OPEN

Obesity Paradox

Conditioning on Disease Enhances Biases in Estimating the Mortality Risks of Obesity

Samuel H. Preston and Andrew Stokes

Background: Many studies have documented an obesity paradox—a survival advantage of being obese—in populations diagnosed with a medical condition. Whether obesity is causally associated with improved mortality in these conditions is unresolved.

Methods: We develop the logic of collider bias as it pertains to the association between smoking and obesity in a diseased population. Data from the National Health and Nutrition Examination Survey (NHANES) are used to investigate this bias empirically among persons with diabetes and prediabetes (dysglycemia). We also use NHANES to investigate whether reverse causal pathways are more prominent among people with dysglycemia than in the source population. Cox regression analysis is used to examine the extent of the obesity paradox among those with dysglycemia. In the regression analysis, we explore interactions between obesity and smoking, and we implement a variety of data restrictions designed to reduce the extent of reverse causality.

Results: We find an obesity paradox among persons with dysglycemia. In this population, the inverse association between obesity and smoking is much stronger than in the source population, and the extent of illness and weight loss is greater. The obesity paradox is absent among never-smokers. Among smokers, the paradox is eliminated through successive efforts to reduce the extent of reverse causality.

Conclusion: Higher mortality among normal-weight people with dysglycemia is not causal but is rather a product of the closer inverse association between obesity and smoking in this subpopulation.

(*Epidemiology* 2014;25: 454–461)

Submitted 24 May 2013; accepted 25 October 2013; posted 6 March 2014.

From the Population Studies Center, University of Pennsylvania, Philadelphia, PA.

Supported by grant number R01AG040212 from the National Institute on Aging.
The authors report no conflicts of interest.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com). This content is not peer-reviewed or copy-edited; it is the sole responsibility of the author.

Correspondence: Samuel H. Preston, Population Studies Center, University of Pennsylvania, 3718 Locust Walk, McNeil Building, Room 239, Philadelphia, PA 19104. E-mail: spreston@sas.upenn.edu.

Copyright © 2014 by Lippincott Williams & Wilkins. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives 3.0 License, where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially.

ISSN: 1044-3983/14/2503-0454
DOI: 10.1097/EDE.0000000000000075

In a wide variety of disease states, obese persons have been shown to experience lower mortality and better survival than that shown by the nonobese. These states include diabetes,[1–3] coronary artery disease,[4,5] heart failure,[6] peripheral arterial disease,[5] hypertension,[7] chronic obstructive pulmonary disease,[8] lung cancer,[9] and esophageal adenocarcinoma.[10] Superior survival among the obese patients has also been demonstrated after myocardial infarction,[11] coronary revascularization,[12] and angiography,[13] and among hemodialysis patients.[14] Better survival for obese patients in these disease states is considered paradoxical because obesity is associated with higher mortality in the vast majority of studies where it has been investigated.[15]

In this article, we argue that the obesity paradox is a product of statistical biases. Although these biases are present in most observational cohort studies of the mortality risks of obesity, they are exaggerated when attention is limited to populations that are conditioned on a disease state. We demonstrate these biases through formal reasoning and by application to a population with diabetes and prediabetes.

BIASES IN STUDIES OF THE MORTALITY RISKS OF OBESITY

Confounding by unmeasured or poorly measured variables is probably the most widely recognized bias in analyses of mortality risk factors. In studies of the mortality consequences of obesity, failure to properly control for smoking is a classic example of confounding.[16] Because smoking is negatively correlated with body mass index (BMI) in the United States,[17] failure to properly control smoking exposure is likely to result in underestimating the mortality risks associated with obesity.[18] Crude categorizations and mismeasurement of smoking allow ample opportunity for residual confounding even when smoking is "controlled."[16,19] Stram et al[18] cite 6 studies in which the correlation[5] between cotinine levels and reported smoking behavior was only 0.40 to 0.70. A frequent suggestion aimed at minimizing confounding by smoking is to restrict a study to never-smokers,[16] although such restrictions may produce only small changes in the shape of the relation between obesity and mortality.[20,21]

A second potential bias producing an underestimation of the mortality risks of obesity has been referred to as reverse

causation. Low weight can be caused by loss of appetite or increased metabolic demands from manifest or occult disease which may in turn raise the risk of death.[16,22] Among those with diabetes, unintentional weight loss commonly accompanies poor glucose control.[23] The most serious condition associated with unintentional weight loss is cachexia, which is an involuntary loss of muscle mass as a result of an increase in circulating inflammatory cytokines.[24] Many diseases can initiate this process, including chronic obstructive pulmonary disease, cancer, heart failure, and chronic renal failure.[24,25] This condition is common; approximately 5 million Americans experience from cachexia.[25] Although these processes have been clearly described, their empirical significance for estimating the relation between obesity and mortality has been questioned.[26]

INCREASED BIAS AMONG PERSONS SELECTED INTO A DISEASE STATE

Reverse causation and statistical confounding by smoking may become greater threats to accurate identification of the mortality risks of obesity when a study is confined to people who have entered a disease state. Consider first the threat of statistical confounding of obesity risks by smoking. Figure 1 presents three 2×2 tables. Figure 1A is the joint distribution of a population on 2 risk factors, which we have labeled Obese (Yes/No) and Smoker (Yes/No). Figure 1B represents the relative risk of developing diabetes for each combination of exposures on these two risk factors. Each cell in Figure 1C is the product of the equivalent cells in Figure 1A and B, so this table represents the joint distribution of the number of new cases of diabetes with respect to smoking and obesity exposure.

Now compare the relation between risk factors among persons with diabetes with the equivalent relation in the population as a whole (the source population). In the source population, the ratio of the odds of smoking among nonobese persons to the odds of smoking among obese persons is:

$$O_P = \frac{n_{12}/n_{22}}{n_{11}/n_{21}} \tag{1}$$

Among people with diabetes, the equivalent odds ratio of smoking among nonobese persons to smoking among obese persons is:

$$O_D = \frac{n_{12}r_{12}/n_{22}r_{22}}{n_{11}r_{11}/n_{21}r_{21}} \tag{2}$$

Thus, the ratio of the odds ratios in the two populations is:

$$O_D/O_P = \frac{r_{12}/r_{22}}{r_{11}/r_{21}} \tag{3}$$

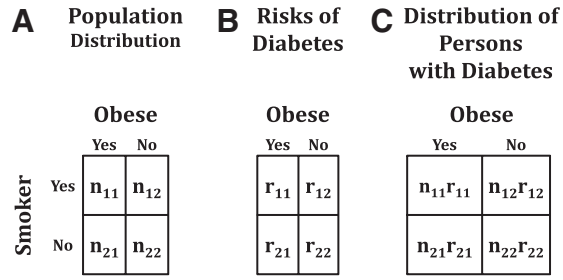

FIGURE 1. Illustration of how the joint distribution of smoking and obesity changes on selection into the state of diabetes.

A ratio of $O_D/O_P > 1.00$ would indicate that the relative prevalence of smoking among nonobese persons compared with the obese persons is greater among those with diabetes than it is in the source population.

Suppose that the risks of developing diabetes associated with smoking and obesity are multiplicative, ie, that the relative risk of developing diabetes if one is a smoker is the same whether or not one is obese. Then,

$$\frac{r_{11}}{r_{21}} = \frac{r_{12}}{r_{22}} \tag{4}$$

In this case, $O_P = O_D$, and the joint distribution of risk factors among those with diabetes will simply replicate that in the population. But if,

$$\frac{r_{11}}{r_{21}} < \frac{r_{12}}{r_{22}}, \tag{5}$$

then $O_D > O_P$. Condition (5) means that the risk of developing diabetes is less than a multiplicative function of being obese and being a smoker. That is, if one is both obese and smoker, the relative risk of developing diabetes is less than the product of the smoking-relative risk for nonobese persons and the obesity-relative risk for nonsmokers. Such a circumstance would apply, for example, when risks are additive rather than multiplicative. Supermultiplicativity would reverse the inequality sign in (5).

If condition (5) prevails, then exposures that are independently distributed in the population become negatively correlated among the diseased. When exposures are negatively associated in the source population, as in the case of obesity and smoking, the negative association will become stronger among those in the diseased state.

The change in the association between two variables on entry into a disease state is a special case of what has been referred to as "conditioning on an effect," "selection bias," "collider bias," and "event index bias."[27–31] The relations among exposures and mortality observed in the source population are conditioned on the "effect" of being diabetic. This effect may change the relation between the two exposures (smoking and obesity) relative to their relation in the source

© 2014 Lippincott Williams & Wilkins

population. The magnitude and direction of any bias induced by conditioning on diabetes is unpredictable without additional information. What is unusual about the application to relations between smoking and obesity is that the two risk factors are negatively correlated in the source population. When condition (5) prevails, this negative correlation becomes exaggerated in the diseased population, enhancing the threat that the mortality risks of obesity will be underestimated because of a failure to fully control the effects of smoking.

Consider now the bias associated with reverse causation. The strength of reverse causal pathways should be greater in a diseased population because the frequency of disease-associated weight loss should be higher. Among people with diabetes, those losing weight have higher mortality than those with stable weight.[32] We expect that smoking will play a key role in patterns of weight loss among those with dysglycemia. A meta-analysis concluded that smokers were 44% more likely than nonsmokers to develop diabetes.[33] Smoking is associated with many diseases that may cause cachexia, including chronic obstructive pulmonary disease, cancer, heart failure, and chronic renal failure. Therefore, we expect that the incidence of weight loss, and the associated biases in estimating the mortality effects of obesity, will be accentuated in the population with diabetes.

We summarize our causal model in the directed acyclic graph (DAG) shown in Figure 2. Both obesity (represented by "body mass index") and smoking are seen as causally related to the development of diabetes. The two exposures "collide" in diabetes, a condition that raises the risk of death. Smoking is seen as producing lower weight. This effect works through two pathways: smoking increases rates of metabolism and reduces appetite, producing the negative cross-sectional association between BMI and smoking found in the United States[17]; and smoking increases the incidence of diseases that produce weight loss and cachexia. The first of these pathways represents a standard version of statistical confounding, whereas the second is typically referred to as reverse causation. Smoking increases the risk of death apart from the pathway operating through diabetes.

On the basis of this DAG and the reasoning and evidence that support it, we hypothesize the following:

(1) The collision of obesity and smoking among those diagnosed with diabetes will strengthen the negative association between obesity and smoking.

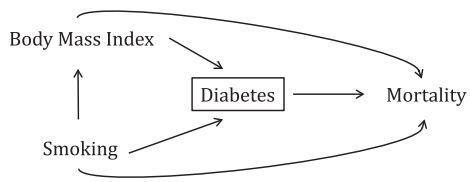

FIGURE 2. Directed acyclic graph representing the causal effects of body mass index on mortality.

(2) The high prevalence of smoking among those diagnosed with diabetes will increase the incidence of illness and weight loss relative to the incidence in the source population.
(3) Because of the two processes hypothesized in (1) and (2), the relatively low mortality of nonobese persons in the source population will be muted or even reversed (producing an "obesity paradox") in the diabetic population.
(4) Among people with diabetes who never smoked, the normal relation of higher mortality among the obese will be maintained. The DAG appropriate for those who never smoked is easily envisioned by eliminating all arrows from smoking in Figure 2. Among never-smokers, there is no collision between smoking and obesity, and the sources of bias associated with smoking are eliminated.

EMPIRICAL APPLICATION

To examine the biases produced by conditioning on diabetes, we require data on the population with diabetes as well as data on the population from which persons with diabetes are selected. Such data are provided by the National Health and Nutrition Examination Survey (NHANES), a nationally representative survey of the noninstitutionalized US population. We investigate the obesity paradox among people with diabetes, a state in which the obesity paradox has been repeatedly demonstrated.[1–3] To increase the number of observations, we include people with prediabetes, a group that we refer to as having "dysglycemia," as well as those with diabetes.

We construct the sample using data from NHANES III (1988–1994) and from NHANES continuous waves (1999–2004). After linkage to the National Death Index through 2006,[34] the period used for mortality analysis is 1988–2006. Mean follow-up is 9.2 years. NHANES provides clinical data on glycemia, as well as measured data on height and weight. We defined diabetes using the hemoglobin A1c (HbA1c) test, which we adopted over other diagnostic markers (such as fasting plasma glucose) because HbA1c exhibits greater stability and lower intraindividual variation.[35] To construct our group with dysglycemia, we adopt an HbA1c threshold of 5.7%, in accordance with the most recent national American Diabetes Association guidelines.[36] We include as dysglycemic those persons who report a prior diagnosis of diabetes even if their HbA1c value is less than 5.7%. We eliminate persons whose reported age at the time of diagnosis was less than 25 years to reduce the number with type 1 diabetes. For the analysis, we focus on people aged 35 to 74 at baseline. After eliminating missing observations on HbA1c, BMI, smoking, education, and mortality status, the final sample consists of 4642 dysglycemic persons, among whom 1162 died during follow-up.

We operationalize weight status using a dichotomous construct of obesity consistent with that used in some previous studies.[1,9,11] We classify people as either normal weight (18.5–24.9 kg/m^2) or overweight/obese (≥25 kg/m^2). We estimate age-standardized mortality rates by weight and smoking status using the US 2000 census as the standard. We

 © 2014 Lippincott Williams & Wilkins

additionally apply Cox proportional hazards models to adjust for covariates. Attained age is used for analysis time. Covariates include sex, race/ethnicity (non-Hispanic black, Hispanic, or other), educational attainment (less than high school, high school or equivalent, or some college or greater), HbA1c, and detailed categories of smoking intensity (never smoker, former smoker <1 pack/day, former smoker 1 to <2 packs/day, former smoker 2+ packs/day, smoker <1 pack/day, current smoker 1 to <2 packs/day, or current smoker 2+ packs/day).

Regression models are implemented under various sample-restriction criteria designed to successively attenuate reverse-causation bias. These measures include delaying onset of exposure by 3 years to reduce the influence of those who were ill at baseline and who may have lost weight as a consequence; eliminating persons with a prior diagnosis of emphysema or a smoking-related cancer; and excluding persons who have lost significant weight in the 10-year period before the survey. The category of smoking-related cancers is based on relative risks in a recent large study[37] and includes cancers of the lung, larynx, mouth/tongue/lip, esophagus, bladder, kidney, and pancreas. Weight loss was assessed using BMI 10 years before baseline, based on recalled weight in combination with measured height at baseline. To reduce bias that may result from inaccurate recall of past weight, we applied a correction factor based on the proportionate error between measured and self-reported BMI at baseline. The threshold for exclusion was 2.2 BMI units.

Demonstration of Greater Potential Biases Among the Dysglycemic Population

Table 1 demonstrates differences in the joint distribution of obesity and smoking status between the source population and the dysglycemic population. Values are age- and sex-standardized to the US 2000 census using 5-year age groups between ages 35 to 74 years.

The table shows that the joint distribution of smoking and obesity is quite different in the dysglycemic population from that in the US population as a whole. In the source population, the ratio of the odds of being a current smoker among nonobese persons to that among obese persons is 1.58. Among those with dysglycemia, the odds ratio increases to 2.88. Clearly, the inverse association between smoking and obesity has greatly strengthened in the dysglycemic population, confirming hypothesis (1) and the operation of condition (5). Such strengthening also occurs for the categories of heavy smoker and long-duration smoker. This stronger inverse association between smoking and obesity increases the risk of statistical confounding in the dysglycemic population relative to the risk in the source population.[38]

Table 2 investigates reverse causation, the second source of bias in studies of the mortality risks of obesity. Persons with dysglycemia have higher rates of illness than the source population, providing support for hypothesis (2). Higher rates of illness are observed not only for dysglycemia itself but also for emphysema, heart attacks, strokes, and self-rated health. The only exception is cancers that are not smoking-related. The greater extent of illness in the dysglycemic population should produce more weight loss than that in the source population. Using retrospectively reported data on height and weight 10 years before baseline, we define substantial weight loss as the top 50th percentile of weight loss in the dysglycemic population during this period. This criterion translates into a loss of 2.2 BMI units. Using this criterion, 13% of the dysglycemic population lost weight compared with 8% of the source population. To the extent that weight loss is a signal of disease, the greater extent of weight loss among persons with dysglycemia increases the risk of reverse causation relative to the risk in the source population.

Effect of Biases on Excess Mortality of Overweight/Obesity Among Persons with Dysglycemia

Table 3 demonstrates that the obesity paradox is present among the dysglycemic population. On the basis of the age-standardized death rates, the relative risk of death for

TABLE 1. Joint Distribution of Population[a] by Weight Status[b] and Smoking: Source Population and Dysglycemic Population[c]

	Source Population				Dysglycemic Population[c]			
	All %	Normal Weight %	Overweight/ Obese %	Odds Ratio[e]	All %	Normal Weight %	Overweight/ Obese %	Odds Ratio[e]
Smoking								
Never	44	41	46	0.84	41	31	43	0.60
Former	31	27	32	0.80	30	19	32	0.52
Current	25	31	22	1.58	29	50	26	2.88
Current heavy smoker[d]	14	18	13	1.53	16	30	14	2.59
Current smoker, 20+ years	21	27	19	1.61	25	43	21	2.72

[a]Proportions are age- and sex-standardized to the US 2000 census using 5-year age groups between ages 35 and 74 years.
[b]Normal weight: BMI = 18.5–24.9 kg/m^2; overweight/obese: BMI ≥25 kg/m^2. BMI is calculated using measured height and weight.
[c]Dysglycemia is defined as hemoglobin A1c (HbA1c) ≥5.7% or previous diagnosis of diabetes.
[d]Current heavy smoker: 1+ packs/day.
[e]Odds ratios compare the odds of being in a particular smoking category among normal-weight persons to the odds among the overweight/obese.
NHANES baseline data are derived from NHANES III (1988–1994) and continuous (1999–2004).

© 2014 Lippincott Williams & Wilkins

normal-weight people compared with that for overweight/obese people is 1.31 (95% confidence interval [CI] = 1.05–1.58). Next, we divide the dysglycemic population into ever-smokers and never-smokers. The table shows that the obesity paradox is strengthened among ever-smokers, with relative risks of death among normal-weight people rising to 1.51 (1.17–1.84). However, the obesity paradox is eliminated among never-smokers, among whom people of normal weight experience a lower death rate compared with that in overweight/obese people (relative risk = 0.66 [95% CI = 0.41–0.91]).

TABLE 2. Frequency of Illness, Poor Health, and Weight Loss in the Source Population Compared with the Dysglycemic Population[a]

	Source Population %	Dysglycemic Population %
Medical condition[b]		
Dysglycemia	20.3	100.0
Emphysema	2.1	2.8
Smoking-related cancer[c]	0.6	0.8
Other cancers	6.1	5.9
Heart attack	4.0	5.5
Coronary heart disease	3.8	5.1
Angina	3.4	4.9
Congestive heart failure	2.1	4.0
Stroke	2.2	3.8
Poor self-rated health[d]	3.4	5.9
Weight loss[e]	8.1	13.2

Estimates are age- and sex-standardized to the US 2000 census using 5-year age groups between ages 35–74.

[a]Dysglycemia is defined as hemoglobin A1c (HbA1c) ≥5.7% or previous diagnosis of diabetes.

[b]Medical conditions at baseline are based on self-reported data of a prior diagnosis. Information on coronary heart disease, angina, and congestive heart failure is elicted only in the NHANES continuous surveys.

[c]The category of smoking-related cancers is based on relative risks in a recent large study[37] and includes cancers of the lung, larynx, mouth/tongue/lip, esophagus, bladder, kidney, and pancreas.

[d]Self-rated health is measured on a 5-point scale with "poor" being the least healthy category. Data on self-rated health were not available in the NHANES 1999–2000 wave.

[e]See text for definition of weight loss.

NHANES baseline data are derived from NHANES III (1988–1994) and continuous (1999–2004).

Table 3 controls only age. In Table 4, we introduce Cox proportional hazards models with various model specifications (varying across the rows) and various exclusions designed to reduce the impact of reverse causation (varying down the columns). The main independent variable in each regression model is a dichotomous indicator of BMI status (normal weight vs. overweight/obese). Model 1 examines the risk of mortality as a function of BMI status, adjusting for sex, race/ethnicity, education, and HbA1c; Model 2 additionally adjusts for detailed smoking intensity; and Model 3 introduces an interaction term between normal weight and an indicator for ever-smoking (former and current smoker).

Using all data and before introducing variables related to smoking, mortality among normal-weight people is substantially higher than that among overweight/obese people (hazard ratio [HR] = 1.28 [95% CI = 1.08–1.52]). When smoking status is introduced in Model 2, the excess risk associated with normal weight declines to 13% (HR = 1.13 [0.93–1.37]).

The third model introduces an interaction term between ever-smoking and weight status. A substantial effect modification emerges: among never-smokers, the HR associated with normal weight becomes 0.69 (95% CI = 0.45–1.06). This result is consistent with hypothesis (4). Among ever-smokers, however, relative mortality is much higher among normal-weight persons (HR = 1.33 [95% CI = 1.07–1.66]). If we did not control for smoking history and intensity, the HR for the nonobese relative to the HR for obese among smokers would be even higher at 1.59.

We now introduce the restrictions aimed at reducing the intensity of reverse causal pathways. The restrictions are

TABLE 3. Age-standardized Mortality Rates (per 1000 Person-years) for Persons with Dysglycemia[a] by Smoking and Weight[b] Status

Population		Never-smokers		Ever-smokers	
Mortality Rate (95% CI)	RR (95% CI)	Mortality Rate (95% CI)	RR (95% CI)	Mortality Rate (95% CI)	RR (95% CI)
		All			
18.82 (16.93–20.71)		11.97 (10.16–13.77)		23.13 (20.24–26.01)	
		Overweight/obese			
17.84 (15.92–19.76)	1.00	12.68 (10.71–14.65)	1.00	21.18 (18.25–24.11)	1.00
		Normal weight			
23.46 (19.05–27.87)	1.31 (1.05–1.58)	8.37 (5.32–11.42)	0.66 (0.41–0.91)	31.91 (25.55–38.28)	1.51 (1.17–1.84)

Mortality rates are age-standardized (using age at exposure) to the US 2000 census with age categories 40–59 and 60+. Relative risks are based on age-standardized mortality rates and are otherwise unadjusted.

[a]Dysglycemia is defined as hemoglobin A1c (HbA1c) ≥5.7% or previous diagnosis of diabetes.

[b]Normal weight: BMI = 18.5–24.9 kg/m^2; overweight/obese: BMI ≥25 kg/m^2. BMI is calculated using measured height and weight.

NHANES baseline data are derived from NHANES III (1988–1994) and continuous (1999–2004) and are linked to mortality follow-up through 2006.

RR indicates relative risk.

© 2014 Lippincott Williams & Wilkins

TABLE 4. Hazard Ratios Associated with Normal Weight Relative to Overweight/Obese,[a] US Adults with Dysglycemia[b]

	No. in Sample	No. Deaths	Model 1: Sociodemographic[c]	Model 2: Sociodemographic and Smoking History[c]	Model 3: Sociodemographic, Smoking History, and Ever-smoking Interaction[c]	
			HR (95% CI)	HR (95% CI)	Never-smokers HR (95% CI)	Ever-smokers HR (95% CI)
Exclusion criteria (cumulative)						
No exclusions	4642	1,162	1.28 (1.08–1.52)	1.13 (0.93–1.37)	0.69 (0.45–1.06)	1.33 (1.07–1.66)
3-year delayed entry[d]	3813	933	1.25 (1.02–1.54)	1.06 (0.84–1.34)	0.65 (0.38–1.10)	1.26 (1.00–1.58)
Baseline condition[e]	3698	878	1.19 (0.95–1.49)	0.98 (0.75–1.30)	0.66 (0.39–1.12)	1.15 (0.86–1.52)
Weight loss[f]	2841	594	1.01 (0.77–1.33)	0.84 (0.60–1.18)	0.61 (0.31–1.21)	0.95 (0.66–1.36)

The HR of the interaction term between normal weight and ever-smoking in Model 3 is not shown. The HR associated with ever-smoking is calculated based on the parameter estimates for normal weight and the interaction term between normal weight and ever-smoking.

[a]Normal weight: BMI = 18.5–24.9 kg/m^2; overweight/obese: BMI ≥25 kg/m^2. BMI is calculated using measured height and weight.

[b]Dysglycemia is defined as hemoglobin A1c (HbA1c) ≥5.7% or previous diagnosis of diabetes.

[c]Model 1 adjusts for social and demographic variables (sex, race/ethnicity, and education) and HbA1c. Model 2 additionally adjusts for detailed smoking intensity; Model 3 introduces an interaction term between normal weight and ever-smoking (the combination of former and current-smokers). Race/ethnicity categories are non-Hispanic black, Hispanic, and other race/ethnicities. Educational attainment categories are less than high school, high school or equivalent, and some college or greater. Detailed smoking categories are never smoker, former smoker <1 pack/day, former smoker 1 to <2 packs/day, former smoker 2+ packs/day, current smoker <1 pack/day, current smoker 1 to <2 packs/day, and current smoker 2+ packs/day.

[d]Three-year delayed entry: exposure to risk begins accumulating after 3 years of follow-up.

[e]Baseline condition: persons who report a prior diagnosis of a smoking-related cancer or emphysema are excluded from the sample.

[f]Weight loss: persons who lost >2.2 units of BMI in the 10 years before the survey are excluded from the sample. This threshold corresponds to the top 50th percentile of weight loss among dysglycemic persons who lost weight in the interval. Individuals with missing information on weight loss are excluded before imposing the weight loss exclusion criterion. This resulted in a loss of 307 observations and 88 deaths.

NHANES baseline data are derived from NHANES III (1988–1994) and continuous (1999–2004) and are linked to mortality follow-up through 2006.

cumulative. In row 2, we exclude exposure time in the first 3 years after baseline. In row 3, we additionally exclude people with emphysema and smoking-related cancers. In row 4, we additionally exclude those in the top 50th percentile of weight loss during the 10 years before baseline.

After these restrictions are imposed in Model 2, normal-weight people have 16% lower mortality compare with that in overweight/obese people (HR = 0.84 [95% CI = 0.60–1.18]). Results from the third model are especially revealing and underscore the importance of smoking-associated reverse causation. The introduction of successive restrictions designed to reduce the impact of reverse causation has a minor effect on the HR of normal-weight nonsmokers, which declines from 0.69 to 0.61. However, it has a larger effect on the relative risk of normal-weight smokers (HR declines from 1.33 to 0.95). The larger effect of these exclusions among smokers than among never-smokers suggests that, as expected, reverse causation is a more common source of bias among smokers than among nonsmokers.

Note that coefficients in Model 3 in Table 4 were estimated using an interaction term between smoking status and obesity. If we had instead stratified the sample into ever-smokers and never-smokers, results would be virtually identical. The HR for nonobesity among never-smokers would be 0.69 instead of 0.68, and for ever-smokers it would be identical at 1.33.

Sensitivity Analyses

We have performed 3 additional analyses to test the robustness of our conclusions (eAppendix; http://links.lww.com/EDE/A769). In the first, we have tightened the criterion for being dysglycemic from a reading of 5.7% to 6.0%. The second sensitivity analysis removes those in the overweight range (BMI between 25 and 30 kg/m^2) from the overweight/obese group. These two sensitivity analyses appear in eTable 1 (http://links.lww.com/EDE/A768). Third, we have repeated all analyses using data from the National Health Interview Survey (eTables 2–5; http://links.lww.com/EDE/A768). In all cases, basic results are unchanged: an obesity paradox is present before statistical controls are imposed, it is eliminated among never-smokers, and it is attenuated among ever-smokers by the introduction of restrictions designed to reduce reverse causation. Using data from NHIS rather than NHANES, we show that conditioning on dysglycemia enhances the risk of bias from improper control of smoking and from reverse causation (eTables 2–3; http://links.lww.com/EDE/A768).

DISCUSSION

The threat of downward biases in estimating the mortality risks of obesity is widely discussed. That the threats are magnified in studies of populations diagnosed with a disease has received little attention. Banack and Kaufman[31] were among the first to suggest that the obesity paradox may be attributable to collider bias. They hypothesize that a collision between obesity and "unmeasured risk factors" occurs in the state of heart failure. At least in the case of diabetes (and we suspect elsewhere), the salient additional risk factor seems to be smoking. We demonstrate both theoretically and empirically that the inverse

© 2014 Lippincott Williams & Wilkins

association between smoking and obesity is strengthened by conditioning on diabetes. We also demonstrate that a variety of forms of illness, as well as weight loss, are more prevalent among persons with dysglycemia, raising the threat of reverse causation.

In multivariate analyses, we identify an obesity paradox for people with dysglycemia. The paradox is attenuated when smoking behavior is controlled, and it is eliminated completely among those who have never smoked. Among smokers, successive restrictions designed to reduce the strength of reverse causal pathways eliminate the obesity paradox. This evidence supports the hypothesis that illness-associated weight loss is more prevalent—and hence a greater source of bias—among smokers than among never-smokers.

Our study has some limitations. One is the relatively small number of persons with dyglycemia in our NHANES study population. Another is that, to construct our weight-change variable, we used recall data on weight to calculate BMI 10 years before baseline. This may have introduced error. However, validation studies of 10-year weight recall support their validity for use in epidemiologic studies.[39] An additional limitation is that, apart from the mortality follow-up, data are cross-sectional and do not permit sequential analysis of intervening pathways of disease, recovery, and weight change that would help illuminate the sources of the obesity paradox.

Clinical recommendations depend on whether the obesity paradox is a product of real biological processes or is produced by statistical biases. In this regard, it is noteworthy that never-smokers do not exhibit an obesity paradox in any of the specifications we have considered, including sensitivity analyses. The absence of a paradox among never-smokers implies that the paradox is not a product of physiological protection afforded by a high BMI unless that protection itself is somehow invoked by smoking. No plausible biological mechanisms have been offered for such a relation. We conclude that the apparent protective effect of obesity among persons with diabetes and prediabetes is not causal.

ACKNOWLEDGMENTS

We are grateful to Jere Behrman, Virginia Chang, Irma Elo, Douglas Ewbank, Neil Mehta, Sari Stenholm, and Lucia Tiererova, for helpful suggestions.

REFERENCES

1. Carnethon MR, De Chavez PJ, Biggs ML, et al. Association of weight status with mortality in adults with incident diabetes. *JAMA*. 2012;308: 581–590.
2. Doehner W, Erdmann E, Cairns R, et al. Inverse relation of body weight and weight change with mortality and morbidity in patients with type 2 diabetes and cardiovascular co-morbidity: an analysis of the PROactive study population. *Int J Cardiol*. 2012;162:20–26.
3. Jackson CL, Yeh HC, Szklo M, et al. Body-mass index and all-cause mortality in US adults with and without diabetes. *J Gen Intern Med*. 2014;29:25–33.
4. Romero-Corral A, Montori VM, Somers VK, et al. Association of bodyweight with total mortality and with cardiovascular events in coronary artery disease: a systematic review of cohort studies. *Lancet*. 2006;368:666–678.
5. Lavie CJ, Milani RV, Ventura HO. Obesity and cardiovascular disease: risk factor, paradox, and impact of weight loss. *J Am Coll Cardiol*. 2009;53:1925–1932.
6. Oreopoulos A, Padwal R, Kalantar-Zadeh K, Fonarow GC, Norris CM, McAlister FA. Body mass index and mortality in heart failure: a meta-analysis. *Am Heart J*. 2008;156:13–22.
7. Carman WJ, Barrett-Connor E, Sowers M, Khaw KT. Higher risk of cardiovascular mortality among lean hypertensive individuals in Tecumseh, Michigan. *Circulation*. 1994;89:703–711.
8. Landbo C, Prescott E, Lange P, Vestbo J, Almdal TP. Prognostic value of nutritional status in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 1999;160:1856–1861.
9. Attaran S, McShane J, Whittle I, Poullis M, Shackcloth M. A propensity-matched comparison of survival after lung resection in patients with a high versus low body mass index. *Eur J Cardiothorac Surg*. 2012;42:653–658.
10. Yoon HH, Lewis MA, Shi Q, et al. Prognostic impact of body mass index stratified by smoking status in patients with esophageal adenocarcinoma. *J Clin Oncol*. 2011;29:4561–4567.
11. Bucholz EM, Rathore SS, Reid KJ, et al. Body mass index and mortality in acute myocardial infarction patients. *Am J Med*. 2012;125: 796–803.
12. Oreopoulos A, Padwal R, Norris CM, Mullen JC, Pretorius V, Kalantar-Zadeh K. Effect of obesity on short- and long-term mortality postcoronary revascularization: a meta-analysis. *Obesity (Silver Spring)*. 2008;16:442–450.
13. Angerås O, Albertsson P, Karason K, et al. Evidence for obesity paradox in patients with acute coronary syndromes: a report from the Swedish Coronary Angiography and Angioplasty Registry. *Eur Heart J*. 2013;34:345–353.
14. Schmidt DS, Salahudeen AK. Obesity-survival paradox-still a controversy? *Semin Dial*. 2007;20:486–492.
15. Whitlock G, Lewington S, Sherliker P; Prospective Studies Collaboration. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet*. 2009;373:1083–1096.
16. Hu F. Obesity and mortality. In: Hu FB, ed. *Obesity Epidemiology*. New York, NY: Oxford University Press; 2008:216–233.
17. Chiolero A, Faeh D, Paccaud F, Cornuz J. Consequences of smoking for body weight, body fat distribution, and insulin resistance. *Am J Clin Nutr*. 2008;87:801–809.
18. Stram DO, Huberman M, Wu AH. Is residual confounding a reasonable explanation for the apparent protective effects of beta-carotene found in epidemiologic studies of lung cancer in smokers? *Am J Epidemiol*. 2002;155:622–628.
19. Renehan AG, Leitzmann MF, Zwahlen M. Re: body mass index and risk of lung cancer among never, former, and current smokers. *J Natl Cancer Inst*. 2012;104:1680–1681.
20. Durazo-Arvizu RA, Cooper RS. Issues related to modeling the body mass index-mortality association: the shape of the association and the effects of smoking status. *Int J Obes (Lond)*. 2008;32(suppl 3):S52–S55.
21. Cooper RS. Which factors confound or modify the relationship between body weight and mortality? *Int J Obes (Lond)*. 2008;32(suppl 3): S47–S51.
22. Habbu A, Lakkis NM, Dokainish H. The obesity paradox: fact or fiction? *Am J Cardiol*. 2006;98:944–948.
23. Pi-Sunyer FX. Weight loss and mortality in type 2 diabetes. *Diabetes Care*. 2000;23:1451–1452.
24. Braun TP, Zhu X, Szumowski M, et al. Central nervous system inflammation induces muscle atrophy via activation of the hypothalamic-pituitary-adrenal axis. *J Exp Med*. 2011;208: 2449–2463.
25. Morley JE, Thomas DR, Wilson MM. Cachexia: pathophysiology and clinical relevance. *Am J Clin Nutr*. 2006;83:735–743.
26. Flegal KM, Graubard BI, Williamson DF, Cooper RS. Reverse causation and illness-related weight loss in observational studies of body weight and mortality. *Am J Epidemiol*. 2011;173:1–9.

 © 2014 Lippincott Williams & Wilkins

27. Ferreira I, Stehouwer CD. Obesity paradox or inappropriate study designs? Time for life-course epidemiology. *J Hypertens.* 2012;30:2271–2275.
28. Dahabreh IJ, Kent DM. Index event bias as an explanation for the paradoxes of recurrence risk research. *JAMA*. 2011;305:822–823.
29. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15:615–625.
30. Cole SR, Platt RW, Schisterman EF, et al. Illustrating bias due to conditioning on a collider. *Int J Epidemiol*. 2010;39:417–420.
31. Banack HR, Kaufman JS. The "obesity paradox" explained. *Epidemiology*. 2013;24:461–462.
32. Gregg EW, Gerzoff RB, Thompson TJ, Williamson DF. Trying to lose weight, losing weight, and 9-year mortality in overweight U.S. adults with diabetes. *Diabetes Care*. 2004;27:657–662.
33. Willi C, Bodenmann P, Ghali WA, Faris PD, Cornuz J. Active smoking and the risk of type 2 diabetes: a systematic review and meta-analysis. *JAMA*. 2007;298:2654–2664.
34. National Center for Health Statistics; Office of Analysis and pidemiology. *The National Health Interview Survey (1986–2004) Linked Mortality Files, Mortality Follow-up Through 2006: Matching Methodology, May 2009*. Hyattsville, MD; 2009. Available at: http://www.cdc.gov/nchs/data/datalinkage/matching_methodology_nhis_final.pdf.
35. Kirkman MS, Kendall DM. Hemoglobin A1c to diagnose diabetes: why the controversy over adding a new tool? *Clin Chem*. 2011;57:255–257.
36. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2010;33(suppl 1):S62–S69.
37. Pirie K, Peto R, Reeves GK, Green J, Beral V. The 21st century hazards of smoking and benefits of stopping: a prospective study of one million women in the UK. *Lancet*. 2012;6736:1–9.
38. Fewell Z, Davey Smith G, Sterne JA. The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. *Am J Epidemiol*. 2007;166:646–655.
39. Perry GS, Byers TE, Mokdad AH, Serdula MK, Williamson DF. The validity of self-reports of past body weights by U.S. adults. *Epidemiology*. 1995;6:61–66.

© 2014 Lippincott Williams & Wilkins