Commentary

Reverse Causation and Illness-related Weight Loss in Observational Studies of Body Weight and Mortality

Katherine M. Flegal*, Barry I. Graubard, David F. Williamson, and Richard S. Cooper

* Correspondence to Dr. Katherine M. Flegal, National Center for Health Statistics, Centers for Disease Control and Prevention, 3311 Toledo Road, Room 4336, Hyattsville, MD 20782 (e-mail: kmf2@cdc.gov).

Initially submitted June 5, 2010; accepted for publication September 10, 2010.

In studies of weight and mortality, the construct of reverse causation has come to be used to imply that the exposure-outcome relation is biased by weight loss due to preexisting illness. Observed weight-mortality associations are sometimes thought to result from this bias. Evidence for the occurrence of such bias is weak and inconsistent, suggesting that either the analytical methods used have been inadequate or else illness-related weight loss is not an important source of bias. Deleting participants has been the most frequent approach to control possible bias. As implemented, this can lead to deletion of almost 90% of all deaths in a sample and to deletion of more overweight and obese participants than participants with normal or below normal weight. Because it has not been demonstrated that the procedures used to adjust for reverse causation increase validity or have large or systematic effects on relative risks, it is premature to consider reverse causation as an important cause of bias. Further research would be useful to elucidate the potential effects and importance of reverse causation or illness-related weight loss as a source of bias in the observed associations between weight and mortality in cohort studies.

bias (epidemiology); body mass index; body weight; confounding factors (epidemiology); epidemiologic methods; mortality; selection bias

Abbreviations: BMI, body mass index; COPD, chronic obstructive pulmonary disease; IRWL, illness-related weight loss.

Many analyses have investigated the association of relative weight or body mass index (BMI) with mortality with broadly consistent results. Studies in young and middle-aged individuals often show a curvilinear relation with increased mortality risk at both low and high BMI values (1–3). Relative risks of mortality in these studies are usually below 2.0 and often below 1.5. Studies in the elderly often show the highest relative risks at low BMI levels, with little excess risk at higher BMI levels (4, 5). It is sometimes suggested that these observations result from reverse causation, namely, bias caused by preexisting illness and attendant weight loss (6–8). In this commentary, we discuss some unresolved methodological issues regarding the concept of reverse causation and the effects of preexisting illness in studies of weight and mortality and make some brief suggestions for directions in future research. This is not intended as a comprehensive literature review of this complex topic.

UNRESOLVED ISSUES

Definition of reverse causation in studies of weight and mortality

Reverse causation ordinarily refers to the situation in which the outcome precedes and causes the exposure instead of the other way around (9–11). In much of the epidemiologic literature, the term “reverse causation” is used in this standard sense in contexts as varied as studies of asthma and antibiotics (are people with asthma prone to infections, or do antibiotics increase the risk of asthma?) (12), cancer and cholesterol levels (is low cholesterol a risk factor for cancer, or does cancer cause low cholesterol levels?) (13), and BMI and depression (are those with high BMI more likely to be depressed, or does depression lead to high BMI?) (14).

In 1999, the term “reverse causation” began to be used in studies of weight and mortality in a different (and somewhat...
The proportion of study participants with IRWL prior to baseline

In analyses using baseline weight as the exposure, the question of interest from an epidemiologic perspective is what proportion of study participants is likely to have had substantial IRWL prior to baseline. Without a clear way to identify such participants, this proportion is difficult to estimate. Patients with extreme weight loss are often very severely ill and tend to have short survival times, often under a year (15). This minimizes the likelihood that many such patients would be included in a baseline sample of participants in an observational study. According to several descriptions in Table 1, weight loss prior to baseline may be due to undiagnosed cancers. Several studies show that undiagnosed cancers may be associated with weight loss and that a substantial minority of patients referred because of weight loss may have undiagnosed cancers (16–21). As noted by Lankisch et al. (17, p. 45), several investigations of weight loss (16–18) “reveal that non-malignant, rather than malignant, diseases are the major cause of weight loss, and that gastrointestinal tract disorders predominate in both categories.”

Available data suggest little effect of undiagnosed neoplasms on baseline weight in population studies. In an Austrian study of 65,000 participants, with weight measured annually for 7 years before baseline and cancer incidence determined for 8 years after baseline, there was no clear association between all-cancer incidence and either weight gain or weight loss (22). In a group who had had weight
measured every 2 months. Kritchevsky et al. (23) found that men who eventually developed cancer weighed 0.1 kg less 6–7 years before diagnosis and 1.2 kg less within the 6 months preceding diagnosis. Wilcosky et al. (24, p. 750) found that “decedents from cancer in the early years of follow-up were no leaner at baseline than were those who died in the later years” and felt that this argued against reverse causation. Peterson and Trell (25) found that men who died of cancer did not differ in weight from survivors. Rose and Shipley (26) addressed “unsuspected sickness” as a possible explanation for low cholesterol and cancer and found that seemingly healthy men who died of cancer within the first few years after baseline tended to have lower cholesterol levels, but not lower weight levels, than men dying of cancer after the first few years. In the Physicians’ Health Study, the relative risk for incident prostate cancer was the same for BMI at baseline as for BMI from 8 years before baseline (27). In the Austrian study (22) and the Health Professionals Follow-up Study (28), the diagnosis of colon cancer was preceded by weight gain rather than by weight loss. Several other studies of weight change prior to cancer diagnoses show an association of weight gain, rather than loss, with cancer incidence (29–33).

Weight loss prior to diagnosis of other conditions has been reported. Chen et al. (34) found a reported loss of 5.2 pounds (2.4 kg) in the 10 years prior to the diagnosis of Parkinson’s disease. Johnson et al. (35) found that the rate of weight loss approximately doubled, from 0.6 pound (0.3 kg) per year to 1.2 pound (0.5 kg) per year, in the year preceding a diagnosis of Alzheimer’s disease and that participants who eventually developed dementia weighed about 8 pounds (3.6 kg) less than controls even at study entry. Stewart et al. (36) similarly found that in late life men who developed incident dementia lost 0.36 kg/year more than men who did not.

Diagnosed preexisting illness may be associated with either weight loss or weight gain, perhaps due to treatment. For example, in the Atherosclerosis Risk in Communities Study (37), participants with preexisting illness (cardiovascular disease, cancer, or self-reported poor health) were about equally likely to have gained 3 or more BMI units (6.4%) as to have lost 3 or more BMI units (4.8%) in the 3 years prior to baseline. Among participants with cancer, 4.4% lost 3 or more BMI units, and 5.9% gained 3 or more units. Breast cancer patients often gain weight after treatment (38, 39). Parkinson’s disease is associated with weight loss, but some treatments for Parkinson’s lead to marked weight gain (40). Treatment of hyperthyroidism led to an average gain of more than 9 kg over a 4-year period (41). Weight changes after a diabetes diagnosis ranged from gain to loss (42). Several drugs for the treatment of migraine are associated with weight gain (43).

Nature and direction of potential bias due to IRWL

Discussions of potential biases due to IRWL do not always distinguish between weight loss and low BMI. Moderate weight loss may not lead to a low BMI, and low BMI is not necessarily the result of weight loss. IRWL is sometimes thought to bias relative risks upward at low BMI levels and downward at higher BMI levels, but the possible mechanisms of such bias have not been completely described. Most discussions have focused on the effects at very low BMI levels. The proposed mechanism of bias is that some severely ill people lose sufficient weight as a result of their illness to fall into the very low BMI category. The prevalence of very low BMI, particularly among nonsmokers, is often so low that the admixture of even a small number of severely ill people at high risk of mortality could potentially increase the apparent risk in this group.

Even though it is sometimes assumed that IRWL will bias relative risks downward at somewhat higher BMI levels, there has been little discussion of the possible type and direction of bias at higher BMI levels. For example, chronically ill obese people might migrate into the overweight category, thus increasing rather than decreasing the apparent risk in the overweight group. The expected effects of IRWL on relative risks in any weight category depend on a number of variables, including the relative prevalence of each weight category, the probabilities of weight change (in either direction), and the absolute mortality risk in each group. As a consequence, any effects of IRWL on relative risks are not readily predictable and might vary in magnitude and even in direction from cohort to cohort.

Sometimes circular reasoning, based on a priori assumptions about the direction of the expected bias from IRWL, is used to infer the presence of IRWL bias. If an analytical maneuver intended to reduce IRWL bias leads to small changes in relative risk in the expected direction, this may be taken as evidence that the assumed bias was present (44). However, if the changes are not in the expected direction, then this may be interpreted as evidence that the adjustment for bias was insufficient, rather than as an absence of bias. For example, Baik et al. (45, p. 270) state: “...we found evidence that reverse causation strongly influenced the shape of the relation between body mass index and mortality in these data. Even with careful attempts to reduce this artifact by excluding persons who reported chronic disease at baseline and those with recent weight loss, we were unable to completely avoid the effects of reverse causation.”

In general, the prevalence of existing disease is the same or higher among those who are overweight and obese as among those of lower weights (44, 46). This suggests that the effect of preexisting disease without IRWL might be to introduce positive confounding, biasing the relative risks upward.

Controlling for IRWL in statistical analysis

If one had ideal data including detailed health status at baseline and subsequent mortality data, then the BMI-mortality dose-response could be defined separately in the several strata defined by baseline health status, age, and smoking. One could look at healthy nonsmokers and determine whether their dose-response to relative weight was different from those in other strata. Perhaps arguments about reverse causation would largely disappear if such data were available. However, there is no clear way to identify healthy participants or participants who have lost weight due to
illness. Various maneuvers have been proposed to compensate for the lack of direct information. These approaches can be applied regardless of whether any of the participants have IRWL.

The method most frequently proposed to control for possible IRWL bias is to create a subgroup by excluding participants from the definitive analyses. Willett et al. (8, p. 428) suggested the following approach to exclude study participants who might have lost weight prior to baseline because they were ill:

Subjects with diagnoses that might affect weight and subjects who report recent weight loss, such as during the previous five years, can be excluded from a prospective study. Deaths that occur during the first several years of follow-up—possibly as a result of conditions that caused lower weights at base line—can also be excluded.

With this indirect approach, those with IRWL are not necessarily excluded and those who are excluded do not necessarily have IRWL. Such exclusions do not specifically target people with low BMI and often result in the deletion of as many or more overweight and obese people as lean people (2, 45, 47). The probability that those excluded actually have IRWL is difficult to estimate and may vary considerably from cohort to cohort. To the extent that those excluded are likely to have preexisting disease but not weight loss, these exclusions become in effect some degree of control for preexisting disease among participants without IRWL.

The scale of these proposed deletions can be quite large. It has been explicitly argued (7) that it is necessary to start with very large data sets because most of the data will have to be excluded in order to get the correct (i.e., unbiased) results. Similar deletions, combined with exclusion of current and former smokers, can result in deletion of almost 90% of the deaths in a sample. For example, in both a report of the Nurses’ Health Study (2) and the National Institutes of Health (NIH)-AARP Study (44), final analyses included only 11% of the original deaths.

A number of studies have applied some variant of this approach by complete exclusion of part of the sample or by stratification (refer to Table 2 for some examples). Of these suggestions, deleting early mortality has been most frequently used. Excluding early mortality is intended to at least partially reduce any potential bias arising from IRWL among participants who were at high mortality risk at baseline due to illness. Studies of the effects of deleting early mortality have shown little effect except perhaps in the first year after baseline (37, 48–52). One of the reasons why the deletion of early mortality tends to make little difference may be because studies with measured weight and height include only participants who are able to attend an examination, thus in practice excluding sicker participants.

Some studies allow for comparison of results before and after exclusions. Careful investigations in a number of cohorts (53–65) have not shown any marked or systematic impact of deletion or adjustment for illness, weight loss, or early mortality, although the approaches used are heterogeneous. Large-scale deletions in some cohorts (2, 44, 45, 47, 66) show small changes in relative risks that are not always in the hypothesized directions. In the Nurses’ Health Study (2), the multivariate relative risk for BMI of 32 or above increased from 1.9 to 2.2 among nonsmokers after excluding women who had gained or lost at least 4 kg in the first 3 years after baseline and excluding the first 4 years of mortality. Among never smokers in the Cancer Prevention Study II, the relative risk for a BMI of 25–26.4 was 1.01 for men without prevalent disease and 1.05 for men with prevalent disease; corresponding values for women were 1.03 and 1.06 (66). In the Health Professionals Follow-up Study (45), the relative risk for a BMI of 30 or above fell from 1.50 to 1.49 after excluding participants with at least 10-pound (4.5-kg) weight losses before baseline and excluding the first 4 years of mortality. In the Physicians’ Health Study (47), the relative risk for a BMI of 30 or above rose from 1.67 to 1.71 after excluding the first 2 years of mortality. In middle-aged adults in the US Health and Retirement Study, the relative risk for a BMI of 35 or above dropped from 1.53 to 1.29 after adjusting for disease history and health status and was 1.52 when analyses were limited to those in good or excellent health (61).

Do these methods eliminate bias or introduce bias?

The use of numerous exclusions and subsets of the data in studies of weight and mortality is, in effect, analysis of subgroups, which can lead to many known methodological problems (67–71). The approaches used in weight and mortality studies are often inconsistent with recommended practices for subgroup analyses, such as using predefined subgroups, presenting results for all subgroups, not only for selected subgroups, and using formal tests of heterogeneity. Wang et al. (71, p. 2193) note that subgroup analyses “can lead to overstated and misleading results” and caution: “Avoid overinterpretation of subgroup differences. Be properly cautious in appraising their credibility, acknowledge the limitations, and provide supporting or contradictory data from other studies, if any.”

Results obtained after extensive exclusions do not necessarily provide a more valid and less biased estimate. They could also be simply an artifact of random variability from excluding such large proportions of the sample (72) or be biased by the effect of the exclusions themselves. The exclusions reduce statistical power, selectively affect certain causes of death, and may introduce additional, unrecognized sources of bias. In the large data sets needed to accommodate large-scale exclusions, weight and height are often self-reported, a characteristic that itself may lead to bias (73). It is possible, particularly with self-reported weight and height, that the deletions lead to a subgroup with measurement errors different from the full sample or to a subgroup that has different confounding characteristics (74).

In general, the effect of these exclusions is to create a subgroup that is poorly characterized. A subgroup may be described as “healthy” after a small number of health conditions are excluded (7, 44, 66); however, generally other health conditions, such as diabetes, are not excluded, and there is no control for risk factors such as hypertension or dyslipidemia. There may be little or no difference in overall health status between those included and those excluded.
Repeated deletions followed by repeated statistical testing increase the likelihood that false positive associations will be detected. In studies of weight and mortality, the data are sometimes analyzed with a sequence of exclusions, each time testing the relative risks in the new subgroup without adjustments for the multiple comparisons, until a monotonic relation is observed. The probability of a false research finding is increased by this procedure, by the small effect sizes (relative risks) characteristic of weight and mortality studies, by the small data sets that may be created by excluding 70%–90% of the deaths, and by nonstandardized approaches that vary between studies (75). Comparisons between the relative risks before and after exclusions are often done by inspection, and point estimates of relative risk may be described as being different between 2 subgroups without any statistical testing (2, 44).

Interpretive biases (76) may arise because of prior assumptions about the direction and importance of the effects of IRWL. Findings of other-than-expected weight–mortality associations have sometimes been attributed to some undetected source of reverse causation bias. For example, after Fontaine et al. (77) found a weaker association between weight and mortality in blacks than in whites, Manson and Bassuk (78) speculated that these differences might be explained by effects of reverse causation affecting blacks more than whites because blacks might be more likely than whites to suffer from undiagnosed disease. Gelber et al. (79) suggested that the curvilinear rather than monotonic relation found by Calle et al. (66) among healthy never-smoking participants with no recent weight loss may have been due to the failure of the investigators to exclude participants with early deaths. Lee and Manson (80) interpreted a modestly

### Table 2. Some Examples of Exclusions Used in Studies of Weight and Mortality

<table>
<thead>
<tr>
<th>First Author (Reference)</th>
<th>Prior Illness</th>
<th>Weight Loss/Change</th>
<th>Early Deaths Excluded</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ajani (47)</td>
<td>Myocardial infarction, angina, stroke, transient ischemic attacks, cancer, liver disease, or renal disease</td>
<td>None</td>
<td>2 years</td>
<td></td>
</tr>
<tr>
<td>Baik (45)</td>
<td>Myocardial infarction, angina, coronary artery bypass grafting or angioplasty, stroke, transient cerebral ischemia, peripheral venous thrombosis, intermittent claudication, pulmonary embolus, heart rhythm disturbances, cancer, renal failure, chronic pulmonary disease</td>
<td>Reported weight loss of 10 pounds (4.54 kg) or more in 5 years prior to baseline</td>
<td>4 years</td>
<td>Statistical adjustment for smoking. Excluded BMI outside the range of 15–50 kg/m²</td>
</tr>
<tr>
<td>Calle (66)</td>
<td>Cancer, heart disease, stroke, respiratory disease, any current illness at baseline</td>
<td>Reported weight loss of 10 pounds or more in year prior to baseline</td>
<td>Not excluded</td>
<td></td>
</tr>
<tr>
<td>Durazo-Arvizu (56)</td>
<td>Cardiovascular disease or cancer</td>
<td>None</td>
<td>4 years</td>
<td></td>
</tr>
<tr>
<td>Gu (58)</td>
<td>Cardiovascular disease, cancer, stroke, chronic obstructive pulmonary disease, end-stage renal disease</td>
<td>None</td>
<td>5 years</td>
<td>Excluded participants with heavy alcohol use</td>
</tr>
<tr>
<td>Hozawa (90)</td>
<td>None</td>
<td>None</td>
<td>5 years</td>
<td>Excluded participants with total cholesterol of &lt;4.1 mmol/L</td>
</tr>
<tr>
<td>Jee (91)</td>
<td>Atherosclerotic cardiovascular disease, cancer, liver disease, diabetes, or a respiratory disease at or before the initial study visit</td>
<td>None</td>
<td>2 years</td>
<td>Excluded participants with BMI of &lt;16 kg/m² or stature of &lt;1.3 m</td>
</tr>
<tr>
<td>Lawlor (110)</td>
<td>None</td>
<td>None</td>
<td>5 years</td>
<td></td>
</tr>
<tr>
<td>Lee (111)</td>
<td>History of coronary heart disease, stroke, or cancer</td>
<td>None</td>
<td>5 years</td>
<td></td>
</tr>
<tr>
<td>Lindsted (112)</td>
<td>Heart disease, stroke, or cancer or severe physical complaints (e.g., chest pain, shortness of breath, loss of appetite)</td>
<td>Loss or gain of &gt;10 pounds in 5 years before baseline</td>
<td>15 years</td>
<td>Excluded deaths due to poisoning, accidents, or congenital malformations</td>
</tr>
<tr>
<td>Manson (2)</td>
<td>Cancer or cardiovascular disease</td>
<td>Gain or loss of 4 kg or more after baseline</td>
<td>4 years</td>
<td></td>
</tr>
<tr>
<td>Whitlock (93)</td>
<td>None</td>
<td>None</td>
<td>5 years</td>
<td>Statistical control for smoking</td>
</tr>
</tbody>
</table>

Abbreviation: BMI, body mass index.

*Except as noted, all the studies listed also excluded current and former smokers.*

*Am J Epidemiol* 2011;173:1–9
Bias or causality?

It is sometimes assumed that, if illness leads to low weight, then the association of low weight with increased mortality is not causal. Low weight itself, however, may also contribute to increased mortality, perhaps from a different cause than the illness causing weight loss. For example, low weight can lead to lower bone density (81, 82), which is a risk factor for mortality from complications subsequent to hip fracture (81–86). A large study of colon cancer patients found that underweight patients were at significantly increased risk of mortality from causes other than colon cancer (87). Studies of patients with chronic obstructive pulmonary disease (COPD) suggest that not only is low weight associated with increased mortality after controlling for disease severity but also weight gain improves mortality outcomes, suggesting that the low weight itself is in part the causative factor (88). A study using Mendelian randomization suggested that the inverse association of obesity and lung cancer may be causal rather than artifactual (89). Large studies in the United States, the United Kingdom, India, Japan, Korea, and China have shown increased mortality at low BMI values among never smokers after deletions for prevalent illness and early mortality (44, 58, 64, 66, 90–93). Thus, at least some of the increase in mortality at low BMI levels does not appear to be artifactual and cannot be explained solely by reverse causation.

The impact of a given weight level on mortality is a function of the effects of weight on disease incidence, the effects of disease on weight, and the effects of weight on survival in patients with that disease. It is difficult to distinguish among these effects, and all may be occurring within a given sample, as discussed, for example, by Jee et al. (91) for COPD. Any effect of weight loss leading to poorer survival should not be considered a form of bias. According to Baik et al. (45, p. 269), “[T]he interpretation of the elevated mortality among the leanest older men depends heavily on whether being lean causes chronic pulmonary disease or is the result of chronic pulmonary disease”; however, these authors overlooked the third possibility that leanness increases mortality in COPD patients. A body of evidence has begun to accumulate suggesting that, in numerous health conditions, low weight is associated with poorer survival even after adjustments for disease severity (94–97). To the extent that this is the case, it is unnecessary to invoke bias to explain the numerous studies showing an inverse BMI-mortality relation in the elderly.

FUTURE DIRECTIONS

It cannot be ruled out that bias due to preexisting illness may affect weight-mortality studies, possibly through mechanisms other than IRWL. Attempts to adjust for biases due to IRWL, however, have provided little evidence to date for the existence of such biases. This might be because the actual prevalence of IRWL prior to baseline is very low in most cohorts or because IRWL is offset by illness-related weight gain. Another possibility is that IRWL produces little or no bias or that the bias is in the opposite direction of that hypothesized. The effects of low weight on mortality may be causal, not artifactual, and thus not be a form of bias. Yet another possibility is that bias occurs, but the crude approach of deleting large numbers of participants is an ineffective way to control bias. Exclusions of large numbers of participants without IRWL may mask the effects of deleting some individuals with IRWL. The deletions themselves may increase rather than reduce bias.

There are no agreed-upon definitions of reverse causation in studies of weight and mortality, no good evidence as to how often this phenomenon occurs, and little theoretical or empirical basis for describing what type of bias it might cause. Furthermore, we have no good way to control for this putative bias, and the proposed methods may actually cause bias rather than correct for bias.

Indirect methods such as deleting large amounts of data do not appear to be likely to lead to further insights. In future research, it would be preferable to avoid the use of vague terms such as “reverse causation” or “healthy” and to create testable hypotheses for specific effects in specific cohorts. Standard methods for assessing subgroups should be used, including statistical testing of interactions (98, 99). More focused and detailed investigations would be useful, such as those carried out by Stevens et al. (37) on weight change and health, as well as the use of more recent methods for causal modeling (100, 101). Cohorts with repeated measures of weight and health status could be exploited for this purpose, as has been done with data on smoking and smoking cessation (102, 103).

CONCLUSIONS

At present, there is little evidence that observed associations between weight and mortality in cohort studies are biased by effects of preexisting illness. Continued application of indirect approaches to adjust for IRWL may introduce new sources of bias. Because it has not been demonstrated that these procedures give more valid results or have large or systematic effects on relative risks, it is premature to consider reverse causation as an important cause of bias. Further research would be useful to elucidate the potential effects and importance of reverse causation or illness-related weight loss as a source of bias in the observed associations between weight and mortality in cohort studies.

ACKNOWLEDGMENTS

Author affiliations: National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, Maryland (Katherine M. Flegal); Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland (Barry I. Graubard); Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, Georgia (David F. Williamson); and Department of Preventive Medicine and Epidemiology,}

Am J Epidemiol 2011;173:1–9
REFERENCES


37. Stevens J, Juhaeri Cai J. Changes in body mass index prior to baseline among participants who are ill or who die during the early years of follow-up. Am J Epidemiol. 2001;153(10):946–953.

38. Irwin ML, McTiernan A, Baumgartner RN, et al. Changes in body fat and weight after a breast cancer diagnosis: influence...


