Flawed methods and inappropriate conclusions for health policy on overweight and obesity: the Global BMI Mortality Collaboration meta-analysis

Katherine M. Flegal[1], John P.A. Ioannidis[1,2,3,4] & Wolfram Doehner[5,6,7]*

[1]Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA, [2]Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA, USA, [3]Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, CA, USA, [4]Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA, [5]Division of Cardiology and Metabolism; Department of Cardiology (CVK), Deutsches Zentrum für Herz-Kreislauf-Forschung (DZHK), Charité-Universitätsmedizin Berlin, Berlin, Germany, [6]Berlin-Brandenburg Center for Regenerative Therapies (BCRT), Charité Universitätsmedizin Berlin, Berlin, Germany, [7]Center for Stroke Research Berlin (CSB) Charité Universitätsmedizin Berlin, Berlin, Germany

Guideline recommendations and health policy decisions rely on evidence from clinical and epidemiological studies. Adequate methodology and appropriate conclusions are essential to support healthcare and health policy decisions. An analysis of body mass index and mortality by the Global BMI Mortality Collaboration (GBMC) concluded that the association of excess body weight with higher mortality was similar worldwide and that overweight and obesity should be combated everywhere. To reach this conclusion, the GBMC used highly selected data, rather than a systematic approach. The GBMC initially chose individual participant data from 239 prospective studies with approximately 10.6 million participants. The GBMC then excluded over 60% of data and over 75% of fatal events by eliminating all cases with any reported disease at baseline or smoking history and all events within the first 5 years of follow-up. After applying these restrictions, the association of overweight with lower mortality was reversed and the association of obesity with higher mortality was increased. Given the major flaws in the selection process, in the adequacy of the data, in the data analysis, and in the interpretation, the GBMC conclusions should be viewed sceptically as a guide to action, either for clinical decisions or for public health in general. The flawed conclusion that overweight is uniformly associated with substantially increased risk of death and thus should be combated in any circumstances may lead not only to unjustified treatment efforts and potential harm in a wide range of clinical conditions but also to a tremendous waste of resources.

Keywords Body weight; Weight change; Health policy; Outcome; Risk factor; Epidemiologic methods; Obesity paradigm

Introduction

In 2016, the Global BMI Mortality Collaboration (GBMC) published a meta-analysis[1] of body mass index (BMI) and all-cause mortality that has already been cited 378 times as of September 2018. This highly visible publication was authored by a writing committee of 61 distinguished epidemiologists and included over 500 collaborators.

The data selection

The publication was a meta-analysis but not a systematic review. Examinations of the GBMC article revealed numerous inconsistencies and flaws in the way the data were selected.[2–4] As we discuss in detail elsewhere,[3,4] the data selection appeared to consist only of articles already known to the senior GBMC author and whose results were also

largely already known. Some data meeting the inclusion criteria were excluded; in turn, some data meeting the exclusion criteria were included. Some of these inappropriate inclusions and exclusions could have changed the essential findings of the meta-analysis. Appropriate data sets, some even containing millions of additional participants, were not sought out or included.[2–4]

The data restriction

The analytic methods used in the GBMC article were flawed. The original sample consisted of 239 studies with a combined sample size of 10 625 411. However, the authors deleted over 60% of the data that they considered and about 75% of the deaths to arrive at their final results. The final sample used for the primary pre-specified analysis after applying the planned restrictions consisted of 189 studies with a sample size of 3 951 455. These analyses are described repeatedly as 'strict' or 'stricter' primary analyses, using these words six times in the article and another six times in the supplement. This word choice implies that carefully selected criteria were rigorously applied. However, the GBMC analyses were far from 'strict' in this regard. The GBMC paper in fact, had very limited ability to standardize covariate or even exposure definitions and could not standardize the data used for the listed exclusion criteria, as discussed below.

The GBMC article claims that 'The primary analysis was restricted to never-smokers without certain known chronic diseases at baseline (e.g., cardiovascular disease, cancer, or respiratory diseases), and omitted the first 5 years of follow-up'. Because of the lack of complete data on smoking or pre-existing disease, however, the primary analysis could not even perform the claimed exclusions adequately. Only 28 of the 239 studies even had complete data for the three chronic diseases mentioned previously (see eTable 2 in GBMC[1]). Twenty studies had no information on any of the pre-existing diseases. Only 15 of the 239 studies provided data on respiratory disease and only 68 had data on cancer. Only 19 studies had complete data on pre-existing heart disease, stroke, and cancer. Only 56 studies had data on both cardiovascular disease and cancer.

A major exclusion criterion of the primary analysis was smoking status. Most studies in the GBMC publication could not identify never-smokers adequately (see eTable 6 in GBMC[1]). Only 98 of the 239 studies were able to distinguish between never-smokers and current or former smokers. Thus, the GBMC analyses were not able to appropriately adjust for smoking or pre-existing disease because the researchers were unable to harmonize either smoking or pre-existing disease data across studies.

The GBMC article asserts: 'To help achieve more valid estimates, prospective studies of BMI and mortality should, when possible, exclude: smokers, participants who already have some chronic disease at recruitment that could affect BMI, and those dying within 5 years of recruitment'.[5–8] To justify the claim that these restrictions help to achieve more valid estimates, the authors cite a letter to the editor[7] and three articles.[5,6,8] None of these citations have any data to support the statement that the restrictions used in the GBMC restrictions will decrease bias. Two[5,7] are simply opinion statements written by several authors of the GBMC article, one is a simulation study[6] with no data and which actually does not reach this conclusion, and one[8] does not address bias at all. Thus, the GBMC citations do not support the claim that these restrictions produce more valid estimates. These 'restrictions' seem to yield estimates that the GBMC senior investigators apparently believe to be correct. Expert opinion is a common category in schemata of levels of evidence used in guidelines and recommendations. However, it is usually considered the lowest level of evidence. It is important to clearly identify conclusions based on expert opinion and not to mask them with inconsistent, selective, and over-interpreted data.

Previous analyses by GBMC authors of the data sets used in the GBMC article did not use these exact restrictions, in particular those involving the deletion of initial years of follow-up. The GBMC analyses deleted the first 5 years of follow-up, but other analyses of subsets of the same data by the same authors used either no deletions for follow-up,[9–13] or less than 5 years.[14–18] Only Chen *et al.*[19] and Whitlock *et al.*[20] deleted 5 years. No justification is given why 5 years should be excluded, and no sensitivity analyses are provided on what the effects would be of using different lengths of exclusion or no exclusion of any follow-up.

There is no evidence that the GBMC methods are successful in controlling bias. First, even their attempts to control for smoking and pre-existing illness were inadequate because most of their studies either lacked or did not report the necessary data. Second, although the restrictions changed the results, there is no evidence that this is because they were effective at controlling bias.

Unaddressed bias

In contrast to the debatable and inconsistent restriction procedures as outlined previously, the GBMC analyses ignored a major obvious source of potential bias, namely, the use of self-reported rather than measured weight and height. Unlike the other sources of potential bias that they tried to control for, there is ample documentation of systematic errors in self-reported weight and height, leading to

systematic bias, errors in BMI, and misclassification into the correct BMI categories relative to measured weight and height.[17,21–26] In the primary GBMC analysis of North American data, 94% of deaths came from 20 studies with self-reported weight and height. The GBMC found that for North America, the hazard ratio for overweight was 1.00 using measured weight and height and 1.14 using self-reported weight and height. BMI calculated from self-reported weight and height is systematically biased relative to BMI calculated from objectively measured data,[22] despite a high correlation between the two. The errors in self-reported weight and height can lead to substantial misclassification into BMI categories and attendant bias with overestimation of hazard ratios.[21,23,24,26]

The errors in BMI classification from self-reported weight and height are known to vary by multiple potential confounding factors such as age, sex, race, socio-economic status, health conditions, region, and language.[27] Unmeasured residual confounding is also likely to vary unpredictably across studies and thus can bias not only individual studies but also comparisons between subgroups and regional comparisons. Because of the high potential for misclassification and unmeasured confounding associated with self-report, an easily accomplished first step, if truly more 'strict' control for bias had been intended, would require excluding studies with self-reported height and weight and limiting analyses to studies with measured data. Unfortunately, the GBMC article excluded large numbers of available data sets with measured height and weight that either were not in the initial group of 239 studies that were considered or were not included in the group of 189 studies used in the final analyses. Their final analyses included far more participants with self-reported weight and height than with measured weight and height. The GBMC purports to draw inferences about BMI categories, but such inferences can be quite inaccurate because the rate of misclassification into the wrong measured BMI categories is high when self-reported weight and height data are used.[25] The effects can be strong enough to change the direction of hazard ratios for some categories and particularly the debated overweight category.[23]

The discussion in the context of other published studies

The GBMC article claims, incorrectly, that 'our findings are also broadly consistent with the stricter analyses done in a 2015 study of 12 million Korean adults' by Yi et al.[28] This is a huge study with larger sample size than all 239 GBMC datasets combined but which was nevertheless not even considered by GBMC. The estimates from the GBMC paper are actually not consistent, either broadly or narrowly, with the Korean study. In fact, Yi et al. found almost no effect of applying the

GBMC restrictions, in sharp contrast to the GBMC claims that the restrictions drastically change ('correct') the results. Yi et al. in particular still found hazard ratios below 1 for overweight subjects, suggesting that in fact, the GBMC findings are misleading and the applied restrictions introduced bias to the GBMC data.[29] The Yi paper is instead consistent with the findings of Flegal et al.[30] In particular, the data from Yi et al. showed that overweight was not associated with an increased risk of death (hazard ratio of 0.85) and grade 1 obesity (BMI 30–35) was associated with only a very small increased risk of death (hazard ratio of 1.06).

Unjustified generalization of findings

The GBMC report claims that its findings support strategies to combat the entire spectrum of excessive adiposity across all populations, but this is not clearly justified. The primary pre-specified analysis excluded (where identifiable) ever-smokers, excluded those with some prevalent disease (again, where identifiable), and required a survival of at least 5 years. The conclusion from these data thus would at best only be applicable to healthy, never-smoking individuals with a life expectancy of more than 5 years. Furthermore, the data are inadequate to even support those exclusions, and the exclusions lack a clear rationale. The findings cannot, and must not, be applied to patients with any chronic disease setting where long survival is uncertain.

In this context, the GBMC authors incorrectly quote previous work on protective effects in supposedly healthy individuals.[31] Contrary to what they claim, the cited study did not address apparently healthy individuals but clearly referred to patients with established cardiovascular disease. There is substantial and constantly growing evidence from multiple studies in wide ranges of disease states and in numerous disease settings and using many ways of analyses, that shows that once a chronic disease is established, overweight (BMI 25–30) or mild obesity (BMI 30–35) is associated with incremental longevity as compared with those patients with a normal BMI (18.5 to <25).[32] The extra weight (if 'extra' is even a proper term here) may protect these patients against ravages of diseases such as heart failure, advanced chronic kidney disease,[33] terminal cancers, as well as sarcopenia of old age.[34] Arguments that this is just an artefact of selection bias are undermined by the observation that the same phenomenon is seen even for chronic diseases that are not obesity-related.[35]

The uncritical translation of findings in healthy non-smoking individuals to patients with established chronic diseases exemplifies a common misconception about the significance of weight reduction in such patients. Promoting weight reduction in patients with established chronic disease goes against clinical evidence and guidelines[36] and may be harmful for patients.

Applying emerging data in clinical decision-making—evolving from paradox to paradigm

The finding that overweight and mild obesity are associated with better outcome in a range of diseases was initially termed 'obesity paradox' because it appeared to be an unexpected and counterintuitive finding. This terminology has led to misunderstandings among researchers and the public alike. There is no precise definition to the term, and numerous and sometimes loosely related observations have been summarized under this seemingly paradoxical finding.[35] The core observation, however, of an inverse association of body weight with outcome in patients with prevalent diseases has been confirmed in multiple reports and in numerous patient cohorts, in different ethnic groups,[37] over a wide spectrum of cardiovascular diseases and disease severity and using various methods. With an increasing bulk of data confirming a survival benefit related to overweight and mild obesity in such diseases, and based on pathophysiologic concepts to explain these findings, there is sufficient scientific evidence to abandon the term 'paradox'. Importantly, a finding considered 'paradoxical' will not be applied in clinical practice as it implies an unexplained if not unscientific statement. Accordingly, in recognition of the robust evidence that has been accumulated and the plausible pathophysiologic concepts, it has been suggested that an *obesity paradigm* should be defined rather than a paradox.[38] This advance in terminology will allow the application of the evidence in clinical practice and promote a highly needed differentiated body-weight management that appreciates individual health conditions.

Potential impact of misleading conclusions on health care policy

In their sweeping conclusions, the GBMC authors claim that their results apply globally, but in fact, of the five regions they studied, the results are quite different for two of the five regions compared with the other three. A core issue is whether the overweight category has increased mortality risk. The 2013 US obesity guidelines,[39] sponsored jointly by the American Heart Association, the American College of Cardiology, and The Obesity Society, assert that overweight is not associated with excess mortality, citing the strength of the evidence as 'moderate'. The European Society of Cardiology Guidelines on diagnosis and treatment

of acute and chronic heart failure state that in patients with heart failure with moderate degrees of obesity (BMI < 35), weight loss cannot be recommended.[36] In contrast to the majority of available evidence, the authors of the GBMC report concluded with a recommendation to combat any degree of excessive body weight in any condition where it is found. This erroneous conclusion has major implications because a huge number of people are in this category. In most countries, people in the overweight category outnumber those in the obesity category. Analyses of burden of disease that were based on the estimates of risk from the GBMC analysis were prominently published in the NEJM.[40] They calculated that 4.0 million deaths annually worldwide are due to high BMI, and nearly 40% of them (1.5 million deaths) are in people who are not obese but only overweight.

It makes a huge difference if the annual number of deaths due to overweight is 1.5 million, as the GBMC estimates would suggest, or close to 0, as the totality of the evidence suggests. In the first scenario, a major investment of public health effort should be devoted to treatment and prevention in the overweight category. In the second scenario, such an investment of public health effort would be catastrophic: no benefit would be gained in the overweight category, and resources would shift away from obesity per se and underweight, where major increased risks of death are clearly documented. The lack of documentation of whether most currently available preventive and therapeutic interventions are sufficiently effective even in obesity,[41] let alone overweight, adds further to the potential for major waste of resources.

Given the major flaws in the selection process, in the adequacy of the data, in the data analysis, and in the interpretation, the GBMC conclusions cannot be trusted as a guide to action, either for clinical decisions or for public health in general.

Conflict of interest

The authors declare that no conflict of interest relevant to this article exists. The authors certify that they comply with the ethical guidelines for publishing in the Journal of Cachexia, Sarcopenia and Muscle: update 2017.[42]

References

1. Global BMI Mortality Collaboration. Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet* 2016;**388**:776–786.

2. Bhupathiraju SN, Di Angelantonio E, Danesh J, Hu FB. Commentary on "A meta-analysis but not a systematic review:

an evaluation of the Global BMI Mortality Collaboration". *J Clin Epidemiol* 2017;**88**: 30–32.

3. Flegal KM, Ioannidis JPA. A meta-analysis but not a systematic review: an evaluation of the Global BMI Mortality Collaboration. *J Clin Epidemiol* 2017;**88**:21–29.

4. Flegal KM, Ioannidis JPA. A meta-analysis of individual participant data constructed to align with prior expert views: comments on Bhupathiraju et al. *J Clin Epidemiol* 2017;**88**:33–36.

5. Manson JE, Stampfer MJ, Hennekens CH, Willett WC. Body weight and longevity. *A reassessment JAMA* 1987; **257**:353–358.

6. Singh PN, Wang X. Simulation study of the effect of the early mortality exclusion on confounding of the exposure-mortality relation by preexisting disease. *Am J Epidemiol* 2001;**154**:963–971.

7. Willett WC, Hu FB, Thun M. Overweight, obesity, and all-cause mortality. *JAMA* 2013;**309**:1681.

8. Bamia C, Trichopoulou A, Lenas D, Trichopoulos D. Tobacco smoking in relation to body fat mass and distribution in a general population sample. *Int J Obes Relat Metab Disord* 2004;**28**:1091–1096.

9. Berrington de Gonzalez A, Hartge P, Cerhan JR, Flint AJ, Hannan L, MacInnis RJ, et al. Body-mass index and mortality among 1.46 million white adults. *N Engl J Med* 2010;**363**:2211–2219.

10. Stevens J, Cai J, Pamuk ER, Williamson DF, Thun MJ, Wood JL. The effect of age on the association between body-mass index and mortality. *N Engl J Med* 1998;**338**:1–7.

11. Patel AV, Hildebrand JS, Gapstur SM. Body mass index and all-cause mortality in a large prospective cohort of White and Black U.S. adults. *PLoS One* 2014;**9**: e109153.

12. Pischon T, Boeing H, Hoffmann K, Bergmann M, Schulze MB, Overvad K, et al. General and abdominal adiposity and risk of death in Europe. *N Engl J Med* 2008;**359**:2105–2120.

13. Wormser D, Kaptoge S, Di Angelantonio E, Wood AM, Pennells L, Thompson A, et al. Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. *Lancet* 2011;**377**:1085–1095.

14. Jackson CL, Yeh HC, Szklo M, Hu FB, Wang NY, Dray-Spira R, et al. Body-mass index and all-cause mortality in US adults with and without diabetes. *J Gen Intern Med* 2014;**29**:25–33.

15. Jee SH, Sull JW, Park J, Lee SY, Ohrr H, Guallar E, et al. Body-mass index and mortality in Korean men and women. *N Engl J Med* 2006;**355**:779–787.

16. Joshy G, Korda RJ, Bauman A, Van Der Ploeg HP, Chey T, Banks E. Investigation of methodological factors potentially underlying the apparently paradoxical findings on body mass index and all-cause mortality. *PLoS One* 2014;**9**:e88641.

17. Ni Mhurchu C, Rodgers A, Pan WH, Gu DF, Woodward M. Body mass index and cardiovascular disease in the Asia-Pacific region: an overview of 33 cohorts involving 310 000 participants. *Int J Epidemiol* 2004;**33**:751–758.

18. Canoy D, Cairns BJ, Balkwill A, Wright FL, Green J, Reeves G, et al. Body mass index and incident coronary heart disease in women: a population-based prospective study. *BMC Med* 2013;**11**:87.

19. Chen Z, Yang G, Offer A, Zhou M, Smith M, Peto R, et al. Body mass index and mortality in China: a 15-year prospective study of 220 000 men. *Int J Epidemiol* 2012;**41**: 472–481.

20. Whitlock G, Lewington S, Sherliker P, Clarke R, Emberson J, Halsey J, et al. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet* 2009;**373**:1083–1096.

21. Chiolero A, Peytremann-Bridevaux I, Paccaud F. Associations between obesity and health conditions may be overestimated if self-reported body mass index is used. *Obes Rev* 2007;**8**:373–374.

22. Connor Gorber S, Tremblay M, Moher D, Gorber B. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obes Rev* 2007;**8**:307–326.

23. Flegal KM, Kit BK, Graubard B. Bias in hazard ratios arising from misclassification by self-reported weight and height in observational studies of body mass index and mortality. *Am J Epidemiol* 2017.

24. James WPT, Jackson-Leach R, Ni Mhurchu C, Kalamara E, Shayeghi M, Rigby NJ, et al. 8. Overweight and obesity (high body mass index). In Ezzati M, Lopez AD, Rodgers A, Murray CJL, eds. *Comparative Quantification of Health Risks*, Vol. **1**. Geneva: World Health Organization; 2004. p 497–596.

25. Maukonen M, Mannisto S, Tolonen H. A comparison of measured versus self-reported anthropometrics for assessing obesity in adults: a literature review. *Scand J Public Health* 2018;**46**:1403494818761971.

26. Ni Mhurchu C, Parag V, Nakamura M, Patel A, Rodgers A, Lam TH. Body mass index and risk of diabetes mellitus in the Asia-Pacific region. *Asia Pac J Clin Nutr* 2006; **15**:127–133.

27. Flegal KM. Body-mass index and all-cause mortality. *Lancet* 2017;**389**:2284–2285.

28. Yi SW, Ohrr H, Shin SA, Yi JJ. Sex-age-specific association of body mass index with all-cause mortality among 12.8 million Korean adults: a prospective cohort study. *Int J Epidemiol* 2015;**44**:1696–1705.

29. Flegal KM, Graubard BI, Yi SW. Comparative effects of the restriction method in two large observational studies of body mass index and mortality among adults. *Eur J Clin Invest* 2017;**47**:415–421.

30. Flegal KM, Kit BK, Orpana H, Graubard BI. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *JAMA* 2013; **309**:71–82.

31. Doehner W, Clark A, Anker SD. The obesity paradox: weighing the benefit. *EurHeart J* 2010;**31**:146–148.

32. Doehner W. Critical appraisal of the obesity paradox in cardiovascular disease: how to manage patients with overweight in heart failure? *Heart Fail Rev* 2014;**19**:637–644.

33. Naderi N, Kleine CE, Park C, Hsiung JT, Soohoo M, Tantisattamo E, et al. Obesity paradox in advanced kidney disease: from bedside to the bench. *Prog Cardiovasc Dis* 2018;**61**:168–181.

34. Ahmadi SF, Streja E, Zahmatkesh G, Streja D, Kashyap M, Moradi H, et al. Reverse epidemiology of traditional cardiovascular risk factors in the geriatric population. *J Am Med Dir Assoc* 2015;**16**:933–939.

35. Flegal KM, Ioannidis JPA. The obesity paradox: a misleading term that should be abandoned. *Obesity (Silver Spring)* 2018; **26**:629–630.

36. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJ, et al. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur J Heart Fail* 2016 2016; **18**:891–975.

37. Kleine CE, Moradi H, Streja E, Kalantar-Zadeh K. Racial and ethnic disparities in the obesity paradox. *Am J Kidney Dis* 2018;**72**:S26–s32.

38. Doehner W, von Haehling S, Anker SD. Protective overweight in cardiovascular disease: moving from 'paradox' to 'paradigm'. *Eur Heart J* 2015;**36**:2729–2732.

39. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, et al. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. *Circulation* 2013;**129**: S102–S138.

40. Afshin A, Forouzanfar MH, Reitsma MB, Sur P, Estep K, Lee A, et al. Health effects of overweight and obesity in 195 countries over 25 years. *N Engl J Med* 2017;**377**: 13–27.

41. Solmi M, Kohler CA, Stubbs B, Koyanagi A, Bortolato B, Monaco F, et al. Environmental risk factors and nonpharmacological and nonsurgical interventions for obesity: an umbrella review of meta-analyses of cohort studies and randomized controlled trials. *Eur J Clin Invest* 2018;e12982.

42. von Haehling S, Morley JE, Coats AJS, Anker SD. Ethical guidelines for publishing in the Journal of Cachexia, Sarcopenia and Muscle: update 2017. *J Cachexia Sarcopenia Muscle* 2017;**8**: 1081–1083.