

Attachment 1: Annotated Bibliography of Peer Reviewed BRFSS Reliability and Validity Studies (2011-2016)

2016

- Jungquist, C. R., et al. (2016). "Validation of the Behavioral Risk Factor Surveillance System Sleep Questions." *Journal of Clinical Sleep Medicine* **12**(3): 301-310.
STUDY OBJECTIVES: Sleep problems may constitute a risk for health problems, including cardiovascular disease, depression, diabetes, poor work performance, and motor vehicle accidents. The primary purpose of this study was to assess the validity of the current Behavioral Risk Factor Surveillance System (BRFSS) sleep questions by establishing the sensitivity and specificity for detection of sleep/ wake disturbance. METHODS: Repeated cross-sectional assessment of 300 community dwelling adults over the age of 18 who did not wear CPAP or oxygen during sleep. Reliability and validity testing of the BRFSS sleep questions was performed comparing to BRFSS responses to data from home sleep study, actigraphy for 14 days, Insomnia Severity Index, Epworth Sleepiness Scale, and PROMIS-57. RESULTS: Only two of the five BRFSS sleep questions were found valid and reliable in determining total sleep time and excessive daytime sleepiness. CONCLUSIONS: Refinement of the BRFSS questions is recommended. Copyright © 2016 American Academy of Sleep Medicine.
- Ward, Z. J., et al. (2016). "Redrawing the US Obesity Landscape: Bias-Corrected Estimates of State-Specific Adult Obesity Prevalence." *PLoS ONE [Electronic Resource]* **11**(3): e0150735.
BACKGROUND: State-level estimates from the Centers for Disease Control and Prevention (CDC) underestimate the obesity epidemic because they use self-reported height and weight. We describe a novel bias-correction method and produce corrected state-level estimates of obesity and severe obesity. METHODS: Using non-parametric statistical matching, we adjusted self-reported data from the Behavioral Risk Factor Surveillance System (BRFSS) 2013 (n = 386,795) using measured data from the National Health and Nutrition Examination Survey (NHANES) (n = 16,924). We validated our national estimates against NHANES and estimated bias-corrected state-specific prevalence of obesity (BMI>30) and severe obesity (BMI>35). We compared these results with previous adjustment methods. RESULTS: Compared to NHANES, self-reported BRFSS data underestimated national prevalence of obesity by 16% (28.67% vs 34.01%), and severe obesity by 23% (11.03% vs 14.26%). Our method was not significantly different from NHANES for obesity or severe obesity, while previous methods underestimated both. Only four states had a corrected obesity prevalence below 30%, with four exceeding 40%-in contrast, most states were below 30% in CDC maps. CONCLUSIONS: Twelve million adults with obesity (including 6.7 million with severe obesity) were misclassified by CDC state-level estimates. Previous bias-correction methods also resulted in underestimates. Accurate state-level estimates are necessary to plan for resources to address the obesity epidemic.
- Yin, S., et al. (2016). "Summarizing health-related quality of life (HRQOL): development and testing of a one-factor model." *Population Health Metrics* **14**: 22.
BACKGROUND: Health-related quality of life (HRQOL) is a multi-dimensional concept commonly used to examine the impact of health status on quality of life. HRQOL is often measured by four core questions that asked about general health status and

number of unhealthy days in the Behavioral Risk Factor Surveillance System (BRFSS). Use of these measures individually, however, may not provide a cohesive picture of overall HRQOL. To address this concern, this study developed and tested a method for combining these four measures into a summary score. METHODS: Exploratory and confirmatory factor analyses were performed using BRFSS 2013 data to determine potential numerical relationships among the four HRQOL items. We also examined the stability of our proposed one-factor model over time by using BRFSS 2001-2010 and BRFSS 2011-2013 data sets. RESULTS: Both exploratory factor analysis and goodness of fit tests supported the notion that one summary factor could capture overall HRQOL. Confirmatory factor analysis indicated acceptable goodness of fit of this model. The predicted factor score showed good validity with all of the four HRQOL items. In addition, use of the one-factor model showed stability, with no changes being detected from 2001 to 2013. CONCLUSION: Instead of using four individual items to measure HRQOL, it is feasible to study overall HRQOL via factor analysis with one underlying construct. The resulting summary score of HRQOL may be used for health evaluation, subgroup comparison, trend monitoring, and risk factor identification.

2015

Al Kazzi, E. S., et al. (2015). "Differences in the Prevalence of Obesity, Smoking and Alcohol in the United States Nationwide Inpatient Sample and the Behavioral Risk Factor Surveillance System." PLoS ONE [Electronic Resource] **10**(11): e0140165. BACKGROUND: The lack of adequate and standardized recording of leading risk factors for morbidity and mortality in medical records have downstream effects on research based on administrative databases. The measurement of healthcare is increasingly based on risk-adjusted outcomes derived from coded comorbidities in these databases. However inaccurate or haphazard assessment of risk factors for morbidity and mortality in medical record codes can have tremendous implications for quality improvement and healthcare reform. OBJECTIVE: We aimed to compare the prevalence of obesity, overweight, tobacco use and alcohol abuse of a large administrative database with a direct data collection survey. MATERIALS AND METHODS: We used the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes for four leading risk factors in the United States Nationwide Inpatient Sample (NIS) to compare them with a direct survey in the Behavioral Risk Factor Surveillance System (BRFSS) in 2011. After confirming normality of the risk factors, we calculated the national and state estimates and Pearson's correlation coefficient for obesity, overweight, tobacco use and alcohol abuse between NIS and BRFSS. RESULTS: Compared with direct participant questioning in BRFSS, NIS reported substantially lower prevalence of obesity ($p < 0.01$), overweight ($p < 0.01$), and alcohol abuse ($p < 0.01$), but not tobacco use ($p = 0.18$). The correlation between NIS and BRFSS was 0.27 for obesity ($p = 0.06$), 0.09 for overweight ($p = 0.55$), 0.62 for tobacco use ($p < 0.01$) and 0.40 for alcohol abuse ($p < 0.01$). CONCLUSIONS: The prevalence of obesity, overweight, tobacco smoking and alcohol abuse based on codes is not consistent with prevalence based on direct questioning. The accuracy of these important measures of health and morbidity in databases is critical for healthcare reform policies.

Alcantara, I., et al. (2015). "Relative validation of fruit and vegetable intake and fat intake among overweight and obese African-American women." Public Health Nutrition

18(11): 1932-1940.

OBJECTIVE: To compare commonly used dietary screeners for fat intake and fruit and vegetable intake with 24 h dietary recalls among low-income, overweight and obese African-American women. **DESIGN:** Three telephone interviews were completed; measures included two 24 h dietary recalls (a weekday and weekend day) using the Nutrition Data System for Research software, the Behavioral Risk Factor Surveillance System's (BRFSS) Fruit and Vegetable Consumption Module and the National Cancer Institute's (NCI) Percentage Energy from Fat Screener. **SETTING:** Participants were recruited from three federally qualified health centers in south-west Georgia, USA. **SUBJECTS:** Participants (n 260) were African-American women ranging in age from 35 to 65 years. About half were unemployed (49.6%) and 58.7% had a high-school education or less. Most were obese (88.5%), with 39.6% reporting a BMI > 40.0 kg/m². **RESULTS:** Mean fruit and vegetable intake reported from the 24 h dietary recall was 2.66 servings/d compared with 2.79 servings/d with the BRFSS measure. The deattenuated Pearson correlation was 0.22, with notable variation by weight status, education level and age. Mean percentage of energy from fat was 35.5% as reported from the 24 h dietary recall, compared with 33.0% as measured by the NCI fat screener. The deattenuated Pearson correlation was 0.38, also with notable variation by weight status, education level and age. **CONCLUSIONS:** Validity of brief dietary intake measures may vary by demographic characteristics of the sample. Additional measurement work may be needed to accurately measure dietary intake in obese African-American women.

Conrad, Z., et al. (2015). "Comparability of dietary data collection programs for U.S. adults, 2007-2011." FASEB Journal. Conference: Experimental Biology **29**(1 Meeting Abstracts).

National public health goals include increasing daily fruit and vegetable (F&V) consumption, and continuous surveillance is critical for assessing progress towards this goal. The National Health and Nutrition Examination Survey (NHANES), the Behavioral Risk Factor Surveillance Survey (BRFSS), and the USDA Food Availability Database (USDA FAD) are routinely used to monitor F&V consumption of adults at the national level, but the discrepancy between the data generated from these programs has not been investigated. The objectives of this study are to 1) estimate the daily per capita consumption of F&V from each of these programs, 2) compare estimates of consumption of individual F&V by processing form using statistical agreement tests, and 3) compare estimates of consumption of F&V groups using descriptive means comparison. Data were collected for 2007-2011, and adjusted to reflect F&V consumption among individuals 18 years of age, measured in mean daily per capita cup-equivalents. The Bland-Altman method and ordinary least products regression indicated proportional bias but not fixed bias. Descriptive means comparison revealed similar consumption estimates between some programs for certain F&V groups, but these similarities were not consistent, and overall variability was apparent. These results indicate that these dietary data collection programs do not yield similar results for daily per capita F&V consumption among adults in the U.S. Caution should be exercised when reporting progress toward achieving national public health nutrition goals related to F&V consumption. The authors report no source of funding for this study.

Courtemanche, C., et al. (2015). "Modeling Area-Level Health Rankings." Health Services Research **50**(5): 1413-1431.

OBJECTIVE: Rank county health using a Bayesian factor analysis model. **DATA**

SOURCES: Secondary county data from the National Center for Health Statistics (through 2007) and Behavioral Risk Factor Surveillance System (through 2009). STUDY DESIGN: Our model builds on the existing county health rankings (CHRs) by using data-derived weights to compute ranks from mortality and morbidity variables, and by quantifying uncertainty based on population, spatial correlation, and missing data. We apply our model to Wisconsin, which has comprehensive data, and Texas, which has substantial missing information. DATA COLLECTION METHODS: The data were downloaded from www.countyhealthrankings.org. PRINCIPAL FINDINGS: Our estimated rankings are more similar to the CHRs for Wisconsin than Texas, as the data-derived factor weights are closer to the assigned weights for Wisconsin. The correlations between the CHRs and our ranks are 0.89 for Wisconsin and 0.65 for Texas. Uncertainty is especially severe for Texas given the state's substantial missing data. CONCLUSIONS: The reliability of comprehensive CHRs varies from state to state. We advise focusing on the counties that remain among the least healthy after incorporating alternate weighting methods and accounting for uncertainty. Our results also highlight the need for broader geographic coverage in health data. Copyright © Health Research and Educational Trust.

Morales-Campos, D. Y., et al. (2015). "Correlates of HPV vaccination intention of Hispanic mothers of 11-17-year-old daughters living along the Texas-Mexico border." Cancer Epidemiology Biomarkers and Prevention. Conference: 7th AACR Conference on the Science of Health Disparities in Racial/Ethnic Minorities and the Medically Underserved San Antonio, TX United States. Conference Start **24**(10 SUPPL. 1). Background: Cervical cancer incidence and mortality are higher for Hispanic women along the Texas-Mexico border than for other female population groups. Incidence could be reduced if teenaged Hispanic girls received the HPV vaccine before they became sexually active. However, few Hispanic girls compared to U.S. girls receive all three HPV vaccine doses (31% vs 36%), which prevent cervical cancer. Parents are crucial to the success of HPV vaccine uptake efforts. The purpose of this study is to examine correlates of mothers' intention to vaccinate their 11-17 year old daughters. Methods: We utilized baseline data from an outreach and education program utilizing promotoras and peer educators to deliver health education to mothers and daughters to increase HPV knowledge and promote HPV immunization. Our analyses utilized data from mothers of never vaccinated girls (n=371). First, we checked the validity of perceived susceptibility, perceived severity, perceived benefits and barriers, self-efficacy, attitudes toward HPV vaccination, and subjective norms scales using Cronbach's alpha. Second, we used Pearson's correlation coefficient to examine if perceived susceptibility, perceived severity, perceived benefits and barriers, self-efficacy, attitudes toward HPV vaccination, and subjective norms (constructs) were correlated with mother's vaccine intention. Results: The majority of mothers were married (84%), spoke only Spanish (89%), had middle-school education (64%), and a household income of less than \$10,000 (82%). Analyses revealed that the scales for susceptibility, severity, and subjective norms were not significantly correlated to intention and had low internal consistency [except for the susceptibility scale (alpha=0.72)]. Only the self-efficacy scale showed high internal consistency (alpha=0.92) and was significantly correlated to intention (r=.346, p=<0.01), as was the perceived benefits and barriers scale (r=.159, p=<0.01). Conclusions: Although we selected English-language survey items from published literature and national surveillance tools (i.e., HINTS, BRFSS) and translated them to Spanish, preliminary findings suggest that most of the scales did not adequately measure the constructs in Spanish-speaking populations. Only the self-efficacy scale, which was previously

validated among Spanish-speaking Hispanics, had high internal consistency. The perceived susceptibility scale also showed good reliability. These findings inform future research on HPV vaccination by providing insight into which constructs are correlated with vaccine intention and which scales have good internal consistency among Spanish-speaking populations.

Phillips, B. R., et al. (2015). "Surveyed Enrollees in Veterans Affairs Health Care: How They Differ From Eligible Veterans Surveyed by BRFSS." *Military Medicine* **180**(11): 1161-1169.

OBJECTIVES: We described differences in demographic and socioeconomic characteristics between Veterans enrolled in the Veterans Health Administration (VHA) and Veterans eligible to enroll for Veterans Affairs health care. Knowledge of these differences is important in planning better services for Veterans who enroll and in encouraging additional enrollment. **METHODS:** We compared characteristics of enrollees and eligible Veterans in 2012. To describe enrollees, we used aggregate data from administrative records and results from VHA's Survey of Veteran Enrollees' Health and Reliance Upon VA. To describe eligible Veterans, we analyzed individual-level data from the Behavioral Risk Factor Surveillance System. **RESULTS:** Elderly individuals are more heavily represented among enrollees than eligible Veterans, and elderly enrollees are less likely to describe their health as good to excellent. Enrollees are more than twice as likely as eligible Veterans to have annual household incomes below \$16,000. Representation of minorities is roughly the same among enrollees as eligible Veterans. **CONCLUSIONS:** Our results are consistent with VHA as a safety net provider with respect to income, age, and disease burden. Reprint & Copyright © 2015 Association of Military Surgeons of the U.S.

2014

Alcantara, I., et al. (2014). "Relative validation of fruit and vegetable intake and fat intake among overweight and obese African-American women." *Public Health Nutrition* **18**(11): 1932-1940.

Objective: To compare commonly used dietary screeners for fat intake and fruit and vegetable intake with 24 h dietary recalls among low-income, overweight and obese African-American women. **Design:** Three telephone interviews were completed; measures included two 24 h dietary recalls (a weekday and weekend day) using the Nutrition Data System for Research software, the Behavioral Risk Factor Surveillance System's (BRFSS) Fruit and Vegetable Consumption Module and the National Cancer Institute's (NCI) Percentage Energy from Fat Screener. **Setting:** Participants were recruited from three federally qualified health centers in south-west Georgia, USA. **Subjects:** Participants (n 260) were African-American women ranging in age from 35 to 65 years. About half were unemployed (49.6%) and 58.7% had a high-school education or less. Most were obese (88.5%), with 39.6% reporting a BMI \geq 40.0 kg/m². **Results:** Mean fruit and vegetable intake reported from the 24 h dietary recall was 2.66 servings/d compared with 2.79 servings/d with the BRFSS measure. The deattenuated Pearson correlation was 0.22, with notable variation by weight status, education level and age. Mean percentage of energy from fat was 35.5% as reported from the 24 h dietary recall, compared with 33.0% as measured by the NCI fat screener. The deattenuated Pearson correlation was 0.38, also with notable variation by weight status, education level and age. **Conclusions:** Validity of brief dietary intake measures may vary by demographic characteristics of the sample. Additional measurement work may be needed to accurately measure dietary intake in obese

Burger, A. E. and E. N. Reither (2014). "Monitoring receipt of seasonal influenza vaccines with BRFSS and NHIS data: Challenges and solutions." *Vaccine* **32**(31): 3950-3954. Despite the availability of vaccines that mitigate the health risks associated with seasonal influenza, most individuals in the U.S. remain unvaccinated. Monitoring vaccination uptake for seasonal influenza, especially among disadvantaged or high-risk groups, is therefore an important public health activity. The Behavioral Risk Factor Surveillance System (BRFSS) - the largest telephone-based health surveillance system in the world - is an important resource in monitoring population health trends, including influenza vaccination. However, due to limitations in the question that measures influenza vaccination status, difficulties arise in estimating seasonal vaccination rates. Although researchers have proposed various methodologies to address this issue, no systematic review of these methodologies exists. By subjecting these methods to tests of sensitivity and specificity, we identify their strengths and weaknesses and advance a new method for estimating national and state-level vaccination rates with BRFSS data. To ensure that our findings are not anomalous to the BRFSS, we also analyze data from the National Health Interview Survey (NHIS). For both studies, we find that restricting the sample to interviews conducted between January and September offers the best balance of sensitivity (>90% on average), specificity (>90% on average), and statistical power (retention of 92.2% of vaccinations from the target flu season) over other proposed methods. We conclude that including survey participants from these months provides a simple and effective way to estimate seasonal influenza vaccination rates with BRFSS and NHIS data, and we discuss potential ways to better estimate vaccination rates in future epidemiologic surveys. © 2014 Elsevier Ltd.

Cheung, F. and R. E. Lucas (2014). "Assessing the validity of single-item life satisfaction measures: results from three large samples." *Quality of Life Research* **23**(10): 2809-2818.

PURPOSE: The present paper assessed the validity of single-item life satisfaction measures by comparing single-item measures to the Satisfaction with Life Scale (SWLS)-a more psychometrically established measure. METHODS: Two large samples from Washington (N = 13,064) and Oregon (N = 2,277) recruited by the Behavioral Risk Factor Surveillance System and a representative German sample (N = 1,312) recruited by the Germany Socio-Economic Panel were included in the present analyses. Single-item life satisfaction measures and the SWLS were correlated with theoretically relevant variables, such as demographics, subjective health, domain satisfaction, and affect. The correlations between the two life satisfaction measures and these variables were examined to assess the construct validity of single-item life satisfaction measures. RESULTS: Consistent across three samples, single-item life satisfaction measures demonstrated substantial degree of criterion validity with the SWLS (zero-order $r = 0.62-0.64$; disattenuated $r = 0.78-0.80$). Patterns of statistical significance for correlations with theoretically relevant variables were the same across single-item measures and the SWLS. Single-item measures did not produce systematically different correlations compared to the SWLS (average difference = 0.001-0.005). The average absolute difference in the magnitudes of the correlations produced by single-item measures and the SWLS was very small (average absolute difference = 0.015-0.042). CONCLUSIONS: Single-item life satisfaction measures performed very similarly compared to the multiple-item SWLS. Social scientists would get virtually identical answer to substantive questions regardless of which measure

they use.

Gundersen, D. A., et al. (2014). "Assessing the feasibility and sample quality of a national random-digit dialing cellular phone survey of young adults." *American Journal of Epidemiology* **179**(1): 39-47.

The majority of adults aged 18-34 years have only cellular phones, making random-digit dialing of landline telephones an obsolete methodology for surveillance of this population. However, 95% of this group has cellular phones. This article reports on the 2011 National Young Adult Health Survey (NYAHS), a pilot study conducted in the 50 US states and Washington, DC, that used random-digit dialing of cellular phones and benchmarked this methodology against that of the 2011 Behavioral Risk Factor Surveillance System (BRFSS). Comparisons of the demographic distributions of subjects in the NYAHS and BRFSS (aged 18-34 years) with US Census data revealed adequate reach for all demographic subgroups. After adjustment for design factors, the mean absolute deviations across demographic groups were 3 percentage points for the NYAHS and 2.8 percentage points for the BRFSS, nationally, and were comparable for each census region. Two-sided z tests comparing cigarette smoking prevalence revealed no significant differences between NYAHS and BRFSS participants overall or by subgroups. The design effects of the sampling weight were 2.09 for the NYAHS and 3.26 for the BRFSS. Response rates for the NYAHS and BRFSS cellular phone sampling frames were comparable. Our assessment of the NYAHS methodology found that random-digit dialing of cellular phones is a feasible methodology for surveillance of young adults. © 2013 The Author.

Le, A., et al. (2014). "The geographic distribution of obesity in the US and the potential regional differences in misreporting of obesity." *Obesity* 22(1): 300-306.

OBJECTIVE: State-level estimates of obesity based on self-reported height and weight suggest a geographic pattern of greater obesity in the Southeastern US; however, the reliability of the ranking among these estimates assumes errors in self-reporting of height and weight are unrelated to geographic region. DESIGN AND METHODS: Regional and state-level prevalence of obesity (body mass index > 30 kg m(-2)) for non-Hispanic black and white participants aged 45 and over were estimated from multiple sources: self-reported from the behavioral risk factor surveillance system (BRFSS 2003-2006) (n = 677,425), self-reported and direct measures from the National Health and Nutrition Examination Study (NHANES 2003-2008) (n = 6,615 and 6,138, respectively), and direct measures from the REasons for Geographic and Racial Differences in Stroke (REGARDS 2003-2007) study (n = 30,239). RESULTS: Data from BRFSS suggest that the highest prevalence of obesity is in the East South Central Census division; however, direct measures suggest higher prevalence in the West North Central and East North Central Census divisions. The regions relative ranking of obesity prevalence differs substantially between self-reported and directly measured height and weight. CONCLUSIONS: Geographic patterns in the prevalence of obesity based on self-reported height and weight may be misleading, and have implications for current policy proposals. Copyright © 2013 The Obesity Society.

McElligott, K., et al. (2014). "Assessment of cardiovascular disease risk factors in the coastal region of South Carolina." *Ethnicity & Disease* **24**(2): 155-161.

OBJECTIVE: To assess risk factors for cardiovascular disease, barriers to health care, and desired health care education topics for Hispanics in the coastal region of South Carolina known as the Low country. METHODS: 174 Hispanic adults were surveyed at visits at the Mexican consulate using a novel interview instrument. The prevalence of

cardiovascular risk factors was compared to the Behavioral Risk Factor Surveillance System (BRFSS), an annual telephone survey, to evaluate the validity of the survey instrument. RESULTS: Results are comparable to the BRFSS telephone study of the Hispanics in the same area. However, participants in our study were older (Age > 35 = 41.4% vs. 34.9%) and reported fewer years of formal education (higher level education = 12.9% vs. 44.2%). Cost of care (72.8%) and language barriers (46.8%) were the main difficulties reported in obtaining health care access. The main educational topics of interest were diabetes (61.5%), hypertension (43.7%), stress (42.5%), and cardiac disease (40.2%). CONCLUSION: Our study supports the evidence that there is a demand and need for cardiovascular disease and diabetes education among Hispanics. Our study also shows that a large proportion of Hispanics experience barriers to health care and that large telephone studies may underrepresent higher risk Hispanic populations.

Short, V. L. and V. Mendy (2014). "A comparison of examination-based and self-reported cardiovascular health indicators using Mississippi delta cardiovascular health examination survey preliminary data and the Mississippi behavioral risk factor surveillance system." Circulation. Conference: American Heart Association's Epidemiology and Prevention/Nutrition, Physical Activity, and Metabolism **129**(no pagination).

Background: Cardiovascular disease (CVD) surveillance at the regional or local level is limited. The Behavioral Risk Factor Surveillance System (BRFSS) frequently provides the primary source of information on CVD and its risk factors at the population level. However, BRFSS data may be limited due to reliability of self-reported information, a lack of serologic and other objective measurements, and non-local level sampling frames. The Mississippi Delta Cardiovascular Health Examination Survey (CHES) is an ongoing CVD surveillance system in the 18-county Mississippi Delta, a predominately rural, disadvantaged region with some of the highest rates of CVD in the nation. Our objective was to compare prevalence estimates of self-reported CVD risk factors from the Mississippi BRFSS with examination-based measures of these risk factors from Delta CHES. Methods: Delta CHES uses an in-home data collection model to collect survey data, anthropometric measures and fasting blood specimens on a representative sample of adults > 18 years of age living in the Mississippi Delta region. Preliminary examination data, collected between October 2012 and October 2013, from 484 Delta CHES participants were compared to weighted 2011 Mississippi BRFSS self-reported data from 1,187 respondents living in the Mississippi Delta region. Prevalence estimates and 95% confidence intervals (CI) were calculated for measures related to obesity, hypertension, dyslipidemia, diabetes, and smoking. In Delta CHES, dyslipidemia was defined as abnormal values for any cholesterol component or for triglycerides; hypertension was defined as systolic blood pressure > 140 mm Hg or diastolic blood pressure > 90 mm Hg; obesity was defined as body mass index of > 30.0 kg/m²; current smoking was defined as serum cotinine level > 10.0 ng/mL; and diabetes was defined as plasma glucose > 126 mg/dl or hemoglobin A1c > 6.5%. For prevalence estimates, non-overlapping 95% CIs indicate statistical significance (at alpha=0.05). Results: Delta CHES provided a significantly greater prevalence estimate of obesity [53.5% (95% CI 48.9-58.1%) vs. 42.9% (95% CI 39.0-46.9%)], diabetes [24.4% (95% CI 20.3-28.5%) vs. 16.3% (95% CI 13.7-18.9%)] and smoking [30.4% (95% CI 26.0-34.9%) vs. 20.2% (95% CI 16.7-23.6%)] compared to BRFSS. Conversely, the prevalence of hypertension was significantly lower in Delta CHES compared to BRFSS [14.4% (95% CI 11.2-17.6%) vs. 51.8% (95% CI 47.9-55.8%)]. Conclusion: Objectively measured

estimates from Delta CHES and self-reported measures from the Mississippi BRFSS differed for key CVD risk factors. These preliminary findings suggest self-reporting may have distorted past estimates of CVD risk factor rates in the Mississippi Delta region. Population-based CVD surveillance systems using examination-based measurements can provide perspective on routinely collected self-reports.

Zhang, X., et al. (2014). "Multilevel regression and poststratification for small-area estimation of population health outcomes: A case study of chronic obstructive pulmonary disease prevalence using the behavioral risk factor surveillance system." *American Journal of Epidemiology* **179**(8): 1025-1033.

A variety of small-area statistical models have been developed for health surveys, but none are sufficiently flexible to generate small-area estimates (SAEs) to meet data needs at different geographic levels. We developed a multilevel logistic model with both state- and nested county-level random effects for chronic obstructive pulmonary disease (COPD) using 2011 data from the Behavioral Risk Factor Surveillance System. We applied poststratification with the (decennial) US Census 2010 counts of census-block population to generate census-block-level SAEs of COPD prevalence which could be conveniently aggregated to all other census geographic units, such as census tracts, counties, and congressional districts. The model-based SAEs and direct survey estimates of COPD prevalence were quite consistent at both the county and state levels. The Pearson correlation coefficient was 0.99 at the state level and ranged from 0.88 to 0.95 at the county level. Our extended multilevel regression modeling and poststratification approach could be adapted for other geocoded national health surveys to generate reliable SAEs for population health outcomes at all administrative and legislative geographic levels of interest in a scalable framework. © 2014 Published by Oxford University Press.

2013

Ansarifar, A., et al. (2013). "Validity & reliability assessment of Persian version of core section of the behavioral risk factor surveillance system (BRFSS) questionnaire." *Iranian Journal of Epidemiology* **9**(1): 1-10.

Background and Objectives: The purpose of this study was to assess the reliability and validity of Persian version of survey instrument for Behavioral Risk Factor Surveillance System (BRFS). **Methods:** Content and Face validity of Behavioral Risk Factor Surveillance System (BRFS) questionnaire were examined with Backward-Forward method. This method consisted of four steps: translation, back-translation, expert review and Pilot study. For evaluation of reliability of questionnaire, 194 adults of 18 years or older from Karaj province were interviewed. For the examination of reliability, Cronbach's alpha was calculated to assess the internal consistency of the questionnaire. The reliability of quantitative variables evaluated with intra class correlation (ICC); ordinal variables with weighted Kappa and nominal variables with Kappa and weighted Kappa were calculated as well. **Results:** Average age of participants were 38.8 ± 15 . Majority of participants (66.5%) were in 18-44 years group, 26.8% in 45-64 years and others 65 years and older. In validity procedure, 26 questions were deleted, 3 questions added and 6 questions were modified. Questions of 'having situations of HIV morbidity' and 'Do you currently use drug abuse' were least reliability and deleted from questionnaire. Others were reliable. (It might be better to say how many questions remained finally) **Conclusion:** Our results indicate that BRFSS Persian version questionnaire has acceptable reliability and criterion validity for surveillance system in Iran.

Fulton, J. E., et al. (2013). "Validity of the aerobic physical activity questions in the behavioral risk factor surveillance system." Circulation. Conference: American Heart Association's Epidemiology and Prevention/Physical Activity, Nutrition and Metabolism **127**(12 Meeting Abstracts).

OBJECTIVE: Physical activity provides health benefits for cardiovascular disease and its risk factors. Surveillance of physical activity in U.S. adults informs national health objectives via self-report questionnaires, although the 2011 lack information on validity. This study examines the criterion validity of the aerobic physical activity questions used in the Behavioral Risk Factor Surveillance System (BRFSS). **METHODS:** Participants answered the current BRFSS physical activity questions (8 questions total) on the frequency and duration of participation in their two most frequent aerobic physical activities during an in-person interview. Responses to the questions were compared to two validation standards: (1) non-occupational physical activity and exercise abstracted from diaries; and (2) physical activity derived from accelerometers (CSA model 7164 WAM). All measures were summarized as moderate- to vigorous-intensity physical activity (minutes/week). Criterion validity was assessed by comparing questionnaire responses to each validation standard (diary and accelerometer) using the Spearman correlation, mean difference in activity estimates, Kappa coefficient, and percent positive agreement. **RESULTS:** Participants in the study were 107 African- American and 122 Hispanic women 40 - 70 years of age with a mean age of 49.3 +/- 7.1 years and a mean BMI of 30.0 +/- 6.5 kg/m². Most (85.9%) women completed high-school. Associations between self-reported physical activity, diary reports, and accelerometer counts are shown in the Table. **CONCLUSION:** In this sample, based on Spearman correlations >0.30 and fair to moderate Kappa coefficients, the BRFSS aerobic physical activity questions show acceptable validity. Additional studies may be warranted to validate the current BRFSS physical activity questions in samples representative of the age, sex, and racial/ethnic distribution of U.S. adults. (Table Presented).

Kamimoto, L., et al. (2013). "Seasonal influenza morbidity estimates obtained from telephone surveys, 2007." *American Journal of Public Health* 103(4): 755-763.

OBJECTIVES: We assessed telephone surveys as a novel surveillance method, comparing data obtained by telephone with existing national influenza surveillance systems, and evaluated the utility of telephone surveys.

METHODS: We used the 2007 Behavioral Risk Factor Surveillance System (BRFSS) and the 2007 National Immunization Survey-Adult (NIS-Adult) to estimate the incidence of influenza-like illness (ILI), medically attended ILI, provider-diagnosed influenza, influenza testing, and treatment of influenza with antiviral medications during the 2006-2007 influenza season. **RESULTS:** With the January-May BRFSS, among persons aged 18 years and older, the cumulative incidence of seasonal ILI and provider-diagnosed influenza was 37.9 and 5.7 adults per 100 persons, respectively. Monthly medically attended ILI and provider-diagnosed influenza among adults were temporally associated with influenza activity, as documented by national surveillance. With the NIS-Adult survey data, estimated provider-diagnosed influenza, influenza testing, and antiviral treatment were 2.8%, 1.4%, and 0.6%, respectively. **CONCLUSIONS:** Our telephone interview-based estimates of influenza morbidity were consistent with those from national influenza surveillance systems. Telephone surveys may provide an alternative method by which population-based influenza morbidity information can be gathered.

Kim, H. J. and K. I. Fredriksen-Goldsen (2013). "Nonresponse to a question on self-identified sexual orientation in a public health survey and its relationship to race and ethnicity." *American Journal of Public Health* 103(1): 67-69.

We examined whether nonresponse to the survey question on self-identified sexual orientation was associated with race and ethnicity, utilizing Washington State Behavioral Risk Factor Surveillance System data. The results of adjusted multinomial logistic regression indicated that the nonresponse rates of Asian Americans, Hispanics, and African Americans are higher than those of non-Hispanic Whites. Innovative ways of measuring sexual orientation to reduce racially and ethnically driven bias need to be developed and integrated into public health surveys.

Lee, C.-J. and D. Kim (2013). "A comparative analysis of the validity of US state- and county-level social capital measures and their associations with population health." *Social Indicators Research* 111(1): 307-326.

The goals of this study were to validate a number of available collective social capital measures at the US state and county levels, and to examine the relative extent to which these social capital measures are associated with population health outcomes. Measures of social capital at the US state level included aggregate indices based on the Annenberg National Health Communication Survey and the Behavioral Risk Factor Surveillance System (BRFSS), Petris Social Capital Index (PSCI), Putnam's index, and Kim et al.'s scales. County-level measures consisted of PSCI, Rupasingha et al.'s social capital index, and a BRFSS-derived measure. These measures, except for the PSCI, showed evidence of acceptable validity. Moreover, we observed differences across the social capital measures in their associations with population health outcomes. The implications of the findings for future research in this area were discussed. (PsycINFO Database Record (c) 2013 APA, all rights reserved) (journal abstract).

Linder, J. A., et al. (2013). "Use of practice-based research network data to measure neighborhood smoking prevalence." *Preventing Chronic Disease* 10: E84.

INTRODUCTION: Practice-Based Research Networks (PBRNs) and health systems may provide timely, reliable data to guide the development and distribution of public health resources to promote healthy behaviors, such as quitting smoking. The objective of this study was to determine if PBRN data could be used to make neighborhood-level estimates of smoking prevalence. METHODS: We estimated the smoking prevalence in 32 greater Boston neighborhoods (population = 877,943 adults) by using the electronic health record data of adults who in 2009 visited one of 26 Partners Primary Care PBRN practices (n = 77,529). We compared PBRN-derived estimates to population-based estimates derived from 1999-2009 Behavioral Risk Factor Surveillance System (BRFSS) data (n = 20,475). RESULTS: The PBRN estimates of neighborhood smoking status ranged from 5% to 22% and averaged 11%. The 2009 neighborhood-level smoking prevalence estimates derived from the BRFSS ranged from 5% to 26% and averaged 13%. The difference in smoking prevalence between the PBRN and the BRFSS averaged -2 percentage points (standard deviation, 3 percentage points). CONCLUSION: Health behavior data collected during routine clinical care by PBRNs and health systems could supplement or be an alternative to using traditional sources of public health data.

MacLennan, P. A., et al. (2013). "Medical record validation of self-reported eye diseases and eye care utilization among older adults." *Current Eye Research* 38(1): 1-8.

PURPOSE: Vision impairment is an important public health concern. Accurate

information regarding visual health and eye care utilization is essential to monitor trends and inform health policy interventions aimed at addressing at-need populations. National surveys provide annual prevalence estimates but rely on self-report. The validity of self-reported information regarding eye disease has not been adequately explored. METHODS: This cross-sectional study compared self-report of eye care utilization and eye disease with information obtained from medical records. The study population was 2001 adults aged 70 years and older who completed the Behavioral Risk Factor Surveillance System's Visual Impairment and Access to Eye Care Module. Cohen's kappa (κ) was used to assess agreement. RESULTS: Agreement between self-report and medical records was substantial for eye care utilization ($\kappa = 0.64$) and glaucoma ($\kappa = 0.73$), moderate for macular degeneration ($\kappa = 0.40$) and diabetic retinopathy ($\kappa = 0.47$) and slight for cataracts ($\kappa = 0.18$). Self-report tended to overestimate the number of subjects who visited an eye care provider in the previous year, and underestimated the prevalence in all but one (glaucoma) of the four eye diseases evaluated. CONCLUSIONS: Though agreement was substantial for self-report of eye care utilization, results of the current study suggest that national estimates based on self-report overestimate eye care utilization.

Simon, P. A., et al. (2013). "Declines in sugar-sweetened beverage consumption among children in Los Angeles County, 2007 and 2011." Preventing Chronic Disease **10**(8). This study assessed changes in consumption of sugar-sweetened beverages (SSBs) among children (aged ≤ 17 years) in Los Angeles County. We analyzed children's data from the 2007 ($n = 5,595$) and 2011 ($n = 5,934$) Los Angeles County Health Survey. The percentage of children who consumed 1 or more SSB per day decreased from 43.3% in 2007 to 38.3% in 2011 ($P < .001$); this decrease was seen across most sociodemographic subgroups. Despite measurable progress in reducing SSB consumption among children in Los Angeles County, consumption remains high, highlighting the need for additional policy and programmatic interventions. Objective Given recent evidence of the link between consumption of sugar-sweetened beverages (SSBs) and obesity (1), reducing SSB consumption has become a focus of childhood obesity prevention efforts. For example, California legislation prohibiting the sale of most SSBs on school campuses became effective in 2007 (2). In Los Angeles County, this policy was accompanied by intensive education on SSBs in schools, preschools, childcare sites, and other community settings, and in the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), which serves more than 50% of children younger than 5 years in the county (Shannon Whaley, PhD, Director of Research and Evaluation, PHFE WIC, Los Angeles County, March 20, 2013). To assess the collective effect of these and other efforts, we examined data on SSB consumption among children in the county from the 2007 and 2011 Los Angeles County Health Survey (LACHS). Methods The LACHS is a periodically administered, random-digit-dial telephone survey of the noninstitutionalized county population and includes both adult and child (aged ≤ 17 years) components (3). We analyzed data from the child components of the 2 most recent surveys, conducted in 2007 and 2011 with 2 independent samples ($n = 5,595$ in 2007 and 5,934 in 2011). The study was approved by the Los Angeles County Department of Public Health's institutional review board. For each survey, 1 child was randomly selected from each sampled household, and data were reported by the parent or other adult in the household who was most knowledgeable about the child's health and well-being (henceforth referred to as "primary caretaker"). The surveys included identically worded questions on sociodemographic characteristics and SSB

consumption ("On an average day, about how many sodas or other sweetened drinks such as Gatorade, Red Bull, or Sunny Delight does [child] drink? Do not include diet sodas or sugar-free drinks. Please count a 12-ounce can, bottle, or glass as 1 drink."). The SSB consumption question was adapted from a question used in the 2005 New York City Community Health Survey. Interviews were conducted in English, Spanish, Cantonese, Mandarin, Vietnamese, and Korean. Among all eligible households contacted, interviews were completed for 40% in 2007 and 64% in 2011. There were 2 methodologic differences between the 2 surveys. First, only the 2011 survey included cellular telephones, which accounted for 4% of the interviews. Second, a more sophisticated raking procedure was adopted in weighting the survey in 2011 (4). These changes were made to maintain survey representativeness and validity and are similar to those made to the Behavioral Risk Factor Surveillance System in 2011 (5). Further details regarding the survey design and changes in weighting methods are reported elsewhere (3,4). The percentage of children who consumed 1 or more SSB per day was calculated for the total study group in each survey and by sex, age group, child's race/ethnicity, annual household income relative to the federal poverty level (FPL), and primary caretaker's level of education. Among Latinos, results were further stratified by the language of interview (English or Spanish). Differences in percentages were assessed for significance by using the χ^2 test, and 95% confidence intervals were calculated for each point estimate. All analyses were conducted using SAS (version 9.2, SAS Institute Inc., Cary, North Carolina) using child sample weights; significance was set using an α of .05. Results In both 2007 and 2011, the percentage of children who consumed 1 or more SSB per day was highest among children aged 12 to 17 years and lowest among those aged 5 years or younger (Table). The percentage of children who consumed 1 or more SSB per day was inversely related to household income and primary caretaker's education level. The percentage was higher among Latino and African American children than among white and Asian/Pacific Islander children. Among all children, the percentage who consumed 1 or more SSB per day decreased from 43.3% in 2007 to 38.3% in 2011 ($P < .001$). Similar declines were seen across sex and age groups and among Latino children in households interviewed in Spanish (Table). Declines were also seen among children whose primary caretaker had less than a high school education or a college or postgraduate degree. A decrease was also observed among children from households whose incomes were below 300% of the FPL ($P < .001$) but not among those from households whose incomes were at or above 300% of the FPL ($P = .61$).

2012

Crews, J. E., et al. (2012). "The variability of vision loss assessment in federally sponsored surveys: Seeking conceptual clarity and comparability." *American Journal of Ophthalmology* **154**(6 SUPPL.): S31-S44.e31.

Purpose: To review U.S. national population-based surveys to evaluate comparability and conceptual clarity of vision measures. Design: Perspective. Methods: The vision questions in 12 surveys were mapped to the World Health Organization's International Classification of Functioning, Disability and Health framework under the domains of condition, impairment, activity limitation, participation, and environment. Surveys examined include the National Health Interview Survey, the Behavioral Risk Factor Surveillance Survey, National Health and Nutrition Examination Survey, the Census, and the Visual Function Questionnaire. Results: Nearly 100 vision measures were identified in 12 surveys. These surveys provided no consistent measure of

vision or vision impairment. Survey questions asked about differing characteristics of vision-related disease, function, and social roles. A question related to ability to read newspaper print was the most commonly asked question in surveys. Conclusions: Limited comparability of data and lack of conceptual clarity in the population-based surveys resulted in an inability to consistently characterize the population of people experiencing vision impairment. Consequently, vision surveillance was limited.

Frankel, M. R., et al. (2012). "When data are not missing at random: implications for measuring health conditions in the Behavioral Risk Factor Surveillance System." *BMJ Open* 2(4).

OBJECTIVES: To examine the effect on estimated levels of health conditions produced from large-scale surveys, when either list-wise respondent deletion or standard demographic item-level imputation is employed. To assess the degree to which further bias reduction results from the inclusion of correlated ancillary variables in the item imputation process. **DESIGN:** Large cross-sectional (US level) household survey. **PARTICIPANTS:** 218 726 US adults (18 years and older) in the 2006 Behavioral Risk Factor Surveillance System Survey. This survey is the largest US telephone survey conducted by the Centers for Disease Control and Prevention. **PRIMARY AND SECONDARY OUTCOME MEASURES:** Estimated rates of severe depression among US adults. **RESULTS:** The use of list-wise respondent deletion and/or demographic imputation results in the underestimation of severe depression among adults in the USA. List-wise deletion produces underestimates of 9% (8.7% vs 9.5%). Demographic imputation produces underestimates of 7% (8.9% vs 9.5%). Both of these differences are significant at the 0.05 level. **CONCLUSION:** The use of list-wise deletion and/or demographic-only imputation may produce significant distortion in estimating national levels of certain health conditions.

Grabner, M. (2012). "BMI trends, socioeconomic status, and the choice of dataset." *Obesity Facts* 5(1): 112-126.

OBJECTIVE: This study is a descriptive investigation of trends in BMI in the USA over time, across race/ethnicity, gender, and socioeconomic status (SES) groups, and across different datasets. **METHODS:** The study analyzes micro-level data from three widely used cross-sectional US health datasets: the National Health and Nutrition Examination Survey (NHANES), the National Health Interview Survey (NHIS), and the Behavioral Risk Factor Surveillance System (BRFSS), from the 1970s to 2008. Consistent race/ethnicity and SES groups are constructed for all datasets. SES is measured by education and income. Focusing on adults aged 20-74 years, the study estimates BMI time trends, distributional shifts, and incremental associations (gradients) with SES. **RESULTS:** SES-BMI gradients are consistently larger for women than for men, differ across race/ethnicity groups, and are similar across datasets. Trends in mean BMI are comparable across White, Black and Hispanic males, while Hispanic females range between White and Black females. Self-reported BMI in the NHANES differs markedly from self-reports in the NHIS and BRFSS. **CONCLUSION:** The NHANES, NHIS, and BRFSS provide similar evidence regarding BMI trends over time and across race/ethnicity, gender, and SES groups. Racial disparities in BMI remain after adjusting for SES and should be studied further. Copyright © 2012 S. Karger GmbH, Freiburg.

Hall, J. P., et al. (2012). "Discrepancy among Behavioral Risk Factor Surveillance System, social security, and functional disability measurement." *Disability and Health Journal* 5(1): 60-63.

We examine here the construct validity of Behavioral Risk Factor Surveillance System (BRFSS) disability items using a sample of adults who met the more stringent Social Security Administration (SSA) definition of disability. We compare responses of this sample to the BRFSS questions to responses to a 7-part functional question. Finally, we discuss our findings, which suggest limitations in using the BRFSS for measuring disability prevalence, and implications for the new federal standards for the measurement of disability status. (PsycINFO Database Record (c) 2014 APA, all rights reserved).

Li, C., et al. (2012). "A comparison of prevalence estimates for selected health indicators and chronic diseases or conditions from the Behavioral Risk Factor Surveillance System, the National Health Interview Survey, and the National Health and Nutrition Examination Survey, 2007-2008." *Preventive Medicine* **54**(6): 381-387. OBJECTIVE: To compare the prevalence estimates of selected health indicators and chronic diseases or conditions among three national health surveys in the United States. METHODS: Data from adults aged 18 years or older who participated in the Behavioral Risk Factor Surveillance System (BRFSS) in 2007 and 2008 (n=807,524), the National Health Interview Survey (NHIS) in 2007 and 2008 (n=44,262), and the National Health and Nutrition Examination Survey (NHANES) during 2007 and 2008 (n=5871) were analyzed. RESULTS: The prevalence estimates of current smoking, obesity, hypertension, and no health insurance were similar across the three surveys, with absolute differences ranging from 0.7% to 3.9% (relative differences: 2.3% to 20.2%). The prevalence estimate of poor or fair health from BRFSS was similar to that from NHANES, but higher than that from NHIS. The prevalence estimates of diabetes, coronary heart disease, and stroke were similar across the three surveys, with absolute differences ranging from 0.0% to 0.8% (relative differences: 0.2% to 17.1%). CONCLUSION: While the BRFSS continues to provide invaluable health information at state and local level, it is reassuring to observe consistency in the prevalence estimates of key health indicators of similar caliber between BRFSS and other national surveys. Copyright Published by Elsevier Inc.

Schneider, K. L., et al. (2012). "Evaluating the impact of non-response bias in the Behavioral Risk Factor Surveillance System (BRFSS)." *Journal of Epidemiology & Community Health* **66**(4): 290-295. BACKGROUND: Response rates of national health surveys are decreasing, which potentially can bias obtained prevalence estimates. The purpose of this study is to evaluate the extent to which non-response impacts the representativeness of the 2000 Behavioral Risk Factor Surveillance System (BRFSS) sample compared to the 2000 Decennial Census. METHODS: The 2000 BRFSS had a median response rate of 48%, while the 2000 Decennial Census had a response rate of 67%. Representativeness of the BRFSS sample was evaluated on gender, race, ethnicity, age, household income and marital status. Prevalence of each factor in the BRFSS was compared to the prevalence found in the US Census on both the state and county levels. Prevalence differences between the BRFSS and Census were calculated and their association with response rates was evaluated using robust OLS regression and polytomous logistic regression. The relationship between prevalence differences and other survey design elements, such as data collection procedure and sampling fraction, was also explored. RESULTS: The BRFSS prevalence estimates diverged from the Census estimates on several sociodemographic factors even after adjustment for non-response/non-coverage. This was found on both the state and county levels; however, smaller absolute differences between the BRFSS and Census prevalence

estimates were found for factors included in the non-response/non-coverage adjustment weight. Lower response rates (<40%) were associated with the under-representation of racial/ethnic minorities, women and younger individuals in the BRFSS survey. CONCLUSION: Future research should examine alternative approaches to increase response rate (e.g., mixed mode) and to adjust for non-response (e.g., multiple imputation).

2011

Balluz, L., et al. (2011). "Noncoverage bias in household landline telephone surveys, BRFSS 2008." *American Journal of Epidemiology* **173**: S108.

The Behavior Risk Factor Surveillance System is a state based telephone survey conducted in all 50 states, DC and the territories. To determine whether exclusion of adults with cell-phone-only may bias estimates from a landline-based survey, a cell phone survey was conducted in parallel with the ongoing, monthly landline data collection in 18 states in 2008. The landline and cell phone samples were weighted at state level. Logistic models were developed for each of 16 health indicators to examine whether survey approach affected estimates after adjusting for the impact of demographic characteristics. The relative biases for estimates were calculated to estimate the potential biases in landline telephone surveys that exclude cell phones. The study found that as noncoverage rate for cell-phone-only adults continued to increase, the biases of 16 health indicators resulting from their exclusion from landline-based health survey were significantly high and cannot be ignored.

Bossarte, R. M., et al. (2011). "Development and validation of a 6-day standard for the identification of frequent mental distress." *Social Psychiatry & Psychiatric Epidemiology* **46**(5): 403-411.

PURPOSE: The goals of the current study were to assess the concurrent validity of a single-item measure of general mental distress with established, multi-item mental health measures used in population-level surveillance and to establish the optimal cutpoint for determining psychological distress (previously identified as frequent mental distress) using recently available data from the Behavioral Risk Factor Surveillance System survey. METHODS: Data for this study were obtained from the core questionnaire and two optional modules available as part of the 2006 and 2007 Behavioral Risk Factor Surveillance System (BRFSS) surveys. Frequent mental distress (FMD) was identified by the number of days of self-reported poor mental health during the last 30 days. Comparisons of the number of days with poor mental health and positive scores for measures of depression and serious mental illness were calculated to identify the most efficient cutpoint for establishing FMD. RESULTS: Comparisons of results obtained from ROC analyses using the PHQ-8 and K6 reported 0.867 (95% CI 0.861-0.872) and 0.840 (95% CI 0.836-0.845) of the area under the curve, respectively, suggesting good accuracy. Using the Youden index, 6 days of poor mental health in the past 30 days, rather than the existing 14-day standard, was identified as the point at which the sum of the sensitivity and specificity was greatest. CONCLUSION: Results from this study suggest that a 6-day standard (FMD-6) can be used as a valid and reliable indicator of generalized mental distress with strong associations to both diagnosable depressive symptomology and serious mental illness.

Hall, T., et al. (2011). "Examining functional content in widely used Health-Related Quality of Life scales." *Rehabilitation Psychology* **56**(2): 94-99.

PURPOSE: Assess extent to which generic Quality of Life (QOL) and Health-Related Quality of Life (HRQOL) scales include function in assessment of health, and identify health assessment items that are free of functional content. METHODS: An expert panel on measurement of health and disability reached consensus on definitions of health, disability, and function. They assessed all items of all generic (non-condition-specific) scales in the 2006 ProQolid database for being important to measuring health as distinct from function. Ratings were summarized as content validity ratios. Retained items were written into standard format and reviewed again by the expert panel and a validity panel with expertise in specific disabilities. RESULTS: Of 85 scales, 21 were retained as containing items important for assessing health. Scales ranged from 100% (BRFSS HRQOL, WHO-5) to only 4% of items rated as important. In further review of "important" items, functional content was identified in many of the items, particularly with regard to mental functioning. CONCLUSIONS: Popular generic scales of QOL and HRQOL vary greatly in the degree to which they include content on function. A pool of items can be identified that are relatively free of function. Distinguishing measurement of function and health is particularly important for people with long-standing functional limitations and for assessing the relationship of health with function.

Hu, S. S., et al. (2011). "Improving public health surveillance using a dual-frame survey of landline and cell phone numbers." *American Journal of Epidemiology* **173**(6): 703-711.

To meet challenges arising from increasing rates of noncoverage in US landline-based telephone samples due to cell-phone-only households, the Behavioral Risk Factor Surveillance System (BRFSS) expanded a traditional landline-based random digit dialing survey to a dual-frame survey of landline and cell phone numbers. In 2008, a survey of adults with cell phones only was conducted in parallel with an ongoing landline-based health survey in 18 states. The authors used the optimal approach to allocate samples into landline and cell-phone-only strata and used a new approach to weighting state-level landline and cell phone samples. They developed logistic models for each of 16 health indicators to examine whether exclusion of adults with cell phones only affected estimates after adjustment for demographic characteristics. The extents of the potential biases in landline telephone surveys that exclude cell phones were estimated. Biases resulting from exclusion of adults with cell phones only from the landline-based survey were found for 9 out of the 16 health indicators. Because landline noncoverage rates for adults with cell phones only continue to increase, these biases are likely to increase. Use of a dual-frame survey of landline and cell phone numbers assisted the BRFSS efforts in obtaining valid, reliable, and representative data. Copyright Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health 2011.

Ibrahimova, A., et al. (2011). "Comparison of 2008 national and state-level self-reported and observed seatbelt use estimates." *Injury Prevention* **17**(3): 201-203.

The objective of the study was to compare national and state-level estimates of self-reported and observed seatbelt use for 2008. Self-reported seatbelt use from the 2008 Behavioral Risk Factor Surveillance System was compared with 2008 observed seatbelt use published by the National Highway Traffic Safety Administration. The ratio of self-reported belt use to observed use was calculated for each state, and the correlation between the two seatbelt measures was examined using the Pearson correlation coefficient. The median state ratio of self-reported to observed belt use was 0.97. Self-reported use was lower than observed use in 38 states. A moderate

association was revealed between the self-reported and observed use ($r=0.71$, $p<0.01$). The findings suggest that, as seatbelt use has increased over time, measures of self-reported and observed use have converged, and any upward bias in self-reported use due to social desirability has substantially declined.

Jiang, Y. and J. E. Hesser (2011). "A comparison of depression and mental distress indicators, Rhode Island Behavioral Risk Factor Surveillance System, 2006." *Preventing Chronic Disease* 8(2): A37.

INTRODUCTION: Depression is a public health concern that warrants accurate population estimates. The patient health questionnaire 8 (PHQ-8) offers high sensitivity and specificity for assessing depression but is time-consuming to administer, answer, and score. We sought to determine whether 1 of 3 simpler instruments - the shorter PHQ-2 or 2 single questions from the health-related quality of life (HRQOL) module of the Behavioral Risk Factor Surveillance System (BRFSS) - could offer accuracy comparable to the PHQ-8. **METHODS:** We compared the depression and mental distress indicators of 2006 Rhode Island BRFSS data by using 4 types of analyses: 1) sensitivity and specificity estimates, 2) prevalence estimates, 3) multivariable logistic regression modeling of the relationship between each of the 4 indicators and 11 demographic and health risk variables, and 4) geographic distribution of prevalence. **RESULTS:** Compared with the PHQ-8, the 3 other measures have high levels of specificity but lower sensitivity. Depression prevalence estimates ranged from 8.6% to 10.3%. The adjusted odds ratios from logistic regression modeling were consistent. Each of the indicators was significantly associated with low income, being unable to work, current smoking, and having a disability. **CONCLUSION:** The PHQ-8 indicator is the most sensitive and specific and can assess depression severity. The HRQOL and PHQ-2 indicators are adequate to obtain population prevalence estimates if questionnaire length is limited.

Kittur, N. D., et al. (2011). "Comparison of contraceptive use between the Contraceptive CHOICE Project and state and national data." *Contraception* 83(5): 479-485. **Background:** We compared contraceptive prevalence reported in the Contraceptive CHOICE Project (CHOICE) at time of enrollment with estimates from representative surveys, the 2006-2008 National Survey of Family Growth (NSFG) and 2006 Missouri Behavioral Risk Factor Surveillance System (BRFSS). **Study design:** We calculated survey weights for CHOICE participants and compared selected demographic characteristics and prevalence estimates of current contraceptive methods being used at the time of enrollment. **Results:** Compared with the NSFG, CHOICE participants at the time of enrollment were less likely to be contraceptive pill users (16.1% vs. 24.0%) and more likely to use condoms (23.8% vs. 13.8%). Compared with the BRFSS, CHOICE participants were more likely to use condoms (20.4% vs. 12.9%) and withdrawal (6.6% vs. 0.4%). **Conclusion:** Despite differences in sampling strategies between CHOICE and state and national surveys, the contraceptive prevalence estimates were largely similar. This information combined with the high rates of long-acting reversible contraception (LARC) use after enrollment by CHOICE participants that have been previously reported by study participants may imply that cost and restricted access to LARC could be essential factors in the low rates of LARC use in the United States. © 2011 Elsevier Inc. All rights reserved.

Kwak, L., et al. (2011). "The repeatability and validity of questionnaires assessing occupational physical activity - a systematic review." *Scandinavian Journal of Work,*

Environment and Health 37(1): 6-29.

OBJECTIVES This study aims to review systematically the repeatability and validity of questionnaires used to assess occupational physical activity among healthy adults. **METHODS** We searched PubMed and Embase using occupational, work-related, job-related, physical activity, motor activity, and questionnaires as keywords. Two reviewers independently performed article selection, data extraction, and quality assessment. The methodological quality and results of the studies were evaluated based on an existing checklist. The level of evidence and repeatability, criterion, and construct validity were rated. **RESULTS.** We included 31 papers describing 30 questionnaires in the review. Repeatability was assessed in 22 studies, 11 used appropriate measures to assess 12 questionnaires. Intra-class correlation coefficients and weighted Cohen's kappa ranged between 0.43-0.95. Six studies used appropriate measures to assess criterion validity of 13 questionnaires. One questionnaire, the Tecumseh Self Administered Occupational Physical Activity Questionnaire (TOQ), showed good criterion validity against a physical activity (PA) record. Eighteen studies used appropriate measures to assess the construct validity of 23 questionnaires. Comparison included those against accelerometers, maximal oxygen uptake, questionnaires, and body composition measures. None showed good construct validity. **CONCLUSIONS** There is strong evidence for good reliability of four questionnaires. None of the reviewed questionnaires showed good criterion validity compared to objective measures. Compared to PA records, moderate-to-good validity was observed for two questionnaires. Objective measures of occupational PA are needed.

Njai, R., et al. (2011). "Misclassification of survey responses and black-white disparity in mammography use, Behavioral Risk Factor Surveillance System, 1995-2006." *Preventing Chronic Disease* 8(3): A59.

INTRODUCTION: The validity of self-reported data for mammography differ by race. We assessed the effect of racial differences in the validity of age-adjusted, self-reported mammography use estimates from the Behavioral Risk Factor Surveillance System (BRFSS) from 1995 through 2006 to determine whether misclassification (inaccurate survey question response) may have obscured actual racial disparities. **METHODS:** We adjusted BRFSS mammography use data for age by using 2000 census estimates and for misclassification by using the following formula: $(\text{estimated prevalence} - 1 + \text{specificity}) / (\text{sensitivity} + \text{specificity} - 1)$. We used values reported in the literature for the formula (sensitivity = 0.97 for both black and white women, specificity = 0.49 and 0.62, respectively, for black and white women). **RESULTS:** After adjustment for misclassification, the percentage of women aged 40 years or older in 1995 who reported receiving a mammogram during the previous 2 years was 54% among white women and 41% among black women, compared with 70% among both white and black women after adjustment for age only. In 2006, the percentage after adjustment for misclassification was 65% among white women and 59% among black women compared with 77% among white women and 78% among black women after adjustment for age only. **CONCLUSION:** Self-reported data overestimate mammography use - more so for black women than for white women. After adjustment for respondent misclassification, neither white women nor black women had attained the Healthy People 2010 objective (> 70%) by 2006, and a disparity between white and black women emerged.

Pierannunzi, C., et al. (2013). "A systematic review of publications assessing reliability and validity of the Behavioral Risk Factor Surveillance System (BRFSS), 2004-2011." *BMC*

Medical Research Methodology 13: 49.

BACKGROUND: In recent years response rates on telephone surveys have been declining. Rates for the behavioral risk factor surveillance system (BRFSS) have also declined, prompting the use of new methods of weighting and the inclusion of cell phone sampling frames. A number of scholars and researchers have conducted studies of the reliability and validity of the BRFSS estimates in the context of these changes. As the BRFSS makes changes in its methods of sampling and weighting, a review of reliability and validity studies of the BRFSS is needed. **METHODS:** In order to assess the reliability and validity of prevalence estimates taken from the BRFSS, scholarship published from 2004-2011 dealing with tests of reliability and validity of BRFSS measures was compiled and presented by topics of health risk behavior. Assessments of the quality of each publication were undertaken using a categorical rubric. Higher rankings were achieved by authors who conducted reliability tests using repeated test/retest measures, or who conducted tests using multiple samples. A similar rubric was used to rank validity assessments. Validity tests which compared the BRFSS to physical measures were ranked higher than those comparing the BRFSS to other self-reported data. Literature which undertook more sophisticated statistical comparisons was also ranked higher. **RESULTS:** Overall findings indicated that BRFSS prevalence rates were comparable to other national surveys which rely on self-reports, although specific differences are noted for some categories of response. BRFSS prevalence rates were less similar to surveys which utilize physical measures in addition to self-reported data. There is very little research on reliability and validity for some health topics, but a great deal of information supporting the validity of the BRFSS data for others. **CONCLUSIONS:** Limitations of the examination of the BRFSS were due to question differences among surveys used as comparisons, as well as mode of data collection differences. As the BRFSS moves to incorporating cell phone data and changing weighting methods, a review of reliability and validity research indicated that past BRFSS landline only data were reliable and valid as measured against other surveys. New analyses and comparisons of BRFSS data which include the new methodologies and cell phone data will be needed to ascertain the impact of these changes on estimates in the future.

Potter, S. J. and D. J. Laflamme (2011). "An assessment of state level sexual assault prevalence estimates." *Maternal and Child Health Journal* **15**(1): 77-86.

The purpose of our research is to compare sexual violence prevalence rates from three sources of state level data. Public health officials, legislators and other policymakers often require state-level sexual assault prevalence estimates to justify funding and rationalize both new and ongoing sexual violence prevention programs, as well as programs for victims. We compared survey design and resulting prevalence rates of the three surveys frequently used at the state level: the Behavioral Risk Factor Surveillance System (BRFSS), the National Violence Against Women Survey (NVAWS) extrapolations, and replications of the NVAWS. Although the specificity of the questions used in the NVAWS provides a clearer picture of the prevalence of sexual assault than the BRFSS questions, the sexual violence module on the BRFSS survey has the advantage that it is used regularly by some states. Currently available female sexual assault prevalence estimates differ widely at the state level but can be used when interpreted with informed caution. The new National Intimate Partner and Sexual Violence Surveillance System holds promise for providing better estimates in the future. © 2010 Springer Science+Business Media, LLC.