PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	The Relationship of Adiposity and Mortality among People with Diabetes in the U.S. General Population: a Prospective Cohort Study
AUTHORS	Menke, Andy; Casagrande, Sarah; Cowie, Catherine

VERSION 1 - REVIEW

REVIEWER	Mykola (Mykolay) Khalangot, MD PhD Department of Endocrinology P.L. Shupyk National Medical Academy of Postgraduate Education
REVIEW RETURNED	19-May-2014

GENERAL COMMENTS	The main result of this unquestionably interesting and important work, according to its authors, is that in US population with diabetes obesity was NOT associated with risk of mortality. It seems that such conclusion should give a new momentum to the discussion about obesity paradox in type 2 diabetes. Today, on one hand, the well-shaped concept of fighting excessive weight and obesity is overshadowed by some “awkward” information, that “Large studies of people with diabetes from Scotland [8] and Ukraine [9] found a U shaped association with the lowest risk of mortality in the range of 25-35 kg/m2”, whereas on the other hand Swedish researchers (Eeg-Olofsson K et al., 2009) indicate a quite acceptable for healthy lifestyle propaganda relation between obesity and mortality: “Adjusted hazard ratios for total mortality of type 2 diabetic patients with overweight were 1.16 (0.94-1.45) and 1.71 (1.36-2.14) with obesity, as compared with normal weight”. Recently, an analysis of the association between body-mass index (BMI) and the risk of death among participants with incident diabetes from two large prospective cohort Tobias at al, 2014 [10] observed a direct linear relationship between BMI and mortality among those who had never smoked. It would be logical to suppose that if we have two types of opposite results, we can expect to see a third type of results that does not conform to the first two. At first, we can consider such type of results to be a study, presented in this manuscript, as it claims that there is no association between obesity and mortality in type 2 patients. However the conclusions made by the authors do not seem to be substantiated enough. Even after a superficial glance at the main tables and figures we can see that the highest mortality rates

(except for cancer-related) are in < 25 kg/m^2 category. Unadjusted mortality risks confirm this observation for total (≥35 kg/m^2), cardiovascular, and diabetes-related mortality (30-34.9 kg/m^2 and ≥35 kg/m^2 BMI categories) with statistical significance. After adjusting for age, race-ethnicity, sex, smoking status, education, income, and diabetes duration, the statistical significance for these risks disappears, which caused the authors to make a conclusion about absence of association between BMI and mortality in T2D population.

Unfortunately the authors did not present intermediate data on sequential adjusting, therefore the reader does not know which one of the variables is responsible for disappearance of risk significance.

The most interesting results of this work were for some reason hidden by the authors in Supplementary Tables:

1. For T2D persons with HbA1c ≥7.1% multivariable adjusted HRs of All-Cause Mortality for overweight and obesity categories vs <25 kg/m^2 = 0.52 (0.33-0.82) and 0.59 (0.38-0.91) respectively (Table S2).Therefore mortality risk in the largest population of T2D patients (mean HbA1c according to population diabetes register data rarely below 8%) is twice lower for overweight and obesity categories! Very similar results (without categorization according to HbA1c levels) were obtained in Ukraine [9].
2. A significant decrease of mortality risk was noted for overweight vs <25 kg/m^2 categories in taking only oral meds diabetics: multivariable adjusted HRs of All-Cause Mortality 0.48 (0.28-0.83) [see table S3]
3. Another important factor, related to the association between moderate obesity and mortality, according to presented data, is age: for diabetics age ≥65 years, and being in BMI category 30-34.9 kg/m^2 means decrease of mortality risk vs < 25 kg /m2 – HR = 0.49 (0.27-0.88) [see table S6]

Unfortunately, the authors comment these results as follows: “we did not find a significant association between BMI …and mortality in a nationally representative sample of people with diabetes. This finding was consistent when we stratified by … A1c level, diabetes medication usage, and other cardiovascular risk factors”.

Thus, this work contains a lot of new data about association between BMI and mortality in T2D patients that were for some reason not sufficiently commented upon. This led to conclusions that do not quite reflect obtained results. The studies, cited by the authors in support of their conclusions [13-15] are hard to consider as a good confirmation. Furthermore, TRIAD Study (the only one of the three mentioned that is based on sufficient number of observations) reports about unadjusted HR (95% CI) for BMI category < 25 vs BMI >=25 to <30 = 1.45 (1.17–1.81) [14].

Hence, this study contains some quite interesting new data that

	need a more accurate presentation and commentary. Additional reference: Eeg-Olofsson K, Cederholm J, Nilsson PM, Zethelius B, Nunez L, Gudbjörnsdóttir S, Eliasson B. Risk of cardiovascular disease and mortality in overweight and obese patients with type 2 diabetes: an observational study in 13,087 patients. Diabetologia. 2009 Jan;52(1):65-73.

REVIEWER	Deirdre Tobias Harvard School of Public Health, US
REVIEW RETURNED	30-May-2014

GENERAL COMMENTS	This paper presents findings addressing the "obesity paradox" among a nationally representative sample of US NHANES participants with diabetes. The major strength of this paper is the use of waist circumference and technician measured body weight. The major limitation is that half of the participants had prevalent diabetes, which is a major concern for bias. Several stratified analyses were presented, however the statistical power appears too low to draw conclusions in the comparison across strata. Additional major and minor comments are outlined below. Major comments: - A major concern with this analysis is that half of the participants had prevalent diabetes, some with a diagnosis >20 years ago. Body weight measured years after the disease onset can be extremely biased for a variety of reasons, including medications, lifestyle changes, wasting due to diabetes and correlated chronic diseases, etc. This measurement error bias cannot be “adjusted” for by including duration of diabetes in the model. “Depletion of susceptibles” is also a major concern when participants who have had diabetes for several decades are included. There is evidence for this bias in Table 1 “diabetes duration”. - The multivariable models do not adjust for any lifestyle factors other than smoking. Physical activity and diet (e.g., overall healthful diet score) are potential confounders and adjustment for these variables should be considered. - Mortality follow-up is through 2006 which is nearly a decade ago. Is it possible to update this? - It would be helpful to include the number of deaths in your abstract. - Waist circumference was evaluated in quartiles while BMI was evaluated in specific pre-defined categories of overweight/obesity. Since this is a sick diabetic population, the reference group might not actually be a “healthy” or “normal” WC, as was defined for the BMI analysis reference group. Could the authors conduct a secondary analysis with a “healthy” WC reference group (e.g., 94cm in men, 80cm in women)? - Participants with BMI<18.5kg/m2 should not be included in the “normal” BMI range reference group. If there is not enough power to evaluate this group separately, then please exclude them altogether. They are a heterogeneous subgroup and there is significant concern for bias including them in the reference category. - Why would you need to include a sensitivity analyses including participants with CVD, cancer, early deaths and probable type 1

	diabetes? It is well appreciated that including these participants will lead to significant concern for bias. Given the lower statistical power of this study, I would not expect the results to be dramatically different enough to draw meaningful conclusions by comparing results with and without these participants; thus it would be ideal to exclude them altogether from all analyses. I recommend excluding Table S8. - Key sensitivity analyses, such as stratification by smoking status (a major confounder) and stratification by prevalent/incident diabetics are useful, however it would be ideal to see several of these analyses conducted simultaneously. For example, among the "undiagnosed diabetes" strata in Table S1, I am still very concerned that there is major residual confounding by smoking status. Similarly, among the never smokers, I am concerned for the bias due to the prevalent diabetes subjects. Ideally, readers would want to see the "least possible biased" analysis conducted, among those with incident diabetes, excluding smokers, excluding first 2 years of follow-up, no additional chronic diseases at baseline, etc. What are the results among this subgroup? Even if the argument can be made that these findings may not be generalizable to other subjects, these results should still be presented for this minimally biased strata. - I disagree with the authors' conclusion in the discussion that "they cannot rule out a modest U-shaped association between BMI and mortality". This is misleading, given that the confidence intervals are so wide that none of the HR's were significantly below 1.0 to suggest a U-shaped association. Similarly, the spline model shading of 95% CI's never falls below 1.0. The authors should provide a p-value for nonlinearity to support this in their conclusions. The authors could have just as easily concluded that a "direct relationship between BMI and mortality could not be ruled out". The low statistical power leaves these findings open to interpretation and should therefore not be overstated. - The authors conclude that "finding was consistent when we stratified by…". P-values for interaction were not provided by the authors for these stratified analyses. Further, the relatively low statistical power may make it unlikely to detect any modest to moderate differences between strata. Rather than conclude that findings were consistent, it would be more appropriate to state that you were unable to draw conclusions from stratified analyses given the relatively low statistical power to investigate effect modification. As an example, results for current and never smokers have quite different trends across BMI categories, but the wide CIs are unlikely to result in significant heterogeneity. Minor: - Can you provide examples of "diabetes mortality" in the methods? It is not immediately clear what is meant by this. - Table 1 should include the number of participants in the top row. - Can Table 1 please include rows for the reference groups for the variables that have more then 2 categories (e.g., smoking, education, etc)? This would be easier for the reader so they do not have to calculate the difference themselves. - Table titles for all-cause mortality could say "All-Cause and Cause-Specific Mortality" instead of just mortality to be clearer. - Tables 2, 3 and the supplementary Hazard Ratio tables should include the number of events for across category. - Is the age in Table S6 age at diagnosis or age at NHANES assessment? Age at diagnosis would be more interesting since age at NHANES is random. - The crude and multivariable models for BMI are quite different. Out

	of curiosity, which covariable(s) in the model were responsible for the majority of this attenuation? - Typo, page 8 line 20: “tears” instead of “years” - Page 8 line 25 – what is meant by “diagnosis status”? - Age is adjusted for rather crudely, and is often a major confounder in mortality analyses. Did you try adjusting for this more finely, or as a continuous variable? - Table S9 (and the accompanying text) refers to the categories as “low risk” and “extremely high risk”. This terminology may be somewhat misleading and/or presumptuous for the readers who strongly believe that the “lowest risk” is actually among those who are overweight or obese (i.e., believers of the “paradox”). Perhaps you can define these categories in another way, such as “lean” through “extremely obese”? - I thought that participants ages <30 were excluded? Were they included back in only for Table S6? Were results different for the youngest strata when they were excluded, as in the main analysis? - Page 13 of the discussion

REVIEWER	Jennifer Logue University of Glasgow UK
REVIEW RETURNED	02-Jun-2014

GENERAL COMMENTS	This study examines the effect of BMI and waist circumference on mortality in patients with diabetes. It is generally well written and clearly presented and while it has a number of limitation, these are mentioned and the conclusions fair. there are a couple of very small points that should be tidied before publication: 1. Table 1 would benefit from an analysis for trend 2. page 10 ln 41 - I think this should be "table 3"

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name Mykola (Mykolay) Khalangot, MD PhD
Institution and Country Department of Endocrinology
P.L. Shupyk National Medical Academy of Postgraduate Education Vyshgorodska 69 Kiev 04114 Ukraine
Please state any competing interests or state 'None declared': none declared

The main result of this unquestionably interesting and important work, according to its authors, is that in US population with diabetes obesity was NOT associated with risk of mortality. It seems that such conclusion should give a new momentum to the discussion about obesity paradox in type 2 diabetes.

Today, on one hand, the well-shaped concept of fighting excessive weight and obesity is overshadowed by some "awkward" information, that "Large studies of people with diabetes from Scotland [8] and Ukraine [9] found a U shaped association with the lowest risk of mortality in the range of 25-35 kg/m2", whereas on the other hand Swedish researchers (Eeg-Olofsson K et al., 2009) indicate a quite acceptable for healthy lifestyle propaganda relation between obesity and mortality: "Adjusted hazard ratios for total mortality of type 2 diabetic patients with overweight were 1.16 (0.94-1.45) and 1.71 (1.36-2.14) with obesity, as compared with normal weight".

Recently, an analysis of the association between body-mass index (BMI) and the risk of death among participants with incident diabetes from two large prospective cohort Tobias at al, 2014 [10] observed a direct linear relationship between BMI and mortality among those who had never smoked.

It would be logical to suppose that if we have two types of opposite results, we can expect to see a third type of results that does not conform to the first two. At first, we can consider such type of results to be a study, presented in this manuscript, as it claims that there is no association between obesity and mortality in type 2 patients.

However the conclusions made by the authors do not seem to be substantiated enough. Even after a superficial glance at the main tables and figures we can see that the highest mortality rates (except for cancer-related) are in < 25 kg/m2 category. Unadjusted mortality risks confirm this observation for total (≥35 kg/m2), cardiovascular, and diabetes-related mortality (30-34.9 kg/m2 and ≥35 kg/m2 BMI categories) with statistical significance. After adjusting for age, race-ethnicity, sex, smoking status, education, income, and diabetes duration, the statistical significance for these risks disappears, which caused the authors to make a conclusion about absence of association between BMI and mortality in T2D population.

Unfortunately the authors did not present intermediate data on sequential adjusting, therefore the reader does not know which one of the variables is responsible for disappearance of risk significance.

We added an intermediate model adjusting for age, race-ethnicity, and sex. Adjusting for age had the biggest impact of any variable in the model. BMI and waist circumference were generally not associated with all-cause mortality after adjustment for age.

The most interesting results of this work were for some reason hidden by the authors in Supplementary Tables:

1. For T2D persons with HbA1c ≥7.1% multivariable adjusted HRs of All-Cause Mortality for overweight and obesity categories vs <25 kg/m2 = 0.52 (0.33-0.82) and 0.59 (0.38-0.91) respectively (Table S2).Therefore mortality risk in the largest population of T2D patients (mean HbA1c according to population diabetes register data rarely below 8%) is twice lower for overweight and obesity

categories! Very similar results (without categorization according to HbA1c levels) were obtained in Ukraine [9].
2. A significant decrease of mortality risk was noted for overweight vs <25 kg/m2 categories in taking only oral meds diabetics: multivariable adjusted HRs of All-Cause Mortality 0.48 (0.28-0.83) [see table S3]
3. Another important factor, related to the association between moderate obesity and mortality, according to presented data, is age: for diabetics age ≥65 years, and being in BMI category 30-34.9 kg/m2 means decrease of mortality risk vs < 25 kg /m2 – HR = 0.49 (0.27-0.88) [see table S6]

Unfortunately, the authors comment these results as follows: "we did not find a significant association between BMI …and mortality in a nationally representative sample of people with diabetes. This finding was consistent when we stratified by … A1c level, diabetes medication usage, and other cardiovascular risk factors".

We thank the reviewer for carefully examining our Supplemental Tables. In response to reviewer 2 comments, we made several changes to the multivariable analysis including adding exercise and dietary variables to the model. As a result, the first two results mentioned above (stratified by A1c level and medication usage) are no longer statistically significant.

For the third result mentioned above (stratified by age), as reviewer 2 mentioned, some of the sensitivity analyses were based on small sample sizes and we don't want to overstate the results. Therefore, we did not add further emphasis of the analysis stratified by age. We removed the statement in the conclusion stating the results were consistent when stratified by smoking etc. Interested readers can view the results in the supplemental tables and interpret as they deem appropriate.

Thus, this work contains a lot of new data about association between BMI and mortality in T2D patients that were for some reason not sufficiently commented upon. This led to conclusions that do not quite reflect obtained results. The studies, cited by the authors in support of their conclusions [13-15] are hard to consider as a good confirmation. Furthermore, TRIAD Study (the only one of the three mentioned that is based on sufficient number of observations) reports about unadjusted HR (95% CI) for BMI category < 25 vs BMI >=25 to <30 = 1.45 (1.17–1.81) [14].

Hence, this study contains some quite interesting new data that need a more accurate presentation and commentary.

Additional reference:

Eeg-Olofsson K, Cederholm J, Nilsson PM, Zethelius B, Nunez L, Gudbjörnsdóttir S, Eliasson B. Risk of cardiovascular disease and mortality in overweight and obese patients with type 2 diabetes: an observational study in 13,087 patients. Diabetologia. 2009 Jan;52(1):65-73.

We thank the reviewer for pointing out this reference. We added a sentence describing it in the discussion on page 12.

"In a study of 13,087 people in the Swedish National Diabetes Registry, BMI categories were positively related to risk of mortality."

Reviewer: 2
Reviewer Name Deirdre Tobias
Institution and Country Harvard School of Public Health, US

Please state any competing interests or state ‘None declared’: None declared

This paper presents findings addressing the "obesity paradox" among a nationally representative sample of US NHANES participants with diabetes. The major strength of this paper is the use of waist circumference and technician measured body weight. The major limitation is that half of the participants had prevalent diabetes, which is a major concern for bias. Several stratified analyses were presented, however the statistical power appears too low to draw conclusions in the comparison across strata. Additional major and minor comments are outlined below.

Major comments:
- A major concern with this analysis is that half of the participants had prevalent diabetes, some with a diagnosis >20 years ago. Body weight measured years after the disease onset can be extremely biased for a variety of reasons, including medications, lifestyle changes, wasting due to diabetes and correlated chronic diseases, etc. This measurement error bias cannot be “adjusted” for by including duration of diabetes in the model. “Depletion of susceptibles” is also a major concern when participants who have had diabetes for several decades are included. There is evidence for this bias in Table 1 “diabetes duration”.

We believe there is value in understanding the adiposity-mortality relationship among people with incident diabetes and among all people with diabetes (i.e., incident or prevalent cases). Although our main analysis includes all cases, we show the results stratified by diagnosis status in Supplement Table 1. The results only among those with undiagnosed diabetes were consistent with the results among all people with diabetes. We discussed the potential bias with using prevalent cases of diabetes on page 13:

“In studies of adiposity and mortality among people with diabetes, the timing of when BMI or waist circumference is measured relative to diagnosis of diabetes may be important. In our study, about half of our participants were previously diagnosed with diabetes. Weight gain or loss may have occurred during the course of diabetes as a result of changes in lifestyle, medication use, or diabetes disease progression.”

- The multivariable models do not adjust for any lifestyle factors other than smoking. Physical activity and diet (e.g., overall healthful diet score) are potential confounders and adjustment for these variables should be considered.

We added physical activity and diet variables (calories consumed and percent of calories from saturated fat based on a 24-hour recall) to multivariable adjusted models. Most results were similar to models without physical activity and diet. One exception is that waist circumference was associated with mortality among men when modeled as a spline (shown in Figure 2).

- Mortality follow-up is through 2006 which is nearly a decade ago. Is it possible to update this?

NHANES has not released more recent follow-up data, so we cannot update the analysis at this time.

- It would be helpful to include the number of deaths in your abstract.

We added the number of deaths to the abstract:

“(n=599 deaths)”

- Waist circumference was evaluated in quartiles while BMI was evaluated in specific pre-defined categories of overweight/obesity. Since this is a sick diabetic population, the reference group might

not actually be a “healthy” or “normal” WC, as was defined for the BMI analysis reference group. Could the authors conduct a secondary analysis with a “healthy” WC reference group (e.g., 94cm in men, 80cm in women)?

We prepared the below table using the cutpoints 94 cm for men and 80 cm for women for the first category, and tertiles above those cutpoints for categories 2-4. Only ~10% of participants fell into the reference category, which limited power to detect association, particularly for cause-specific mortality. Similar to the main analysis, there were no significant associations in the table (see below).

Hazards Ratio (95% Confidence Interval) of Mortality Associated With Sex-Specific* Categorization of Waist Circumference
Category 1 Category 2 Category 3 Category 4
All-cause mortality
Mortality rate (SE)† 30 (6.4) 32 (5.2) 30 (3.8) 29 (3.5)
Hazard ratios
Unadjusted 1.00 1.07 (0.64-1.81) 1.01 (0.60-1.70) 0.97 (0.59-1.58)
Multivariable adjusted 1‡ 1.00 1.18 (0.65-2.12) 0.96 (0.50-1.82) 1.33 (0.72-2.46)
Multivariable adjusted 2§ 1.00 1.29 (0.77-2.15) 1.02 (0.55-1.89) 1.37 (0.77-2.43)
Cardiovascular mortality
Mortality rate (SE)† 6 (2.2) 13 (2.3) 11 (2.3) 8 (1.7)
Hazard ratios
Unadjusted 1.00 2.31 (0.99-5.41) 1.91 (0.80-4.58) 1.41 (0.55-3.60)
Multivariable adjusted 1‡ 1.00 2.37 (0.96-5.88) 1.68 (0.72-3.92) 2.07 (0.78-5.52)
Multivariable adjusted 2§ 1.00 2.17 (0.84-5.59) 1.61 (0.69-3.76) 1.94 (0.71-5.30)
Cancer mortality
Mortality rate (SE)† 6 (2.6) 10 (3.4) 8 (1.9) 7 (2.0)
Hazard ratios
Unadjusted 1.00 1.62 (0.53-4.88) 1.23 (0.49-3.12) 1.06 (0.38-2.95)
Multivariable adjusted 1‡ 1.00 1.84 (0.55-6.21) 1.22 (0.41-3.66) 1.46 (0.49-4.32)
Multivariable adjusted 2§ 1.00 2.01 (0.60-6.71) 1.19 (0.38-3.73) 1.53 (0.45-5.12)
Diabetes mortality
Mortality rate (SE)† 9 (3.3) 12 (2.5) 13 (2.2) 6 (1.1)
Hazard ratios
Unadjusted 1.00 1.33 (0.56-3.14) 1.35 (0.61-2.98) 0.67 (0.29-1.56)
Multivariable adjusted 1‡ 1.00 1.31 (0.50-3.40) 1.14 (0.45-2.87) 0.78 (0.30-2.05)
Multivariable adjusted 2§ 1.00 1.65 (0.53-5.11) 1.43 (0.47-4.37) 0.88 (0.31-2.53)
SE, standard error
*For men: <94.0 cm, 94.0-103.7 cm, 103.8-112.7 cm, and ≥112.8 cm; for women: <80.0 cm, 80.0-96.2 cm, 96.3-108.4 cm, and ≥108.5 cm
†Per 1000 person-years
‡Adjusted for age, race-ethnicity, sex
§Adjusted for age, race-ethnicity, sex, smoking status, education, income, diabetes duration, exercise, diet

- Participants with BMI<18.5kg/m2 should not be included in the “normal” BMI range reference group. If there is not enough power to evaluate this group separately, then please exclude them altogether. They are a heterogeneous subgroup and there is significant concern for bias including them in the reference category.

We did not have enough power to investigate the <18.5 kg/m2 group separately (n=13), so we excluded them from the analysis.

- Why would you need to include a sensitivity analyses including participants with CVD, cancer, early deaths and probable type 1 diabetes? It is well appreciated that including these participants will lead to significant concern for bias. Given the lower statistical power of this study, I would not expect the results to be dramatically different enough to draw meaningful conclusions by comparing results with and without these participants; thus it would be ideal to exclude them altogether from all analyses. I recommend excluding Table S8.

We think this table has value in illustrating the effect our exclusions had on the analysis, and we believe some readers may be interested in seeing these results. However, we would be willing to remove this table if the editor prefers we do not show it.

- Key sensitivity analyses, such as stratification by smoking status (a major confounder) and stratification by prevalent/incident diabetics are useful, however it would be ideal to see several of these analyses conducted simultaneously. For example, among the "undiagnosed diabetes" strata in Table S1, I am still very concerned that there is major residual confounding by smoking status. Similarly, among the never smokers, I am concerned for the bias due to the prevalent diabetes subjects. Ideally, readers would want to see the "least possible biased" analysis conducted, among those with incident diabetes, excluding smokers, excluding first 2 years of follow-up, no additional chronic diseases at baseline, etc. What are the results among this subgroup? Even if the argument can be made that these findings may not be generalizable to other subjects, these results should still be presented for this minimally biased strata.

We agree this would be an interesting analysis, but power is very limited for never smokers with undiagnosed diabetes in our study (N=230; n=64 deaths). The multivariable adjusted HR (95% CI) compared to a BMI 18.5-24.9 were 1.29 (0.32-5.28) for 25-29.9, 1.34 (0.30-5.96) for 30-34.9, and 1.92 (0.57-6.43) for ≥35 kg/m2. We chose not to present these results due to the limited sample size.

- I disagree with the authors' conclusion in the discussion that "they cannot rule out a modest U-shaped association between BMI and mortality". This is misleading, given that the confidence intervals are so wide that none of the HR's were significantly below 1.0 to suggest a U-shaped association. Similarly, the spline model shading of 95% CI's never falls below 1.0. The authors should provide a p-value for nonlinearity to support this in their conclusions. The authors could have just as easily concluded that a "direct relationship between BMI and mortality could not be ruled out". The low statistical power leaves these findings open to interpretation and should therefore not be overstated.

We removed the statement saying we "cannot rule out a modest U-shaped association" and reference to a non-significant association in spline figures in the discussion.

- The authors conclude that "finding was consistent when we stratified by…". P-values for interaction were not provided by the authors for these stratified analyses. Further, the relatively low statistical power may make it unlikely to detect any modest to moderate differences between strata. Rather than conclude that findings were consistent, it would be more appropriate to state that you were unable to draw conclusions from stratified analyses given the relatively low statistical power to investigate effect modification. As an example, results for current and never smokers have quite different trends across BMI categories, but the wide CIs are unlikely to result in significant heterogeneity.

We removed the statement in the conclusion stating the results were consistent when stratified by smoking etc. Those findings are addressed earlier in the manuscript and readers can view the results in the supplemental tables and interpret as they deem appropriate.

Minor:
- Can you provide examples of "diabetes mortality" in the methods? It is not immediately clear what is

meant by this.

For diabetes mortality, we included those with diabetes listed anywhere on the death certificate. To clarify this, we revised text in the methods section on page 7.

“We were also interested in deaths in which diabetes played any role leading to death; to investigate this, diabetes mortality (Ninth Revision codes 250; Tenth Revision codes E10-E14) was defined using any cause of death listed on the death certificates (among the underlying cause and up to 20 contributing causes).”

- Table 1 should include the number of participants in the top row.

We added the number of participants in the top row as suggested.

- Can Table 1 please include rows for the reference groups for the variables that have more then 2 categories (e.g., smoking, education, etc)? This would be easier for the reader so they do not have to calculate the difference themselves.

We added those rows in Table 1.

- Table titles for all-cause mortality could say “All-Cause and Cause-Specific Mortality” instead of just mortality to be clearer.

We added “All-Cause and Cause-Specific” to the table titles as the reviewer suggested.

- Tables 2, 3 and the supplementary Hazard Ratio tables should include the number of events for across category.

We added number of deaths to the tables.

- Is the age in Table S6 age at diagnosis or age at NHANES assessment? Age at diagnosis would be more interesting since age at NHANES is random.

Table S6 stratifies by age at baseline. We revised the title to make it clearer:

“Hazards Ratio (95% Confidence Interval) of All-Cause Mortality by Age at Study Baseline”

- The crude and multivariable models for BMI are quite different. Out of curiosity, which covariable(s) in the model were responsible for the majority of this attenuation?

Adjusting for age had the largest effect on hazard ratios. In response to this comment and a comment from reviewer 1, we added the results of an intermediate model adjusting for age, race-ethnicity, and sex to the tables.

- Typo, page 8 line 20: “tears” instead of “years”

Thank you for catching this typo. We changed it to “years”.

- Page 8 line 25 – what is meant by “diagnosis status”?

We were referring to whether the participant had a previous diagnosis. We changed the text to “diagnosed versus undiagnosed diabetes” for better clarity.

- Age is adjusted for rather crudely, and is often a major confounder in mortality analyses. Did you try adjusting for this more finely, or as a continuous variable?

We now adjust for age as a continuous variable in all models. Results were generally similar to equivalent models including age as a categorical variable.

- Table S9 (and the accompanying text) refers to the categories as “low risk” and “extremely high risk”. This terminology may be somewhat misleading and/or presumptuous for the readers who strongly believe that the “lowest risk” is actually among those who are overweight or obese (i.e., believers of the “paradox”). Perhaps you can define these categories in another way, such as “lean” through “extremely obese”?

These are previously defined categories based on an NHLBI report entitled “Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults” (see reference 19). This report was prepared by an expert panel to summarize evidence regarding the effects of treatment on overweight and obesity. We prefer to use terminology consistent with the NHLBI publication.

- I thought that participants ages <30 were excluded? Were they included back in only for Table S6? Were results different for the youngest strata when they were excluded, as in the main analysis?

We excluded participants <20 years of age from all analyses. Participants 20-29 were included in all analyses (except age subgroups in Table S6).

Reviewer: 3
Reviewer Name Jennifer Logue
Institution and Country University of Glasgow
UK
Please state any competing interests or state ‘None declared’: None declared

This study examines the effect of BMI and waist circumference on mortality in patients with diabetes. It is generally well written and clearly presented and while it has a number of limitation, these are mentioned and the conclusions fair.

there are a couple of very small points that should be tidied before publication:

1. Table 1 would benefit from an analysis for trend

We added p-trends to Table 1 as suggested.

2. page 10 ln 41 - I think this should be "table 3"

We thank the reviewer for catching this error. We changed the text to “Table 3”.

VERSION 2 – REVIEW

REVIEWER	Mykolay Khalangot P.L. Shupyk National Medical Academy of Postgraduate Education Vyshgorodska 69 Kiev, Ukraine
REVIEW RETURNED	13-Jul-2014

GENERAL COMMENTS	In their commentary, which came before the new version of the manuscript, the authors declared their intention to keep balance in considering variable comments of the reviewers. Unfortunately the outcome of their work is far from being balanced. Thus, after a somewhat strange expression of gratitude to us for a detailed examination of the tables from the appendix, the authors do not provide any explanation of the fact of ignoring the results given in them. Instead, the authors only inform that after doing additional adjusting for two new variables -- physical activity and diet, the differences in mortality risks related to BMI disappear. Cardiorespiratory fitness significantly alters the obesity paradox (McAuley et al, 2012, McAuley, 2014), however this does not likely give reason to doubt the existence of the phenomenon itself. Besides, we should note, that it is cardiorespiratory fitness, that affects the obesity paradox, and it was quantified in metabolic equivalents using objective treadmill exercise test (McAuley et al, 2012). The authors had used self-reported data about “physical activity”. To prove the reasonability of including these data, when creating multivariate regression model, we must consider risk value and corresponding level of significance (p), related namely to physical activity. The authors for sure have these data. If the corresponding risk is far from the statistical significance (e.g. p near 0.5), it becomes obvious, that adding exercise variable to the model is useless, and even harmful. The same can be said for diet. When there is a relatively small amount of observations, the infinitely increasing number of variables can ruin any evaluation model, which we think was successfully demonstrated in this case. In the first version of the manuscript the authors presented what seems to be new data indicating an association between the level of hyperglycemia, type of treatment, and association between BMI and mortality, however these results did not receive an adequate commentary. Furthermore the results section of the revised version is missing the following statement: "among participants with an A1c ≥7.1%, those with a BMI 25-29.9 kg/m2 and 30-34.9 kg/m2 had a significantly lower risk of mortality than those with a BMI <25 kg/m2". As a result, the reason of dissimilarity of conclusions, some of which confirm the obesity paradox in diabetes patients (Jackson et al., 2013), while others do not (Tobias at al, 2014), remains unclear. Refering to data of Swedish researchers (Eeg-Olofsson K et al., 2009) by the authors does not resolve the situation, as it is unclear why according to Swedish data, mortality increases together with BMI, and the authors of this manuscript have not revealed that.

	Such distrust to their own data about possible influence of blood glucose level on association between BMI and mortality surprises us to say the least, however maybe the authors have serious reasons for this. If this distrust is connected with a relatively small amount of observations and a wide range of CIs of corresponding risks, in the case of sufficient significance ($p < 0.05$), there is no reason to doubt the fact of such an association. We may only think about possible insufficient precision of its quantitative evaluation. Thus, we still believe the authors may have obtained interesting results that need a more thorough analysis. Reference Paul A. McAuley, Nancy S. Smith, Brian T. Emerson, and Jonathan N. Myers, "The Obesity Paradox and Cardiorespiratory Fitness," Journal of Obesity, vol. 2012

REVIEWER	Jennifer Logue University of Glasgow, Scotland.
REVIEW RETURNED	14-Jul-2014

GENERAL COMMENTS	I felt this paper was sufficient at the first draft stage and my minor comments have all been addressed. The main issue is lack of power and differing duration of diabetes but the discussion does cover these issues.

REVIEWER	Deirdre Tobias Harvard School of Public Health
REVIEW RETURNED	17-Jul-2014

GENERAL COMMENTS	Thank you to the authors for taking the time to consider the reviewers' comments and incorporate additional changes to the manuscript. The manuscript is significantly improved and will be of value to this body of literature. Additional major concerns still exist. Major comments: 1. It is surprising that the authors do not mention a similar previous analysis published earlier this year also using the NHANES dataset to evaluate BMI and mortality among participants with diabetes (Preston, Stokes; Epidemiology 2014). After careful review of this previous publication, I have some major questions for the authors of this manuscript: - This submitted manuscript includes participants from NHANES III (1988-1994). The analysis by Preston, et al includes NHANES III (1988-1994), plus NHANES waves 1999-2004 and have almost double the number of deaths. Could the authors of this paper include these participants as well? This may significantly improve their statistical power, which is of concern. - The analysis by Preston, et al, highlight the major concern for smoking status as a source for confounding and bias in the same population that you are analyzing. Rather than going into too much detail here, I suggest you read this paper carefully, if you have not done so already. Therefore, the implications of the Preston paper with your findings should be carefully addressed in your discussion

	section, since this paper that is already published on the same dataset essentially highlights why your findings could potentially be biased. It is always of concern when different analyses of the same dataset draw such different conclusions. - It would be ideal to include smoking status in the models as Preston et al defined it, with 7 categories, since this will attempt to better control for smoking status: never, former <1 pack/d, former 1-<2 packs/d, former 2+ packs/d, current <1 pack/d, current 1-<2 packs/d, current 2+ packs/d. - If the authors collapsed BMI exposure groups into just 2 categories (normal vs. overweight/obese), as the Preston paper did, and stratify by smoking status as ever/never (See their Table 4)…are there similar results to theirs, such that BMI>25 is associated with significantly higher mortality among smokers and significantly lower mortality among never smokers (i.e. is there an interaction by smoking status that was also seen in our Tobias et al NEJM 2014 paper)? Your supplemental Table S7 suggests the trend is consistent with their observation, but CI's are so wide. Attempting to replicate their analysis would be interesting since it is the same dataset – it would be concerning if results were vastly different. 2. For Table 2 with the main results and the additional table on WC in the comments to reviewers, it is surprising to see how much "diabetes mortality" influenced the overall all-cause mortality HR. In Table 1, all-cause mortality HR=1.01 for >35 vs. normal weight, while CVD=1.08 and cancer=1.13. In the WC results, HR=1.37 for category 4 vs 1 for all-cause mortality, while HR=1.94 for CVD mortality and HR=1.53 for cancer mortality. Why was the diabetes mortality HR=0.88 so influential here? As per your definition of diabetes mortality, these causes of death were not mutually exclusive from CVD or cancer, correct? Also, what are the other causes of mortality included in "all cause" and what are the HRs for this hodge-podge endpoint? Out of curiosity, what did the authors make of this finding? Or is the power too low to read into it? 3. The findings that the authors presented in their response stratified both by smoking status and diabetes diagnosis status are very interesting. Although power is limited, as they suggest, there is a suggested linear trend of increased mortality across BMI categories among never smokers, which is actually consistent with what was found in the Tobias et al NHS/HPFS NEJM paper. The authors state they chose not to present these results because of the limited sample size (64 deaths), but they did choose present results among other strata with even fewer deaths (e.g., only 29 deaths in the lowest age category). Please include a supplementary table stratified by both smoking and diabetes diagnosis status simultaneously, despite the limited power. This analysis will be very interesting to many readers. 4. The authors describe in their response that they included participants 20+ years old, yet their exclusion criteria explicitly states that they excluded participants with likely type 1 diabetes (diagnosis <30 years of age). So how can there be participants with diabetes younger than age 30, wouldn't these be excluded b/c of their likely type 1 diabetes status? These younger participants should be excluded from the analysis, because as the methods indicate, it is likely to be type 1 diabetes. 5. The authors conclude that "further research should investigate possible reasons why the adiposity-mortality association may differ

	in people with diabetes compared to the general population." This conclusion is inappropriate for this paper. Comparing BMI-mortality associations for diabetics vs. general population was not the aim of this analysis and this results for this comparison were not done. 6. Given the limitations that are presented in the discussion section, do the authors truly believe that there is no association between adiposity and mortality? Is it possible that the validity to these findings are overwhelmed by bias, or are the authors comfortable with how they addressed their limitations? The reader should be left with some impression of how relevant these limitations are to interpretation of the results as a whole, even if it does not impact the overall conclusions. 7. Supplemental Table 1 by diagnosis status (pre-existing vs. newly diagnosed) is very helpful. I agree with the authors' response that there is value in understanding the relationship between adiposity and mortality among all diabetics, not just incident. I do want to make one point that the authors do not necessarily need to respond to, but it is a recurring theme in the discussion of the obesity paradox. Please remember that internal validity is a cornerstone of epidemiologic research. Stratifying one's dataset or using strict exclusion criteria certainly can reduce external validity (generalizability), but internal validity (i.e., minimal confounding, bias, etc), is foremost (purely crude descriptive statistics and prediction models aside). As I'm sure you agree, discussing the generalizability of findings is moot if there is concern that they are not even valid to begin with. Think of a randomized clinical trial. RCTs enroll the most compliant, most eager, most willing participants. Exclusion criteria often leave researchers with only a fraction of participants they approached for the study. They know from the beginning that the generalizability of their findings to the general population will be highly questionable, but they sacrifice this limitation for the sake of increasing their internal validity and ensuring the least biased result. Similarly, it is ok to stratify observational data to explore the potential for bias and to explore potential threats to internal validity (e.g., smoking, diabetes duration). Including everyone doesn't mean you have achieved generalizability if the results themselves are biased. Minor comments: 1. Supplemental Table S8: Could you please include a row for # of deaths? How many additional deaths are included when all of the excluded participants are allowed back in? The authors point out that results were similar for this analysis compared with the main findings, but if few additional deaths were actually included, then it would not be surprising that results did not change much. 2. How is diabetes duration modeled for those with new onset diabetes? Were they lumped in the <5 years category? I suggest making newly diagnosed in their own category. 3. Table 1: Please include the following variables o 1) % with dm diagnosed at NHANES visit (undiagnosed diabetes) , o 2) % with prevalent diabetes, o among the prevalent, provide 3) % with dm dx <5 years ago, 5-9, 10-14, 15-19, 20+ years, o 4) mean A1c concentrations, o 5) % in tertiles of A1c as presented in Supp table S2: <5.9%, 5.9-7.0%, 7.1%+ o 6) % not taking diabetes medications

	o 7) % taking any oral medications o 8) % taking insulin 4. In your statistical methods, tables, and elsewhere, please change the definition of your lowest BMI category from <25 to 18.5-25.0. 5. What is the correlation between BMI and WC in your sample for men and women? (i.e., how unique is WC from BMI in assessing adiposity?) This might be useful to mention in the results or discussion. What are the results if the BMI model is also adjusted for WC? And similarly, what are the results if the WC model is also adjusted for BMI? 6. An additional limitation in the discussion should be that diabetics were enrolled from 1988-1994, and diagnostic criteria and treatment regimens have changed since then. 7. Table S8 is fine to include in the supplement if the authors feel it would interest other readers.

VERSION 2 – AUTHOR RESPONSE

Reviewer Name Mykolay Khalangot
Institution and Country P.L. Shupyk National Medical Academy of Postgraduate Education Vyshgorodska 69 Kiev, Ukraine
Please state any competing interests or state 'None declared': None declared

In their commentary, which came before the new version of the manuscript, the authors declared their intention to keep balance in considering variable comments of the reviewers. Unfortunately the outcome of their work is far from being balanced.

Thus, after a somewhat strange expression of gratitude to us for a detailed examination of the tables from the appendix, the authors do not provide any explanation of the fact of ignoring the results given in them. Instead, the authors only inform that after doing additional adjusting for two new variables -- physical activity and diet, the differences in mortality risks related to BMI disappear.

Cardiorespiratory fitness significantly alters the obesity paradox (McAuley et al, 2012, McAuley, 2014), however this does not likely give reason to doubt the existence of the phenomenon itself. Besides, we should note, that it is cardiorespiratory fitness, that affects the obesity paradox, and it was quantified in metabolic equivalents using objective treadmill exercise test (McAuley et al, 2012).

The authors had used self-reported data about "physical activity". To prove the reasonability of including these data, when creating multivariate regression model, we must consider risk value and corresponding level of significance (p), related namely to physical activity. The authors for sure have these data. If the corresponding risk is far from the statistical significance (e.g. p near 0.5), it becomes obvious, that adding exercise variable to the model is useless, and even harmful. The same can be said for diet. When there is a relatively small amount of observations, the infinitely increasing number of variables can ruin any evaluation model, which we think was successfully demonstrated in this case.

We agree that taking into account the relationship between the diet/exercise and mortality, and being cautious not to include too many variables in the model are important factors to consider. We also looked at how adjusting for diet and exercise affected our analysis and our main results do not qualitatively change with or without diet and exercise variables in our models, suggesting they were not important confounders. Therefore, to alleviate the reviewers concerns regarding too many variables in our models given our sample size, we have removed physical activity and diet from the analysis.

In the first version of the manuscript the authors presented what seems to be new data indicating an association between the level of hyperglycemia, type of treatment, and association between BMI and mortality, however these results did not receive an adequate commentary. Furthermore the results section of the revised version is missing the following statement: "among participants with an A1c ≥7.1%, those with a BMI 25-29.9 kg/m2 and 30-34.9 kg/m2 had a significantly lower risk of mortality than those with a BMI <25 kg/m2".

As a result, the reason of dissimilarity of conclusions, some of which confirm the obesity paradox in diabetes patients (Jackson et al., 2013), while others do not (Tobias at al, 2014), remains unclear.

Although the main results of our analysis did not find a significant association between adiposity and mortality, we found a few significant associations when stratifying the results by sex, age, and A1c level in the supplement. We discuss these findings in the results section of our manuscript. On page 10 in the results, we added the sentences:

“Among participants in the higher tertile of A1c (≥7.1%), those with a BMI 25-29.9 kg/m2 had a significantly lower risk of mortality than those with a BMI 18.5-24.9 kg/m2. Although there were very few deaths among participants 20-44 years of age, those with a BMI ≥25 kg/m2 had a significantly lower risk of mortality than with a BMI 18.5-24.9 kg/m2.”

Refering to data of Swedish researchers (Eeg-Olofsson K et al., 2009) by the authors does not resolve the situation, as it is unclear why according to Swedish data, mortality increases together with BMI, and the authors of this manuscript have not revealed that.

It remains unclear why previously published studies are inconsistent with each other and our study. In the discussion section, we address this issue on page 13:

“These studies vary in terms of location, population, BMI categorization, BMI assessment, and timing of BMI measurement relative to diabetes diagnosis, but it is unclear why they are inconsistent.”

Such distrust to their own data about possible influence of blood glucose level on association between BMI and mortality surprises us to say the least, however maybe the authors have serious reasons for this.
If this distrust is connected with a relatively small amount of observations and a wide range of CIs of corresponding risks, in the case of sufficient significance ($p < 0.05$), there is no reason to doubt the fact of such an association. We may only think about possible insufficient precision of its quantitative evaluation.

Although the main results of our analysis did not find a significant association between adiposity and mortality, we found a few significant associations when stratifying the results by sex, age, and A1c level in the supplement. We discuss these findings in the results and discussion sections of our manuscript. For example, on page 12 of the discussion, we added the statement:

“In sensitivity analyses in which we stratify the analysis by important characteristics, there were significant results suggesting a U-shaped association between BMI and mortality among men and among people in the highest tertile of A1c (≥7.1%); among people 20-44 years of age, those with a healthy weight based on BMI had a higher risk of mortality than those with higher BMI.”

Thus, we still believe the authors may have obtained interesting results that need a more thorough analysis.

Reference
Paul A. McAuley, Nancy S. Smith, Brian T. Emerson, and Jonathan N. Myers, “The Obesity Paradox and Cardiorespiratory Fitness,” Journal of Obesity, vol. 2012

Reviewer Name Jennifer Logue
Institution and Country University of Glasgow, Scotland.
Please state any competing interests or state ‘None declared’: None declared

I felt this paper was sufficient at the first draft stage and my minor comments have all been addressed. The main issue is lack of power and differing duration of diabetes but the discussion does cover these issues.

Reviewer Name Deirdre Tobias
Institution and Country Harvard School of Public Health
Please state any competing interests or state ‘None declared’: None declared

Thank you to the authors for taking the time to consider the reviewers’ comments and incorporate additional changes to the manuscript. The manuscript is significantly improved and will be of value to this body of literature. Additional major concerns still exist.

Major comments:
1. It is surprising that the authors do not mention a similar previous analysis published earlier this year also using the NHANES dataset to evaluate BMI and mortality among participants with diabetes (Preston, Stokes; Epidemiology 2014). After careful review of this previous publication, I have some major questions for the authors of this manuscript:
- This submitted manuscript includes participants from NHANES III (1988-1994). The analysis by Preston, et al includes NHANES III (1988-1994), plus NHANES waves 1999-2004 and have almost double the number of deaths. Could the authors of this paper include these participants as well? This may significantly improve their statistical power, which is of concern.

We added the 1999-2004 NHANES data with follow-up through 2006. We now have 668 deaths. Please note that the number of deaths in their manuscript was different largely due to a different approach to exclusions aimed at reducing reverse causation. We excluded deaths in the first 2 years of follow-up and those with previously diagnosed CVD and cancer upfront before any analyses. The Preston manuscript had similar exclusions only in their final table.

- The analysis by Preston, et al, highlight the major concern for smoking status as a source for confounding and bias in the same population that you are analyzing. Rather than going into too much detail here, I suggest you read this paper carefully, if you have not done so already. Therefore, the implications of the Preston paper with your findings should be carefully addressed in your discussion section, since this paper that is already published on the same dataset essentially highlights why your findings could potentially be biased. It is always of concern when different analyses of the same dataset draw such different conclusions.

The paper by Preston was published while our paper was previously under review at the journal. We took a different approach to investigating this topic than Preston et al, but our results were similar to their results when the exclusions were similar (see Table 4 of the Preston manuscript). Both studies found no significant association between BMI and mortality among the overall population or by smoking strata. We added discussion of this article to our discussion section on page 13:

"The authors of a previous study suggested that collider bias may play a role in the inverse BMI-mortality association found in some previous studies of people with diabetes; they used NHANES data to investigate the issue finding an inverse association between BMI and mortality only among smokers, an association that was attenuated and not significant after introducing restrictions aimed at reducing the intensity of reverse causal pathways. Our study similarly excluded participants to limit the impact of reverse causal pathways and we found no association between multiple measures of adiposity and mortality. The non-significant positive association among never smokers suggests the possibility that there is some residual bias due to reverse causation in our analysis of the overall population and completely eliminating the bias may result in a positive association between adiposity and mortality."

- It would be ideal to include smoking status in the models as Preston et al defined it, with 7 categories, since this will attempt to better control for smoking status: never, former <1 pack/d, former 1-<2 packs/d, former 2+ packs/d, current <1 pack/d, current 1-<2 packs/d, current 2+ packs/d.

We changed our smoking categorization to match what was done in the Preston manuscript. The results were similar to results using our previous smoking categorization.

- If the authors collapsed BMI exposure groups into just 2 categories (normal vs. overweight/obese), as the Preston paper did, and stratify by smoking status as ever/never (See their Table 4)…are there similar results to theirs, such that BMI>25 is associated with significantly higher mortality among smokers and significantly lower mortality among never smokers (i.e. is there an interaction by smoking status that was also seen in our Tobias et al NEJM 2014 paper)? Your supplemental Table S7 suggests the trend is consistent with their observation, but CI's are so wide. Attempting to replicate their analysis would be interesting since it is the same dataset – it would be concerning if results were vastly different.

Our results are comparable to the "baseline condition" row in Table 4 of the Preston manuscript, which was the row that used exclusion criteria most similar to our analysis. Although we use a different definition of diabetes (they included prediabetes), age range (they restricted it to ages 35-74), covariates (we include income and diabetes duration), and exclusion criteria (we excluded a history of CVD and only the first 2 years of follow-up), we found consistent results when we model BMI the way Preston did. Never smokers with a healthy weight (BMI 18.5-25 kg/m2) had a non-significantly lower risk of mortality (HR: 0.85 [0.51-1.40]) and ever smokers with a healthy weight had a non-significantly higher risk of mortality (HR: 1.32 [0.91-1.90]).

2. For Table 2 with the main results and the additional table on WC in the comments to reviewers, it is surprising to see how much "diabetes mortality" influenced the overall all-cause mortality HR. In Table 1, all-cause mortality HR=1.01 for >35 vs. normal weight, while CVD=1.08 and cancer=1.13. In the WC results, HR=1.37 for category 4 vs 1 for all-cause mortality, while HR=1.94 for CVD mortality and HR=1.53 for cancer mortality. Why was the diabetes mortality HR=0.88 so influential here? As per your definition of diabetes mortality, these causes of death were not mutually exclusive from CVD or cancer, correct? Also, what are the other causes of mortality included in "all cause" and what are the HRs for this hodge-podge endpoint? Out of curiosity, what did the authors make of this finding? Or is the power too low to read into it?

We could only speculate as to why we saw those differences. Ultimately, we believe the sample size is too small to draw any conclusions based on those findings. Diabetes mortality could include any underlying cause of death such as CVD, cancer, other chronic diseases, infectious diseases, and accidents as long as the person filling out the death certificate felt diabetes played a role in the sequence of events leading up to the death (e.g., hypoglycemic event leading to a car accident).The most common underlying cause of death among people with diabetes mortality (diabetes listed as

anywhere on the death certificate) was diabetes followed by heart disease, stroke, and pneumonia. All-cause mortality included all deaths regardless of the cause. The most common underlying cause of death among everyone in the study who died was heart disease followed by cancer, diabetes, stroke, pneumonia, and other respiratory diseases.

3. The findings that the authors presented in their response stratified both by smoking status and diabetes diagnosis status are very interesting. Although power is limited, as they suggest, there is a suggested linear trend of increased mortality across BMI categories among never smokers, which is actually consistent with what was found in the Tobias et al NHS/HPFS NEJM paper. The authors state they chose not to present these results because of the limited sample size (64 deaths), but they did choose present results among other strata with even fewer deaths (e.g., only 29 deaths in the lowest age category). Please include a supplementary table stratified by both smoking and diabetes diagnosis status simultaneously, despite the limited power. This analysis will be very interesting to many readers.

We added this table as Supplement Table 8 and describe the results in the sensitivity analysis section of the results.

4. The authors describe in their response that they included participants 20+ years old, yet their exclusion criteria explicitly states that they excluded participants with likely type 1 diabetes (diagnosis <30 years of age). So how can there be participants with diabetes younger than age 30, wouldn't these be excluded b/c of their likely type 1 diabetes status? These younger participants should be excluded from the analysis, because as the methods indicate, it is likely to be type 1 diabetes.

In the methods we define type 1 diabetes as "previous diagnosis before 30 years of age, current insulin use, and first insulin use within 1 year of diagnosis". To further clarify the definition, we revised the text to read "having all 3 of the following criteria: previous diagnosis before 30 years of age, current insulin use, and first insulin use within 1 year of diagnosis".

5. The authors conclude that "further research should investigate possible reasons why the adiposity-mortality association may differ in people with diabetes compared to the general population." This conclusion is inappropriate for this paper. Comparing BMI-mortality associations for diabetics vs. general population was not the aim of this analysis and this results for this comparison were not done.

We deleted that sentence from the conclusion.

6. Given the limitations that are presented in the discussion section, do the authors truly believe that there is no association between adiposity and mortality? Is it possible that the validity to these findings are overwhelmed by bias, or are the authors comfortable with how they addressed their limitations? The reader should be left with some impression of how relevant these limitations are to interpretation of the results as a whole, even if it does not impact the overall conclusions.

We revised text in the discussion paragraph on limitations attempting to describe the importance of the limitations and give due consideration to how likely the limitation would impact the results.

7. Supplemental Table 1 by diagnosis status (pre-existing vs. newly diagnosed) is very helpful. I agree with the authors' response that there is value in understanding the relationship between adiposity and mortality among all diabetics, not just incident. I do want to make one point that the authors do not necessarily need to respond to, but it is a recurring theme in the discussion of the obesity paradox. Please remember that internal validity is a cornerstone of epidemiologic research. Stratifying one's dataset or using strict exclusion criteria certainly can reduce external validity (generalizability), but internal validity (i.e., minimal confounding, bias, etc), is foremost (purely crude descriptive statistics

and prediction models aside). As I'm sure you agree, discussing the generalizability of findings is moot if there is concern that they are not even valid to begin with. Think of a randomized clinical trial. RCTs enroll the most compliant, most eager, most willing participants. Exclusion criteria often leave researchers with only a fraction of participants they approached for the study. They know from the beginning that the generalizability of their findings to the general population will be highly questionable, but they sacrifice this limitation for the sake of increasing their internal validity and ensuring the least biased result. Similarly, it is ok to stratify observational data to explore the potential for bias and to explore potential threats to internal validity (e.g., smoking, diabetes duration). Including everyone doesn't mean you have achieved generalizability if the results themselves are biased.

We agree with the reviewer's comments, which is why we included several stratified tables in the supplement.

Minor comments:
1. Supplemental Table S8: Could you please include a row for # of deaths? How many additional deaths are included when all of the excluded participants are allowed back in? The authors point out that results were similar for this analysis compared with the main findings, but if few additional deaths were actually included, then it would not be surprising that results did not change much.

We added the number of deaths in the table. Without exclusions, there were almost twice as many deaths (n=1285). The results were similar to the main analysis and not significantly associated with mortality after multivariable adjustment.

2. How is diabetes duration modeled for those with new onset diabetes? Were they lumped in the <5 years category? I suggest making newly diagnosed in their own category.

They are included in their own category.

3. Table 1: Please include the following variables
o 1) % with dm diagnosed at NHANES visit (undiagnosed diabetes) ,
o 2) % with prevalent diabetes,
o among the prevalent, provide 3) % with dm dx <5 years ago, 5-9, 10-14, 15-19, 20+ years,
o 4) mean A1c concentrations,
o 5) % in tertiles of A1c as presented in Supp table S2: <5.9%, 5.9-7.0%, 7.1%+
o 6) % not taking diabetes medications
o 7) % taking any oral medications
o 8) % taking insulin

We added these rows to Table 1.

4. In your statistical methods, tables, and elsewhere, please change the definition of your lowest BMI category from <25 to 18.5-25.0.

We changed the text to include the lower interval: 18.5-24.9.

5. What is the correlation between BMI and WC in your sample for men and women? (i.e., how unique is WC from BMI in assessing adiposity?) This might be useful to mention in the results or discussion. What are the results if the BMI model is also adjusted for WC? And similarly, what are the results if the WC model is also adjusted for BMI?

The correlation was 0.923 for men and 0.893 for women. We added this information to the results section on page 9.

When we included both BMI and waist circumference in the same model, results were similar and not significant (p=0.99 for BMI and p=0.74 for waist circumference).

6. An additional limitation in the discussion should be that diabetics were enrolled from 1988-1994, and diagnostic criteria and treatment regimens have changed since then.

We previously mentioned that changes in medication use may cause weight gain or loss in the discussion on page 14:

“Weight gain or loss may have occurred during the course of diabetes as a result of changes in lifestyle, medication use, or diabetes disease progression.”

This was also mentioned it in the Strengths and Limitations section after the abstract:

“About 60% of the study participants were previously diagnosed with diabetes. Weight gain or loss may have occurred during the course of diabetes as a result of changes in lifestyle, medication use, or diabetes disease progression.”

We added a sentence addressing the generalizability in the discussion on page 14:

“The diagnostic criteria and treatment regimens commonly used in the US may have changed since the baseline of our study, which may affect the generalizability of our results to people who currently have diabetes.”

7. Table S8 is fine to include in the supplement if the authors feel it would interest other readers.

We decided to include this table thinking that some readers would be interested in the data since it illustrates the effect our exclusions had on the analysis.

VERSION 3 - REVIEW

REVIEWER	Mykola (Mykolay) Khalangot, MD PhD Department of Endocrinology P.L. Shupyk National Medical Academy of Postgraduate Education
REVIEW RETURNED	12-Sep-2014

GENERAL COMMENTS	"Under "minor revision" I mean correcting the abstract, that was not changed by the authors despite new data in manuscript's third version, which I have mentioned (points 1-3 of my last revision)."

REVIEWER	Deirdre Tobias Brigham and Women's Hospital, Harvard Medical School, Harvard School of Public Health, US
REVIEW RETURNED	24-Sep-2014

GENERAL COMMENTS	I sincerely thank the authors for addressing the reviewers' comments. It is clear that they took the comments seriously and have put a lot of work into the revisions. I have some additional comments based on authors' responses in the latest revision. Assuming these are reasonably addressed, I should have no further comments. - Please include p-values for interaction for stratified analyses. - What caused the attenuation in your WC results in this version compared with the previous version? - "The authors of a previous study suggested that collider bias may play a role in the inverse BMI-mortality association found in some previous studies of people with diabetes; they used NHANES data to investigate the issue finding an inverse association between BMI and mortality only among smokers, an association that was attenuated and not significant after introducing restrictions aimed at reducing the intensity of reverse causal pathways. Our study similarly excluded participants to limit the impact of reverse causal pathways and we found no association between multiple measures of adiposity and mortality. The non-significant positive association among never smokers suggests the possibility that there is some residual bias due to reverse causation in our analysis of the overall population and completely eliminating the bias may result in a positive association between adiposity and mortality." o Thank you for elaborating on this important point in your discussion. As a point of clarification, reverse causation is confounding, rather than collider bias, which are distinct forms of bias. - Your response: "We could only speculate as to why we saw those differences. Ultimately, we believe the sample size is too small to draw any conclusions based on those findings. Diabetes mortality

	could include any underlying cause of death such as CVD, cancer, other chronic diseases, infectious diseases, and accidents as long as the person filling out the death certificate felt diabetes played a role in the sequence of events leading up to the death (e.g., hypoglycemic event leading to a car accident).The most common underlying cause of death among people with diabetes mortality (diabetes listed as anywhere on the death certificate) was diabetes followed by heart disease, stroke, and pneumonia. All-cause mortality included all deaths regardless of the cause. The most common underlying cause of death among everyone in the study who died was heart disease followed by cancer, diabetes, stroke, pneumonia, and other respiratory diseases." o Very interesting. Since CVD (heart disease/stroke) HRs were all above the combined diabetes mortality HRs for BMI and WC, then it would likely be pneumonia/respiratory diseases largely driving the inverse direction of the association for "diabetes deaths"….two conditions highly prone to wasting/unintentional preclinical weight loss, increased mortality, and thus, reverse causation. I am even more convinced now that reverse causation is a major issue in your results, but as you mention, your low power will really prohibit any meaningful analyses to support this. Please add a "respiratory illness" (including pneum.) endpoint to Tables 2 and 3 to facilitate this discussion among readers. Also please include somewhere in the text, either the discussion or methods, these top contributors to "diabetes mortality" (% from CVD, % resp/pneum, etc) as you mentioned in your reviewer response above. - Exercise and diet are excluded from your methods in this updated version in response to another reviewer. I suggest keeping these details in your methods and including these variables in Table 1 (it seems diet was included previously b/c there is a footnote to the 24-hour recall?). Also mention in your results that including these variables in your multivariable model did not change the effect estimates. This will appease readers who will wonder why these confounders were omitted. - We added this table as Supplement Table 8 and describe the results in the sensitivity analysis section of the results. o Is the inclusion of "Diabetes Duration" a typo for Supplemental Table 8? If they all are incident diabetics then their duration should all be 0? - The conclusion states "the BMI-mortality association was similar when stratified by previous diabetes diagnosis, suggesting the timing of BMI measurement relative to diagnosis may not have affected our results". o I don't believe you have the statistical power to conduct interaction a test to make this conclusion. It would be more conservative to tone down this sentence. o Additionally, if you are going to make conclusive sentences about there being differences (e.g., you conclude that those with higher

	hba1c are different), then these statements should be accompanied by a p-interaction. o The ideal approach would be to discuss suggested trends, but ultimately that low power is unable to confirm these. - Paragraph in the discussion starting with “our findings in the overall population…”; do you mean among our total analytical population with diabetes, rather than the overall population? - In the discussion, the sentence with “…make obesity appear to have a relatively lower mortality risk”, could instead be written as “…make obesity appear to have a relatively lower mortality risk, compared with normal weight individuals”. - Statistical power needs to be mentioned as a limitation in the discussion. - In the methods we define type 1 diabetes as “previous diagnosis before 30 years of age, current insulin use, and first insulin use within 1 year of diagnosis”. To further clarify the definition, we revised the text to read “having all 3 of the following criteria: previous diagnosis before 30 years of age, current insulin use, and first insulin use within 1 year of diagnosis”. o I strongly feel that participants <30 years of age should not be included in your analysis. These are highly likely to be type 1 diabetics. With only 32 deaths in the 20-44 age range, I can’t imagine that it will change your findings, but would be more scientifically appropriate. - “Our results were robust in that we found similar results for both markers of adiposity.” o I wouldn’t consider this “robust” necessarily, given the very high correlation between BMI and WC in your sample. - “In conclusion, measures of adiposity were generally not associated with mortality in a nationally representative sample of people with diabetes.”, I suggest including “…people with prevalent and incident diabetes”.

VERSION 3 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:
Reviewer: 1
Reviewer Name Mykolay Khalangot
Institution and Country Department of Endocrinology
P.L. Shupyk National Medical Academy of Postgraduate Education Vyshgorodska 69 Kiev
04114 Ukraine
Please state any competing interests or state ‘None declared’: None declared

"Under "minor revision" I mean correcting the abstract, that was not changed by the authors despite new data in manuscript's third version, which I have mentioned (points 1-3 of my last revision)."

We added the following text to our abstract describing the significant findings in our sensitivity analysis:

“Several sensitivity analyses were conducted and most found no significant association between measures of adiposity and mortality, but there were significant results suggesting a U-shaped

association among people in the highest tertile of A1c (≥7.1%), and there was an inverse association between BMI and mortality among people 20-44 years of age."

Reviewer: 2
Reviewer Name Deirdre Tobias
Institution and Country Brigham and Women's Hospital, Harvard Medical School, Harvard School of Public Health, US
Please state any competing interests or state 'None declared': None declared

I sincerely thank the authors for addressing the reviewers' comments. It is clear that they took the comments seriously and have put a lot of work into the revisions. I have some additional comments based on authors' responses in the latest revision. Assuming these are reasonably addressed, I should have no further comments.

- Please include p-values for interaction for stratified analyses.

We calculated p-values for interaction and now address them in the results section.

- What caused the attenuation in your WC results in this version compared with the previous version?

For all-cause mortality, quartiles 2-4 had adjusted HR's (95% CI) of 1.20 (0.83-1.74), 0.92 (0.59-1.44), and 1.24 (0.79-1.96) in the previous version. In the most recent version, the analogous HR's (95% CI) were 1.03 (0.77-1.37), 1.02 (0.73-1.42), and 1.12 (0.77-1.61). We view these estimates as generally similar. We believe the minor differences may be random variation largely due to adding NHANES 1999-2004 data.

- "The authors of a previous study suggested that collider bias may play a role in the inverse BMI-mortality association found in some previous studies of people with diabetes; they used NHANES data to investigate the issue finding an inverse association between BMI and mortality only among smokers, an association that was attenuated and not significant after introducing restrictions aimed at reducing the intensity of reverse causal pathways. Our study similarly excluded participants to limit the impact of reverse causal pathways and we found no association between multiple measures of adiposity and mortality. The non-significant positive association among never smokers suggests the possibility that there is some residual bias due to reverse causation in our analysis of the overall population and completely eliminating the bias may result in a positive association between adiposity and mortality."
o Thank you for elaborating on this important point in your discussion. As a point of clarification, reverse causation is confounding, rather than collider bias, which are distinct forms of bias.

We revised the text to suggest there are multiple biases that may be affecting our results:

"The non-significant positive association among never smokers suggests the possibility that there are some residual biases in our analysis of the overall population and completely eliminating the biases may result in a positive association between adiposity and mortality."

- Your response: "We could only speculate as to why we saw those differences. Ultimately, we believe the sample size is too small to draw any conclusions based on those findings. Diabetes mortality could include any underlying cause of death such as CVD, cancer, other chronic diseases, infectious diseases, and accidents as long as the person filling out the death certificate felt diabetes played a role in the sequence of events leading up to the death (e.g., hypoglycemic event leading to a car accident).The most common underlying cause of death among people with diabetes mortality (diabetes listed as anywhere on the death certificate) was diabetes followed by heart disease, stroke,

and pneumonia. All-cause mortality included all deaths regardless of the cause. The most common underlying cause of death among everyone in the study who died was heart disease followed by cancer, diabetes, stroke, pneumonia, and other respiratory diseases."
o Very interesting. Since CVD (heart disease/stroke) HRs were all above the combined diabetes mortality HRs for BMI and WC, then it would likely be pneumonia/respiratory diseases largely driving the inverse direction of the association for "diabetes deaths"….two conditions highly prone to wasting/unintentional preclinical weight loss, increased mortality, and thus, reverse causation. I am even more convinced now that reverse causation is a major issue in your results, but as you mention, your low power will really prohibit any meaningful analyses to support this. Please add a "respiratory illness" (including pneum.) endpoint to Tables 2 and 3 to facilitate this discussion among readers. Also please include somewhere in the text, either the discussion or methods, these top contributors to "diabetes mortality" (% from CVD, % resp/pneum, etc) as you mentioned in your reviewer response above.

We added respiratory illness mortality as an endpoint in Tables 2 and 3. The analysis is underpowered and is not significantly associated with BMI or waist circumference. However, the hazard ratios are below 1 and may partially explain the lower hazard ratios for diabetes mortality. We also added the top underlying causes of death among those with "diabetes mortality" in the methods section:

"the most common underlying causes of death among these participants were diabetes, heart disease, stroke, and pneumonia"

- Exercise and diet are excluded from your methods in this updated version in response to another reviewer. I suggest keeping these details in your methods and including these variables in Table 1 (it seems diet was included previously b/c there is a footnote to the 24-hour recall?). Also mention in your results that including these variables in your multivariable model did not change the effect estimates. This will appease readers who will wonder why these confounders were omitted.

We added those variables to Table 1 and state in the discussion that including them in multivariable adjusted models did not affect our findings:

"Finally, the results were similar when we additionally adjusted for exercise and dietary variables (calories consumed and percent of calories from saturation fat; data not shown)."

- We added this table as Supplement Table 8 and describe the results in the sensitivity analysis section of the results.
o Is the inclusion of "Diabetes Duration" a typo for Supplemental Table 8? If they all are incident diabetics then their duration should all be 0?

Yes, we have corrected that typo.

- The conclusion states "the BMI-mortality association was similar when stratified by previous diabetes diagnosis, suggesting the timing of BMI measurement relative to diagnosis may not have affected our results".
o I don't believe you have the statistical power to conduct interaction a test to make this conclusion. It would be more conservative to tone down this sentence.

We revised that sentence removing the suggestion that the timing of BMI measurement may not have affected our results:

"However, when stratified by previous diabetes diagnosis, the BMI-mortality association was similar

and not significant for those previously diagnosed and those who were undiagnosed at baseline."

o Additionally, if you are going to make conclusive sentences about there being differences (e.g., you conclude that those with higher hba1c are different), then these statements should be accompanied by a p-interaction.
o The ideal approach would be to discuss suggested trends, but ultimately that low power is unable to confirm these.

We now refer to p-values for interaction in the results section. Only the analysis stratified by A1c had a significant p-value for interaction (p=0.003).

- Paragraph in the discussion starting with "our findings in the overall population…"; do you mean among our total analytical population with diabetes, rather than the overall population?

Yes, we revised the phrase to read: "overall diabetes population".

- In the discussion, the sentence with "…make obesity appear to have a relatively lower mortality risk", could instead be written as "…make obesity appear to have a relatively lower mortality risk, compared with normal weight individuals".

We revised the text as the reviewer suggested.

- Statistical power needs to be mentioned as a limitation in the discussion.

We added a sentence in the limitations paragraph of the discussion:

"Another limitation was the limited power to detect associations in some analyses of cause-specific mortality and in some sensitivity analyses in which we stratified results."

- In the methods we define type 1 diabetes as "previous diagnosis before 30 years of age, current insulin use, and first insulin use within 1 year of diagnosis". To further clarify the definition, we revised the text to read "having all 3 of the following criteria: previous diagnosis before 30 years of age, current insulin use, and first insulin use within 1 year of diagnosis".
o I strongly feel that participants <30 years of age should not be included in your analysis. These are highly likely to be type 1 diabetics. With only 32 deaths in the 20-44 age range, I can't imagine that it will change your findings, but would be more scientifically appropriate.

We agree with the reviewer that it will not have a substantial effect on our results. For example, the HR (95% CI) for fully adjusted all-cause mortality associated with each BMI category would change from 0.85 (0.60-1.21), 0.87 (0.57-1.33), and 1.05 (0.72-1.53) to 0.85 (0.60-1.21), 0.87 (0.57-1.32), and 1.05 (0.72-1.53) after excluding participants <30 years of age.

However, we respectfully disagree with the suggestion that people <30 years of age with diabetes are highly likely to have type 1 diabetes. It is well established in studies such as SEARCH that type 2 diabetes occurs commonly at a young age. We believe current age is not as important as age of diagnosis, insulin use, and the time between diabetes diagnosis and insulin initiation, which are the criteria we used to exclude people with likely type 1 diabetes. After excluding those people, the remaining participants 20-29 years of age were unlikely to have type 1 diabetes (only 3.6% were taking insulin).

- "Our results were robust in that we found similar results for both markers of adiposity."
o I wouldn't consider this "robust" necessarily, given the very high correlation between BMI and WC in

your sample.

We changed that sentence to read:

“Our results were consistent in that we found similar results for both markers of adiposity.”

- “In conclusion, measures of adiposity were generally not associated with mortality in a nationally representative sample of people with diabetes.”, I suggest including “…people with prevalent and incident diabetes”.

We changed the text as the reviewer suggested.

VERSION 4 - REVIEW

REVIEWER	Mykola (Mykolay) Khalangot, MD PhD Department of Endocrinology P.L. Shupyk National Medical Academy of Postgraduate Education
REVIEW RETURNED	22-Oct-2014

GENERAL COMMENTS	" After quite a lengthy review the manuscript obtained a Supplementary Figure S1 (Adjusted Relative Hazard of All-Cause Mortality Associated With Body Mass Index Among Men), from which we can conclude that mortality risk in T2D men with overweight and moderate obesity is twice lower, comparing to diabetic men, with normal BMI. Supplementary Table S2 (Hazards Ratio (95% Confidence Interval) of All-Cause Mortality) confirms the same phenomenon for both genders with HbA1c over 7.1%. In the cohort, examined by the authors this corresponds to the highest tertile, which looks quite unusual, as in most large T2D cohorts, mean HbA1c is at least 8%. Whereas 7% is the target level, which is rarely reached for a large number of patients. Thus, we can conclude, that the authors in reality have confirmed the fact of obesity paradox in patients with chronic hyperglycemia (HbA1c over 7.1%). They have also discovered several gender and age specifications. It seems that the authors see the results of their work somewhat differently (see conclusion). However we believe that publishing all of the results, obtained by the authors will encourage clarification of the obesity paradox problem and explain the nature of ambiguousness of previous results.

REVIEWER	Deirdre Tobias US, Brigham and Women's Hospital and Harvard Medical School
REVIEW RETURNED	29-Oct-2014

GENERAL COMMENTS	I have no further comments.

Reviewer: 1
Reviewer Name Mykolay Khalangot
Institution and Country Department of Endocrinology
P.L. Shupyk National Medical Academy of Postgraduate Education Vyshgorodska 69 Kiev
04114 Ukraine
Please state any competing interests or state 'None declared': None declared

"Under "minor revision" I mean correcting the abstract, that was not changed by the authors despite new data in manuscript's third version, which I have mentioned (points 1-3 of my last revision)."

We added the following text to our abstract describing the significant findings in our sensitivity analysis:

"Several sensitivity analyses were conducted and most found no significant association between measures of adiposity and mortality, but there were significant results suggesting a U-shaped association among people in the highest tertile of A1c (≥7.1%), and there was an inverse association between BMI and mortality among people 20-44 years of age."

Reviewer: 2
Reviewer Name Deirdre Tobias
Institution and Country Brigham and Women's Hospital, Harvard Medical School, Harvard School of Public Health, US
Please state any competing interests or state 'None declared': None declared

I sincerely thank the authors for addressing the reviewers' comments. It is clear that they took the comments seriously and have put a lot of work into the revisions. I have some additional comments based on authors' responses in the latest revision. Assuming these are reasonably addressed, I should have no further comments.

- Please include p-values for interaction for stratified analyses.

We calculated p-values for interaction and now address them in the results section.

- What caused the attenuation in your WC results in this version compared with the previous version?

For all-cause mortality, quartiles 2-4 had adjusted HR's (95% CI) of 1.20 (0.83-1.74), 0.92 (0.59-1.44), and 1.24 (0.79-1.96) in the previous version. In the most recent version, the analogous HR's (95% CI) were 1.03 (0.77-1.37), 1.02 (0.73-1.42), and 1.12 (0.77-1.61). We view these estimates as generally similar. We believe the minor differences may be random variation largely due to adding NHANES 1999-2004 data.

- "The authors of a previous study suggested that collider bias may play a role in the inverse BMI-mortality association found in some previous studies of people with diabetes; they used NHANES data to investigate the issue finding an inverse association between BMI and mortality only among smokers, an association that was attenuated and not significant after introducing restrictions aimed at reducing the intensity of reverse causal pathways. Our study similarly excluded participants to limit the impact of reverse causal pathways and we found no association between multiple measures of adiposity and mortality. The non-significant positive association among never smokers suggests the possibility that there is some residual bias due to reverse causation in our analysis of the overall population and completely eliminating the bias may result in a positive association between adiposity and mortality."
o Thank you for elaborating on this important point in your discussion. As a point of clarification, reverse causation is confounding, rather than collider bias, which are distinct forms of bias.

We revised the text to suggest there are multiple biases that may be affecting our results:

"The non-significant positive association among never smokers suggests the possibility that there are some residual biases in our analysis of the overall population and completely eliminating the biases may result in a positive association between adiposity and mortality."

- Your response: “We could only speculate as to why we saw those differences. Ultimately, we believe the sample size is too small to draw any conclusions based on those findings. Diabetes mortality could include any underlying cause of death such as CVD, cancer, other chronic diseases, infectious diseases, and accidents as long as the person filling out the death certificate felt diabetes played a role in the sequence of events leading up to the death (e.g., hypoglycemic event leading to a car accident).The most common underlying cause of death among people with diabetes mortality (diabetes listed as anywhere on the death certificate) was diabetes followed by heart disease, stroke, and pneumonia. All-cause mortality included all deaths regardless of the cause. The most common underlying cause of death among everyone in the study who died was heart disease followed by cancer, diabetes, stroke, pneumonia, and other respiratory diseases.”
o Very interesting. Since CVD (heart disease/stroke) HRs were all above the combined diabetes mortality HRs for BMI and WC, then it would likely be pneumonia/respiratory diseases largely driving the inverse direction of the association for “diabetes deaths”….two conditions highly prone to wasting/unintentional preclinical weight loss, increased mortality, and thus, reverse causation. I am even more convinced now that reverse causation is a major issue in your results, but as you mention, your low power will really prohibit any meaningful analyses to support this. Please add a “respiratory illness” (including pneum.) endpoint to Tables 2 and 3 to facilitate this discussion among readers. Also please include somewhere in the text, either the discussion or methods, these top contributors to “diabetes mortality” (% from CVD, % resp/pneum, etc) as you mentioned in your reviewer response above.

We added respiratory illness mortality as an endpoint in Tables 2 and 3. The analysis is underpowered and is not significantly associated with BMI or waist circumference. However, the hazard ratios are below 1 and may partially explain the lower hazard ratios for diabetes mortality. We also added the top underlying causes of death among those with “diabetes mortality” in the methods section:

“the most common underlying causes of death among these participants were diabetes, heart disease, stroke, and pneumonia”

- Exercise and diet are excluded from your methods in this updated version in response to another reviewer. I suggest keeping these details in your methods and including these variables in Table 1 (it seems diet was included previously b/c there is a footnote to the 24-hour recall?). Also mention in your results that including these variables in your multivariable model did not change the effect estimates. This will appease readers who will wonder why these confounders were omitted.

We added those variables to Table 1 and state in the discussion that including them in multivariable adjusted models did not affect our findings:

“Finally, the results were similar when we additionally adjusted for exercise and dietary variables (calories consumed and percent of calories from saturation fat; data not shown).”

- We added this table as Supplement Table 8 and describe the results in the sensitivity analysis section of the results.
o Is the inclusion of “Diabetes Duration” a typo for Supplemental Table 8? If they all are incident diabetics then their duration should all be 0?

Yes, we have corrected that typo.

- The conclusion states “the BMI-mortality association was similar when stratified by previous diabetes diagnosis, suggesting the timing of BMI measurement relative to diagnosis may not have affected our results”.
o I don’t believe you have the statistical power to conduct interaction a test to make this conclusion. It would be more conservative to tone down this sentence.

We revised that sentence removing the suggestion that the timing of BMI measurement may not have affected our results:

"However, when stratified by previous diabetes diagnosis, the BMI-mortality association was similar and not significant for those previously diagnosed and those who were undiagnosed at baseline."

o Additionally, if you are going to make conclusive sentences about there being differences (e.g., you conclude that those with higher hba1c are different), then these statements should be accompanied by a p-interaction.
o The ideal approach would be to discuss suggested trends, but ultimately that low power is unable to confirm these.

We now refer to p-values for interaction in the results section. Only the analysis stratified by A1c had a significant p-value for interaction (p=0.003).

- Paragraph in the discussion starting with "our findings in the overall population…"; do you mean among our total analytical population with diabetes, rather than the overall population?

Yes, we revised the phrase to read: "overall diabetes population".

- In the discussion, the sentence with "…make obesity appear to have a relatively lower mortality risk", could instead be written as "…make obesity appear to have a relatively lower mortality risk, compared with normal weight individuals".

We revised the text as the reviewer suggested.

- Statistical power needs to be mentioned as a limitation in the discussion.

We added a sentence in the limitations paragraph of the discussion:

"Another limitation was the limited power to detect associations in some analyses of cause-specific mortality and in some sensitivity analyses in which we stratified results."

- In the methods we define type 1 diabetes as "previous diagnosis before 30 years of age, current insulin use, and first insulin use within 1 year of diagnosis". To further clarify the definition, we revised the text to read "having all 3 of the following criteria: previous diagnosis before 30 years of age, current insulin use, and first insulin use within 1 year of diagnosis".
o I strongly feel that participants <30 years of age should not be included in your analysis. These are highly likely to be type 1 diabetics. With only 32 deaths in the 20-44 age range, I can't imagine that it will change your findings, but would be more scientifically appropriate.

We agree with the reviewer that it will not have a substantial effect on our results. For example, the HR (95% CI) for fully adjusted all-cause mortality associated with each BMI category would change from 0.85 (0.60-1.21), 0.87 (0.57-1.33), and 1.05 (0.72-1.53) to 0.85 (0.60-1.21), 0.87 (0.57-1.32), and 1.05 (0.72-1.53) after excluding participants <30 years of age.

However, we respectfully disagree with the suggestion that people <30 years of age with diabetes are highly likely to have type 1 diabetes. It is well established in studies such as SEARCH that type 2 diabetes occurs commonly at a young age. We believe current age is not as important as age of diagnosis, insulin use, and the time between diabetes diagnosis and insulin initiation, which are the criteria we used to exclude people with likely type 1 diabetes. After excluding those people, the remaining participants 20-29 years of age were unlikely to have type 1 diabetes (only 3.6% were taking insulin).

- "Our results were robust in that we found similar results for both markers of adiposity."
o I wouldn't consider this "robust" necessarily, given the very high correlation between BMI and WC in your sample.

We changed that sentence to read:

"Our results were consistent in that we found similar results for both markers of adiposity."

- “In conclusion, measures of adiposity were generally not associated with mortality in a nationally representative sample of people with diabetes.”, I suggest including “…people with prevalent and incident diabetes”.

We changed the text as the reviewer suggested.