

Predicting geographical variations in behavioural risk factors: an analysis of physical and mental healthy days

H Jia, P Muennig, E I Lubetkin, M R Gold

J Epidemiol Community Health 2004;58:150-155

See end of article for authors' affiliations

Correspondence to: Dr H Jia, Department of Community Medicine, Mercer University School of Medicine, 1550 College Street, Macon, GA, USA; haomia@yahoo.com

Accepted for publication 22 April 2003

Study objectives: To determine the validity of physical and mental unhealthy days as summary measures for county health status and to forward a method for examining county level health trends using a single year of data from the Behavioral Risk Factor Surveillance System (BRFSS).

Design: The study analysed geographical variation in physical and mental unhealthy days at the state and county level using the 2000 BRFSS. Whereas state level analyses used individual level data, this research conducted multilevel regression analysis using county level data as independent variables and individual level reports of physical and mental unhealthy days as dependent variables.

Setting: Population based samples of non-institutionalised civilian adult residents from each of the 50 states and the District of Columbia in the United States.

Main results: Socioeconomic variables predicted similar mean numbers of physical and mental unhealthy days at both the state and county level, validating the county level analyses. County level disability rates were strongly associated with county mean unhealthy days. Using the regression method we forward, it is possible to analyse county level trends using a single year of BRFSS data.

Conclusions: Physical and mental unhealthy days may be used as valid summary measures of county health status. Regression models may be used to assist local decision makers in assessing the needs of their communities and may be used to improve health resource allocation within states.

While there are rich and varied sources of national health data, few standardised sources of data are available at the local level. Local health departments and community based organisations are often faced with the task of collecting data on their own or attempting to make inferences based on mortality data. The Behavioral Risk Factor Surveillance System (BRFSS) was designed to help improve local decision making by providing state level data, but there exists no consistent source of county level health data capable of assessing health trends.¹⁻⁵

Health related quality of life (HRQOL) is a useful measure for local planners that allows for examination of the local burden of non-fatal diseases. The Centers for Disease Control and Prevention (CDC) have developed a set of survey instruments for measuring HRQOL that includes respondents' perceived health status and activity limitations.⁶⁻⁸ These HRQOL measures, termed physical and mental healthy days, have been included in the BRFSS since 1993 and are conducted on an annual basis.^{7,9} Because of the absence of other temporal and spatial data that measure population HRQOL, the CDC have promoted the BRFSS as being capable of filling an important data void.⁸

The CDC's "healthy days" serves as a proxy measure of HRQOL and has the advantages of both simplicity and brevity. Compared with the Medical Outcomes Study Short Form 36 questions (SF-36) for example,¹⁰ healthy days questions are easier to answer.⁸ While healthy days are not as rich a source of HRQOL information, they measure perceptions of physical and mental health using one question each, so complex weighting factors do not need to be developed to calculate summary scores.¹⁰

A person's HRQOL is associated with the presence of chronic disease, and serves as a good summary measure of health.^{8,10} Many investigators have hypothesised associations between the characteristics of communities, such as socioeconomic status and HRQOL.^{4,8,11,12} It is not known, however, whether county characteristics predict HRQOL or whether

the healthy days measure in BRFSS serves as a valid tool for informing local policymakers of the overall health of their communities. Moreover, current methods for estimating county level characteristics require many years of BRFSS data to be pooled, limiting their usefulness in estimating trends in health status and risk factors for disease for planning purposes.⁵

The purpose of this study was to examine the validity of physical and mental unhealthy days as an indicator for county health status, and to introduce a method for analysing BRFSS data that requires a single year of data to assess county level trends. We provide a working model that may help local health officials and other planners understand the needs of the communities they serve, to track health trends from year to year, and make comparisons between their counties and neighbouring counties. Finally, it allows for the analysis of BRFSS questions that are asked only intermittently, and allows for temporal analysis of the effects of local health interventions.

METHODS

We examined geographical variation in unhealthy days at the state and county level and identified county characteristics that predict HRQOL. We hypothesised that physical and mental unhealthy days would vary predictably with state and county socioeconomic statistics, mortality rates, and disability rates.

Data

We used the 2000 BRFSS data to examine physical and mental unhealthy days among persons over 18 years of age.

Abbreviations: BRFSS, Behavioral Risk Factor Surveillance System; HRQOL, health related quality of life; CDC, The Centers for Disease Control and Prevention

survey coordinated by the CDC.¹³ This survey uses the random digit dialling method to generate population based samples of non-institutionalised civilian adult residents from each of the 50 states and the District of Columbia.

Physical and mental unhealthy days measures are based on two questions:

- 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?
- 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?

We included the following BRFSS variables in analyses: the respondent's age, race, gender, income, education, employment status, self rated health status (excellent, very good, good, fair, poor), and county of residence.

To obtain a more robust set of outcome measures we merged county sociodemographic descriptors from the 2000 decennial census¹³ and 1996–1998 age adjusted mortality data from the CDC¹⁴ with the BRFSS dataset using the county of residence for each respondent in the BRFSS.

Data analyses

The overall mean physical and mental unhealthy days and standard errors of estimates were calculated using the SURVEYMEAN procedure of SAS software package to adjust for complexities of the BRFSS sampling design.¹⁵ We calculated age adjusted mean unhealthy days for selected sociodemographic subgroups for all states.

Small sample size precluded direct calculation of county specific mean unhealthy days from the BRFSS. County variations in mean unhealthy days were examined through multilevel regression by identifying county level variables that were related to unhealthy days.^{4 16 17} The individual level of multilevel model takes the form:

$$y_{ij} = \beta_{0i} + \beta_{1i} \times age_{ij} + \beta_{2i} \times race_{ij} + \beta_{3i} \times sex_{ij} + \varepsilon_{ij}$$

where y_{ij} is the unhealthy days for respondent j in county i , β_{0i} , β_{1i} , β_{2i} , and β_{3i} are county specific regression coefficients for baseline, age, race, and sex respectively; and ε_{ij} is the error of prediction.

The county level conditional model is

$$\beta_{ki} = \gamma_{0k} + \gamma_{1k}x_i + \mu_{ki} \quad k = 0, 1, 2, 3$$

where each county's coefficient, β_{ki} , is expressed as regression parameter estimation by county level covariates x_i , and μ_{ki} refers to county specific error. In our analysis, county level covariates included poverty rate, unemployment rate, mortality rate, the proportion of adults without high school diploma, physical disability rate, mental disability rate, self care disability rate, and employment disability rate. In the unconditional model, β_{ki} is expressed as the sum of an intercept and a random error associated with the i^{th} county.

The relative difference of the variance of μ_{ki} terms between the unconditional and conditional model is the proportion of between county variations in unhealthy days explained by covariates x_i .¹⁷

We then used the resulting regression estimates to make predictions for mean number of unhealthy days for all counties in the US. To examine for a non-linear relation

Spline regressions rather than ordinal linear regressions. The Spline regression uses a piecewise polynomial function that fits data locally.¹⁸ This method corrects for bias, particularly at boundary regions, fitting data better than a straight line when independent variables are skewed or have outliers.¹⁸

The relations between county level variables and unhealthy days were examined via scatterplots of county level variables versus the regression predicted unhealthy days for corresponding counties. A steeper slope of the plotting curve indicates a stronger relation between a county level variable and number of unhealthy days. We used the 99% curve growth to measure the slope. The curve growth is calculated using:

$$\frac{E_{99} - E_1}{E_1} \times 100\%$$

where E_1 and E_{99} are the 1st and 99th centile of estimated values of a given variables, respectively.

As a validation measure of the physical and mental unhealthy days measures at the county level, we plotted these measures against self reported physical, mental, self care disability, and employment disability from the 2000 census.¹³ Physical disability was defined as having limits in basic physical activities such as walking, climbing stairs, or reaching, lifting, or carrying objects. Mental disability was having any difficulty in learning, remembering, or concentration. Self care disability was having any difficulty in dressing, bathing, or getting around inside the home. Employment disability was having difficulty working because of a physical, mental, or emotional condition.

RESULTS

In 2000, the BRFSS interviewed a total of 180 244 adults. About one third (34%) of adult respondents aged 18 and older reported at least one physical unhealthy day and one third (34%) reported at least one mental unhealthy day. For those who reported at least one unhealthy day, about a fifth reported between 1–7 physical (22%) and mental (21%) unhealthy days. The maximum number of physical and mental unhealthy days, 30 days, was reported by 5.8% and 4.4% of the population respectively. The average number of physical unhealthy days was 3.3 days and the average number of mental unhealthy days was 3.2 days.

Physical unhealthy days increased consistently with increasing age (table 1). By contrast, mental unhealthy days decreased after age 55. We also identified significant differences among other sociodemographic subgroups. In particular, being female, black, American Indian or Alaska Native, Hispanic, being out of work or unable to work, not having finished high school, and having a lower household income were associated with a higher number of mean physical and mental unhealthy days. Persons reporting lower levels of self rated health also reported more physical and mental unhealthy days.

State level

Accurate state level data were obtained directly from the BRFSS. Figure 1 presents a scatterplot of state level age adjusted mean physical unhealthy days compared with mental unhealthy days. The plot illustrates the strong positive association between the state level mean physical and mental unhealthy days, with a Spearman rank correlation of 0.73. For example, West Virginia and Kentucky ranked the highest and the third highest in mean physical unhealthy days, respectively. These states respectively ranked the second and first highest in mental unhealthy days.

	Count	Mean physical unhealthy days*	Mean mental unhealthy days*
Overall	180244	3.34 (SE=0.03)	3.21 (SE=0.03)
Age			
18–24	16100	1.9	4.1
25–34	32425	2.1	3.4
35–44	40021	2.8	3.5
45–54	34357	3.5	3.5
55–64	23115	4.6	2.7
65–74	18922	4.9	2.0
75+	14199	6.3	2.0
Sex			
Male	73225	3.0	2.5
Female	107019	3.7	3.9
Race/ethnicity			
Hispanic	12263	4.0	3.3
Not Hispanic			
White	143144	3.2	3.2
Black	14293	4.0	3.5
Asian/Pacific islander	4326	2.7	2.0
American Indian/Alaska Native	3276	5.3	4.8
Other	1396	4.4	4.8
School			
Less than high school	21055	5.2	4.4
High school graduate	56955	3.6	3.5
Some college	49819	3.4	3.3
College graduate	52415	2.2	2.3
Employment status			
Employed for wages	100188	2.2	2.7
Self employed	15605	2.3	2.7
Out of work ≥ 1 year	2229	6.3	5.7
Out of work < 1 year	3676	5.0	5.2
Homemaker	12768	3.3	3.3
Student	5221	2.3	3.4
Retired	32511	4.3	3.0
Unable to work	7719	14.7	9.3
Income			
< \$15000	18405	7.1	6.1
\$15000–\$24999	30094	4.5	4.1
\$25000–\$34999	25639	3.3	3.3
\$35000–\$49999	29709	2.7	2.8
≥ \$50000	50679	2.2	2.4
Self rated health			
Excellent	40021	1.0	1.8
Very good	60363	1.7	2.4
Good	52100	2.7	3.4
Fair	19596	8.0	6.1
Poor	7817	19.8	10.7

SE, standard error. *Age adjusted to 2000 US population for all characteristics except age groups.

Similarly, Hawaii ranked the lowest and the second lowest in mean physical and mental unhealthy days respectively.

As would be expected, state mean physical and mental unhealthy days were strongly associated with state socio-demographic and health related data. States with the highest mean physical or mental unhealthy days were among those with the highest mortality and disability rates, and with the lowest income. For example, West Virginia ranked first and second in age adjusted mean physical and mental unhealthy days, respectively. It also ranked the fifth highest in mortality, the highest in disability rate, and the second lowest in per capita income. The correlations between mortality, disability, and income and physical unhealthy days were 0.50, 0.68, and -0.30 respectively. For mental unhealthy days the correlations were 0.39, 0.50, and -0.14 respectively.

County level

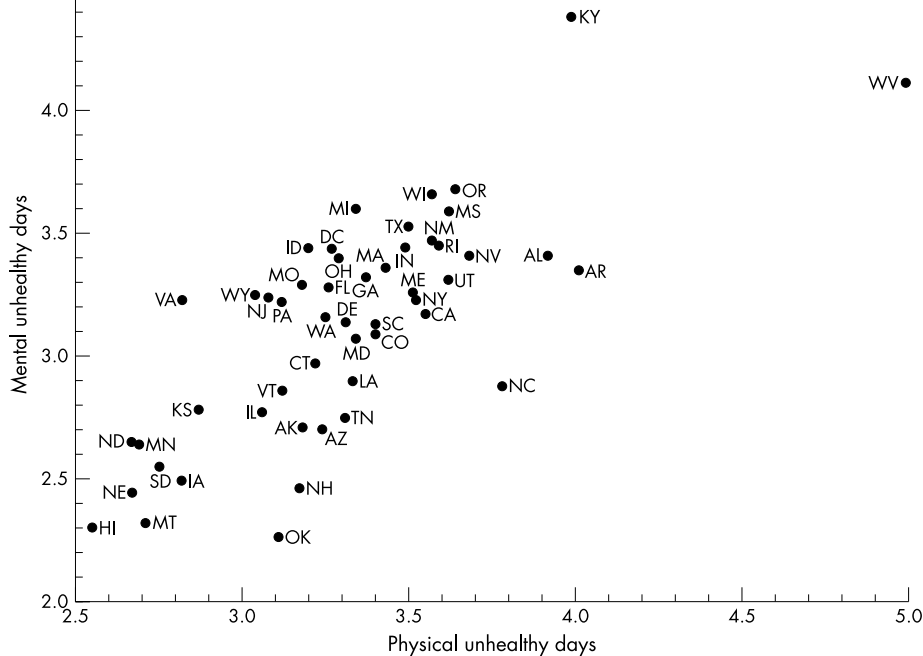
The multilevel model incorporated individual level survey responses of unhealthy days as a dependent variable from 3067 counties (98%). The analysis of county level variation of unhealthy days was performed by examining the percentage of within county variations explained by individual level

variables and between county variations explained by county level variables. Table 2 presents the results from these multilevel analyses. Respondents' age had the strongest association with unhealthy days measures, explaining 8.8% variation of within counties physical unhealthy days and 5.2% variation of within counties mental unhealthy days. Gender and race accounted for only about 1%–2% variation of unhealthy days.

With respect to county level effects on mean county unhealthy days, the included covariates explained a larger proportion of between county variation than the individual level analysis, with each measure of disability accounting for nearly 30% of the physical unhealthy days, and socio-economic covariates accounting for 10%–20% of the variation. All county level covariates explained smaller proportion of between county variation of mental unhealthy days than physical unhealthy days. The mental disability rate had the strongest association with county mean mental unhealthy days, followed by employment disability rates.

We investigated the magnitude of association between county level covariates and county mean unhealthy days for counties with at least one observation (figs 2 and 3). In these figures, a steeper slope indicates a stronger relation between

Figure 1 Mean physical and mental unhealthy days among adults by state (age adjusted)—BRFSS, 2000



a county level variable and number of unhealthy days. In figure 2, a consistent positive association between these variables and physical (black line) and mental unhealthy (grey line) days can be seen. Physical unhealthy days demonstrated a stronger association than mental unhealthy days, as indicated by steeper slopes for physical curves. The curve growth for the unemployment rate, the proportion without a high school diploma, the poverty rate, and the mortality rate was 62.5%, 78.5%, 63.2%, and 55.5%, respectively for physical unhealthy days. Curve growth rates were 35.4%, 39.5%, 38.2%, and 37.4% respectively for mental unhealthy days.

We also validated the two measures by examining the relation between mean unhealthy days and the prevalence of physical, mental, self care, and employment disability (fig 3). As expected, physical unhealthy days increased greatly with an increase in the physical disability rate, and mental unhealthy days increased moderately with an increase in the physical disability rate. The physical curve growth for physical disability rate was 94.2% and mental curve growth was 25.0%. In comparison, both physical and mental

unhealthy days increased greatly with an increase in the mental disability rate. The physical curve growth was 119% and mental curve growth was 66.0%.

Self care disability was strongly related to physical unhealthy days and moderately related to mental unhealthy days, as indicated by 94.6% of curve growth for the physical unhealthy days and 49.9% for the mental unhealthy days. Similarly, employment disability was strongly related to both physical and mental unhealthy days. The physical and mental curve growths were 102.1% and 62.5%, respectively.

DISCUSSION

Socioeconomic and health related data show a strong and consistent relation with the number of physical and mental unhealthy days at both the state and county levels. This finding confirmed previous surveillance reports by the CDC, which suggested that state HRQOL measures were related to disability and disease prevalence rates.^{2 4 8 11 12} Self reported physical and mental unhealthy days are also strongly associated with standard measures of disability at the county

Table 2 Multilevel regression analysis of physical and mental unhealthy days—BRFSS, 2000

2000 BRFSS	Unhealthy days	
	Physical	Mental
Percentage variation within counties explained by respondents'		
Age	8.8	5.2
Race	1.1	1.1
Sex	1.3	1.8
Age, race, and sex together	9.5	6.5
Percentage explainable variation between counties explained by county's		
Poverty rate	13.4	1.7
Unemployment rate	13.2	1.9
Mortality rate	10.8	3.8
Less than high school rate	19.6	2.1
Physical disability rate	27.7	5.0
Mental disability rate	26.9	6.3
Self care disability rate	26.8	3.9
Employment disability rate	25.6	6.1

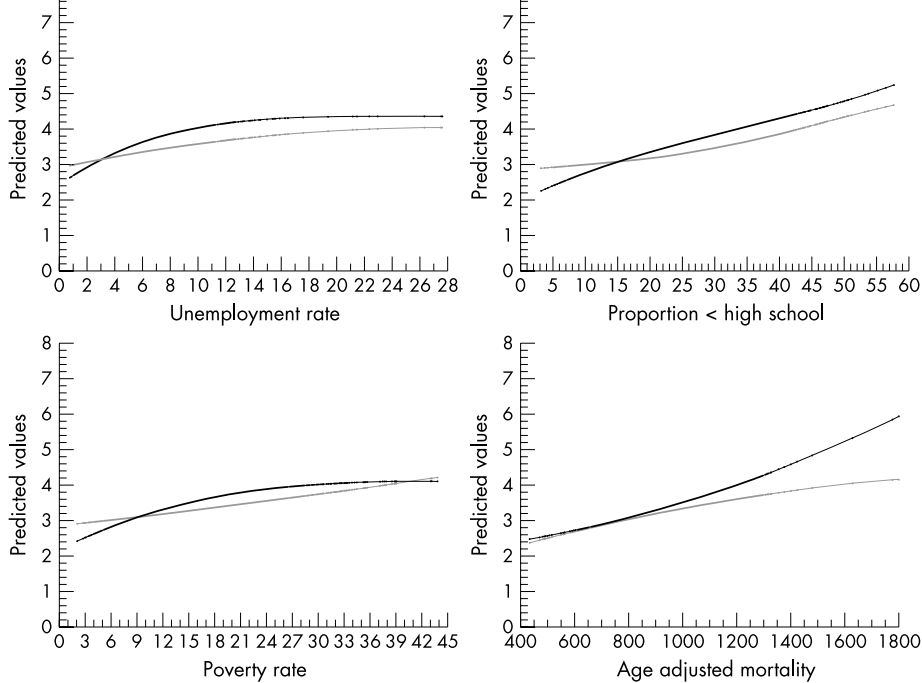


Figure 2 County level socioeconomic data compared with regression predicted county mean physical and mental unhealthy days. Black line, physical unhealthy days; grey line, mental unhealthy days.

level, indicating that these measures may serve as valid summary measures of local health status assessments. Counties reporting higher mean number of physical and mental unhealthy days were likely to have higher unemployment rates, a higher poverty rate, a higher percentage of adults who did not complete high school, higher mortality rates, and a higher prevalence of disability.

We also found that physical unhealthy days had a much stronger association with all county level variables than mental unhealthy days. As expected, while age adjusted mortality was moderately correlated with mental unhealthy days, the association was much stronger with respect to

physical unhealthy days. The association between mental health and physical health we noted adds to a growing body of literature.^{19, 20}

One improvement of this study over previous correlation studies^{5, 11, 12} was the use of plots to examine the relation between county level variables and BRFSS outcome measures (in this case, unhealthy days). This allowed us to predict the impact of county level variables on unhealthy days for each covariate.^{4, 16} For example, if a county with a poverty rate of 10% decreases to 5%, we would expect that its adjusted mean physical unhealthy days would decrease from 3.2 days to 2.7 days.

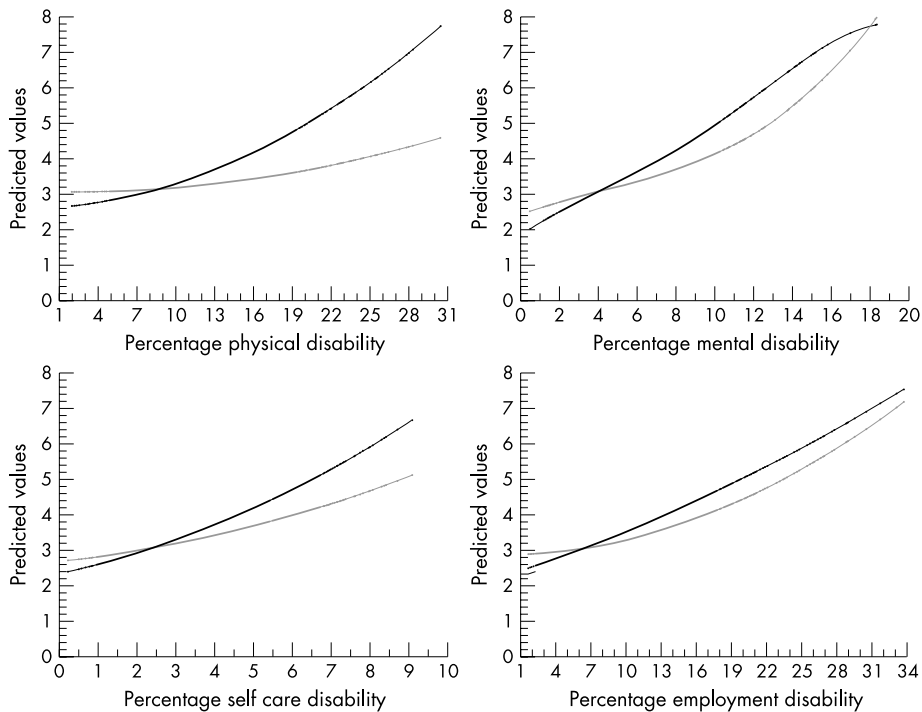


Figure 3 County level disability data compared with regression predicted county mean physical and mental unhealthy days. Black line, physical unhealthy days; grey line, mental unhealthy days.

Spline regression, which allowed us to test for non-linear relations. This was particularly important when the county level variables were highly skewed or had outliers. Estimates based on ordinal linear regression are usually biased near boundary regions.¹⁸ Yuma County, Arizona, for instance had 130 respondents, enough to provide accurate estimation. Using ordinal linear regression, the estimated adjusted mean unhealthy days was 5.8 days, while the use of Spline regression yielded an estimate of 4.3 days, which was much closer to the observed value of 4.0 days.

The findings in this study are subject to a number of limitations. Firstly, the BRFSS is based on self reported data elicited by a telephone survey. Data are therefore subject to recall bias and selection bias. Moreover, the sample includes only the civilian, non-institutionalised population, excluding persons living in long term care facilities or other institutions, which probably leads to underestimates of unhealthy days.

Secondly, about 50 counties were not sampled in the BRFSS. These counties were excluded. Other methods, such as geographical smoothing techniques, do little to improve predictions for these smaller counties as the smaller counties are rarely sampled in the BRFSS.⁵ Moreover, other techniques require multiple years of data, reducing the ability of policymakers to track trends or capture the effects of rapidly changing health outcomes.

In conclusion, it is possible to predict the mean number of physical and mental unhealthy days at the county level. These data may be used to improve local health decision making, and may allow trends to be tracked over time. The mean number of physical and mental unhealthy days seems to be a valid indicator for county health status and therefore, perhaps a useful summary measure for monitoring the health of a local population.

Authors' affiliations

H Jia, Department of Community Medicine, Mercer University School of Medicine, Macon, GA, USA

P Muennig, **E I Lubetkin**, **M R Gold**, Department of Community Health and Social Medicine, City University of New York Medical School, New York, USA

- 1 **Remington PL**, Smith MY, Williamson DF. Design, characteristics, and usefulness of state-based Behavioral Risk Factor Surveillance: 1981–87. *Public Health Rep* 1988;**103**:366–75.
- 2 **Centers for Disease Control and Prevention**. 1996–2001 BRFSS summary prevalence report. <http://www.cdc.gov/brfss/pubrfdat.htm> (accessed 17 Apr 2003).
- 3 **Kim I**, Keppel KG. Priority data needs: sources of national, state, and local-level data and data collection systems. *Healthy People 2000 Statistical Notes* 1997;**15**:1–15.
- 4 **Borawski E**, Jia H, Wu GW. *The use of the Behavioral Risk Factor Surveillance System (BRFSS) to estimate the prevalence of state and substate Disability*. Atlanta, GA: Centers for Disease Control and Prevention, 1999.
- 5 **Pickle LW**, Su Y. Within-state geographic patterns of health insurance coverage and health risk factors in the United States. *Am J Prev Med* 2002;**22**:75–83.
- 6 **Hennessy CH**, Moriarty DG, Zack MM, *et al*. Measuring health-related quality of life for public health surveillance. *Public Health Rep* 1994;**109**:665–72.
- 7 **Centers for Disease Control and Prevention**. Quality of life as a new public health measure—Behavioral Risk Factor Surveillance System, 1993. *MMWR* 1994;**43**:375–80.
- 8 **Centers for Disease Control and Prevention**. *Measuring Healthy Days*. Atlanta, GA: Centers for Disease Control and Prevention, Nov 2000.
- 9 **Centers for Disease Control and Prevention**. Health-related quality of life measures—United States, 1993. *MMWR* 1995;**44**:195–00.
- 10 **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**:473–83.
- 11 **Kanarek N**, Sockwell D, Jia H. Community indicators of health-related quality of life—United States, 1993–1997. *MMWR* 2000;**49**:281–5.
- 12 **Centers for Disease Control and Prevention**. State differences in reported healthy days among adults—United States, 1993–1996. *MMWR* 1998;**47**:239–44.
- 13 **US Census Bureau**. Census 2000 data, prepared by the US Census Bureau, 2002. <http://www.census.gov/main/www/cen2000.html> (Accessed 17 Apr 2003).
- 14 **Centers for Disease Control and Prevention**. *Mortality file summary 1996–1998*. <http://wonder.cdc.gov/mortSQL.shtml> (accessed 17 Apr 2003).
- 15 **SAS Institute**. *SAS/Stat user's guide, version 8*. Cary, NC: SAS Institute, 1999.
- 16 **Ericksen EP**. A regression method for estimating population changes of local areas. *J Am Stat Assoc* 1974;**69**:867–75.
- 17 **Singer JD**. Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth modes. *Journal of Educational and Behavioral Statistics* 1998;**24**:323–55.
- 18 **Simonoff JS**. *Smoothing methods in statistics*. New York, NY: Springer-Verlag, 1996.
- 19 **Wells KB**, Stewart A, Hays RD, *et al*. The functioning and well-being of depressed patients. Results from the Medical Outcomes Study. *JAMA* 1989;**262**:914–19.
- 20 **Centers for Disease Control and Prevention**. Health-related quality of life and activity limitations—eight states, 1995. *MMWR* 1998;**47**:134–40.