,263	88%
Black	4792	7%
Other race	3214	5%
Gender		
Men	26,158	40%
Women	40,111	61%

Note. GED = General Equivalency Diploma.

**Table 2**

Descriptive Statistics for CDC Healthy Days Measures and Subscales of the SF-36

	Mean	Standard deviation	Range
Number of unhealthy days			
Physically unhealthy	4.38	8.75	0–30
Mentally unhealthy	2.06	5.80	0–30
SF-36 subscales			
Physical functioning	65.53	27.29	0–100
Role limitations, physical	59.66	42.63	0–100
Bodily pain	63.83	25.03	0–100
General health	65.71	20.16	0–100
Vitality	58.80	21.12	0–100
Social functioning	82.75	23.16	0–100
Role limitations, emotional	79.07	35.72	0–100
Mental health	79.27	16.43	0–100

Note. SF-36 scores represent the beneficiary's transformed 0–100 scale score.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3**  
 Model 1: standardized factor loadings from a multi-group confirmatory factor analyses of physical health and mental health factors, by gender, number of limitations in activities of daily living (ADL), and number of medical conditions

Indicators	Fully constrained model			Gender model			ADLs			Multiple chronic conditions		
	$\lambda$	Men		Women	Men	No lim.		1 lim.	No or one condition	Two or more conditions		
		$n$	$\lambda$			$n$	$\lambda$			$n$	$\lambda$	
<i>N</i> = 66,269		<i>n</i> = 26,158	$\lambda$	<i>n</i> = 40,111	$\lambda$	<i>n</i> = 39,092	$\lambda$	<i>n</i> = 24,983	$\lambda$	<i>n</i> = 29,454	$\lambda$	<i>n</i> = 29,880
<b>Physical health factor</b>	<b><math>\lambda</math></b>	<b><math>\lambda</math></b>	<b><math>\lambda</math></b>	<b><math>\lambda</math></b>	<b><math>\lambda</math></b>	<b><math>\lambda</math></b>	<b><math>\lambda</math></b>	<b><math>\lambda</math></b>	<b><math>\lambda</math></b>	<b><math>\lambda</math></b>	<b><math>\lambda</math></b>	<b><math>\lambda</math></b>
SF-36 subscales												
Physical functioning	0.78	0.76	0.78	0.78	0.70	0.60	0.72	0.75				
Role limitations, physical	0.79	0.78	0.79	0.68	0.65	0.77	0.75					
Bodily Pain	0.77	0.74	0.78	0.64	0.63	0.72	0.75					
Vitality	0.65	0.81	0.56	0.43	0.58	0.52	0.64					
General health	0.66	0.79	0.59	0.50	0.58	0.53	0.67					
CDC physically unhealthy days	0.64	0.64	0.64	0.54	0.63	0.54	0.59					
<b>Mental health factor</b>												
SF-36 subscales												
Social functioning	0.85	0.86	0.85	0.81	0.78	0.79	0.84					
Role limitations, emotional	0.69	0.68	0.69	0.64	0.63	0.66	0.65					
Mental health	0.74	0.73	0.74	0.63	0.75	0.71	0.72					
Vitality	0.19	0.02	0.26	0.32	0.17	0.31	0.18					
General health	0.15	-0.01	0.22	0.17	0.16	0.20	0.11					
CDC mentally unhealthy days	0.60	0.59	0.60	0.51	0.63	0.54	0.57					

Note. Fully constrained model:  $\chi^2(32) = 21525.5$ ; RMSEA = 0.10; CFI = 0.92; SRMR = 0.05;  $\lambda$  = standardized factor loading.

All SF-36 subscales were reverse scored.

**Table 4** Model 2: standardized factor loadings from a multi-group confirmatory factor analyses of physical health, general wellbeing, and mental health factors, by gender, number of limitations in activities of daily living (ADL), and number of medical conditions

Indicators	Gender model				ADLs			Multiple chronic conditions		
	Fully constrained Model		Men	Women	No lim.	1 lim.	No or one condition	Two or more conditions		
	$\lambda$	$\lambda$	$\lambda$	$\lambda$	$\lambda$	$\lambda$	$\lambda$	$\lambda$	$\lambda$	$\lambda$
Physical health factor										
SF-36 subscales										
Physical functioning	0.78	0.77	0.80	0.69	0.61	0.73	0.76			
Role limitations, physical	0.80	0.79	0.80	0.65	0.65	0.77	0.76			
Bodily Pain	0.78	0.75	0.79	0.62	0.65	0.73	0.76			
CDC physically unhealthy days	0.65	0.64	0.63	0.54	0.64	0.56	0.61			
General well-being factor										
SF-36 subscales										
Vitality	0.83	0.84	0.82	0.77	0.76	0.78	0.81			
General health	0.82	0.83	0.81	0.76	0.75	0.79	0.80			
Mental health factor										
SF-36 subscales										
Social functioning	0.84	0.84	0.83	0.82	0.77	0.81	0.84			
Role limitations, emotional	0.68	0.66	0.68	0.61	0.63	0.67	0.66			
Mental health	0.74	0.73	0.75	0.62	0.77	0.72	0.74			
CDC mentally unhealthy days	0.61	0.57	0.61	0.52	0.65	0.56	0.60			

Note. Fully constrained model:  $\chi^2(32) = 17113.94$ ;  $p < 0.001$ ; CFI = 0.93; RMSEA = 0.09; SRMR = 0.05;  $\lambda$  = standardized factor loading. All SF-36 subscales were reverse scored.



**Table 5**

Associations between SF-36 health Subscales and CDC Healthy Days items

	Physically unhealthy day			Mentally unhealthy days				
	<i>b</i>	S.E.	<i>p</i>	I.R.R.	<i>b</i>	S.E.	<i>p</i>	I.R.R.
SF-36 physical health subscales								
Physical functioning	0.03	0.02	0.095	1.03	0.01	0.03	0.852	1.01
Role limitations, physical	0.28	0.01	<0.000	1.32	-0.08	0.02	0.001	0.93
Bodily Pain	0.46	0.02	<0.000	1.59	0.06	0.03	0.012	1.07
General health	0.40	0.02	<0.000	1.50	0.08	0.03	<0.0003	1.09
SF-36 mental health subscales								
Vitality	0.01	0.00	<0.000	1.01	0.01	0.00	<0.000	1.01
Social functioning	0.25	0.01	<0.000	1.29	0.30	0.02	<0.000	1.36
Role limitations, emotional	-0.09	0.01	<0.000	0.91	0.20	0.02	<0.000	1.22
Mental health	0.02	0.02	0.281	1.02	1.03	0.03	<0.000	2.79

Note. The SF-36 subscales were reverse scored so that incident rate ratios (IRR) greater than 1 indicate a positive association between the SF-36 subscale and unhealthy days and a value less than 1 indicates a negative association. The SF-36 was also divided by 10 in order to put it on a 1–10 scale. The CDC Healthy Days items served as the dependent variables. *b* = regression coefficient. S.E. = standard error the regression coefficient.