

Nonlinearity in Demographic and Behavioral Determinants of Morbidity

Jean C. Norris, Mark J. van der Laan, Sylvia Lane, James N. Anderson, and Gladys Block

Objective. To examine nonlinearity of determinants of morbidity in the United States

Data Sources. A secondary analysis of data on individuals with dietary data from the Cancer Epidemiology Supplement and National Health Interview Survey (NHIS) 1987, a cross-sectional, stratified random sample of the U.S. population ($n = 22,080$).

Study Design. A statistical exploration using additive multiple regression models.

Methods. A Morbidity Index (0–30 points), derived from 1987 National Health Interview Survey data, combines number of conditions, hospitalizations, sick days, doctor visits, and degree of disability. Behavioral (health habits) variables were added to multivariate models containing demographic terms, with Morbidity Index and Self-assessed Health outcomes ($n = 17,612$). Tables and graphs compare models of morbidity with self-assessed health models, with and without behavioral terms. Graphs illustrate curvilinear relationships.

Principal Findings. Morbidity and health are associated nonlinearly with age, race, education, and income, as well as alcohol, diet change, vitamin supplement use, body mass index (BMI), marital status/living arrangement, and smoking. Diet change and supplement use, education, income, race/ethnicity, and age relate differently to self-assessed health status than to morbidity. Morbidity is strongly associated with income up to about \$15,000 above poverty. Additional income predicts no further reduction in morbidity. Better health is strongly related to both higher income and education. After controlling for income, black race does not predict morbidity, but remains associated with lower self-assessed health.

Conclusions. Good health habits, as captured in these models, are associated with a 10–20-year delay in onset and progression of morbidity.

Key Words. Morbidity, self-assessed health, generalized additive models, nonlinear, behavioral factors

Socioeconomic status (SES) indicators, including income and education, are consistently related to increased mortality and mental and physical morbidity, as well as to biological risk factors and health-related behaviors (Liberatos, Link, and Kelsey 1988; Adler and Ostrove 1999). These relationships typically follow a gradient in which lower SES is associated with higher incidence of mortality, greater prevalence of morbidity, or greater prevalence of risk factors such as smoking (Adler and Ostrove 1999).

Risk factors, such as smoking, alcohol, and dietary patterns, all contribute not only to the leading fatal diseases (Kumanyika 1990), and accidental and violent deaths (Adler and Ostrove 1999), but also to nonfatal diseases, such as diabetes, osteoarthritis, chronic respiratory disease, and hypertension, as well as to accidents and injuries. Risk factors for disease, like morbidity itself, are unevenly distributed in the population, with many (but not all) more prevalent among minority populations (Myers 1995; Heckler 1985; Polednak 1989; U.S. Department of Health and Human Services 2000; Winkleby et al. 1992).

Several of the relationships between predictors, or determinants, and morbidity appear to be nonlinear, or curved. Curved relationships have been reported between socioeconomic status and morbidity or mortality outcomes (House et al. 1990; Backlund, Sorlie, and Johnson 1996; McDonough et al. 1997). Curved relationships have also been reported between chronic diseases, mortality, or wellness and health behavior variables such as alcohol intake (Liao et al. 1990; Shaper and Wannamethee 1998), and body mass index (BMI) (Manson 1995).

Curvilinear relationships mean that one additional year of age, or dollar of income, or smoking pack-year makes more difference for morbidity, mortality, or health at some levels of age, income, or smoking than others. Models designed to estimate, evaluate, and control for curvature can answer questions such as, "What levels of income, tobacco use, dietary factors or obesity are associated with the least amount of morbidity or the highest levels of health?" and "How does self-assessed health or morbidity differ for blacks, whites, and Hispanics, when income (and other factors) are estimated as curved relationships?"

The research reported here had three objectives: (1) examining the shape of relationships between general morbidity (not mortality) and contributing factors, such as socioeconomic status and behavioral factors, and (2) comparing models of morbidity with models of health for a U.S.

This research was supported, in part, by a Wellness Foundation Fellowship, a Better Health Foundation Fellowship, a Public Health Service Traineeship, two Grossman Grants, and an anonymous gift.

Address correspondence to Jean C. Norris, M.S., R.D., Dr.P.H., Senior Research Scientist, Public Health Institute, 555 12th Street, 10th Floor, Oakland, CA 94607-4046. Mark J. van der Laan, Ph.D., is a Professor of Biostatistics, Sylvia Lane, Ph.D., is Professor Emeritus, James N. Anderson, Ph.D., is Professor Emeritus, and Gladys Block, Ph.D., is a Professor of Epidemiology and Public Health Nutrition, all at the University of California at Berkeley.

population sample. The final objective (3) was evaluating the relative importance of sociodemographic and behavioral factors in multivariate models. We were particularly interested in examining the relationship of dietary behavior to morbidity and health.

METHODS

Study Sample

These analyses use data for the adult respondents in the National Health Interview Survey (NHIS) 1987 Cancer Epidemiology Supplement ($n = 22,080$). In addition to NHIS core demographic variables and morbidity measures for individual respondents, the 1987 Cancer Epidemiology Supplement (National Center for Health Statistics [NCHS] 1989) included a uniquely rich set of questions on usual dietary behavior, including vitamin supplement use and diet change.¹ These data provided the first and only opportunity to include a number of “usual diet” variables in models with morbidity or health outcomes for the U.S. population.

Because of small numbers, respondents who reported race/ethnicity other than white, black, or Hispanic were dropped from analysis. Persons with seriously flawed dietary data were excluded. Women with hospitalization for childbirth during the past 13 months were dropped, because their food intake over the past year was likely to differ from their usual diet. The data set used for modeling contained complete information on 17,612 individual adults.

Approach

To explore known and suspected nonlinearity in determinants of morbidity, this analysis uses an additive regression modeling system designed to estimate and evaluate curvilinear and nonlinear relationships in multivariate models. Multivariate models begin with demographic variables then add health habits variables to evaluate the relative importance of socioeconomic and behavioral contributions to morbidity or health in the U.S. population. All continuous variables, including demographic terms such as age, income, and education, were assumed to have nonlinear relationships to morbidity or health, until the shapes and statistical importance of curvature, if any, could be evaluated.

Measures

Two dependent variables were used in the analysis, one representing morbidity and the other representing health. Morbidity in the NHIS is

represented by about 20 measures, no one of which describes the total morbidity of an individual. Several NHIS variables describe morbidity over *the past year*—sick days in bed, nights in the hospital, and doctor (or other outpatient medical care) visits. Some represent *cumulative burden* of disease or disability—number of conditions (chronic and acute), disability activity limitation, and need for help with personal care. One measure, days of activity restriction due to illness *in the past two weeks*, may indicate relative susceptibility to severe illness from transient, acute conditions.

The seven most highly correlated NHIS measures were incorporated into a composite morbidity scale intended to rank individuals by relative morbidity. After examining the results of principal components and factor analyses (data not shown), the seven NHIS morbidity variables were combined in a nonparametric manner by assigning cut points along their ranges and summing the points. This shortened the tails of highly skewed distributions and weighted the variables according to the authors' perception of their relative importance, resulting in a Morbidity Index ranging from 0 to 30 points. Table 1 shows the original range for each variable contributing to the Morbidity Index and the population distribution in the collapsed ranges. Morbidity Index derivation is described in greater detail elsewhere (Norris 1996). Reliability of the final Morbidity Index is indicated by a measure of internal consistency (Cronbach's $\alpha = .8$) considered acceptable for group comparisons (Streiner and Norman 1995). Distribution of respondents in each quintile range of the Morbidity Index, by race and sex, is shown in Table 2a.

The second dependent variable, representing health, is the single five-value NHIS variable, self-assessed health status, coded Excellent = 1 to Poor = 5. Self-assessed health has been shown to be "a very adequate general health measure" (Davies and Ware 1981), incorporating aspects of both physical and mental health (Davies and Ware 1981; Stewart and Ware 1992; McDowell and Newell 1996).

Demographic Variables

Race/ethnicity of self-identified Hispanic respondents was changed from black or white to Hispanic. Education was categorized according to common programs in the U.S. school system. "Income < > poverty" is the amount of income above (or below) poverty, calculated by subtracting the 1987 poverty guideline (based on family size) from reported family income (midpoint of range). Missing income data was imputed as the mean income of subsets (cells) defined by all demographic variables (10.5 percent imputed)². Because the

Table 1: Morbidity Index, Construction and Sample Distribution ($N=19,431$)*

<i>NHIS Variable Description</i>	<i>Original Range</i>	<i>Scale Points</i>	<i>Sample N</i>	<i>Sample %</i>
Restricted activity days in past 2 weeks, at least half a day, due to illness	0	0	17,318	89.1
	1-4	1	1,180	6.1
	5-13	2	437	2.2
	14	3	496	2.6
Doctor visits in past year, all medical care providers, including telephone contacts and allied health practitioners	0	0	9,062	46.6
	1-2	1	4,614	23.7
	3-5	2	2,854	14.7
	6-11	3	1,752	9.0
	12+	4	1,139	5.9
Days in bed due to illness, past year	0	0	10,973	56.5
	1-3	1	4,186	21.5
	4-7	2	1,941	10.0
	8-14	3	971	5.0
	15-30	4	721	3.7
	31-90	5	376	1.9
	91-80	6	160	.8
	181-365	7	103	.5
	Nights hospitalized in past 13 months	0	0	17,765
1-3		1	555	2.9
4-7		2	544	2.8
8-14		3	309	1.6
15-30		4	193	1.0
31+		5	65	.3
Number of conditions, acute or chronic, from four sources: 1. Caused restricted activity or doctor visit past 2 weeks 2. Caused activity limitation 3. Caused hospitalization in past 13 months 4. Ever had condition for more than 3 months, asked for 1 of 6 randomly assigned lists	0	0	9,974	51.3
	1	1	4,805	24.7
	2	2	2,368	12.2
	3	3	1,128	5.8
	4	4	568	2.9
	5	5	437	2.2
	6+	6	151	.8
	Activity limitation in daily work, due to disability	None	0	15,839
Other activities		1	1,253	6.4
Limited in work		2	1,373	7.1
Unable to work		3	966	5.0
Need for help with personal care		None	0	18,109
	IADL	1	1,127	5.8
	ADL	2	195	1.0

Table 1: *Continued*

<i>NHIS Variable Description</i>	<i>Original Range</i>	<i>Scale Points</i>	<i>Sample N</i>	<i>Sample %</i>
<i>Morbidity Index:</i>	0	N/A	4,824	24.8
Sum of scale points for 7 variables listed above, for each individual respondent with no missing values	1		3,270	16.8
	2		2,377	12.2
	3		1,849	9.5
	4		1,498	7.7
	5		1,139	5.9
	6		823	4.2
	7-9		1,616	8.3
	10-12		925	4.8
	13-16		649	3.3
	17-30		461	2.4

*Excludes: Other race/ethnicity, food frequency fatal errors, women with deliveries in past 13 months, and missing morbidity variables.

majority of the sample was “married” and “living with spouse,” marital status and living arrangement variables were combined into a single nine-level indicator.

Behavioral Variables

Diet change describes duration of intentional “lasting and major” dietary change (none, less than one year, less than five years, or longer than five years). Frequency of taking vitamin/mineral supplements (any type) was dichotomized into less than twice a week (or none) versus twice a week or more often. Body mass index (BMI) [weight (kg)/height (m)²] is a measure of body size, or obesity. Alcoholic beverages per week is the sum of frequency of use of beer, wine, and hard liquor. Smoking pack-years was derived by multiplying usual packs per day by number of years since beginning smoking. Table 2a describes the sample distribution of behavioral characteristics, by race/ethnicity and sex.

Modeling System

S-Plus Generalized Additive Models (GAM) was used for model selection, final model analyses, and graphical output (StatSci 1993; Norris 1996), following data preparation using *SAS* (SAS Institute 1989). Continuous

Table 2b: *Continued*

	<i>n</i>	<i>Male</i>	<i>Female</i>	<i>Male</i>	<i>Female</i>	<i>Male</i>	<i>Female</i>	
		6,789	8,782	954	1,643	543	710	
	<i>Percent</i>	35.0	45.2	4.9	8.5	2.8	3.7	
	<i>Percent by race</i>	43.6	56.4	36.7	63.3	43.3	56.7	
	<i>n</i>	<i>%</i>	<i>% by race and sex</i>	<i>% by race and sex</i>	<i>% by race and sex</i>	<i>% by race and sex</i>	<i>% by race and sex</i>	
None	5,844	30.1	21.9	33.6	29.0	44.7	19.7	40.7
Two/day or less	11,802	60.8	61.9	62.5	54.3	51.5	63.2	57.0
More than two/day	1,779	9.2	16.2	4.0	16.7	3.8	17.1	2.3
Body weight/ideal	719	missing						
Under 90%	2,467	13.2	6.4	19.5	5.8	11.0	6.8	17.9
Normal, +/- 10%	8,237	44.0	44.0	46.0	44.3	32.9	47.4	41.8
Heavy, to < 130%	5,282	28.2	36.2	21.2	34.3	28.1	37.6	25.6
Obese, 130%+	2,726	14.6	13.3	13.3	15.6	28.0	8.2	14.7
Duration of diet change								
No change	11,889	61.2	61.8	58.1	69.7	64.2	71.1	68.0
Past 1 year	3,073	15.8	14.2	16.9	13.1	18.6	12.7	17.3
Past 5 years	3,054	15.7	16.3	17.4	12.2	11.0	10.5	8.9
Longer than 5 years	1,415	7.3	7.7	7.6	5.0	6.2	5.7	5.8
Vitamin supplement use								
Never or < 1 month	9,965	51.3	55.6	43.6	67.4	59.6	62.6	55.8
Less than 2/week	2,220	11.4	11.4	11.7	9.8	10.8	11.1	12.4
Twice/week or more	7,246	37.3	33.0	44.7	22.9	29.6	29.3	31.8

*Excludes: Other race/ethnicity, food frequency fatal errors, women with deliveries in past 13 months, and missing morbidity variables.

variable curves were estimated with a nonparametric lowess (locally weighted estimate) smoother, which calculates pointwise regression functions, giving more weight to data nearer to the point being estimated and less weight to more distant data (Hastie and Tibshirani 1990). Lowess functions tend to be robust against undue influence by outlier values and to give smooth curves, particularly in sparse data tails (Roosen 1996). Statistical significance of curvature was assessed with an F-test comparing models with linear and nonlinear versions of the variable of interest (Hastie and Tibshirani 1990; Venables and Ripley 1994).

Generalized additive models functions produce plots of entire models, or of selected model terms, showing the relationship of each independent variable to the outcome, adjusted for all other model variables. Plots for curved response functions reveal thresholds, critical points where the relationship between an independent variable and the outcome changes, and minima or maxima (best- and worst-case points) along curves.

Model Selection

There were a number of ways that sociodemographic and behavioral or health habits terms could be constructed. Selection of alternative terms was based on change in explained variance and statistical significance ($p < .01$) as terms were added to the base model, or later dropped from a larger Morbidity Index model. Variable selection was also influenced by interpretation of graphical displays of models, observing the effect of one variable on another, and reasoned choices between alternative versions of variables with similar meanings. For example, although the curve shapes and significance were very similar, we retained household income relative-to-poverty instead of total household income. Biological interpretation was also considered. While current smoking status and smoking pack-years explained the same variance in preliminary models, smoking pack-years was retained because it attempts to capture cumulative biological insult over time. Outlier values in the tails of the continuous behavioral variables—smoking pack-years, alcohol drinks per week, and BMI—were truncated, or moved in as far as the 99th percentile, to avoid obscuring detail in plotted figures.

A six-term multivariate base model composed of sociodemographic variables—age, sex, race/ethnicity, income, education, and marital status/living arrangement—was defined and evaluated before adding behavioral variables. When education and income are both in the base model (correlation $r = 0.44$), each is reduced in importance, but both remain significant and the change in curve shape when each term is adjusted for the other is interesting in its own right.

The final Morbidity Index full model contained 11 terms after the 5 behavioral terms were added—smoking pack-years, alcoholic beverages per week, body mass index (BMI), duration of intentional diet change, and vitamin supplements per week. Two models for self-assessed health parallel those for the Morbidity Index, a six-term demographic base model and an 11-term full model. For comparison, the health models contain the same terms selected for the morbidity models.

RESULTS

Table 2a presents descriptive health and demographic data by race/ethnicity and sex for all persons with complete morbidity data ($n = 19,431$). Self-assessed health was higher for males than females. Blacks reported poorer health than whites or Hispanics. Morbidity, measured by the 30-point

Morbidity Index, was higher for females than males. Blacks and Hispanics were more likely than whites to report no morbidity (a score of zero on the morbidity index 30-point scale), but blacks were also most likely to score high, 7 or more points. Blacks and Hispanics in the sample are younger than whites, with a higher proportion in the 18–29 age category and relatively fewer 65 and older. Blacks and Hispanics were two to three times more likely than whites to have income below poverty guidelines, and income was lowest for black women.

Table 2b presents the sample distribution by race/ethnicity and sex of the behavioral factors included in full models. Blacks and Hispanics, both males and females, were more likely than whites never to have smoked, but they were also less likely to have quit (Novotny et al. 1988; Hatziafreu et al. 1995), shown by lower prevalence of former smokers and higher prevalence of current smokers. Nevertheless, blacks and Hispanics had accumulated fewer smoking pack-years than whites (Table 3). Regarding alcohol intake, over half of persons in all race/sex categories drank moderately, two drinks per day or less, while blacks of both sexes and Hispanic females are more likely than whites not to drink at all. About 40 percent of sample respondents were overweight (heavy or obese), and black females were most likely to be overweight. Well over half the sample had made no intentional change in diet and blacks and Hispanics were less likely than whites to have made a diet change. Males were less likely than females to take vitamin supplements, while blacks and Hispanics were less likely than whites to take vitamins at all, and less likely to take them twice a week or more.

Demographic Base Models

Statistics from the six-term demographic models for the Morbidity Index and self-assessed health are presented side by side in Table 3, including an R^2 value for the entire model (last row) and an approximation of the unique contribution of each term (the change in R^2 that occurs when that term is *dropped* from the model), labeled $\sim Rsq$. In both demographic base models all six terms—sex, age, race/ethnicity, education, family income, and marital status/living arrangement—are statistically significant ($p < .001$). Most of the explained variance in morbidity is attributable to two terms—age and income. In the self-assessed health base model three terms are primarily responsible for the explained variance—age, income, and education.

Including age, race/ethnicity, education, income, and marital status/living arrangement in multivariate models alters the bivariate patterns seen in

Table 3: Morbidity Index $\hat{}$ and Health $\hat{}$, Six-term Sociodemographic Base Models

Multivariate generalized additive regression models (GAM) are designed to estimate and evaluate curvilinearity. Each term is controlled for all others in the model. Data are from the National Health Interview Survey 1987 Cancer Epidemiology Supplement ($N=17,612$)

1 Variables	2 Levels	Morbidity Model					Health Model					
		Variables			Nonlinearity		Variables			Nonlinearity		
		3 Coef* ~ Rsq	4 sig [†]	5 nl.sig [†]	6 nl.df	7 nl.sig [†]	3 Coef* ~ Rsq	4 sig [†]	5 nl.sig [†]	6 nl.df	7 nl.sig [†]	
1 Sex	Male											
	Female	.66	.005	0	-	-	.07	.001	7.4E-6	-	-	
2 Race	White						.23					
	Black	-.01										
	Hispanic	-.78	.002	4.4E-8	-	-	.00	.004	0	-	-	
3 Marital status:	Never:Alone											
	Never:Others	-.31					.10					
Living arrangement	Mar:Spouse	-.14					.11					
	Div:Alone	.51					-.00					
	Div:Others	1.10					.08					
	Sep:Alone	.09					.08					
	Sep:Others	.12					-.10					
	Wid:Alone	.40	.003	4.8E-9	-	-	.17	.003	1.1E-10	-	-	
	Wid:Others											
4 Education	Elementary											
	Jr. high	-.34					-.12					
	HS grad	-.76					-.34					
	Any college	-.43					-.44					
	College grad.+	-.65	.002	1.5E-9	-	-	-.60	.018	0	-	-	
5 Income \diamond			.015	0	3.1	1.0E-9		.021	0	3.1	0	
6 Age \diamond			.026	0	2.3	.002		.055	0	2.3	4.4E-16	
Base model	Intcpt [‡]	~ Rsq					Intcpt [‡]	~ Rsq				
		3.92	.112	0			2.40	.212	0			

Mar = married; Div = divorced; Sep = separated; Wid = widowed.

$\hat{}$ Morbidity Index (none to 30 points) is a composite of seven NHIS variables. Health is self-assessed (Excellent = 1 to Poor = 5). See Methods section in text.

*Categorical term coefficients contrast the first with subsequent levels.

~ Rsq Approx. R-squared estimated from change in deviance residuals when variable of interest is dropped from 11-term model.

**Significance estimates are from F-tests comparing the 6-term model to a five-term model, dropping the variable of interest.

[†]P-values less than 1.0×10^{-4} are written as 1.0E-4; p-values less than 10E-16 are given as zero.

[‡]Predicted mean value of outcome, the model intercept, is plotted as zero on Y-axis in plots of model terms.

***Nonlinearity significance indicates whether a curved term is statistically preferable to a linear version.

\diamond Continuous functions estimated with *S-Plus* GAM *lo()* smoother (span = 1/2 of data).

[‡]Model intercept, predicted mean value of outcome, is zero on Y-axis in plots of model terms. See Figures.

Table 2a. Higher morbidity and poorer health of females relative to males appears greatly reduced (see coefficients columns, labeled 3, in Table 3). In the health model, poorer self-assessed health reported by blacks compared to whites remains substantial and significant. In the morbidity model, adjusted morbidity for Hispanics is substantially lower than that of whites (see coefficients column in Table 3). After controlling for *income*, black race/ethnicity ceased to be associated with increased morbidity. Kington and Smith (1997) also reported elimination of black–white differences in morbidity after controlling for income and wealth, and McDonough et al. (1997) reported similar findings for mortality, after controlling for income. In all models, standard errors are very wide for the categories of marital status/living arrangement variable, rendering interpretation inconclusive.

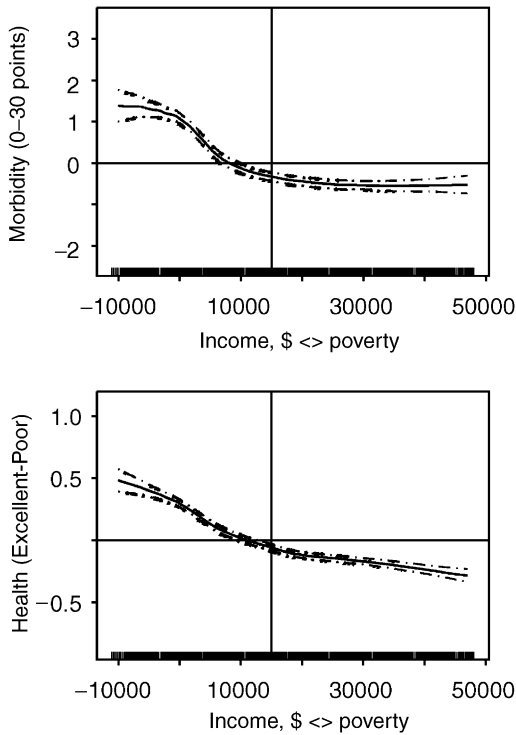
The income curves from Morbidity Index and self-assessed health models differ (see Figure 1). The estimated function is shown in each plot by a solid line, and pointwise standard error estimates are shown by dashed lines. Curve shapes range above and below the mean value of outcome variable, indicated by zero on the vertical Y-axis, and an optional horizontal line intended to help with spatial orientation. A rug on the horizontal X-axis identifies regions of sparse data by appearing threadbare (better seen in subsequent plots). The upper plot in Figure 1 reveals a region of vulnerability to morbidity, from below poverty to about \$15,000 above poverty (marked with a vertical line), after which additional income has no further association with changes in morbidity. Backlund et al. (1996), and McDonough et al. (1997) reported similar curves relating income to mortality. In contrast, increasing income is associated with better health throughout the range of income, although the slope decreases after about \$20,000 above poverty.

The education step function in each model takes on a similar shape to the income curve. Increasing education is related nonlinearly to decreasing morbidity only through high school graduation (see coefficients column in Table 3). There is a notable uptick in morbidity for persons who have completed some college, which would include dental and medical assistant training, and many trades. In contrast, every higher level of education is associated with better self-assessed health.

Behavioral Variables

Five behavioral variables—body mass index (BMI), smoking pack-years, alcoholic beverage intake, intentional diet change, and vitamin/mineral

Figure 1: Income in Morbidity and Health Models



Multivariate models also include sex, age, race/ethnicity, and marital status/living arrangement. Y-axes show zero at sample mean. Continuous functions (solid lines) and pointwise standard errors (dashed lines) estimated with *S-Plus* GAM *lo()* smoother (span = .5). See Table 3 for model statistics ($N = 17,612$).

supplement use—were added to each six-term demographic base model—sex, age, race/ethnicity, education, income, and marital status/living arrangement—resulting in full models with 11 independent variables. Changes in the coefficients, significance, and shapes of the six demographic terms, comparing full models to base models, reflect the cumulative impact of the second set of covariates (seen by comparing the same term for each model in Tables 3 and 4).

The continuous health habit variables—BMI, smoking pack-years, and alcoholic beverage intake—are statistically significant in full models for both outcomes, self-assessed health and the Morbidity Index. All three are curved

Table 4: Morbidity Index $\hat{\alpha}$ and Health $\hat{\alpha}$ models, after Adding Health Behavior Variables
 Multivariate generalized additive regression models (GAM) are designed to estimate and evaluate curvilinearity. Each term is controlled for all others in the model. Data are from the National Health Interview Survey 1987 Cancer Epidemiology Supplement ($N = 17,612$)

Variables	1	2	Morbidity Model				Health Model					
			Variables		Nonlinearity		Variables		Nonlinearity			
			3	4	5	6	7	3	4	5	6	7
Sex			Coef*	~ Rsq	sig [†]	nLdf	nLsig [†]	Coef*	~ Rsq	sig [†]	nLdf	nLsig [†]
1	Female	Male	.53	.002	6.9E-12	-	-	.08	.0006	.0003	-	-
2	Race	White	.04									
		Black						.23				
		Hispanic	-.61	.001	2.1E-5	-	-	.04	.004	0	-	-
3	Marital status:	Never:Alone	-.28					.10				
	Living arrangement	Never:Others	-.17					.09				
		Mar:Spouse	.41					-.03				
		Div.:Alone	-.06					-.05				
		Div.:Others	1.13					-.08				
		Sep.:Alone	-.06					.13				
		Sep.:Others	.18					-.09				
		Wid.:Alone	.37	.003	1.1E-8	-	-	.15	.002	6.8E-9	-	-
4	Education	Wid.:Others										
		Elementary	-.46					-.14				
		Jr. high	-.86					-.34				
		HS grad	-.64					-.44				
		Any college	-.83	.002	1.5E-10	-	-	-.58	.015	0	-	-
		Col. grad.+										

continued

Table 4: Continued

Variables	Morbidity Model							Health Model									
	Variables			Nonlinearity				Variables			Nonlinearity						
	1	2	3	4	5	6	7	3	4	5	6	7	3	4	5	6	7
5 Income \diamond				.014	0	3.1	1.0E-10		.018	0	3.1	1.0E-10				3.1	1.0E-10
6 Age \diamond				.016	0	2.3	2.5E-7		.033	0	2.3	1.3E-6				2.3	1.3E-6
7 Smoking pack-years \diamond				.004	0	3.0	7.8E-4		.007	0	3.0	7.0E-12				3.0	7.0E-12
8 Alcoholic beverages/week \diamond			.002	6.0E-8	3.8	4.7E-7		.004	3.3E-16		3.8	1.0E-10				3.8	1.0E-10
9 Body mass index (BMI) \diamond			.007	0	2.5	1.0E-10		.01	0		2.5	1.0E-10				2.5	1.0E-10
10 Diet change:																	
<1 year			1.80										.24				
<5 years			1.15										.17				
5+ years			1.08		.025	0	-		.14	.007	0	-					
LT 2/wk.																	
2+/week			.47	.003	4.2E-13	-	-	.03	.0002	.06	-	-					
Vitamin supplements:			Intcpt [‡]	~Rsq	0			Intcpt [‡]	~Rsq	0							
Full model			3.39	.156				2.33	.241								

Mar = married; Div = divorced; Sep = separated; Wid = widowed.

[^]Morbidity Index (none to 30 points) is a composite of seven NHIS variables. Health is self-assessed (Excellent = 1 to Poor = 5). See Methods section in text.

*Categorical term coefficients contrast the first with subsequent levels.

~Rsq Approx. R-squared estimated from change in deviance residuals when variable of interest is dropped from 11-term model.

**Significance estimates are from F-tests comparing the 11-term model to a 10-term model, without the variable of interest.

[†]P-values less than 1.0×10^{-4} are written as 1.0E-4; p-values less than 10E-16 are given as zero.

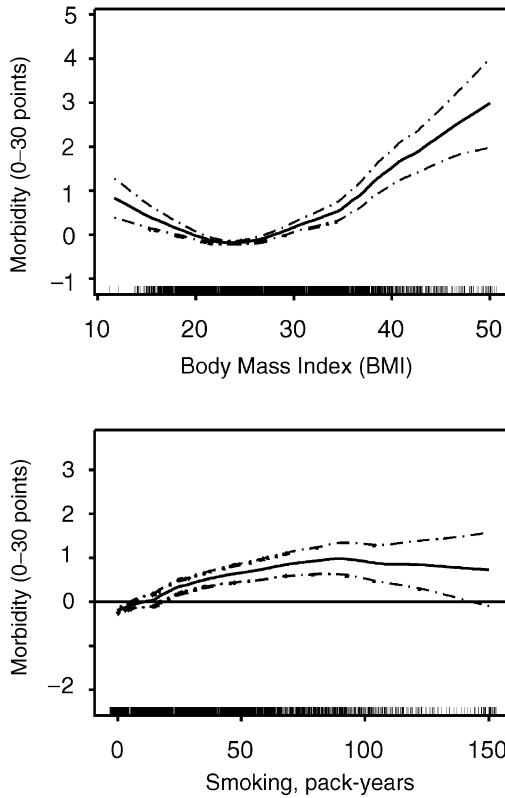
[‡]Predicted mean value of outcome, the model intercept, is plotted as zero on Y-axis in plots of model terms.

***Nonlinearity significance indicates whether a curved term is statistically preferable to a linear version.

\diamond Continuous functions estimated with S-Plus GAM lo0 smoother (span = 1/2 of data).

[‡]Model intercept, predicted mean value of outcome, is zero on Y-axis in plots of model terms. See Figures.

Figure 2: BMI and Smoking in Morbidity Full Model



Body Mass Index and smoking pack-years from multivariate model with nine other variables. Y-axes show zero at sample mean. Continuous functions (solid lines) and pointwise standard errors (dashed lines) estimated with *S-Plus* GAM *lo()* smoother (span = .5). See Table 4 for model statistics ($N=17,612$).

terms, and the curvature is statistically significant ($p < .01$ or smaller), shown in the columns labeled 7 in Table 4.

The shapes of two of the curves, for smoking pack-years and BMI, are very similar in both morbidity and health models. They are shown for only the morbidity model in Figure 2. Body mass index (upper curve) has an almost parabolic shape with a low point or minima. The healthiest people in this sample, on average, are those reporting BMI, or “body size,” in the range of 22 to 24 kg/m^2 , which coincides with current optimum (BMI) recommendations

from the Centers for Disease Control (2003). Manson et al. (1995) reported a similar parabolic curve relating BMI to mortality among women. High BMI is a proxy for long-term habits of food intake and physical activity, as well as a risk factor for chronic disease. The low end of the BMI curve should be interpreted with caution, as it seems likely to represent involuntary weight loss following the onset of morbidity.

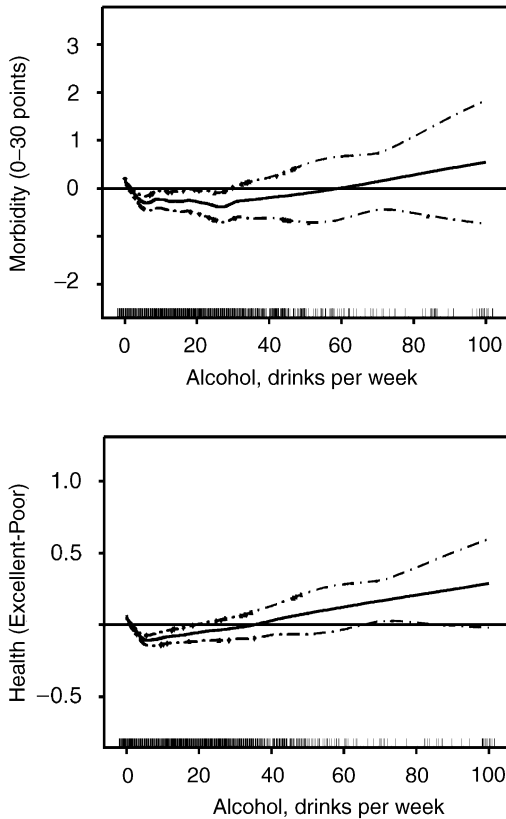
In both models, the smoking pack-years curve (Figure 2, lower curve) increases almost linearly to a point of maximal impact and then flattens out, indicating no further effect of smoking. In the morbidity model the maximal association is at about 80 pack-years, and in the health model, the maximal association occurs at about 50 pack-years (curve not shown).³

The alcohol intake term takes on a U- or J-shape in both models (see Figure 3 for both curves). Individuals who consume 5 to 25 drinks per week (upper curve) experienced the least morbidity, and persons consumed 5 drinks per week (crisply defined in the lower curve) reported the best health. Higher or lower intake is associated in both models with worse morbidity or health. Others have reported, discussed, and debated the true meaning of alcohol J-curves related to morbidity and mortality from cardiovascular and other conditions (reviewed by Shaper and Wannamethee 1998), and to all-cause mortality (Liao et al. 2000).

Diet change is strongly related to morbidity ($\sim Rsq = .021$ in Table 3), equaling or exceeding the explanatory power of any other behavioral variable. However, the most recent diet change (past year) is associated with higher morbidity and lower health, suggesting behavior change after onset of morbidity. Vitamin supplement use (twice a week or more often) is also significantly related to higher morbidity, but is unrelated to self-assessed health.

The mean (or median) values of the behavioral variables in these models represent not smoking, very moderate alcohol intake, normal body weight, and dietary patterns that do not require changing or use of vitamin supplements, which could be collectively described as "good health habits." Adding behavioral variables to demographic base models increased the explained variance in the Morbidity Index by 39 percent ($R^2 = 0.112$ to 0.156, last row in Tables 4). The contribution of behavioral variables to morbidity is similar to the impact of chronic disease risk factors on Medicare expenditures (28–40 percent) in the Framingham population (Schauffler 1989) and the impact of health practices on disability (33–50 percent) in the Alameda County study (Breslow and Breslow 1993). Behavioral terms added less explained variance (14 percent) to the Self-assessed Health full model

Figure 3: Alcohol in Morbidity and Health Full Models

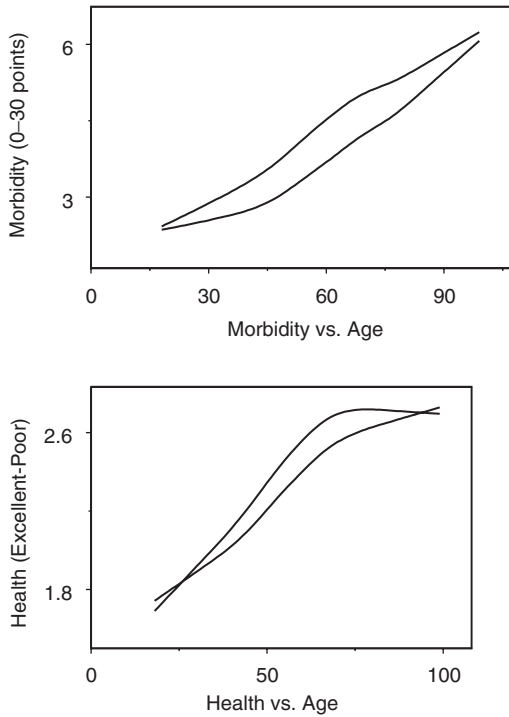


Adjusted for six demographic terms and four other behavioral variables. Y-axes show zero at sample mean. Continuous functions (solid lines) and pointwise standard errors (dashed lines) estimated with *S-Plus* GAM *lo()* smoother (span = .5). See Table 4 for model statistics ($N = 17,612$).

($R^2 = 0.212$ to 0.241 , last row in Table 4), similar to the contribution of health behaviors to mortality (12–13 percent) in the American’s Changing Lives Survey (Lantz et al. 1998).

Adding behavioral variables did not reduce the importance of income in the Morbidity Index model (compare unique $\sim Rsq$ values in Tables 3 and 4). In the self-assessed health model, after adding behavioral variables education and income remained prominent (see Tables 3 and 4). Moreover, plotted

Figure 4: Age versus Morbidity and Health, Before and After Adding Health Behavior



Age terms from 6-term demographic base model (left curve) and 11-term full model (right curve) are overlaid for the Morbidity Index outcome (upper graph) and self-assessed health (lower). Narrow standard errors are not shown. See Tables 3 and 4 for model statistics ($N=17,612$).

shapes and ranges of the step functions for sex, race/ethnicity, and education and the income curves from the base and full models for each outcome are virtually identical when plots are overlaid (not shown due to space limitations).

In both health and morbidity models only one term was substantially changed by the addition of behavioral variables, the age term. In the Morbidity Index full model, adding behavioral variables reduced the age contribution by half (compare Tables 3 and 4). Figure 4 shows age curves, overlaid and without standard errors, before and after the addition of health habits, for Morbidity Index (top graph) and self-assessed health models. Both age curves for each outcome are plotted on the same scale, shown on the

Y-axis in natural units of the outcome variable.⁴ After adding behavioral variables, the full-model age curve (right-most curve in both graphs) slumps to the right. Comparing Morbidity Index age curves (upper graph), 80-year-olds reporting “good health habits” reported morbidity at a level associated with 65-year-olds. The corresponding curves for self-assessed health (lower graph) show about a 10-year delay in perceived declining health associated with “good health habits.”

DISCUSSION

The first purpose of this research was exploratory examination of nonlinearity in the relationships of demographic determinants and behavioral factors with morbidity and health outcomes. Every one of the five continuous variables in the 11-term full models exhibited significant and meaningful curvilinearity, and categorical terms with multiple levels, such as education, exhibited nonlinear step patterns suggestive of curves. The shapes of the curves for age, income, smoking pack-years, alcohol intake, and BMI, as well as the step-function categorical variable for education, are consistent with previously reported findings from studies of health, morbidity, and mortality. The new information from this study is the revelation that all of these curves can be found in U.S. population data, in a single multivariate additive model, using a nonparametric smoother.

The second purpose of this study was to compare determinants of morbidity and of health in the U.S. population. In this study, full models explain more of health than morbidity (R^2 0.241 versus 0.156). The total contribution of behavioral factors is similar in morbidity and health models, and the shapes of corresponding model terms are similar. It is the impact of socioeconomic factors—income and education—that differs for morbidity and health. The relationship between health and socioeconomic status is such that every higher level of education and every additional dollar of income are associated with better self-assessed health. In contrast, morbidity is concentrated in the portion of the population with less than high school education and with income below poverty to about \$15,000 above poverty, that is, among those with household income below the median.

Nonlinearity in the relationship of income and education to morbidity and health is important from a policy perspective, because it suggests that one could channel public dollars or policy investments to individuals likely to return the highest marginal health benefit. From a research perspective,

correctly specified models are necessary for obtaining unbiased results from every analysis in which income–health relationships are important.

Limitations of the Study

The Morbidity Index composite variable appears to be a reasonable measure of general morbidity, which is related to known demographic determinants of morbidity, as expected (Aday 1993). Creating the Morbidity Index by assigning scale points permitted selective emphasis on contributing variables, thereby limiting the weight of NHIS services utilization measures—doctor visits and hospitalizations—known to be constrained by lack of health insurance (Keith and Jones 1990). Even so, the Morbidity Index may underestimate morbidity for blacks and Hispanics, who have lower utilization of health care services associated with less access to health insurance (LaPlante 1993).

All morbidity and behavioral variables were self-reported and thus subject to measurement error from failure of memory or selective reporting, including those used to construct the Morbidity Index. Wide standard errors in the tails of behavioral variables probably represent a combination of measurement error and wide variability in natural susceptibility to biological insult. Underreporting of either morbidity or behavioral measures would result in underestimation of the effects of behavioral factors, so that these results probably represent a lower bound for the extent of relationships between behavioral variables and health or morbidity.

We selected the 1987 NHIS Cancer Epidemiology Supplement data because of the unique range of variables, particularly diet change measures, which have not been repeated in subsequent national surveys. The dataset lacks several variables likely to be important in health and morbidity models, including additional indicators of income adequacy (such as expenses for housing, transportation, and medical costs), health insurance status, and physical activity measures. Lack of important determinants of morbidity and health care utilization may contribute to the modest amount of variance explained by the morbidity model, in particular.

Changes in the NHIS since 1987, including increasing the sample size for Hispanics, should have little affect on interpretation of the results of this study. Both blacks and Hispanics were oversampled in 1987, analogous to current sampling procedures, and questions used for identifying Hispanic respondents remain the same. The findings presented here can be compared to subsequent analyses of NHIS 1998 and 2000 Prevention and Cancer Control Supplement data, providing a basis for reexamining over time the major nonlinear

relationships between morbidity, or health, and socioeconomic variables. How will the relationship of income to morbidity change, in the context of decreasing coverage and value of welfare supports, including Medicaid, decreasing wages and job security, and increasing income inequality?

This study assumes that demographic factors and health-related behaviors remain stable over long periods of time and that health behavior patterns reported in the NHIS preceded development of morbidity. While assumptions of precedence and continuity seem valid for behavioral terms like smoking pack-years, diet change, and at least the high side of the BMI curve, they may not be equally valid for alcohol intake or vitamin use. Diet change, in particular, may be not a cause, but rather a result, of declining health or perceived morbidity. A more complex model might be able to isolate predictors of beginning to make diet changes from the presumably beneficial effects of diet change. The additive models presented here do not attempt to deal conclusively with such endogeneity; instead, the nonlinear estimation functions reveal and describe problem areas where data are sparse, or standard errors are wide, or curves go the wrong way, or the existence of curves undoes linear model assumptions.

We chose the nonparametric lowess smoother because it reveals natural relationships in the data, but the trade-off is that estimates generated by these models may not be quite right. Smoother calculations are, practically speaking, incompatible with variance correction software and population weights necessary for calculation of precise population estimates (NCHS 1989). However, having found the shapes and critical points of the curvilinear terms with nonparametric estimations, we could redefine the models with parametric terms or splined regression terms, and use the “simplified” models with appropriate weights and variance correction to derive true population values. For example, Backlund et al. (1996) used a natural log curve to approximate a curve relating income to mortality, analogous to the income curve in our Morbidity Index model. Nevertheless, the magnitude and shapes of determinants of morbidity and health in these models are consistent with findings from longitudinal studies investigating mortality and a wide range of morbidity outcomes.

We now return to the third purpose of this study, examining the relative contributions of demographic and behavioral factors to morbidity and health. The two most important determinants of morbidity and health in these models are age and income, as reported in many other studies. The curve for income reveals the worst morbidity among those with inadequate income, at and below poverty, and the adverse relationship continues until income exceeds

about \$15,000 above poverty, or about median income for the population. The education curve shows a similar adverse relationship between morbidity and a less than high school education. Despite social welfare programs—social security, Temporary Assistance for Needy Families, unemployment insurance, Medicaid, Medicare, SSI/ SSDI disability programs, WIC, food stamps, and food pantries—poverty and inequality in morbidity persist. For most families receiving public assistance, total income (including assistance) was still below poverty (Ohls and Beebout 1993).

The second goal listed in *Healthy People 2010* (U.S. Department of Health and Human Services 2000) is to eliminate disparities in health among segments of the U.S. population. These models appear to show a delay in the effects of aging associated with adherence to “good health habits,” including not smoking, low alcohol use, avoiding obesity, and adequate dietary intake (not in need of change) without the use of vitamin supplements. Health habits themselves are also unequally distributed by socioeconomic status (Goldstein 1992; Winkleby et al. 1992), with some adverse behaviors more prevalent among low-income subpopulations and others not. While findings from these models support the public health utility of continuing efforts to improve health behavior in all segments of the population, our findings, like those of Lantz et al. (1998), indicate that improving health behavior is not sufficient to offset the pervasive and powerful effects of socioeconomic factors.

Commonly reported black–white differences in morbidity and health care utilization are among the largest and most persistent disparities in the United States. In these models, income, not health habits, “explains away” black–white differences in morbidity. This interaction between black race/ethnicity and the income term reflects the fact that virtually the entire income distribution for blacks was lower than the median income for whites (data not shown). These models suggest that the underlying causes of health disparities are disparities in the distribution of income and education, which provide the means to avoid situations resulting in morbidity and to take up practices contributing to a high level of self-assessed health (Guralnik et al. 1993; Guralnik and Leveille 1997).

Both the shape and the magnitude of the income term in the Morbidity Index model suggest that we must concentrate on the low end of the income distribution to address inequality in morbidity. Those with a high school education or less, less than median income, and nonwhite race/ethnicity are at greatest risk of morbidity and ill health, and are most in need of supportive public health and social welfare policy. Reduction or elimination of disparities in health and morbidity may depend on elimination of poverty, establishment

of universal health insurance with coverage for health promotion, behavioral health treatment, and occupational health services. These models suggest that a successful attack on disparities in health may further require universal access to higher education and reduction in disparities in income.

NOTES

1. The NHIS 1987 Cancer Epidemiology Supplement questionnaire (NCHS 1988) included the NCI/Block 59-item food frequency and vitamin/mineral supplement questionnaire. Similar data from the 1992 NHIS/CES survey were available for about one-half the sample size, but did not include the diet change questions. The 1998 Adult Prevention Supplement dietary behavior questions were limited to intent to lose or maintain weight, weight control methods, and frequency of purchasing “low salt” or “low sodium” food products, reading nutrition labels, and salting food. The 2000 Cancer Control Supplement again collected food frequency and vitamin supplement data, but did not repeat the diet change questions.
2. This nonparametric method generated fewer erroneous estimates than regression of income on the same demographic predictors, assessed by comparing results for respondents who failed to give detailed income information, but did report whether their income was less than or greater than \$20,000. In more than 80% of missing income cases for which the \$20,000 comparison was possible, household income was lower than the expected (mean) value for the demographic subset.
3. Two alternatives smoking variables were considered for these models, smoking status (never, former, current) and smoking pack-years. Morbidity increased, and health status decreased, step-wise, from lowest for never smokers to highest for current smokers. Curve-shapes for other model terms were the same with either smoking variable.
4. The full model curve is positioned (offset) to correspond to the model intercept (the adjusted overall mean of the response variable) from the sociodemographic base model. The adjusted, overlaid age plots were accomplished outside the GAM system, by a series of steps similar to the processes inside the system, by extracting the additive predictor for age from each model object, adding the appropriate constant term to obtain uncentered response value units, then adding the difference between the model intercepts to the full model additive predictor. The adjusted additive predictors were plotted against the original age variable with a different smoother (Friedman’s Supersmoother), which closely approximates the appearance and placement of the curves produced by the GAM loess smoother which estimated the original additive predictor functions for the age terms.

REFERENCES

- Aday, L. A. 1993. “Indicators and Predictors of Health Services Utilization.” In *Introduction to Health Services*, edited by S. J. Williams and P. R. Torrens, chap. 3, pp. 46–70. Albany, NY: Delmar.

- Adler, N. E., and J. M. Ostrove. 1999. "Socioeconomic Status and Health: What We Know and What We Don't." *Annals of the New York Academy of Sciences* 896: 3-15.
- Backlund, E., P. D. Sorlie, and N. J. Johnson. 1996. "The Shape of the Relationship between Income and Mortality in the United States: Evidence from The National Longitudinal Mortality Study." *Annals of Epidemiology* 6: 16-20.
- Breslow, L., and N. Breslow. 1993. "Health Practices and Disability: Some Evidence from Alameda County." *Preventive Medicine* 22 (3): 86-95.
- Centers for Disease Control (CDC). 2003. "BMI: Body Mass Index for Adults." Available at <http://www.cdc.gov/nccdphp/dnpa/bmi/bmi-adult.htm> [accessed June 20, 2003].
- Davies, A. R., and J. E. Ware, Jr. 1981. *Measuring Health Perceptions in the Health Insurance Experiment*. Health Insurance Experiment Series report no. R-2711-HHS. Santa Monica, CA: Rand.
- Goldstein, M. A. 1992. *The Health Movement: Promoting Fitness in America*. New York: Twayne.
- Guralnik, J. M., K. C. Land, D. Blazer, G. G. Fillenbaum, and L. G. Branch. 1993. "Educational Status and Active Life Expectancy among Older Blacks and Whites." *New England Journal of Medicine* 329 (2): 110-16.
- Guralnik, J. M., and S. G. Leveille. 1997. "Annotation: Race, Ethnicity, and Health Outcomes: Unraveling the Mediating Role of Socioeconomic Status." *American Journal of Public Health* 87 (5): 5-6.
- Hastie, T. T., and R. J. Tibshirani. 1990. *Generalized Additive Models. Statistical Models in S*. London: Chapman and Hall.
- Hatziandreu, E. J., J. P. Pierce, M. Lefkopoulou M, M. C. Fiore, S. L. Mills, T. E. Novotny, G. A. Giovino, and R. M. Davis. 1995. "Quitting Smoking in the United States in 1986." *Journal of the National Cancer Institute* 82 (17): 1402-6.
- Heckler, M. M. 1985. *Report of the Secretary's Task Force on Black and Minority Health, Volume I: Executive Summary*. GPO no. 017-090-00078-0. Washington, DC: U.S. Department of Health and Human Services.
- House, J. S., R. C. Kessler, A. R. Herzog, R. P. Mero, A. M. Kinney, and M. J. Breslow. 1990. "Age, Socioeconomic Status, and Health." *Milbank Quarterly* 68 (3): 383-411.
- Keith, V. M., and W. Jones. 1990. "Determinants of Health Services among the Black and White Elderly." *Journal of Health and Social Policy* 1 (3): 73-89.
- Kington, R. S., and J. P. Smith. 1997. "Socioeconomic Status and Racial and Ethnic Differences in Functional Status Associated with Chronic Diseases." *American Journal of Public Health* 87 (5): 805-10.
- Kumanyika, S. 1990. "Diet and Chronic Disease Issues for Minority Populations." *Journal of Nutrition Education* 22 (2): 89-96.
- LaPlante, M. L. 1993. *Disability, Health Insurance Coverage, and Utilization of Acute Health Services in the United States*. Washington, DC: U.S. Department of Health and Human Services (National Institute on Disability and Rehabilitation Research).
- Lantz, P. M., J. S. House, J. M. Lepkowski, D. R. Williams, R. P. Mero, and J. Chen. 1998. "Socioeconomic Factors, Health Behaviors, and Mortality: Results from a

- Nationally Representative Prospective Study of U.S. Adults." *Journal of the American Medical Association* 279 (21): 1703–8.
- Liao, Y., D. L. McGee, G. Cao, and R. S. Cooper. 2002. "Alcohol Intake and Mortality: Findings from the National Health Interview Surveys (1999 and 1990)." *American Journal* 151 (7): 651–9.
- Liberatos, P., B. G. Link, and J. L. Kelsey. 1988. "The Measurement of Social Class in Epidemiology." *Epidemiology Reviews* 10: 87–121.
- McDonough, P., G. J. Duncan, D. Williams, and J. House. 1997. "Income Dynamics and Adult Mortality in the United States, 1972 through 1989." *American Journal of Public Health* 87: 1476–83.
- McDowell, I., and C. Newell. 1996. *Measuring Health: A Guide to Rating Scales and Questionnaires*, 2d ed. New York: Oxford University Press.
- Manson, J. E., W. C. Willett, M. J. Stampfer, G. A. Colditz, D. J. Hunter, S. E. Hankinson, C. H. Hennekans, and F. E. Speizer. 1993. "Body Weight and Mortality among Women." *New England Journal of Medicine* 333 (11): 677–85.
- Myers, H. F., M. Kagawa-Singer, S. K. Kumanyika, B. W. Lex, and K. S. Markides. 1995. "Behavioral Risk Factors Related to Chronic Diseases in Ethnic Minorities." *Health Psychology* 14 (7): 613–21.
- National Center for Health Statistics (NCHS), 1989. *Design and Estimation for the Health Interview Survey, 1985–94. Vital and Health Statistics, Series 2: Data Evaluation and Methods Research*, no. 110 (U.S. DHHS publication no. [PHS] 89-1384). Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Center for Health Statistics.
- Norris, J. C. 1996. "Nutrition, Morbidity and Lost Productive Time in the 1987 National Health Interview and Cancer Epidemiology Supplement." Dissertation, University of California, Berkeley.
- Novotny, T. E., K. E. Warner, J. S. Kendrick, and P. L. Remington. 1988. "Smoking by Blacks and Whites: Socioeconomic and Demographic Differences." *American Journal of Public Health* 78 (9): 1187–9.
- Ohls, J. C., and H. Bebout. 1993. *The Food Stamp Program: Design, Trade-offs, Policy and Impacts: Mathematica Policy Research Study*. Washington, DC: Urban Institute Press.
- Polednak, A. P. 1989. *Racial and Ethnic Differences in Disease*. New York: Oxford University Press.
- Roosen, C. 1996. *S-Plus Technical Support Telephone Conversation*. Seattle, WA: StatSci.
- SAS Institute. 1989. *SAS* (version 6.09). Cary, NC: SAS Institute.
- Schauffler, H. H. 1989. "Modifiable Risk Factors and Medicare Reimbursements." Dissertation, Brandeis University, the Florence Heller Graduate School for Advanced Studies in Social Welfare, Waltham, MA.
- Shaper, A. G., and S. G. Wannamethee. 1998. "The J-Shaped Curve and Changes in Drinking Habit." *Novartis Foundation Symposia* 216: 173–88.
- StatSci. 1993 and 2001. *S-Plus* (version 3.2) for Unix and *S-Plus* (version 6.1) for Windows. Seattle, WA: A Division of MathSoft (now called Insightful Corporation).

- A. L. Stewart, and J. E. Ware Jr., eds. 1992. *Measuring Functioning and Well-Being: The Medical Outcomes Study Approach*. Durham, NC: Duke University Press.
- Streiner, D. L., and G. R. Norman. 1995. *Health Measurement Scales: A Practical Guide to Their Development and Use*, 2d ed. New York: Oxford University Press.
- U.S. Department of Health and Human Services. 2000. *Healthy People 2010: Understanding and Improving Health*, 2d ed. Washington, DC: U.S. Government Printing Office. Available at http://www.health.gov/healthypeople/document/html/uih/uih_bw/uih_2.htm#goals.
- Venables, W. N., and B. D. Ripley. 1994. *Modern Applied Statistics with S-Plus*, 2d ed. New York: Springer-Verlag.
- Winkleby, M. A., D. E. Jatulis, E. Frank, and S. P. Fortmann. 1992. "Socioeconomic Status and Health: How Education, Income, and Occupation Contribute to Risk Factors for Cardiovascular Disease." *American Journal of Public Health* 82 (6): 816–20.