Alameda County Redux: Replication in Michigan[1,2]

BRUCE M. BROCK, PH.D., M.P.H.,* DON P. HAEFNER, PH.D.,†[,3] AND
DAVID S. NOBLE, M.P.H.‡

*Information Transfer Systems, Inc., 2451 S. Industrial Hwy., Ann Arbor, Michigan 48104,
†Department of Health Behavior and Health Education, School of Public Health,
The University of Michigan, Ann Arbor, Michigan 48109, and ‡Market Opinion Research, Inc.,
550 Washington Blvd., Detroit, Michigan 48226

This article describes an attempt to reproduce, in Michigan, cross-sectional findings concerning the relationship between health habits and health status previously obtained in Alameda County, California by Belloc and Breslow. Data for this study were gathered by a telephone interview of a state-wide sample of 3,259 adult Michigan residents. The ridit (relative to an identified distribution) analyses presented include a comparison of findings from the Michigan and Alameda County studies. In Michigan, as in Alameda County, health status was found to be associated with various health practices, both individually and in combination. Consistent relationships were found between physical health status and individual health practices regarding hours of sleep, eating breakfast, eating between meals, cigarette smoking, weight for height, and physical activity. Physical health status was also linked to the overall number of health practices individuals engaged in. When the confidence interval for each ridit value was taken into account, however, only some of the findings proved statistically reliable. © 1988 Academic Press, Inc.

INTRODUCTION

In 1965, the Human Population Laboratory of the California State Department of Public Health surveyed almost 7,000 individuals representing a probability sample of adult residents of Alameda County, California by mail. Based on these survey findings, Belloc and Breslow published a now-classic article on the relationship between seven lifestyle-related health practices and physical health status (2). The health practices included number of hours of sleep, physical activity, alcohol consumption, cigarette smoking, eating breakfast, snacking between meals, and weight in relation to height. Physical health status was classified along a spectrum ranging from severe physical disability at one end through various levels of health to physical vigor at the other end. From the survey findings, the authors concluded that there was indeed a link between health status and the various health practices both separately and cumulatively.

Since that 1965 cross-sectional survey, the Alameda County sample has been followed as a panel. Based on these latter efforts, findings have been reported concerning the relationship of health practices to subsequent morbidity and mortality within that sample (1, 4, 14). The longitudinal results support the notion that for Alameda County residents, health practices had a causal impact on health outcomes.

---

[1] An earlier version of this paper was presented at the 110th Annual Meeting of the American Public Health Association, November 15, 1982, Montreal, Quebec, Canada.

[2] This research was supported in part by the Michigan Department of Public Health.

[3] To whom reprint requests should be addressed.

To date, only a few articles have appeared describing attempts to generalize the original Alameda County cross-sectional findings by studying other populations (6, 10, 11, 12, 15). These studies have yielded mixed results. On one hand, the total number of favorable health habits practiced has usually been found to be positively associated with physical health level, although not always to a statistically significant extent. The separate health habits, however, with the exception of controlling one's weight, have yielded inconsistent relationships to physical health.

It should be noted that these studies typically employed constructive rather than exact replication: They used different measures of the same concepts and/or different analytic procedures than those used in the Belloc and Breslow study rather than attempting to duplicate its procedures as exactly as possible. This creates some problems in assessing the implications of the findings, particularly those that are at variance. Such disparate findings could represent genuine differences in relationships between variables, differences in outcome due to different methods of measurement and/or analysis, or a combination of these factors.

This study sought to determine whether the cross-sectional findings concerning health habits and health status originally obtained in Alameda County could be reproduced elsewhere when the measurement and analytic procedures employed by Belloc and Breslow were duplicated as closely as possible. This report compares findings of a 1978 state-wide telephone survey in Michigan with those from the original 1965 Alameda County study.

METHODS

The Sample

The sample was developed in conjunction with a series of studies concerning the impact of polybrominated biphenyl (PBB) contamination of the food chain in Michigan. Selection for telephone interviewing was accomplished using the two-stage Waksberg method of random digit dialing (RDD) (13). In the first stage, computer-generated lists of random numbers were combined with telephone exchange prefixes; clusters were then built around residential household telephone numbers identified in the first stage. These second-stage numbers within clusters were eligible for interviews. The RDD approach was used for several reasons: usefulness in collecting data during inclement weather, availability of expertise in its application through the Survey Research Center of the University of Michigan, and existence of residential telephone service estimated to cover 95% of the Michigan population.

One adult was selected randomly from each of the 3,297 cooperating households through the use of Kish tables (8). Adults selected into the sample represented a response rate of about 83% of attempted telephone interviews. Of the total Michigan sample, 3,259 yielded sufficient information for inclusion in the data analyses. To compensate for the different individual probabilities of selection due to sampling procedures, all of the Michigan analyses used weighted data.

Given the large sample N's for both Michigan and Alameda County, even trivial demographic differences of a couple of percentage points were found to be sta-

tistically significant ($P < 0.05$), using a z test. We chose to treat only differences of 5% or more as substantial enough to warrant mention.

Both the Michigan and the Alameda County samples had nearly identical sex distributions, each containing slightly more women than men. Age distributions for the two samples were also quite similar, with one exception: the percentage of Michigan respondents in the youngest age group, 24 years or less, was nearly twice that of the Alameda County sample (22.4% vs 12.2%). In the Michigan study, the overall objectives of the PBB research led to including as adults persons age 18 or older or who had ever been married, whereas in the Alameda County study, persons 20 or older or 16–19 years old and ever married were included. In preliminary analyses of the Michigan data, minimal differences were found for comparisons involving one age cutpoint vs the other. Therefore, all available cases were included in order to increase the precision of estimates for the youngest age group.

It was not feasible to make a direct comparison of the two samples on income level. The income measure used in Alameda County would have required major modifications for telephone administration, and its response options were too skewed toward lower income levels in view of the inflation that had occurred in the 13 years between the two studies.

The Alameda County and Michigan samples showed both similarities and differences in other demographic factors. The two samples were generally similar in employment status, each containing slightly over 50% employed full-time, and about 25% who classified themselves as homemakers. Other demographic variables, however, yielded clear-cut differences between the two locations. The populations from which the samples were drawn differed markedly in their urban–rural composition. In the U.S. decennial censuses of both 1960 and 1970, bracketing the 1965 Alameda County study, approximately 99% of that county's residents lived in urban areas. In contrast, the 1980 U.S. census, conducted 2 years after our study, found only 70% of Michigan residents living in urban areas. Computation of confidence intervals taking into account population values and sample sizes yielded a likelihood of better than 99.9% that the two samples differed by 24% or more on this characteristic. With regard to education, the Alameda County sample contained substantially more respondents who completed grade school only and fewer who completed high school than was true of the Michigan sample. With respect to marital status, Alameda County contained substantially more currently married respondents and fewer never-married ones than did Michigan.

In view of the geographic and demographic differences between the two samples, the authors felt that if the Michigan study yielded results consistent with those from Alameda County, it would lend substantial support to the generalizability of the original findings.

The Physical Health Spectrum: The Measure of Health Status

In an article prior to the one by Belloc and Breslow on health habits and health status, Belloc et al. (3) discussed the rationale underlying development of the physical health spectrum. What they sought was the ability to array respondents

across a spectrum from invalidism at one end through various health levels to physical vigor at the other end. Their seven-category operationalization of this spectrum, employed in the Alameda County study, also served as the measure of health status in the Michigan study.

The health spectrum as defined by Belloc *et al.* has intuitive appeal, particularly since it attempts to assess differences that might influence an individual's ability to function productively in society. The spectrum does have certain limitations as acknowledged by its authors; their measures by no means exhaust the data that could be incorporated into a health index since they do not include information from screening tests. Furthermore, severity and control status of chronic conditions included in the spectrum have not been taken into account. But even with such limitations, the spectrum serves as a useful measure of health status, although not the only possible one.

In order to permit meaningful comparisons between the Michigan and original Alameda County studies, the original form of health status questions was preserved wherever possible. In a very few instances, slight modification of item wording was required and a few rare chronic conditions (e.g., epilepsy) were omitted. Researchers from the Human Population Laboratory provided helpful advice in making these minor changes.

Data Analyses

In the original Alameda County study, a technique known as ridit analysis was employed to assess the relationships between health practices and health status. A ridit (relative to an identified distribution) is a form of standardized score (5, 7). A ridit value of 0.50 represents average health in the reference group (the total sample), while lower values indicate better than average health and higher values denote worse than average health. Ridit values can be utilized as weights in calculating mean ridits, permitting comparison between a subgroup score and the total group value (3). They can also be adjusted for differences in specified characteristics of the population studied. In the Belloc and Breslow study, both crude ridit values and ridits adjusted for age and sex were analyzed using, apparently, the adjustment procedures described by Kantor and Winkelstein (7). We therefore employed these procedures in calculating the age-adjusted ridit values for the Michigan sample. Because of the appropriateness of age adjustments to the analysis of data from this study, tables presented in this article include only adjusted ridits.

The analysis presented includes a reexamination of the original Alameda County findings in light of comparable data for the State of Michigan. Findings from the two studies are interpreted taking into account the confidence interval for each ridit value. Each subgroup ridit value is tested for a possible significant departure from 0.50, the average value for the total sample. The emphasis on statistically significant results is based on the present authors' preference for a more conservative approach to interpretation of data than that used by Belloc and Breslow. They occasionally gave interpretive emphasis to statistically nonsignificant ridits that were in the expected direction. Dependence upon statistically

reliable findings, particularly on the consistency of these findings across the two studies, is central to our preferred approach.

RESULTS

In examining possible relationships between health practices and physical health, the health practices were analyzed individually and in combination, as in the original Belloc and Breslow article (2). Results presented in this section include those for both the Michigan and Alameda County samples.[4]

Hours of Sleep

Alameda County men and women and Michigan women reportedly sleeping 6 hr or less per night had significantly worse than average health (Table 1). The ridit for Michigan men was in the same direction but failed to reach statistical significance. Respondents obtaining 7 or 8 hr of sleep tended to have better than average health, whereas those sleeping 9 hr or more tended to have worse than average health. At each of these sleep levels, however, only one out of four subgroups yielded a significant finding: Michigan men sleeping 7 hr, Alameda County women sleeping 8 hr, and Michigan men sleeping 9 hr.

Breakfast Eating

In the Alameda County survey, respondents were asked whether they ate breakfast almost every day, sometimes, or rarely/never. In the Michigan survey, response options included every day, usually, sometimes, rarely, or never. In analyzing the Michigan data, responses of every day and usually were combined, as were answers of rarely and never, to maximize comparability.

One of the popularly publicized findings from the Alameda County study was the importance of eating breakfast, based on ridit values indicating better than average health for men and women who reported eating breakfast almost every day. The Michigan survey also yielded such findings for both sexes (Table 1). However, none of these findings for either location was statistically significant (i.e., given the confidence intervals and the associated ridit value, none was different from a 0.50 ridit representing average health status). In the same manner, eating breakfast only sometimes or rarely/never was associated with poorer than average health ridit values, but only the ridit for Alameda County women attained statistical significance.

Eating between Meals

Respondents were asked how often they ate between regular meals. The same response options were provided as for breakfast eating, and the Michigan data were combined in the same way.

Table 1 presents findings concerning the relationship between eating between meals and physical health status. Alameda County and Michigan data are consis-

---

TABLE 1

AGE–SEX-RELATED PHYSICAL HEALTH RIDITS BY USUAL HOURS OF SLEEP AND EATING HABITS IN MICHIGAN (M) AND ALAMEDA COUNTY(A)

Item	Men						Women					
	Number		Age–sex-related ridit		95% confidence interval		Number		Age–sex-related ridit		95% confidence interval	
	A	M	A	M	A	M	A	M	A	M	A	M
Usual hours of sleep												
6 or less hr	482	530	0.56*	0.51	±0.03	±0.03	568	423	0.59*	0.55*	±0.02	±0.03
7 hr	1,314	449	0.49	0.46*	±0.02	±0.03	1,368	434	0.49	0.47	±0.02	±0.03
8 hr	1,135	455	0.48	0.51	±0.02	±0.03	1,473	684	0.47*	0.48	±0.02	±0.03
9 or more hr	220	81	0.52	0.57*	±0.04	±0.06	356	188	0.51	0.50	±0.03	±0.04
How often eat breakfast?												
Almost every day (Every day, usually)	2,096	843	0.49	0.49	±0.01	±0.02	2,414	950	0.49	0.48	±0.01	±0.02
Sometimes	563	258	0.52	0.51	±0.02	±0.04	761	319	0.52	0.51	±0.02	±0.03
Rarely or never	489	408	0.51	0.51	±0.03	±0.03	585	464	0.53*	0.53	±0.02	±0.03
How often eat between meals?												
Rarely or never	880	445	0.49	0.49	±0.02	±0.03	840	418	0.49	0.50	±0.02	±0.03
Once in a while	1,414	492	0.48	0.49	±0.02	±0.03	1,741	657	0.49	0.49	±0.01	±0.03
Almost every day (Every day, almost every day)	853	570	0.54*	0.51	±0.02	±0.03	1,176	649	0.52	0.50	±0.02	±0.03
Weight in relation to height												
10% or more under desirable weight	206	124	0.57*	0.53	±0.04	±0.05	624	332	0.49	0.49	±0.02	±0.03
5.0–9.99% underweight	261	127	0.51	0.50	±0.04	±0.05	542	201	0.47*	0.43*	±0.02	±0.04
Within ±4.99% of desirable weight	798	438	0.49	0.49	±0.02	±0.03	960	415	0.48	0.47	±0.02	±0.03
5.0–9.99% overweight	521	220	0.49	0.48	±0.03	±0.04	359	178	0.48	0.51	±0.03	±0.04
10.0–19.99% overweight	767	321	0.47*	0.50	±0.02	±0.03	535	209	0.50	0.50	±0.02	±0.04
20.0–29.99% overweight	377	159	0.51	0.54	±0.03	±0.05	290	120	0.53	0.56*	±0.03	±0.05
30% or more overweight	204	114	0.57*	0.55	±0.04	±0.05	406	221	0.59*	0.60*	±0.03	±0.04

* Denotes ridit values that are significantly different from 0.50 based on 95% confidence levels.

tent for both men and women: individuals who snacked once in a while or rarely/ never had ridit values indicating average or better health status while those who snacked every day/almost every day had ridit values in the direction of worse than average health. The sole significant ridit characterized Alameda County men who snacked almost every day as having worse than average health.

Deviation from Ideal Weight

As another indicator of eating habits, both samples were categorized on the basis of a combination of reported height and weight, using a scale of desirable weights prepared by the Metropolitan Life Insurance Co. (9).

As shown in Table 1, Michigan women 20% or more overweight and Alameda County women 30% or more overweight had significantly poorer than average health. Both Michigan and Alameda County women between 5 and 10% under-weight had significantly better than average health. Among Alameda County men, those in the extreme weight categories (30% or more overweight, or 10% or more underweight) displayed significantly worse than average health status. Results for comparable Michigan men were in the same direction but were not statistically significant. Alameda County men between 10 and 19.99% overweight had signif-icantly better than average health, but that finding was not duplicated for comparable Michigan men.

Physical Activity

Respondents in both studies were asked how often they engaged in a variety of physical activities. The same response options were provided as for breakfast eating and for snacking between meals; this was true also for combinations of categories used in the data analyses.

Table 2 presents the findings for five separate activities. For the Michigan sample, all activities except hunting or fishing yielded one or more statistically significant findings concerning physical health status. Both Michigan men and women who reported never participating in active sports, in swimming, in taking long walks, or in working in the garden, had ridits clearly indicating worse than average health status. The same was also true for Michigan women who never exercised. The Alameda County sample yielded a similar pattern of findings, but the absence of information concerning ridit confidence levels for the separate activities precluded establishing the statistical significance of those results.

Findings concerning overall physical activity and physical health appear in Table 3. Similar findings were obtained for the Michigan and the Alameda County samples: for both men and women, higher levels of physical activity are associ-ated with better than average health, while lower levels are associated with poorer than average health. For both samples, individuals who reported never engaging in any of the specific physical activities had health ridits that were significantly worse than average. Other amounts of physical activity yielded ridits that were consistent in direction across samples, but of varying statistical significance.

Alcohol Usage

Both the Michigan and Alameda County questionnaires included items con-

TABLE 2

AGE–SEX RELATED PHYSICAL HEALTH RIDITS BY PHYSICAL ACTIVITIES IN MICHIGAN (M) AND ALAMEDA COUNTY (A)

	Age-sex-related ridits											
	Men						Women					
Item	Often		Sometimes (sometimes, rarely)		Never		Often		Sometimes (sometimes, rarely)		Never	
	A	M	A	M	A	M	A	M	A	M	A	M
Here is a list of active things that people do in their free time. How often do you do any of these things?												
Active sports	0.44[a]	0.47*	0.49	0.48	0.54	0.55*	0.48	0.47	0.47	0.44*	0.52	0.54*
Swimming or taking long walks	0.47	0.49	0.49	0.48	0.53	0.56*	0.46	0.49	0.48	0.47	0.55	0.57*
Working in the garden	0.48	0.48	0.51	0.50	0.52	0.53*	0.46	0.49	0.50	0.46*	0.54	0.55*
Doing physical exercises	0.46	0.51	0.50	0.49	0.52	0.50	0.46	0.48	0.49	0.48	0.53	0.53*
Hunting or fishing	0.49	0.51	0.50	0.48	0.51	0.50	0.50	0.53	0.50	0.52	0.50	0.50

[a] Alameda significant findings could not be determined from original tables since confidence interval data was not specified.
* Denotes ridit values that are significantly different from 0.50 based on 95% confidence levels.

TABLE 3

Age–Sex–Related Physical Health Ridits by Amount of Physical Activity, Alcohol, and Cigarette Usage in Michigan (M) and Alameda County (A)

Item	Men						Women					
	Number		Age–sex-related ridit		95% confidence interval		Number		Age–sex-related ridit		95% confidence interval	
	A	M	A	M	A	M	A	M	A	M	A	M
Amount of physical activity												
Often or sometimes engage in active sports, swim, or take long walks, or often garden or do physical exercises	2,775	1,418	0.49	0.49	±0.01	±0.02	3,009	1,543	0.48*	0.49	±0.01	±0.02
Sometimes garden or do physical exercises or did not answer	214	40	0.58*	0.58	±0.04	±0.09	439	79	0.54*	0.56	±0.03	±0.07
Never do any of these	169	61	0.58*	0.60*	±0.04	±0.07	322	116	0.64*	0.64*	±0.03	±0.05
Alcohol consumption												
Never drink	495	263	0.52	0.57*	±0.03	±0.04	1,084	501	0.50	0.54*	±0.02	±0.03
Amount at one time												
1–2 drinks	1,567	597	0.49	0.48	±0.01	±0.03	2,073	887	0.49	0.47*	±0.01	±0.02
3–4 drinks	767	394	0.49	0.47	±0.02	±0.03	486	262	0.51	0.50	±0.03	±0.04
5 or more	319	235	0.54*	0.50	±0.03	±0.04	110	65	0.60*	0.60*	±0.06	±0.07
Frequency												
Less than 1/week	937	484	0.49	0.49	±0.02	±0.03	1,493	825	0.51	0.48	±0.01	±0.02
1–2 per week	769	380	0.49	0.47	±0.02	±0.03	593	272	0.49	0.47	±0.02	±0.04
More than 2/week	951	356	0.50	0.50	±0.02	±0.03	587	113	0.49	0.52	±0.02	±0.05
Cigarette consumption												
Never	918	496	0.47*	0.45*	±0.02	±0.03	1,806	896	0.48*	0.48	±0.01	±0.02
Formerly	666	423	0.51	0.52	±0.02	±0.03	422	248	0.51	0.53	±0.03	±0.04
Currently	1,554	592	0.51	0.58*	±0.01	±0.03	1,497	590	0.52*	0.52	±0.01	±0.03

* Denotes ridit values that are significantly different from 0.50 based on 95% confidence levels.

cerning frequency of consumption of wine, beer, and liquor, and the number of drinks per sitting of whichever alcoholic beverage they consumed in the largest amount. They were categorized based on the frequency with which they consumed this type of alcoholic beverage. Some respondents who drank more than one type of liquor may be underrated by this method, but such would be the case for both samples.

In Alameda County, physical health ridits for men and women who never drink did not differ significantly from the average value of 0.50 (Table 3). In Michigan, however, not drinking was associated with significantly worse than average health status for both men and women. At the other end of the scale, for both sexes in Alameda County and for women in Michigan, consuming five or more drinks per sitting yielded significant ridit values indicating worse than average health. This was not the case for comparable Michigan men, whose health ridit value was exactly average. Among those consuming intermediate amounts of alcohol at one sitting, the picture is blurred: while each of the four subgroups consuming one to two drinks had an above-average health ridit, only that for Michigan women was statistically significant, and ridits for those consuming three to four drinks per sitting were inconsistent and nonsignificant.

Reported frequency of alcohol consumption was not significantly related to physical health status in either sample.

Cigarette Smoking

In both studies, cigarette smoking status of respondents was assessed by asking whether they never smoked, formerly smoked, or were currently smoking. For current and former smokers, the amount smoked, number of years smoked, and depth of inhalation were also ascertained. Since the original Alameda County Study concentrated on the issue of smoking status and employed it as the sole indicator of smoking in a composite index of health practices, only health ridit findings for the smoking status variable will be presented in the analyses that follow.

As shown in Table 3, for both sexes in both samples, never smoking cigarettes was consistently associated with better than average physical health status, with all subgroups except Michigan women yielding statistically significant ridits. Former and current smokers had ridits consistently suggesting poorer than average health; however, only those for two subgroups of current smokers, Michigan men and Alameda County women, reached statistical significance.

Health Practice Summary Score

Respondents were also categorized by the overall extent to which they reported a desirable level of the seven health practices described above. This overall score gave equal weight to each "good" health habit as defined by Belloc and Breslow.

As shown in Table 4, the relationship between the combined health practice score and the physical health spectrum was quite similar for both samples: There was a strong trend toward better physical health status as the number of health practices increased. Those with the fewest (zero and/or one) health practices had consistently poorer than average health status, but small N's and large variability kept most of their ridits from reaching statistical significance. Those with two to

TABLE 4
AGE–SEX-RELATED PHYSICAL HEALTH RIDITS BY COMPOSITE SCORE ON SELECTED HEALTH
PRACTICES, WITH AVERAGE AGES IN MICHIGAN (M) AND ALAMEDA COUNTY (A)

Health practice score	Number		Average age		Age–sex-related ridit		95% confidence interval	
	A	M	A	M	A	M	A	M
Total	6,928ᵃ	3,233	44	41	0.50	0.50	±0.01	±0.01
7	668	219	46	43	0.41*	0.42*	±0.02	±0.04
6	1,175	687	45	42	0.46*	0.44*	±0.01	±0.02
5	2,105	932	44	42	0.50	0.49	±0.01	±0.02
4	1,518	825	42	41	0.54*	0.53*	±0.01	±0.02
3	622	422	41	38	0.58*	0.54*	±0.02	±0.03
2	202	135	41	37	0.64*	0.61*	±0.04	±0.05
1	35		39		0.63*		±0.10	
None	3		41		0.82		±0.33	
None, 1		14		46		0.60		±0.15

ᵃ Note that in the Belloc and Breslow article, separate breakdowns for men and women were not provided; only the frequency distribution for the total sample was given.
* Denotes ridit values that are significantly different from 0.50 based on 95% confidence levels.

four health practices were significantly below average in health status, those with five health practices were essentially average, and those engaging in six or seven health practices were significantly above average in physical health.

DISCUSSION

In comparing the results of our Michigan study with those originally obtained in Alameda County, there is a clear similarity of findings. For the overwhelming majority of comparisons, both samples yielded ridit values in the same direction although not necessarily at a statistically significant level. In only three instances did the two studies yield ridit values in opposite directions, and for each such comparison both ridits were nonsignificantly different from 0.50. This similarity of findings is especially noteworthy when one considers the many differences between the two studies: 1978 vs 1965, Midwest vs Far West, state-wide vs county-wide, telephone interview vs mail questionnaire, etc. The consistency in findings strengthens our confidence in their generalizability over time and location. It suggests that the cross-sectional relationships between health habits and health status are generally stable, rather than being unique to Alameda County at a single point in time.

The most pronounced difference between the two studies turns out to be in the interpretation placed on data by the respective investigators. Belloc and Breslow were more willing to use a holistic approach, emphasizing consistency of findings as a key basis for their interpretations concerning the associations between health habits and health status. We are less willing to take such an approach. Instead, we opt for emphasizing tests of statistical significance in deciding whether combinations of variables are demonstrably related. In our view, taking into account the statistical significance as well as the direction of results across the two studies helps to illuminate the relationships between health habits and health status. It

suggests to us that while some of the relationships appear to be firmly substantiated, the evidence for others is still tentative.

There are, on one hand, several instances in which consistency and statistical significance of findings from the two studies were both present to a considerable degree. For example, the sleep level of 6 hr or less per night yielded a preponderance of findings of worse than average health that were both consistent and significant. Being slightly (5–9.99%) underweight was significantly associated with above-average health for women in both studies; being 30% or more overweight was consistently associated with below-average health in both studies, with three out of four results statistically significant and the fourth just missing significance. Never participating in the various physical activities studied was consistently and significantly associated with worse-than-average health for both sexes in both studies. The marked trend toward better physical health as the overall number of health practices increased involved a large number of ridits that were statistically significant (11 of 15). We regard such findings as rather solid evidence of relationships between health practices and health status, especially in view of the substantial sample sizes on which these findings are based and the demographic and other differences between the two samples.

On the other hand, there were instances in which consistency of outcome was not bolstered by consistency of statistical significance. As cases in point, the consistent trends for those obtaining 7 or 8 hr of sleep to have above-average health, and for those sleeping 9 hr or more to have worse-than-average health, contained too few instances of statistical significance for us to regard them as definitive. Along the same lines, we do not regard the data as providing clear-cut evidence that either breakfast eating or snacking between meals is related to health status in view of the paucity of statistically significant results, although we recognize the consistency of results across the two studies. And although engaging in greater amounts of physical activity was consistently associated with better health, it was only occasionally accompanied by statistical significance. In instances such as those just cited we find ourselves unwilling to give as much weight to the findings as would be the case if statistical significance were also present. We recognize, however, that the use of statistical testing as a decision rule for evaluating the reliability of findings is controversial and that ours is not the only defensible position.

Beyond the methodological differences across the two studies, other factors complicate analysis of the relationship between health practices and physical health status. As Wilson and Elinson note, "the major issue is the potential for circularity between certain practices and health status measures" (15). The temporal relationship between practices and health is impossible to understand from either of the two cross-sectional studies. Hence, it is not appropriate to infer from either set of cross-sectional data that practice x causes better or worse health. Rather, the cross-sectional approaches used require that the phrase "is associated with" replace the word "causes." This distinction is not trivial, since it allows for the real possibility that poor health may lead to poor practices or that good health may lead to more desirable behavior. Only subsequent follow-up on the same sample, coupled with cross-lagged panel or path analysis could justify making causal inferences from correlational data. Subsequent longitudinal studies of the

Alameda County sample cited earlier (1, 4, 14) suggest that among its members, health habits are precursors rather than consequences of physical, mental, and social health. Those prospective findings, together with the consistency of cross-sectional findings in Alameda County and the State of Michigan, make it appear reasonable that there may be a similar directional relationship between health habits and health status in Michigan as well. The only way to settle the issue definitively, however, would be to follow the Michigan sample over time.

Another important issue is the degree to which the health practices measured can reasonably be assumed to be associated (or not associated) with physical health status. It is possible that the various behaviors or practices under examination may be related to various specific and categorical aspects of physical health in a way that cannot be easily detected from the current analyses using the physical health spectrum measure. For example, the fact that reported frequency of alcohol consumption did not prove related to physical health status in either Alameda County or Michigan does not necessarily demonstrate the absence of negative health consequences; rather, the physical health spectrum may not be sensitive to such effects. Further, the health habits we studied do not exhaust the category of preventive health behaviors.

The foregoing considerations suggest the desirability of studying the relationships between a wide variety of preventive health behaviors and a broad range of possible measures of physical health status, using a longitudinal research design that allows issues of causality to be directly addressed.

REFERENCES

1. Belloc, N. B. Relationship of health practices and mortality. *Prev. Med.* 2, 67–81 (1973).
2. Belloc, N. B., and Breslow, L. Relationship of physical health status and health practices. *Prev. Med.* 1, 409–421 (1972).
3. Belloc, N. B., Breslow, L., and Hochstim, J. R. Measurement of physical health in a general population survey. *Amer. J. Epidemiol.* 93, 328–336 (1971).
4. Breslow, L., and Enstrom, J. E. Persistence of health habits and their relationship to mortality. *Prev. Med.* 9, 469–483 (1980).
5. Bross, I. How to use ridit analysis. *Biometrics* 14, 18–38 (1958).
6. Gottlieb, N. H., and Green, L. W. Life events, social network, life-style, and health: An analysis of the 1979 National Survey of Personal Health Practices and Consequences. *Health Educ. Quart.* 11, 91–105 (1984).
7. Kantor, S., and Winkelstein, W. The rationale and use of ridit analysis in epidemiologic studies of blood pressure. *Amer. J. Epidemiol.* 90, 201–213 (1969).
8. Kish, L. A procedure of objective respondent selection within the household. *J. Amer. Stat. Assoc.* 44, 380–387 (1949).
9. Metropolitan Life Insurance Company. "Overweight, Its Prevention and Significance," p. 5. 1960.
10. Metzner, H. L., Carman, W. J., and House, J. Health practices, risk factors, and chronic disease in Tecumseh. *Prev. Med.* 12, 491–507 (1983).
11. Pope, C. R. Life-styles, health status and medical care utilization. *Med. Care* 20, 402–413 (1982).
12. Reed, W. L. Physical health status as a consequence of health practices. *J. Commun. Health* 8, 217–228 (1983).
13. Waksberg, J. Sampling methods for random digit dialing. *J. Amer. Stat. Assoc.* 73, 40–46 (1978).
14. Wiley, J. A., and Camacho, T. C. Life-style and future health: Evidence from the Alameda County Study. *Prev. Med.* 9, 1–21 (1980).
15. Wilson, R. W., and Elinson, J. National Survey of Personal Health Practices and Consequences: Background, conceptual issues, and selected findings. *Public Health Rep.* 96, 218–225 (1981).