PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The informational value of percent body fat with body mass index for the risk of abnormal blood glucose: a nationally representative cross-sectional study
AUTHORS	Jo, Ara; Mainous III, Arch

VERSION 1 – REVIEW

REVIEWER	Lan T Ho-Pham Ton Duc Thang University, Ho Chi Minh City, Vietnam
REVIEW RETURNED	07-Sep-2017

GENERAL COMMENTS	This study is interesting analysis that may contribute to the better identification of abnormal blood glucose (ABG) in the general population. However, I have some concerns on the threshold for defining "high" vs "low" percent body fat (PBF), and that could affect the authors' conclusion. Specific Comments 1. It is not clear to me whether the authors included people on antidiabetic medication in the analysis. Drugs could have confounded the analysis. 2. While I agree that the use of ABG as an outcome is reasonable, I have concerns on the analysis of data. All analysis variables – HbA1c, BMI and PBF – are continuous in nature. Why not treat them as continuous variables? Why categorize the data. We all know that categorization of data could lead to loss of information and inadequate adjustment. 3. What is the rationale for classifying individuals into 2 age groups (eg 40-70 and >70)? 4. The rationale for classifying low and high of PBF is not clear and not convincing. Based on the authors' criteria (25% for men and 32% for women), almost 88% of participants were classified as "high fat mass", but the proportion of overweight was ~68%. However, there is no scientific criteria for the thresholds (25% and 32%) at all. From the preliminary studies, a sex-specific threshold of %BP was 23.1% for men and 33.3% for women in study of Romeo-Corral et al 2010, and 21% for men and 35% for women in study of Shea J et al 2012; but the threshold of %BP in this study was adopted as 25% for men and 32% for women. While in other recent study also using an NHANES (1999-2006) sample of US adults (Martinez KE et al: J Diabetes Res 2017), the threshold of %BP that based on mean ± SD was 27.8% for men and 40.5% for women. 5. Given the comment #4, it is inappropriate to state that "among normal weight population […] 64% were misclassified as normal …"

	(page 7). 6. Without an analysis of area under the ROC curve, it is not possible to say whether PBF helped improve the prediction of ABG.

REVIEWER	Charlotte Evans University of Leeds UK
REVIEW RETURNED	06-Oct-2017

GENERAL COMMENTS	This is an interesting and informative study which raises important issues of the limitations of BMI. It needs strengthening in a number of areas. Firstly, the interaction term between percent body fat and BMI needs to be included in the analysis to determine whether the association of percent body fat with abnormal blood glucose is different at different for different BMI categories. Secondly, the discussion does not really discuss the main issues that are raised by these findings and is mostly just a summary of the results and a repeat of the introduction. Title – this needs to be changed so it is grammatically correct (missing a verb). Abstract Objective line 3 hypothesis more likely to have ABG than who? The normal weight population with lower percent BF? Make it clear. Outcomes- you may need to include other units such as mmol/mol used in the UK/Europe as this is an international journal. Results – this needs to include some of the descriptive results. how many (unweighted) and percent were in each of the 4 categories. Include odds ratio for all 3 groups compared with the reference category. At the moment the results don't match what you said you would do in the methods. You cannot report p value as 0. Please update all your p values throughout. Introduction There are some grammatical errors that need to be corrected. Add years when stating age. Methods In the section on anthropometric assessment be more precise. Are values above 25% defined as high % bf or does it include 25%? Outcomes – add other units such as mmol/mol at all times (you have included it some times but not others) I would argue that physical activity is a true confounder so wonder if you should change from potential confounder. The information on statistical analysis needs to be improved. You haven't said what you are going to do exactly and it is not clear that you have a reference category and are comparing the other 3 groups to the reference. You also need to include an interaction term to see if the influence of %bf varies by bmi/bmi category. There are also grammatical errors that need to be addressed. Results It is useful to include the number and percent (unweighted) in the text and table 2. The second sentence is not clear. You didn't say in your methods what the definition of misclassified was and how you calculated it or that you would be calculating and reporting it. In the last sentence of the results you say that the risks of abnormal glucose were great for overweight and low %bf but they were not statistically different for the latter as indicated in table 3. This is confusing and needs clarification. Discussion The discussion needs some more indepth discussion including

	comparisons with existing work and whether it is different or similar to the results seen here. Is percent body fat more important than bmi? Should we only look at percent body fat (below bmi of 30) or does bmi also help identify those most at risk? Policy implications are also important to mention. DEXA machines are really expensive so we can't measure everyone's body fat. It is cheaper to have a blood test to measure blood glucose so what is the benefit of measuring body fat? Include strengths as well as limitations.

REVIEWER	Dr Martine J. Barons University of Warwick, UK
REVIEW RETURNED	04-Dec-2017

GENERAL COMMENTS	This is an interesting study that deserves to be made public. In its present form, however, it is reported in insufficient detail, has inadequate sensitivity analysis and is insufficiently well written, and to be suitable for prestigious journal like BMJ Open. People with a BMI category of overweight or obese may be wrongly classified as unhealthy (Bacon & Araphamor 2011 Weight science: Evaluating the evidence for a paradigm shift Nutrition Journal201110:9 https://doi.org/10.1186/1475-2891-10-9) and improving the sensitivity of risk screening may reduce discrimination. This study sets out to examine whether the addition of % body fat to BMI will improve the latter's poor sensitivity performance in identifying those who have an elevated risk of abnormal blood glucose and who are therefore sent for diabetes screening. 1. The statistics used are elementary but seem to be appropriate, but the lack of detail makes it hard to be sure. 2. The results section of the abstract needs to be re-written for clarity, noting the comments below. 3. Define your terms carefully and use them consistently throughout. BMI is well-known and widely used to define overweight and obesity, but %BF cut-offs are not, so how these were arrived at must be described within the paper, page 5 line 56-8, as well as being referenced. The category names must also be given with the definition p5 line 55-7. This definition must be made when %BF is first introduced on p4. This categorisation must also become part of the discussion. 4. It is not clear when you are talking about risk category misclassification or other misclassification. P2 line 35; p4 line 44, p5 line 8, 20; p6 line 15, 26; p7 line 37-8; 5. Avoid pejorative language p4 line 44, 52; p7 line 44; p9 line 3, 48-52; Use low risk and raised / high risk throughout. 6. Missing data rates must be given for all variables. 7. Numbers excluded must also be given, usually in a flow chart p6 line 32-35 and 42-46. 8. It is not clear why you chose these variables. Usually these are the result of a forward or backward model building with a specific measure of fit. This must be reported. 9. P7 line 5 explain light seating. Discuss the reliability of self-reported physical activity. Justify this choice of definition of 'vigorous activity'. 10. P 7 line 46-50: a comparative percentage appears to be missing here. 11. P8 line 18-9: Use of BMI only may misclassify segments of the population in terms of RISK OF abnormal glucose. 12. P9 line 37-9 of RISK OF abnormal glucose 13. Sensitivity analyses:

	Since the %BF cut-offs are not standard, how sensitive are your results to different choices of these? It makes sense that, as in BMI, there would be a 'normal' range, but in this study you used only High and Low, although on p6 line 21 you suggest that Low is normal. How would your results differ if you defined a normal range, a low range and a high range? 14. If %BF improves risk stratification and BMI is poor, does %BF alone provide a better risk stratification than BMI alone? How sensitive is this to categorisation of %BF? 15. Discussion should include other potential confounders, such as poverty, diet quality, sleep (Cappuccio, F. P., D'Elia, L., Strazzullo, P., & Miller, M. A. (2010). Quantity and quality of sleep and incidence of type 2 diabetes: A systematic review and meta-analysis. Diabetes Care, 33(2), 414–420. http://doi.org/10.2337/dc09-1124 Cappuccio, F. P., Taggart, F. M., Kandala, N.-B., Currie, A., Peile, E., Stranges, S., & Miller, M. A. (Physician). (2008). Meta Analysis of Short Sleep Duration and Obseity in Children and Adults. Sleep, 31(5).), smoking, alcohol…. 16. Discussion should include the cost of DXA measurement and the accuracy of alternatives and how this impacts on the practicability of such an approach in the clinical setting. If DXA is expensive and BMI categorisation is a poor risk stratification, would it be cost-effective to make HbA1c testing routine? 17. Discussion should include how many additional / fewer people will be tested under the proposed regime – if the overweigh low group are no longer tested and the normalweight high group are tested, how many additional cases of ABG will be detected and how many missed that would have been detected under the BMI-only risk stratification? 18. P8 line 40-44: you suggest that obesity is a cause of inflammation. Causal claims must be accompanied by robust evidence, typically fulfilment of the Bradford-Hill criteria. Since you exclude individuals with a BMI category of obese in your study, this statement seems irrelevant. 19. Page 9 line 28-9 discuss how similar / different the population is in the latest measurement particularly in the covariates you have selected. 20. Table 1. It is unclear what you mean by BMI misclassification here. It is sufficient to name the 4 categories according to the two measures e.g. overweight low. A 2 x 2 table would be clearer here, with number (out of 9790) as well as percentages in each box e.g. '979 (10%)' 21. Table 2 should also report number and percentage of 9,790 in each section, e.g. age, being careful that all add up to 9,790. Row 2 does not add up to 101,098,270. 22. Table 2 p-values should be given with more precision <.00 is not sufficient. Either p<2.2 x10-16 or p<0.001 format. 23. Table 2 the statistical test to which P-values relate should be stated. If this is the same throughout, then this may be given in the table caption as well as the text. Otherwise, an additional column giving the statistical test should appear next to the p-value. 24. Table 3 the statistical test used should be given. 25. Since there is already a body of literature showing that BMI is a poor indicator of individual health, you could use this to strengthen your conclusion including that the proposal of the EEOC is not supported by these results.

VERSION 1 – AUTHOR RESPONSE

Reviewer Comments:

Reviewer #1: Lan Ho-Pham
This study is interesting analysis that may contribute to the better identification of abnormal blood glucose (ABG) in the general population. However, I have some concerns on the threshold for defining "high" vs "low" percent body fat (PBF), and that could affect the authors' conclusion.

We thank the reviewer for the comments and feedback to improve the quality of the manuscript. We hope that our responses below would address the reviewer's concerns.

1. It is not clear to me whether the authors included people on antidiabetic medication in the analysis. Drugs could have confounded the analysis.

Response: We did not include people on antidiabetic medications in the study. Our study population focused on people who have never been diagnosed with diabetes by doctors or who did not report having been prescribed antidiabetic medications either. Consequently, we feel that diabetic medications would be unlikely to confound the results.

2. While I agree that the use of ABG as an outcome is reasonable, I have concerns on the analysis of data. All analysis variables – HbA1c, BMI and PBF – are continuous in nature. Why not treat them as continuous variables? Why categorize the data. We all know that categorization of data could lead to loss of information and inadequate adjustment.

Response: The reviewer makes an important point. It is true that categorizing continuous data does lose information. However, there are important clinical cut points that represent obesity and diabetes. Thus we used commonly agreed upon clinical thresholds that allow us to define abnormal blood glucose and normal weight/overweight. Also using clinical thresholds could show misclassification and clinical gap occurred by BMI in clinical practice.

3. What is the rationale for classifying individuals into 2 age groups (eg 40-70 and >70)?

Response: The US Preventive Services Task Force (USPSTF) recommends the screening guideline for adults who are overweight or obesity aged between 40 and 70. Also abnormal blood glucose has been critical for older population who are not main target population in clinical guideline. Thus we classified them into two groups.

4. The rationale for classifying low and high of PBF is not clear and not convincing. Based on the authors' criteria (25% for men and 32% for women), almost 88% of participants were classified as "high fat mass", but the proportion of overweight was ~68%. However, there is no scientific criteria for the thresholds (25% and 32%) at all. From the preliminary studies, a sex-specific threshold of %BP was 23.1% for men and 33.3% for women in study of Romeo-Corral et al 2010, and 21% for men and 35% for women in study of Shea J et al 2012; but the threshold of %BP in this study was adopted as 25% for men and 32% for women. While in other recent study also using an NHANES (1999-2006) sample of US adults (Martinez KE et al: J Diabetes Res 2017), the threshold of %BP that based on mean ± SD was 27.8% for men and 40.5% for women.

Response: We thank you for making a point of this. In revising the manuscript we discovered that we had made a typographical error regarding the cut point. We have made corrections in thresholds. We used cutoffs, 25% of men and 35% of women according to the World Health Organization (WHO) Expert Committee. (Reference: World Health Organization (1995) Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. World Health Organ Tech Rep Ser 854: 1–452.World Health Organization1995Physical status: the use and interpretation of

anthropometry. Report of a WHO Expert Committee. World Health Organ Tech Rep Ser8541452) The WHO recommended these cutoffs to define obesity if one was to use %BF. A result of misclassification in normal weight and overweight population is still the same.

5. Given the comment #4, it is inappropriate to state that "among normal weight population […] 64% were misclassified as normal …" (page 7).

Response: We have rephrased this statement as “Table 1 showed that among normal weight population, approximately 64% of normal weight population who had high %BF were misclassified as normal weight population despite a high level of %BF.”

6. Without an analysis of area under the ROC curve, it is not possible to say whether PBF helped improve the prediction of ABG.

Response: The reviewer makes an important observation. We have added the ROC curves of BMI only, %BF only and combined form of BMI and %BF as figure 1 at the end of the manuscript. We also have added interpretation in results section.

Reviewer #2: Charlotte Evans

This is an interesting and informative study which raises important issues of the limitations of BMI. It needs strengthening in a number of areas. Firstly, the interaction term between percent body fat and BMI needs to be included in the analysis to determine whether the association of percent body fat with abnormal blood glucose is different at different for different BMI categories. Secondly, the discussion does not really discuss the main issues that are raised by these findings and is mostly just a summary of the results and a repeat of the introduction.

We thank the reviewer for the overall positive review and comments.

1. Title – this needs to be changed so it is grammatically correct (missing a verb).

Response: We have changed the title-“The Informational Value of Percent Body Fat with Body Mass Index for the Risk of Abnormal Blood Glucose”.

2. Abstract

a. Objective line 3 hypothesis more likely to have ABG than who? The normal weight population with lower percent BF? Make it clear.

Response: Thank you for this comment. We have modified the manuscript in line with this suggestion. We have clarified that the reference group is the normal weight group who has lower %BF.

b. Outcomes- you may need to include other units such as mmol/mol used in the UK/Europe as this is an international journal.

Response: The reviewer is correct that additional units will be helpful to an international audience. We have added other units in HbA1c according to the UK/Europe journal guideline. Main outcome includes prediabetes and undiagnosed diabetes (HbA1c ≥5.7%, ≥39mmol/mol).

c. Results – this needs to include some of the descriptive results. how many (unweighted) and percent were in each of the 4 categories. Include odds ratio for all 3 groups compared with the

reference category. At the moment the results don't match what you said you would do in the methods.

Response: We have modified the manuscript in line with this suggestions. It is important to note that the NHANES uses a complex survey design such as clustering, stratification and multiple oversampling and so all population estimates are based not just on the weighted data but also account for the complex design. Thus we did not unweighted sample size in Table 1, because proportions were computed from weighted sample size. We have stated total unweighted sample size in the title of Table 1. Table 3 presented odds ratio for all three groups compared to the reference group-normal weight group who had lower %BF.

3. Introduction

a. There are some grammatical errors that need to be corrected.

Response: We have corrected grammatical errors in introduction.

b. Add years when stating age.

Response: We have incorporated this suggestions. We have stated years after every ages throughout the manuscript.

4. Methods

a. In the section on anthropometric assessment be more precise. Are values above 25% defined as high % bf or does it include 25%?

Response: Yes, this cutoff included 25% and 35% in higher %BF. We have added language to the manuscript to clarify this for the reader. In revising the manuscript we discovered that we had made a typographical error regarding the cut point. We have made corrections in thresholds. We used cutoffs, 25% of men and 35% of women according to the World Health Organization (WHO) Expert Committee. (World Health Organization (1995) Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. World Health Organ Tech Rep Ser 854: 1–452.World Health Organization1995Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee.World Health Organ Tech Rep Ser8541452) The WHO recommended these cutoffs to define obesity if one was to use %BF.

b. Outcomes – add other units such as mmol/mol at all times (you have included it some times but not others)

Response: Based on previous comments, we have added other units in HbA1c according to the UK/Europe journal guideline throughout the entire manuscript.

c. I would argue that physical activity is a true confounder so wonder if you should change from potential confounder.

Response: We have modified language in the manuscript. We have added justification of this variable in methods section. Vigorous physical activity and muscle strengthening activity might be confounder, because these activities play roles in increasing muscle mass and weighing more. It may result in overweight and reducing risk of abnormal blood glucose due to lower fat mass. Therefore we decided to have physical activity as a confounder.

d. The information on statistical analysis needs to be improved. You haven't said what you are going to do exactly and it is not clear that you have a reference category and are comparing the other 3 groups to the reference. You also need to include an interaction term to see if the influence of %bf varies by bmi/bmi category. There are also grammatical errors that need to be addressed.

Response: In response to this comment, we have added language for clarity. We have added more information on statistical analysis. We have clarified that reference group was the normal weight population who had low %BF.
For the interaction term, bmi and %BF has highly correlated each other. According to Barreira and colleagues, the correlation between %BF and BMI was 0.8 (Barreira, T. V., Staiano, A. E., Harrington, D. M., Heymsfield, S. B., Smith, S. R., Bouchard, C., & Katzmarzyk, P. T. (2012). Anthropometric correlates of total body fat, abdominal adiposity, and cardiovascular disease risk factors in a biracial sample of men and women. Mayo Clin Proc, 87(5), 452-460. doi: 10.1016/j.mayocp.2011.12.017). Therefore, the interaction term may not be appropriate to be included in the model.

5. Results
a. It is useful to include the number and percent (unweighted) in the text and table 2.

Response: As we responded to a previous comment, the NHANES uses a complex survey design such as clustering, stratification and multiple oversampling and so all population estimates are based not just on the weighted data but also account for the complex design (Reference: National Center for Health Statistics. National Health and Nutrition Examination Survey Sample Design, 2011-2014. Retrieved from https://wwwn.cdc.gov/nchs/data/series/sr02_162.pdf). Thus unweighted numbers will not add up with the weighted percentages. We did not add the unweighted numbers in Table 2.

b. The second sentence is not clear. You didn't say in your methods what the definition of misclassified was and how you calculated it or that you would be calculating and reporting it.

Response: Thank you for your point. We have clarified this interpretation. 64% of normal weight population who had higher %BF was misclassified as normal weight population despite a high level of %BF. It is proportion in normal weight population.

c. In the last sentence of the results you say that the risks of abnormal glucose were great for overweight and low %bf but they were not statistically different for the latter as indicated in table 3. This is confusing and needs clarification.

Response: In line with your suggestion, we have clarified this sentence. "Risks of abnormal glucose were significantly greater in population with normal weight and high %BF as well as the overweight with high %BF (Table 3)."

6. Discussion
a. The discussion needs some more in-depth discussion including comparisons with existing work and whether it is different or similar to the results seen here. Is percent body fat more important than bmi? Should we only look at percent body fat (below bmi of 30) or does bmi also help identify those most at risk? Policy implications are also important to mention. DEXA machines are really expensive so we can't measure everyone's body fat. It is cheaper to have a blood test to measure blood glucose so what is the benefit of measuring body fat? Include strengths as well as limitations.

Response: We have modified the manuscript in line with this comment. We have added more discussion about value of %BF and risk of ABG and potential policy implications. We also have added cost of the use of a DXA for the purpose of screening in limitation section. In terms of cost, a DXA

scan may not be appropriate for the purpose of screening. As a cost-effective alternative, we have suggested Bioelectrical Impedance Analysis (BIA) to measure %BF. In addition, we have expanded our discussion about benefits of measuring body fat in discussion section.

Reviewer #3: Martine J. Barons
This is an interesting study that deserves to be made public. In its present form, however, it is reported in insufficient detail, has inadequate sensitivity analysis and is insufficiently well written, and to be suitable for prestigious journal like BMJ Open.
People with a BMI category of overweight or obese may be wrongly classified as unhealthy (Bacon & Araphamor 2011 Weight science: Evaluating the evidence for a paradigm shift Nutrition Journal201110:9 https://urldefense.proofpoint.com/v2/url?u=https-3A__doi.org_10.1186_1475-2D2891-2D10-2D9&d=DwIFaQ&c=pZJPUDQ3SB9JplYbifm4nt2lEVG5pWx2KikqINpWlZM&r=YIfrX0aS4RKuFuyipEmo5Q&m=aAfpQdmkVtGBv1grlYyEQloZJo82qSCn9W_gdvIrNPc&s=Ux1ZDXEF4r4SSNd3RUyjk2IMg9uVnpPmclKy42tO25U&e=) and improving the sensitivity of risk screening may reduce discrimination. This study sets out to examine whether the addition of % body fat to BMI will improve the latter's poor sensitivity performance in identifying those who have an elevated risk of abnormal blood glucose and who are therefore sent for diabetes screening.

We thank the reviewer for the comments to improve the quality of our manuscript. We hope that our responses below would address the reviewer's concerns.

1. The statistics used are elementary but seem to be appropriate, but the lack of detail makes it hard to be sure.

Response: We have modified the manuscript in line with this comment. We have addressed details of statistical methods by responding to the questions below.

2. The results section of the abstract needs to be re-written for clarity, noting the comments below.

Define your terms carefully and use them consistently throughout. BMI is well-known and widely used to define overweight and obesity, but %BF cut-offs are not, so how these were arrived at must be described within the paper, page 5 line 56-8, as well as being referenced. The category names must also be given with the definition p5 line 55-7. This definition must be made when %BF is first introduced on p4. This categorisation must also become part of the discussion.

Response: We thank you for making a point of this. In revising the manuscript we discovered that we had made a typographical error regarding the cut point. We have made corrections in thresholds. We used cutoffs, 25% of men and 35% of women according to the World Health Organization (WHO) Expert Committee. (Reference: World Health Organization (1995) Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. World Health Organ Tech Rep Ser 854: 1–452.World Health Organization1995Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. World Health Organ Tech Rep Ser8541452) The WHO recommended these cutoffs to define obesity if one was to use %BF. Currently, there are no gold standard cutoffs for %BF in practice. To secure validity of the use of %BF, we used -25% for men and 35% for women. We have clarified this point under anthropometric assessment section. We also have added this definition to the introduction on page 4. Also study groups have been defined in methods section.

3. It is not clear when you are talking about risk category misclassification or other misclassification. P2 line 35; p4 line 44, p5 line 8, 20; p6 line 15, 26; p7 line 37-8;

Response: We thank you for your point. We have clarified the classified groups. Four groups of participants were determined based on the combined metric of BMI and %BF. We predicted risks of abnormal blood glucose of each group in statistical analyses (Table 3).

4. Avoid pejorative language p4 line 44, 52; p7 line 44; p9 line 3, 48-52; Use low risk and raised / high risk throughout.

Response: The reviewer makes an important point. We have modified the manuscript according to this suggestion.

5. Missing data rates must be given for all variables.

Response: We thank you for making a point of this. In revising the manuscript we discovered that we had made a typographical error regarding the total unweighted sample size. We have updated study population from 9,790 to 6,335.We also have added missing data rates in the results. “No variable had more than 3% unweighted missing data and none of the demographics had any missing data.”

6. Numbers excluded must also be given, usually in a flow chart p6 line 32-35 and 42-46.

Response: As we responded to a previous comment, we discovered that we had made a typographical error regarding the total unweighted sample size. We have updated study population from 9,790 to 6,335.We also have added missing data rates in the results. “No variable had more than 3% unweighted missing data and none of the demographics had any missing data.”
We have added numbers excluded based on inclusion criteria such as age and BMI. Initial population was 39,352. After limiting to adults aged 40 years old or older, 27,273 was excluded. In addition, respondents who were missed in BMI or %BF, who were underweight or obesity who have ever been diagnosed diabetes by a doctor were excluded (missing=5,744). Thus, final study population was 6,335.

7. It is not clear why you chose these variables. Usually these are the result of a forward or backward model building with a specific measure of fit. This must be reported.

Response: All covariates were selected from the literature. We have added further explanation of the reason for selecting a family history of diabetes and physical activities in methods section. This study focused on the association between combined body composition assessment and the risk of abnormal blood glucose instead of selecting optimal risk factors of abnormal blood glucose. These variables were chosen based on theoretical framework from preliminary study. It is well documented that demographics, first degree relative diabetes and physical activities are associated with risk of abnormal blood glucose and amount of %BF. The goal of the forced inclusion analysis was to control for variables not model building.

8. P7 line 5 explain light seating. Discuss the reliability of self-reported physical activity. Justify this choice of definition of ‘vigorous activity’.

Response: We thank you for making a point. As we responded to a previous comment, we discovered that we had made an error regarding the definition of physical activity. We have removed light seating in text. The NHANES adapted the Global Physical Activity Questionnaire developed by the World Health Organization. This questionnaire also has been adopted by the other population based survey

worldwide, Health Survey England and Korean NHANES. Based on these facts, we believe that this self-reported physical activity is reliable.
Vigorous activity may be a key factor to increase muscle mass and reduce fat mass. As can be seen in our findings, the overweight population with low %BF are more likely to perform vigorous activity compared to normal weight population with high %BF. In particular, intensive exercise may result in increased muscle mass that leads to increased body weight. It may misclassify overweight population as unhealthy group.

9. P 7 line 46-50: a comparative percentage appears to be missing here.

Response: We have added the percentage of the reference group.

10. P8 line 18-9: Use of BMI only may misclassify segments of the population in terms of RISK OF abnormal glucose.

Response: In line with this suggestions, we have revised the wording of the manuscript.

11. P9 line 37-9 of RISK OF abnormal glucose

Response: In line with this suggestions, we have revised the wording of the manuscript.

12. Sensitivity analyses:
a. Since the %BF cut-offs are not standard, how sensitive are your results to different choices of these? It makes sense that, as in BMI, there would be a 'normal' range, but in this study you used only High and Low, although on p6 line 21 you suggest that Low is normal. How would your results differ if you defined a normal range, a low range and a high range?

Response: The reviewer makes an important point on the lack of an agreed upon cut point for %BF. In reading the analysis to look at multiple levels we discovered a typographical error form the previous manuscript. We have made a correction on %BF cut-off from 32% of women to 35% of women. A result is still the same. This cut off was recommended by the WHO expert committee (Reference: World Health Organization (1995) Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. World Health Organ Tech Rep Ser 854: 1–452.World Health Organization1995Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee.World Health Organ Tech Rep Ser8541452) The WHO recommended these cutoffs to define obesity if one was to use %BF.
With this cut-off, we have added sensitivity analyses using ROC curve. Figure has been attached at the end of page. Area Under the Curve (AUC) of combined form of BMI and %BF was higher than areas of BMI only or %BF only.
%BF has one clinical threshold that distinguish low and high body fat. It is not able to define a low, normal and high range according to %BF.

13. If %BF improves risk stratification and BMI is poor, does %BF alone provide a better risk stratification than BMI alone? How sensitive is this to categorisation of %BF?

Response: As we responded to a previous comment, we have added sensitivity analysis at the end of the manuscript. We have plotted ROC curves of BMI only, %BF only and combined form of BMI and %BF. As a result, Area Under the Curve (AUC) of combined form was higher than areas of BMI only or %BF only.

14. Discussion should include other potential confounders, such as poverty, diet quality, sleep (Cappuccio, F. P., D'Elia, L., Strazzullo, P., & Miller, M. A. (2010). Quantity and quality of sleep and

incidence of type 2 diabetes: A systematic review and meta-analysis. Diabetes Care, 33(2), 414–420. https://urldefense.proofpoint.com/v2/url?u=http-3A__doi.org_10.2337_dc09-2D1124&d=DwIFaQ&c=pZJPUDQ3SB9JplYbifm4nt2lEVG5pWx2KikqINpWlZM&r=YIfrX0aS4RKuFuyipEmo5Q&m=aAfpQdmkVtGBv1grlYyEQloZJo82qSCn9W_gdvIrNPc&s=lKNjH7mAAOjA_UFw73GLFyINZhCyYXMvruGEX5leEvI&e= Cappuccio, F. P., Taggart, F. M., Kandala, N.-B., Currie, A., Peile, E., Stranges, S., & Miller, M. A. (Physician). (2008). Meta Analysis of Short Sleep Duration and Obseity in Children and Adults. Sleep, 31(5).), smoking, alcohol….

Response: We thank you for your further insights with regard to additional covariates. Those potential confounders may be considerable factors for further study. Although our current study has focused on integrated form of %BF and BMI as a tool of detecting ABG, we will consider those potential factors in further study. We have added language to that effect in discussion section.

15. Discussion should include the cost of DXA measurement and the accuracy of alternatives and how this impacts on the practicability of such an approach in the clinical setting. If DXA is expensive and BMI categorisation is a poor risk stratification, would it be cost-effective to make HbA1c testing routine?

Response: We agreed your point. While DXA scan is the most accurate measurement, it is expensive to use in primary care setting. We have addressed this point in limitation section and suggested alternative assessment in limitation.

16. Discussion should include how many additional / fewer people will be tested under the proposed regime – if the overweigh low group are no longer tested and the normal weight high group are tested, how many additional cases of ABG will be detected and how many missed that would have been detected under the BMI-only risk stratification?

Response: In line with your suggestion, we have added number of additional cases of ABG by using combined form of assessment at the end of discussion. “This strategy may detect more than 303,000 US adults who are normal weight and who usually miss an opportunity to receive preventive care service on time due to the use of BMI only.”

17. P8 line 40-44: you suggest that obesity is a cause of inflammation. Causal claims must be accompanied by robust evidence, typically fulfilment of the Bradford-Hill criteria. Since you exclude individuals with a BMI category of obese in your study, this statement seems irrelevant.

Response: We thank you for your point. We have removed the statement about inflammation in text.

18. Page 9 line 28-9 discuss how similar / different the population is in the latest measurement particularly in the covariates you have selected.

Response: We have addressed similarity and differences among groups in terms of covariates. “Regardless of BMI, people with high %BF were older, female and Non-Hispanic White. While proportion of family history representing a first degree of relative ever being told by a doctor that they had diabetes was similar across four groups, physical activity was shown significantly different.”

19. Table 1. It is unclear what you mean by BMI misclassification here. It is sufficient to name the 4 categories according to the two measures e.g. overweight low. A 2 x 2 table would be clearer here, with number (out of 9790) as well as percentages in each box e.g. ‘979 (10%)’

Response: According to your suggestion, we have modified the table 1 to a 2x2 table. Percentages are column % in each BMI category.

20. Table 2 should also report number and percentage of 9,790 in each section, e.g. age, being careful that all add up to 9,790. Row 2 does not add up to 101,098,270.

Response: The reviewer makes an important point. The NHANES uses a complex survey design such as clustering, stratification and multiple oversampling and so all population estimates are based not just on the weighted data but also account for the complex design (Reference: National Center for Health Statistics. National Health and Nutrition Examination Survey Sample Design, 2011-2014. Retrieved from https://wwwn.cdc.gov/nchs/data/series/sr02_162.pdf). Thus unweighted numbers will not add up with the weighted percentages.
Moreover, in revising the manuscript we discovered that we made a typographical error regarding to the sample size. We have updated study sample size. 6,335 is unweighted sample size and 65,705,694 is weighted sample size. The heading of Table 2 stated unweighted and weighted sample size.

21. Table 2 p-values should be given with more precision <.00 is not sufficient. Either p<2.2 x10-16 or p<0.001 format.

Response: In line with this suggestions, we have revised all p-values from <.00 to <.001 in Table 2 of the manuscript.

22. Table 2 the statistical test to which P-values relate should be stated. If this is the same throughout, then this may be given in the table caption as well as the text. Otherwise, an additional column giving the statistical test should appear next to the p-value.

Response: Methods section stated that the statistical test was chi-square test for table 2. The statistical test for table 2 was given in text on the basis of format in published studies in BMJ open.

23. Table 3 the statistical test used should be given.

Response: Methods section stated that the statistical test for table 3 was logistic regression. The statistical test for table 3 was given in text on the basis of format in published studies in BMJ open.

24. Since there is already a body of literature showing that BMI is a poor indicator of individual health, you could use this to strengthen your conclusion including that the proposal of the EEOC is not supported by these results.

Response: Thanks to your suggestion, we have addressed that BMI may not be the optimal tool to assess health outcomes for employees and the proposal should be modified based on our findings in last discussion section. “In addition, according to the rule offered by EEOC, employees who are classified as overweight or obesity with high muscle mass and lower body fat may get penalized (6). Our findings indicated that BMI may not be the optimal tool to assess health outcomes for employees and new rule of the EEOC should be modified to consider body fat instead of body weight.”

VERSION 2 – REVIEW

REVIEWER	Lan T Ho-Pham Ton Duc Thang University, Vietnam
REVIEW RETURNED	25-Jan-2018

GENERAL COMMENTS	The authors responded that "We have made corrections in thresholds. We used cutoffs 25% of men and 35% of women

	according to the WHO Expert Committee (1995)...". Unfortunately, the WHO report did not set any threshold of PBF for defining obesity (Lan T Ho-Pham, Mayo Clin Proc 2011:584), only several authors have made the misquotation of the PBF threshold from the article of Oreopulos et al (Oreopulos A, Mayo Clin Proc 2010:607)

REVIEWER	Martine J. Barons University of Warwick, UK
REVIEW RETURNED	05-Feb-2018

GENERAL COMMENTS	The informational value of percent body fat with body mass index for the risk of abnormal blood glucose This paper has been significantly improved following the reviewers' comments. However, a number of points have not been addressed. 1. Still using the words 'misclassification', 'reclassified' or 'misclassified' without qualification or explanation. This was point 4 in the original review. For example, Page 2 line 33 since normal is a BMI category, '64% of population with Normal BMI classification, also had a high %BF' would be more accurate and less confusing than 'misclassified as normal'. 2. Page 2 line 36 give also the percentage of the population with Overweight BMI classification and low %BF to compare with the 13.5%. Also state the odds ratios as in line 45. 3. Still using inappropriate, pejorative language like 'unhealthy' and 'excessive' which do not have precise definitions. Use 'low risk of ABG' and 'raised or high risk of ABG' throughout instead. This was point 5 in the original review. 4. Page 6 line 22 were CLASSIFIED AS underweight etc. 5. Points 9, 11, 14, 17, 19, 20, 23 & 25 in the first review have not yet been addressed.

VERSION 2 – AUTHOR RESPONSE

Reviewer Comments:

Reviewer #1: Lan Ho-Pham
COMMENT
Please leave your comments for the authors below The authors responded that "We have made corrections in thresholds. We used cutoffs 25% of men and 35% of women according to the WHO Expert Committee (1995)...". Unfortunately, the WHO report did not set any threshold of PBF for defining obesity (Lan T Ho-Pham, Mayo Clin Proc 2011:584), only several authors have made the misquotation of the PBF threshold from the article of Oreopulos et al (Oreopulos A, Mayo Clin Proc 2010:607)
RESPONSE: The reviewer made an important point. We have replaced the reference using clinical cut points of body fat percentage given by the guideline of the American Association of Clinical Endocrinologists (AACE) and the American College of Endocrinology (ACE). (Reference: Dickey RA, Bartuska DG, Bray GW, Callaway CW, Davidson ET, Feld S, Ferraro RT, Hodgson SF, Jellinger PS, Kennedy FP, Lawrence AM. AACE/ACE Position statement on the prevention, diagnosis, and treatment of obesity (1998 revision). Endocr Pract. 1998 Sep;4(5):297-350.)

Reviewer #3: Martine J. Barons
Please leave your comments for the authors below The informational value of percent body fat with body mass index for the risk of abnormal blood glucose

This paper has been significantly improved following the reviewers' comments. However, a number of points have not been addressed.

COMMENT
1.Still using the words 'misclassification', 'reclassified' or 'misclassified' without qualification or explanation. This was point 4 in the original review. For example, Page 2 line 33 since normal is a BMI category, '64% of population with Normal BMI classification, also had a high %BF' would be more accurate and less confusing than 'misclassified as normal'.
RESPONSE: We thank you for pointing them out again. In line with your suggestions, we have modified the language throughout the manuscript.

COMMENT
2.Page 2 line 36 give also the percentage of the population with Overweight BMI classification and low %BF to compare with the 13.5%. Also state the odds ratios as in line 45.
RESPONSE: According to your suggestions, we have added the percentage of the population with overweight BMI and low %BF (10.5%) to the page 2, line 36. To clarify the odds ratios in line 45, we have stated the odds ratio in line 45.

COMMENT
3.Still using inappropriate, pejorative language like 'unhealthy' and 'excessive' which do not have precise definitions. Use 'low risk of ABG' and 'raised or high risk of ABG' throughout instead. This was point 5 in the original review.
RESPONSE: In line with your suggestions, we have modified the wordings, 'unhealthy' and 'excessive' to 'high risk of ABG' throughout the manuscript.

COMMENT
4.Page 6 line 22 were CLASSIFIED AS underweight etc.
RESPONSE: We have modified the language according to your suggestion.

COMMENT
5.Points 9, 11, 14, 17, 19, 20, 23 & 25 in the first review have not yet been addressed.
COMMENT-Point 9: P 7 line 46-50: a comparative percentage appears to be missing here.
RESPONSE: We have modified the language according to your suggestion.

COMMENT-Point 11: P9 line 37-9 of RISK OF abnormal glucose.
RESPONSE: In line with this suggestion, we have revised the wording of the manuscript.

COMMENT-Point 14: Discussion should include other potential confounders, such as poverty, diet quality, sleep (Cappuccio, F. P., D'Elia, L., Strazzullo, P., & Miller, M. A. (2010). Quantity and quality of sleep and incidence of type 2 diabetes: A systematic review and meta-analysis. Diabetes Care, 33(2), 414–420. https://urldefense.proofpoint.com/v2/url?u=http-3A__doi.org_10.2337_dc09-2D1124&d=DwIFaQ&c=pZJPUDQ3SB9JplYbifm4nt2lEVG5pWx2KikqINpWlZM&r=YIfrX0aS4RKuFuyipEmo5Q&m=aAfpQdmkVtGBv1grlYyEQloZJo82qSCn9W_gdvIrNPc&s=lKNjH7mAAOjA_UFw73GLFyINZhCyYXMvruGEX5leEvI&e= Cappuccio, F. P., Taggart, F. M., Kandala, N.-B., Currie, A., Peile, E., Stranges, S., & Miller, M. A. (Physician). (2008). Meta Analysis of Short Sleep Duration and Obseity in Children and Adults. Sleep, 31(5).), smoking, alcohol….
RESPONSE: We thank you for your further insights with regard to additional covariates. We have added more information to the discussion about this. We mention the importance of these covariates and their implications for future interventions and how a focus on them could be useful in future research.

COMMENT-Point 17: P8 line 40-44: you suggest that obesity is a cause of inflammation. Causal claims must be accompanied by robust evidence, typically fulfilment of the Bradford-Hill criteria. Since you exclude individuals with a BMI category of obese in your study, this statement seems irrelevant.
RESPONSE: This point was made by the reviewer in the previous review. We had modified the manuscript to remove the sentence with that language. The current version of the manuscript does not include language of a causal relationship between inflammation and obesity.

COMMENT-Point 19: Table 1. It is unclear what you mean by BMI misclassification here. It is sufficient to name the 4 categories according to the two measures e.g. overweight low. A 2 x 2 table would be clearer here, with number (out of 9790) as well as percentages in each box e.g. '979 (10%)'.
RESPONSE: Table 1 had been modified by 2X2 table in the previous revision.
The NHANES uses a complex survey design with clustering, stratification and multiple oversampling. This complex survey design has a huge advantage over many other surveys in that it allows us to compute population estimates (Reference: National Center for Health Statistics. National Health and Nutrition Examination Survey Sample Design, 2011-2014. Retrieved from https://wwwn.cdc.gov/nchs/data/series/sr02_162.pdf).
Consequently, to use the NHANES as recommended we need to use software that accounts for the complex design and make population estimates based on the weighted data. The percentages in Table 1 represent the weighted proportions. It is not recommended to use unweighted numbers.

COMMENT-Point 20: Table 2 should also report number and percentage of 9,790 in each section, e.g. age, being careful that all add up to 9,790. Row 2 does not add up to 101,098,270.
RESPONSE: As with our response to comment #19, the NHANES uses a complex survey design which allows us to provide population estimates. All of the numbers are based on the weighted population estimates that account for the complex survey design. Thus unweighted numbers will not add up with the weighted percentages and are misleading to the reader. For context, we have presented the unweighted and weighted sample size in Table 2 but all of the percentages are based on the weighted numbers.

COMMENT-Point 23: Table 3 the statistical test used should be given.
RESPONSE: We have added 'Logistic Regression' in the title of Table 3 according to your suggestion.

COMMENT-Point 25: Since there is already a body of literature showing that BMI is a poor indicator of individual health, you could use this to strengthen your conclusion including that the proposal of the EEOC is not supported by these results.
RESPONSE: I apologize for missing this comment in previous revision. The reviewer made an important point. As you addressed in the comment, many literature demonstrated that the use of BMI only is not a perfect indicator of individual health. A weakness of EEOC proposal is caused by the use of BMI only. To improve accuracy, we focused on utilizing integrated form of BMI and %BF. Currently, there are limited studies to utilize integrated form of BMI and %BF in detecting risk of chronic diseases.