PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) 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)	Gender Differences in the Association of Individual Social Class and Neighborhood Unemployment Rate With Prevalent Type 2 Diabetes Mellitus: A Cross-sectional Study from The DIAB-CORE Consortium
AUTHORS	Müller, Grit; Hartwig, Saskia; Greiser, Karin Halina; Moebus, Susanne; Pundt, Noreen; Schipf, Sabine; Völzke, Henry; Maier, Werner; Meisinger, Christa; Tamayo, Teresa; Rathmann, Wolfgang; Berger, Klaus

VERSION 1 - REVIEW

REVIEWER	Andrew Hinde Division of Social Statistics and Demography University of Southampton SOUTHAMPTON SO17 1BJ United Kingdom
REVIEW RETURNED	31-Jan-2013

THE STUDY	On p. 6, the authors write that they has 9,399 study participants, of whom 520 were eliminated because of missing information. This left 8,870. But 9,399 - 520 = 8,879. What happened to the other 9 participants? The elimination of 520 participants because of missing information is a cause of concern. I should like the authors to check (and probably to present evidence) that the elimination of the 520 does not bias the sample. In other words, was the distribution among the 520 of those characteristics for which they did provide information similar to the distribution of the same characteristics among the 8,870? I think the graph in the Appendix should be incorporated into the main body of the text. The methods and results are already summarised in the text in the same way that the methods and results of the other analyses are. I do not see why the graph is relegated to be .supplementary on-line material'. In the abstract, ll. 16-17 the statement 'variation was explained by explanatory variables in men, but not in women' is vague. Variation is explained by explanatory variables by definition! What you mean is that the covariates you included in your models helped to 'explain' the variation among women, but not among men.
RESULTS & CONCLUSIONS	The text needs revision to take the reader through the results more gently.
GENERAL COMMENTS	The use of the word 'explained' is problematic. The authors should emphasise that they use this word in a statistical rather than a causal sense (e.g. p. 5, l. 10).

	On the question of causality, then I think you need to remind readers that your analysis does not allow you to infer cause and effect. For example the association between employment status in men and the age-adjusted T2DM prevalence (p. 9, ll. 24-25) could be causal in either direction or be spurious (i.e. due to a third factor). On p.10, l. 10 who are the 'others' who are not employed, unemployed or retired? Given that rural participants were excluded from the analysis in two areas, is this really an urban study? If so, I think this point should be made somewhere. Why did you not use a neighbourhood deprivation score rather than the unemployment rate to define the socio-economic status of neighbourhoods. The unemployment rate might be sensitive to local, temporary factors, such as the closure of a large factory, whereas a deprivation score taken into account a broader range of characteristics and so is less sensitive. In the write up of the results, you move between the three sets of analyses you have done without any indication to the reader that you have moved. For example, the paragraph on p. 9, ll. 1-18 describes the results from the bivariate age-adjusted analysis in the first two sentences, but the final sentence refers to the results of the multiple logistic regression stratified by sex. It is all rather breathless. Can you take the reader through your results in a more leisurely fashion? The same problem characterises the paragraph on p. 10, ll. 5-18 in which the results of three pieces of analysis are discussed in the same paragraph. On p.10, l. 28 insert 'age-adjusted' before 'prevalence'. It is also worth noting that the effect of living with a partner becomes statistically insignificant when other factors are controlled.

REVIEWER	Brendan Smith, MSc, PhD Candidate Dalla Lana School of Public Health University of Toronto Canada I have no competing interests.
REVIEW RETURNED	13-Feb-2013

REPORTING & ETHICS	Around research ethics, no discussion is present in the manuscript. For redundant publication, I have a major concern around whether the results presented in this study are sufficiently different from those in previously published works. In my review I have asked that the author(s) please clarify how the manuscripts in reference 14, 15, 16 (from the submitted manuscript) are different than your proposed study? Is it just that they have analyzed the data in sex subgroups? If this is the case these results could have easily been included in one of these manuscripts and does not warrant a new publication.
GENERAL COMMENTS	The objective of the manuscript was to analyze gender differences in the relationship of individual and neighborhood social factors with present type 2 diabetes mellitus (T2D).

	Before starting the review of this manuscript, I have a major concern around whether the results presented in this study are sufficiently different from those in previously published works. Can the author(s) please clarify how the manuscripts in reference 14, 15, 16 are different than your proposed study? Is it just that you have now analyzed the data in sex subgroups? If this is the case these results could have easily been included in one of these manuscripts and does not warrant a new publication. Also, why in some studies does the DIAB-CORE include six population-based studies (e.g. reference 14) and in others 5 population based studies? **Manuscript Review** I have outlined specific concerns below that I would encourage the authors to address in any subsequent version of this manuscript. These are not listed by importance, but by progression through the manuscript. Overall, the submitted manuscript could improve clarity for readers by applying consistent (names for exposures: Around individual socioeconomic status and neighbourhood characteristics) and appropriate language throughout the text (e.g. odds instead of chance of developing T2D). It would also benefit by a more organized presentation of the results. Some suggestions are provided below. **Article summary** It should be emphasized that there was an association between SES and T2D in both men and women, along with the fact that it was stronger in women than men. This is an important distinction and should be recognized in the manuscript. The term “spatial variation” is both non-specific and not consistent with the language in the text. **Introduction** The introduction does not provide the reader with a background necessary to introduce the study. E.g. Line 7-11 Since it is a major focus of your manuscript, a more complete discussion regarding sex differences in the association between socioeconomic position and T2DM is necessary. Is the association stronger in women, only present in women? You have not referenced a major systematic review on the topic (see Agardh et al, 2011 IJE) and also missed a number of papers that have found an association in men (Smith et al, 2013 Anns of Epidemiol, Ross et al, 2010 Health Reports or Kumari et al, 2004 Arch Intern Med 2004 for example). This section would be strengthened by updating it to

	include relevant manuscripts on this topic. Line 25 – Is it gender differences or sex differences? Any differences between men and women should be labeled as sex differences. Please fix throughout the text. Somewhere in the introduction the concept of residential environment should be defined. Is it neighborhood deprivation, employment rates, the built environment, access to services? …Is it unclear what about residential environments increase risk of T2DM. Line 43 – The aim of the study should clearly state individual and neighborhood social determinants are being investigated? Also, aim 3 does not seem to be adequately address in the manuscript, and I would consider moving it to a web-appendix. **Methods** Outcome Undiagnosed diabetes can be a source of bias. Is there any validation of the self-report diabetes reporting in your sample? The DIAB-CORE Consortium has published many studies now on T2D. This information also could be updated in the limitations section (p. 13, lines 50-52). Do you have information about diabetes diagnosed during pregnancy? How was this accounted for in your analyses? Exposure p. 7 Lines 54 Can you please elaborate on the social status measure (the Winkler-Index of Socioeconomic status). How were household income and education categorized? How was this index created? How was the index modified? This is important information for the reader and should be included in the manuscript. The reference provided is incorrect and points to a manuscript that used occupation (which you do not have in this study) in the development of their Winkler-Index of Socioeconomic status variable. The occupation variable should be discussed at the beginning of page 8. It is unclear why you have only included employment status

	and not occupational achievement. Social inequalities across level of occupation are common, and therefore it is confusing to the reader why you have chosen to classify employment in this fashion. This is not mentioned again until the limitations section. Covariates Why is a lifestyle index used instead of examining the impact of risk factors individually? This index assumes that there is equal risk of developing diabetes among each of these factors, which is likely not the case. For example, obesity has a much strong effect on risk of T2DM compared to alcohol consumption. Also, the name of the categories is inappropriate. Would the authors consider being obese healthy? The presentation of neighbourhood unemployment rate should be included below the individual socioeconomic status measures. A clear rationale should be provided for why neighbourhood unemployment rate was chosen and the degree to which it is an adequate measure of neighbourhood socioeconomic status. Were any other neighbourhood measures available in the dataset? An improved rationale would be beneficial for the choice of equally-sized tertiles of neighbourhood unemployment rate. Is this categorization meaningful and consistent across studies/areas? Statistical Analyses The presentation of the analyses undertaken for the study could be improved. In the statistical models, the distinction needs to made between confounders and mediators along with a description of the authors approach. Along with the description, the tables in the text could include

	additional information of the appropriate statistical models. It would also be beneficial for the reader to present the results in stages: unadjusted/age-adjusted individual level variables, confounder adjusted individual variables, confounder adjusted individual level variables and area level variables and then all variables. These results should be communicated in coordination with those from Table 4 (the fully adjusted models). A main finding of the study is that the effect of SES on T2D is different by sex. This was confirmed by finding a statistically significant multiplicative interaction between SES and sex. Given this finding, it is not clear why some models are sex-specific and others are not. Based on your findings it seems that all models should be sex-specific. Sex-pooled models should be removed from the manuscript, or a clear description of how the results are being interpreted to include the effect of the interaction term from sex and SES. For example, do the results in Figure 1 include the portion of the association between social status and T2D relationship that is modeled through the interaction term? Why are any results being presented in a sex-pooled models if there is interaction by sex with the main exposure? Please provide evidence of ethical approval for the study, and surveys used in the manuscript. **Results** Page 10, line 2-9. The description in this section is not clear and should be revised. Page 10, line 11. The word chance is not appropriate. It should read odds. Please fix throughout the text. In Table 3, the label “Adjusted Prevalence” should be more specific to improve clarity. Also in the title, it refers to neighbourhood socio-demographic variables. Why not simply stated neighbourhood unemployment rate?

	In Table 4 “VA” in is not clear. A more specific labeled would improve clarity. Also, somewhere in Table 4 please include information on which confounders have been adjusted in the presented models? Page 10, lines 33-48 Can be moved to web material. Page 11, line 7-21. The analyses for the presented findings have not been discussed and the results are not presented in text. **Discussion** Page 12, lines 39-58. The discussion around the importance of employment status should be deemphasized. While the explanation is not necessarily inappropriate, given the limitations around the employment measure it is difficult to assess the role of employment on the results from this study. p. 13, lines 3-15. Please provide a more detailed discussion around the difference between neighbourhood unemployment rate and the effect of neighbourhoods in men and women. Why do you believe that high unemployment rate is associated with increase odds of developing diabetes in men and not women, but that the residential environment is more important in women than men? A more complete discussion around the effects of neighbourhoods on T2D is warranted. For example, see Krishnan S, Cozier YC, Rosenberg L, Palmer JR. Socioeconomic status and incidence of type 2 diabetes: results from the Black Women’s Health Study. Am J Epidemiol 2010;171(5):564e70. p. 13, lines 36-40 The statement around “no prior studies…” as suggested in the last comment is not appropriate.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

1. On p. 6, the authors write that they has 9,399 study participants, of whom 520 were eliminated because of missing information. This left 8,870. But 9,399 - 520 = 8,879. What happened to the other 9 participants?

Respond: Additional 9 participants with missing information on their marital status were excluded from the analysis, thus, a total of 529 participants were eliminated. We corrected the corresponding sentence in the methods' section on page 7.

2. The elimination of 520 participants because of missing information is a cause of concern. I should like the authors to check (and probably to present evidence) that the elimination of the 520 does not bias the sample. In other words, was the distribution among the 520 of those characteristics for which they did provide information similar to the distribution of the same characteristics among the 8,870?

Response: Overall, 529 individuals were excluded from the analysis due to missing information on social class, employment status and marital status. The social class variable is a summary measure combining information on educational and professional training and net household income. The missing information on social class was solely due to missing information on the net household income. Moreover, six subjects with missing information on income also missed information on their employment status.

First, we reviewed the characteristics of the two groups: (1) with information on net household income and (2) without information on net household income. The latter group had a higher mean age (61.1 years) than the participants with full information (59.7 years) (P≤0.000) and an uneven sex ratio with a higher fraction of women (62.9%) compared to the participants with full information (48.9%) (P≤0.000). We could observe significant differences according to the educational status and employment status. Among the participants without information on income, a higher proportion had a low educational level (16.7%) compared to participants with income information (10.9%) (P≤0.000). The percentage of retired individuals (46.9% versus 44.9% (with income information)) and individuals with other employment status (17.3% versus 8.6% (with income information)) was also higher (P≤0.000). Furthermore, the groups were disproportional distributed across the tertiles of neighborhood unemployment rate (P≤0.000): study participants without information on income resided more often in neighborhoods with low unemployment rate (39.6% versus 32.7% with income information).

In a second step, we conducted a sensitivity analysis with the sample of 9,381 individuals. We had 18 drop outs due to missing information on the employment status and marital status. Education was applied as a measure of social class (educational score ranging between 1 (lowest) and 7 (highest)). The analyses showed similar results compared to the results presented in the manuscript. In women and men, we found a significant age effect on the odds of type 2 diabetes (men: OR=1.04 (95% CI: 1.02-1.06); women: OR=1.05 (95% CI: 1.02-1.07)). We observed a social gradient in the odds of type 2 diabetes, which was more pronounced in women (OR of educational score: 0.81 (95% CI: 0.74-0.89)) than in men (OR of educational score: 0.89 (95% CI: 0.82-0.96)). We found an increased odds to have type 2 diabetes in men residing in high unemployment neighborhoods (OR=1.45 (95% CI: 1.14-1.85)) compared to men in low unemployment neighborhoods, which could not be observed in women (OR=1.14 (95% CI: 0.82-1.60)).

Among the covariates employment status and marital status, we found deviating effects on type 2 diabetes. Against our previous results, we observed an increased odds of type 2 diabetes in retired men (OR=1.21 (95% CI: 0.87-1.67)), which were, however, not significant. Among women, the estimates were not different from the estimates presented before. In women and men, we found a significant association between living without a partner and the odds of type 2 diabetes (men: OR=1.34 (95% CI: 1.02-1.76); women: OR=1.33 (95% CI: 1.04-1.70)).

In conclusion, the group of excluded study participants without information on net household income

was older on average, had a lower educational level and was more often out of employment (e.g. retirement, housewife/-men). This could have resulted in an underestimation of the social gradient in the odds of type 2 diabetes in our analyses. However, overall, the results from the sensitivity analysis were similar to the presented results in the manuscript.
We included a sentence highlighting the potential underestimation of the social gradient on type 2 diabetes in the section on limitations on page 13-14.

3. I think the graph in the Appendix should be incorporated into the main body of the text. The methods and results are already summarized in the text in the same way that the methods and results of the other analyses are. I do not see why the graph is relegated to be .supplementary on-line material'.

Response: We moved the graph, labeled as figure 1, into the main text.

4. In the abstract, ll. 16-17 the statement 'variation was explained by explanatory variables in men, but not in women' is vague. Variation is explained by explanatory variables by definition! What you mean is that the covariates you included in your models helped to 'explain' the variation among women, but not among men.

Response: We changed the terminology on page 2 according to the reviewers' suggestions.

5. The use of the word 'explained' is problematic. The authors should emphasize that they use this word in a statistical rather than a causal sense (e.g. p. 5, l. 10).

Response: According to the suggestions of the reviewer, we changed the expression to "explained statistically". We also clarified that through the whole text.

6. On the question of causality, then I think you need to remind readers that your analysis does not allow you to infer cause and effect. For example the association between employment status in men and the age-adjusted T2DM prevalence (p. 9, ll. 24-25) could be causal in either direction or be spurious (i.e. due to a third factor).

Response: In the section on limitations on page 13, we pointed out that our results were based on cross-sectional analysis and causal conclusions were limited.

7. On p.10, l. 10 who are the 'others' who are not employed, unemployed or retired?

Response: The group 'others' included individuals in occupational retraining, housewives and housemen. We added this information to the methods' section on page 7.

8. Given that rural participants were excluded from the analysis in two areas, is this really an urban study? If so, I think this point should be made somewhere.

Response: Since the participants in the rural areas of KORA and SHIP could not be considered in our analysis, our study was limited to urban areas. We highlighted that in the methods section on page 6 and also referred to that in the abstract (page 2) and results section (page 9).

9. Why did you not use a neighbourhood deprivation score rather than the unemployment rate to define the socio-economic status of neighbourhoods. The unemployment rate might be sensitive to local, temporary factors, such as the closure of a large factory, whereas a deprivation score taken into account a broader range of characteristics and so is less sensitive.

Response: The authors' agree with the reviewer on that point. Neighborhood unemployment rate is a sensitive indicator. Unfortunately, the pool of available administrative data differs significantly by region and city in Germany. Thus, the selection of context measures is difficult, especially on the low city level and when ensuring comparability across regions. In our previous work (Mueller et al. 2013), we evaluated the association between five available context variables (unemployment rate, number of migrants, married residents, residents 0-17 years old, residents over age 65 relative to the total number of residents) and prevalent type 2 diabetes and only found for unemployment rate significant associations. A number of studies applied unemployment rate as a measure of deprivation (Dragano et al. 2007, Dragano et al. 2009, van Lenthe et al. 2005, Cummins et al. 2005). Campbell et al. 1991 concluded that unemployment rates are a good and simple indicator for deprivation. We added this to the description on variables in the methods' section and two supporting references on page 7.

References:

- Müller G, Kluttig A, Greiser KH, et al. Regional and Neighborhood Disparities in the Risk of Type 2 Diabetes: Results from Five Population-Based Studies in Germany (DIAB-CORE Consortium). AM J EPIDEMIOL 2013; in press.
- Dragano N, Bobak M, Wege N, et al. Neighbourhood socioeconomic status and cardiovascular risk factors: a multilevel analysis of nine cities in the Czech Republic and Germany. BMC PUBLIC HEALTH 2007;7(1):255.
- Dragano N, Hoffmann B, Stang A, et al. Subclinical Coronary Atherosclerosis And Neighbourhood Deprivation in an Urban Region. EUR J EPIDEMIOL 2009;24(1):25-35.
- van Lenthe FJ, Borrell LN, Costa G, et al. Neighbourhood unemployment and all cause mortality: a comparison of six countries. Journal of Epidemiology and Community Health 2005;59(3):231-37.
- Cummins S, Stafford M, Macintyre S, Marmot M, Ellaway A. Neighbourhood environment and its association with self rated health: evidence from Scotland and England. J EPIDEMIOL COMMUNITY HEALTH 2005;59(3):207-13.
- Campbell DA, Radford JM, Burton P. Unemployment rates: an alternative to the Jarman index? BRIT MED J 1991;303(6805):750-55

Van Lenthe et al. 2005 and Campbell et al. 1991 were added as references to the manuscript.

10. In the write up of the results, you move between the three sets of analyses you have done without any indication to the reader that you have moved. For example, the paragraph on p. 9, ll. 1-18 describes the results from the bivariate age-adjusted analysis in the first two sentences, but the final sentence refers to the results of the multiple logistic regression stratified by sex. It is all rather breathless. Can you take the reader through your results in a more leisurely fashion? The same problem characterizes the paragraph on p. 10, ll. 5-18 in which the results of three pieces of analysis are discussed in the same paragraph.

Response: The authors' restructured the results' section according to the type of analysis, starting with the descriptive analysis, followed by the interaction analysis and the stratified regression analysis.

11. On p.10, l. 28 insert 'age-adjusted' before 'prevalence'.

Response: As suggested by the reviewer, we added the term "age-adjusted" to the sentence.

12. It is also worth noting that the effect of living with a partner becomes statistically insignificant when other factors are controlled.

Response: We added that information to the description of the multivariable results on page 10.

Reviewer 2

1. Before starting the review of this manuscript, I have a major concern around whether the results presented in this study are sufficiently different from those in previously published works. Can the author(s) please clarify how the manuscripts in reference 14, 15, 16 are different than your proposed study? Is it just that you have now analyzed the data in sex subgroups? If this is the case these results could have easily been included in one of these manuscripts and does not warrant a new publication.

Response: The referred publications by Schipf et al. 2012, Maier et al. 2013 and Müller et al. 2013 are independent works. The paper by Schipf et al. 2012 presented age-standardized estimates for the prevalence of type 2 diabetes for the five regions as well as Germany as a whole derived from the regional and nation-wide studies. Representative epidemiological data on the prevalence and incidence of type 2 diabetes on a regional level is still rare in Germany.
The article by Maier et al. 2013 investigated the differences in the prevalence of type 2 diabetes between municipalities (n=30, below regional level) applying the new small-area German Index of Multiple Deprivation, developed by Maier et al. 2012. This work included also the rural areas of SHIP and KORA, which could not be considered in this present analysis.
The authors' previous work (Müller et al. 2013) dealt with the evaluation of regional and between neighborhood variations in the prevalence of type 2 diabetes and the contribution of individual and contextual variables (five neighborhood characteristics were available) in the explanation of these variations. The consideration of gender-specific analysis would have gone beyond the scope of this previous work. A major drawback of the existing literature on the impact of social determinants, particularly characteristics of the neighborhood, is the assumption that all population subgroups are affected in the same way. Our work wants to address this research gap.

References:

- Schipf S, Werner A, Tamayo T, et al. Regional differences in the prevalence of known Type 2 diabetes mellitus in 45–74 years old individuals: Results from six population-based studies in Germany (DIAB-CORE Consortium). DIABET MED 2012;29(7):88-95.
- Maier W, Holle R, Hunger M, et al. The impact of regional deprivation and individual socio-economic status on the prevalence of Type 2 diabetes in Germany. A pooled analysis of five population-based studies. Diabet Medicine 2013;30(3):e78-e86.
- Maier W, Fairburn J, Mielck A. [Regional deprivation and mortality in Bavaria. Development of a community-based index of multiple deprivation]. Gesundheitswesen (Bundesverband der Arzte des Oeffentlichen Gesundheitsdienstes (Germany)) 2012;74(7):416-25.
- Müller G, Kluttig A, Greiser KH, et al. Regional and Neighborhood Disparities in the Risk of Type 2 Diabetes: Results from Five Population-Based Studies in Germany (DIAB-CORE Consortium). AM J EPIDEMIOL 2013; in press.

2. Also, why in some studies does the DIAB-CORE include six population-based studies (e.g. reference 14) and in others 5 population based studies?

Response: The DIAB-CORE includes six population-based studies, one nation-wide and five regional studies. Only the study participants of the five regional studies were linked to their area of residence (e.g. municipalities, neighborhoods) and therefore, only their data could be included in the analysis on the influence of spatial characteristics on type 2 diabetes.

Manuscript Review

3. Overall, the submitted manuscript could improve clarity for readers by applying consistent (names for exposures: Around individual socioeconomic status and neighbourhood characteristics) and

appropriate language throughout the text (e.g. odds instead of chance of developing T2D).

Response: We corrected the text according to the consistency of terms used for the exposure variables. According to the reviewers' suggestions, we replaced the term 'chance' with 'odds' through the whole text.

4. It would also benefit by a more organized presentation of the results. Some suggestions are provided below.

Response: The authors reorganized the results' section.

Article summary
5. It should be emphasized that there was an association between SES and T2D in both men and women, along with the fact that it was stronger in women than men. This is an important distinction and should be recognized in the manuscript.

Response: We added this fact to the articles' abstract (page 2), the summary (page 3) and results section in the main text (page 10).

6. The term "spatial variation" is both non-specific and not consistent with the language in the text.

Response: We replaced the term "spatial variation" with the terminology "between-study and between-neighborhood variance" on page 3 to be more specific and congruent with the manuscript.

Introduction

7. The introduction does not provide the reader with a background necessary to introduce the study. E.g. Line 7-11 Since it is a major focus of your manuscript, a more complete discussion regarding sex differences in the association between socioeconomic position and T2DM is necessary. Is the association stronger in women, only present in women? You have not referenced a major systematic review on the topic (see Agardh et al, 2011 IJE) and also missed a number of papers that have found an association in men (Smith et al, 2013 Anns of Epidemiol, Ross et al, 2010 Health Reports or Kumari et al, 2004 Arch Intern Med 2004 for example). This section would be strengthened by updating it to include relevant manuscripts on this topic.

Response: The authors revised the introduction and added additional relevant publications on page 5 according to the suggestions of the reviewer. We differentiated between studies that found associations between social status and type 2 diabetes only in women, studies that found associations in both men and women, whereby the association was stronger in women, and studies that showed contrasting results with a stronger association in men.

References added:

- Agardh E, Allebeck P, Hallqvist J, Moradi T, Sidorchuk A. Type 2 diabetes incidence and socio-economic position: a systematic review and meta-analysis. International Journal of Epidemiology 2011;40(3):804-18.
- Ross NA, Gilmour H, Dasgupta K. 14-year diabetes incidence: The role of socio-economic status. Health Reports Statistics Canada 2010;21(3).
- Kumari M, Head J, Marmot M. Prospective Study of Social and Other Risk Factors for Incidence of Type 2 Diabetes in the Whitehall II Study. Arch Intern Med 2004;164(17):1873-80.
- Maty SC, Everson-Rose SA, Haan MN, Raghunathan TE, Kaplan GA. Education, income, occupation, and the 34-year incidence (1965–99) of Type 2 diabetes in the Alameda County Study. International Journal of Epidemiology 2005;34(6):1274-81.

8. Line 25 – Is it gender differences or sex differences? Any differences between men and women should be labeled as sex differences. Please fix throughout the text.

Response: Since the focus of our study was not on biological differences between women and men and the impact on type 2 diabetes, but rather on social and cultural originated differences, the authors prefer to use the term “gender”.
Please see: Krieger N. Genders, sexes, and health: what are the connections—and why does it matter? International Journal of Epidemiology 2003;32(4):652-57.

9. Somewhere in the introduction the concept of residential environment should be defined. Is it neighborhood deprivation, employment rates, the built environment, access to services? …Is it unclear what about residential environments increase risk of T2DM.

Response: Residential environment is a broad concept, including the built environment, supply of services and neighborhood resources, environmental strains (e.g. air pollution, noise exposure) as well as social and material characteristics. Since the main focus of our work is on socio-economic characteristics of the neighborhood, we specified that in the text and replaced the term residential environment.
According to the relationship between socioeconomic characteristics of the neighborhood and type 2 diabetes, so far, two potential interrelated mechanisms were identified as potential pathways between neighborhood socioeconomic characteristics and the development of type 2 diabetes: the adoption and maintenance of risky health behavior and psychosocial factors such as chronic stress (Diez-Roux et al. 2002).

Reference:
Diez-Roux AV, Jacobs DR, Kiefe CI. Neighborhood characteristics and components of the insulin resistance syndrome in young adults. Diabetes Care 2002; 25(11):1976-1982.

10. Line 43 – The aim of the study should clearly state individual and neighborhood social determinants are being investigated?

Response: We specified the exposure variables investigated in our study on page 5.

11. Also, aim 3 does not seem to be adequately address in the manuscript, and I would consider moving it to a web-appendix.

Response: We revised the section on study objectives to put the focus on the main research questions, aim 1 and 2. The authors’ prefer to keep the study-specific analysis as a sub analysis in the manuscript.

Methods

Outcome

12. Undiagnosed diabetes can be a source of bias. Is there any validation of the self-report diabetes reporting in your sample? The DIAB-CORE Consortium has published many studies now on T2D. This information also could be updated in the limitations section (p. 13, lines 50-52).

Response: In the description of the variables on page 7, we clearly stated that our outcome is solely based on self-reports. This was also highlighted in the section on limitations of our study on page 13. Unfortunately, we were unable to validate the self-reported diabetes or discover undetected diabetes

cases. We specified that in the text and added a very recent publication by the authors Jackson et al.. The authors concluded that self-reported diabetes is a valid outcome for observational studies.

Reference added:
Jackson JM, DeFor TA, Crain AL, et al. Self-reported diabetes is a valid outcome in pragmatic clinical trials and observational studies. Journal of Clinical Epidemiology 2013;66(3):349-50.

13. Do you have information about diabetes diagnosed during pregnancy? How was this accounted for in your analyses?

Response: Unfortunately, we had no information on diabetes diagnosed during pregnancy and could also not account for that in your analysis. However, the authors believe that gestational diabetes only played a minor role.

Exposure

14. p. 7 Lines 54 Can you please elaborate on the social status measure (the Winkler-Index of Socioeconomic status). How were household income and education categorized? How was this index created? How was the index modified? This is important information for the reader and should be included in the manuscript. The reference provided is incorrect and points to a manuscript that used occupation (which you do not have in this study) in the development of their Winkler-Index of Socioeconomic status variable.

Response: Individual social class was coded based on the Winkler Index of Socio-Economic Status. The Winkler Index is a multidimensional, non-weighted additive index, which has been applied in a large number of studies in health research. Originally, the index summarizes information on education, income and occupation. The education dimension is solely based on the individual educational and professional attainment, whereas income is measured as net household income and occupation as the occupational position of the main earner of the household. Each dimension is transformed to an ordinal scale ranging from 1 to 7 and summed up to an index with a scale from 3 to 21 points.
Unfortunately, we could not use information on occupational status, since the assessment of occupational position was not comparable between studies. Thus, the index we built was solely based on the ordinal scales of education and income, ranging between 2 and 14 points. We described the coding of the social class variable in more detail on page 7. We kept the reference, but clarified that it is referring to the original version of the index.

15. The occupation variable should be discussed at the beginning of page 8. It is unclear why you have only included employment status and not occupational achievement. Social inequalities across level of occupation are common, and therefore it is confusing to the reader why you have chosen to classify employment in this fashion. This is not mentioned again until the limitations section.

Response: Information on occupational position could not be used for our analysis, unfortunately, because the information on occupational status was not comparable across studies. We added this information in the description of the coding of the social class variable on page 7.
Moreover, the authors believe that the employment status is as well an important social variable, since the participation in the labor market determines e.g. social networks, material and social resources but also the exposure to and the importance of the neighborhood context (e.g. daily time spend in the neighborhood).

Covariates

16. Why is a lifestyle index used instead of examining the impact of risk factors individually? This index assumes that there is equal risk of developing diabetes among each of these factors, which is likely not the case. For example, obesity has a much strong effect on risk of T2DM compared to alcohol consumption. Also, the name of the categories is inappropriate. Would the authors consider being obese healthy?

Response: The authors focus was to evaluate how the introduction of a life style measure modifies the relationship between the neighborhood unemployment rate and type 2 diabetes. We agree that the risk factors have different importance for type 2 diabetes. We repeated the analysis using the single risk factors to identify important mediators in the group of risk factors, which could potentially vary between women and men.
We performed the mediation analysis in four steps: 1) we tested if type 2 diabetes was associated with social class, employment status or neighborhood unemployment rate; 2) we tested if type 2 diabetes was associated with the risk factors; and 3) we evaluated the association between the risk factors and social class, employment status and neighborhood unemployment rate, all in separate age-adjusted gender-stratified models. In a final step, we gave the risk factors in the fully adjusted multi-level models and evaluate potential effect modifications among social class, employment status and neighborhood unemployment rate.
Type 2 diabetes was associated with social class, employment status and neighborhood unemployment rate, with one exception: employment status was not associated with type 2 diabetes in men. Type 2 diabetes was associated with an increasing BMI, no physical activity in both women and men and moreover, with smoking in men.
In women, a higher BMI was associated with belonging to middle and low social class, being retired and living in neighborhoods with a medium or high level of unemployment rate. Belonging to middle and low social class or being a resident in a neighborhood with medium or high level of unemployment was associated with an increasing odds to be physical inactive among women.
Among men, BMI was associated with social class, being retired and living in high unemployment neighborhoods. A high odds of being physical inactive was associated with middle and low social class, being unemployed or having another employment status and living in high unemployment neighborhoods. Being an ex-smoker or current smoker was more likely in middle and low social class men, in retired and unemployed men. Moreover, being a current smoker was associated with residing in a high unemployment rate neighborhood.
Based on these results, we introduced the risk factors BMI and physical activity in the models for men and women and additionally, smoking status in the models for men. In both men and women, we observed effect modification due to the introduction of the corresponding risk factors. These results were described in the results section on page 11.

17. The presentation of neighbourhood unemployment rate should be included below the individual socioeconomic status measures. A clear rationale should be provided for why neighbourhood unemployment rate was chosen and the degree to which it is an adequate measure of neighbourhood socioeconomic status. Were any other neighbourhood measures available in the dataset?

Response: We placed the description of neighborhood unemployment rate below the individual social variables, social class and employment status. In Germany, administrative data differs significantly by region and city and thus, the selection of context measures is difficult, especially when ensuring comparability across regions. According to the low city level, it is particularly difficult to provide a larger pool of context measures. For the five considered studies, we were able to collect five context measures: unemployment rate, number of migrants, married residents, residents 0-17 years old, residents over age 65 relative to the total number of residents. Their importance for the prevalence of type 2 diabetes was evaluated in our previous work, as described in point 1. Significant associations were only found between neighborhood unemployment rate and type 2 diabetes.
Neighborhood unemployment rate was applied as a measure of deprivation in a number of studies:

Dragano et al. 2007, Dragano et al. 2009, van Lenthe et al. 2005, and Cummins et al. 2005. The work by van Lenthe et al. 2005 was added to the manuscript. Campbell et al. 1991 compared unemployment rates as a single indicator of deprivation in comparison to a British deviation index and concluded that unemployment rates are a good and simple indicator for deprivation. We also added this reference to the manuscript.

References:

- Dragano N, Bobak M, Wege N, et al. Neighbourhood socioeconomic status and cardiovascular risk factors: a multilevel analysis of nine cities in the Czech Republic and Germany. BMC PUBLIC HEALTH 2007;7(1):255.
- Dragano N, Hoffmann B, Stang A, et al. Subclinical Coronary Atherosclerosis And Neighbourhood Deprivation in an Urban Region. EUR J EPIDEMIOL 2009;24(1):25-35.
- van Lenthe FJ, Borrell LN, Costa G, et al. Neighbourhood unemployment and all cause mortality: a comparison of six countries. Journal of Epidemiology and Community Health 2005;59(3):231-37.
- Cummins S, Stafford M, Macintyre S, Marmot M, Ellaway A. Neighbourhood environment and its association with self rated health: evidence from Scotland and England. J EPIDEMIOL COMMUNITY HEALTH 2005;59(3):207-13.
- Campbell DA, Radford JM, Burton P. Unemployment rates: an alternative to the Jarman index? BRIT MED J 1991;303(6805):750-55.

Van Lenthe et al. 2005 and Campbell et al. 1991 were added as references to the manuscript.

18. An improved rationale would be beneficial for the choice of equally-sized tertiles of neighbourhood unemployment rate. Is this categorization meaningful and consistent across studies/areas?

Response: Neighborhood unemployment rate was coded into equally-sized tertiles. The coding was done separately for each study. Thus, we had summarized the neighborhoods with the lowest level of unemployment rate, the neighborhoods with a medium unemployment rate level and neighborhoods with the highest level of unemployment rate for each region to get a relative measure of deprivation. We applied neighborhood unemployment rate not as a continuous measure, since we cannot assume that the relationship between neighborhood unemployment and type 2 diabetes prevalence is linear. We decided to use a small number of categories (tertiles), since the stratified analysis reduces the sample size and number of type 2 diabetes cases in the analysis. Furthermore, the categorization in tertiles allows us to look on dose-response relationships. The latter point, we added to the description in the methods section on page 7.

Statistical Analyses
19. The presentation of the analyses undertaken for the study could be improved. In the statistical models, the distinction needs to made between confounders and mediators along with a description of the authors approach. Along with the description, the tables in the text could include additional information of the appropriate statistical models. It would also be beneficial for the reader to present the results in stages: unadjusted/age-adjusted individual level variables, confounder adjusted individual variables, confounder adjusted individual level variables and area level variables and then all variables. These results should be communicated in coordination with those from Table 4 (the fully adjusted models).

Response: As mentioned above, we restructured the results section according to the type of analysis, starting with the descriptive analysis, followed by the interaction analysis and the stratified regression analysis.
According to the suggestions of the reviewer, we also restructured the section on statistical analyses. The covariates were classified in confounder and potential mediators on page 8.
We applied a stepwise model strategy. The revised table 4 shows the estimated odds ratios for the three exposure variables of main interest: social class, employment status and neighborhood

unemployment rate. In model 1, the unadjusted estimates were displayed, in model 2, the age-adjusted, in model 3, adjusted for the confounders and in the final model 4, adjusted for potential mediators. We revised the description of results according to this on page 10-11.

20. A main finding of the study is that the effect of SES on T2D is different by sex. This was confirmed by finding a statistically significant multiplicative interaction between SES and sex. Given this finding, it is not clear why some models are sex-specific and others are not. Based on your findings it seems that all models should be sex-specific. Sexpooled models should be removed from the manuscript, or a clear description of how the results are being interpreted to include the effect of the interaction term from sex and SES. For example, do the results in Figure 1 include the portion of the association between social status and T2D relationship that is modeled through the interaction term? Why are any results being presented in a sex-pooled models if there is interaction by sex with the main exposure?

Response: We agree with the reviewer on that point. The pooled analysis were done to prove gender differences in the association between individual social class/ employment status/ neighborhood unemployment rate and type 2 diabetes. Another advantage of a pooled analysis is the larger sample size, however, the gender-stratified analysis is more straightforward in its interpretation. We decided to report the results of the interaction analysis (page 10) before presenting the results of the gender-stratified analysis (since the interaction analysis is a good reason for conducting stratified analyses) but to remove the derived estimates and its interpretation as well as its graphical illustration (Figure 1) from the manuscript.

21. Please provide evidence of ethical approval for the study, and surveys used in the manuscript.

Response: The five studies were approved by the local ethics committees and informed written consent was obtained from the study participants. We added that information to the method section on page 6.
The CARLA study was approved by the Ethics Committee of the Medical Faculty of the Martin-Luther-University Halle-Wittenberg and by the State Data Privacy Commissioner of Saxony-Anhalt (Greiser et al. 2005). The HNR research protocol was approved by the local ethics committee of the University Duisburg-Essen (Wege et al. 2008). SHIP was approved by the ethics committee of the University of Greifswald and KORA by the ethics committee of the Bavarian Chamber of Physicians (Hannemann et al. 2011). The DHS was approved by the local ethics committee of the medical faculty at the University of Münster (Khil et al. 2012).

References:
– Greiser K, Kluttig A, Schumann B, et al. Cardiovascular disease, risk factors and heart rate variability in the elderly general population: Design and objectives of the CARdiovascular disease, Living and Ageing in Halle (CARLA) Study. BMC CARDIOVASC DISORD 2005;5(1):33.
– Wege N, Dragano N, Erbel R, et al. When does work stress hurt? Testing the interaction with socioeconomic position in the Heinz Nixdorf Recall Study. Journal of Epidemiology and Community Health 2008;62(4):338-41.
– Hannemann A, Meisinger C, Bidlingmaier M, et al. Association of plasma aldosterone with the metabolic syndrome in two German populations. European Journal of Endocrinology 2011;164(5):751-58.
– Khil L, Pfaffenrath V, Straube A, Evers S, Berger K. Incidence of migraine and tension-type headache in three different populations at risk within the German DMKG headache study. Cephalalgia 2012;32(4):328-36.

Results

22. Page 10, line 2-9. The description in this section is not clear and should be revised.

Response: We revised this section for more clarity on page 9.

23. Page 10, line 11. The word chance is not appropriate. It should read odds. Please fix throughout the text.

Response: According to the suggestions of the reviewer, we replaced the term “chance” by “odds” throughout the whole text.

24. In Table 3, the label “Adjusted Prevalence” should be more specific to improve clarity. Also in the title, it refers to neighbourhood socio-demographic variables. Why not simply stated neighbourhood unemployment rate?

Response: We changed “Adjusted Prevalence” to “Age-Adjusted Prevalence” to clarify that and replaced “neighbourhood socio-demographic variables” with “neighborhood unemployment rate” to specify that.

25. In Table 4 “VA” in is not clear. A more specific labeled would improve clarity.

Response: In the revised table 4, the estimates for variance are not longer displayed. The estimates were only reported in the text. The concept of area-level variation was explained in the section on statistical analysis.

26. Also, somewhere in Table 4 please include information on which confounders have been adjusted in the presented models?

Response: We added a legend to the table describing the adjustment set for each model with emphasizes on confounder and mediator variables.

27. Page 10, lines 33-48 Can be moved to web material.

Response: The authors prefer to keep this sub-analysis in the manuscript and thus, we also moved the supplemental graph from web material to the manuscript. This sub-analysis provides evidence for the homogeneity and the high degree of comparability of the five regional studies.

28. Page 11, line 7-21. The analyses for the presented findings have not been discussed and the results are not presented in text.

Response: In this section, the authors reported the results of the interaction analyses. We shortened this paragraph, took out the estimated coefficients and only reported whether or not significant multiplicative interactions were found between gender and social class or employment status or neighborhood unemployment rate, since the focus was on the gender-stratified analyses.

Discussion

29. Page 12, lines 39-58. The discussion around the importance of employment status should be deemphasized. While the explanation is not necessarily inappropriate, given the limitations around the employment measure it is difficult to assess the role of employment on the results from this study.

Response: We agree with the reviewer on that point and added a sentence to this section highlighting that potential interpretations are limited due to the fact that we were not able to consider the occupational position in our analyses (page 12). Moreover, we shortened this section to reduce its importance within the discussion.

30. p. 13, lines 3-15. Please provide a more detailed discussion around the difference between neighbourhood unemployment rate and the effect of neighbourhoods in men and women. Why do you believe that high unemployment rate is associated with increase odds of developing diabetes in men and not women, but that the residential environment is more important in women than men?

Response: We provided more details on the association between neighborhood unemployment rate and type 2 diabetes in the discussion on page 12. In the authors' opinion, the deviating effects of neighborhood unemployment rate on type 2 diabetes between men and women might be explained by the stronger engagement of men in employment and their consequential higher dependence on the regional labor market and its employment opportunities. According to the "differential exposure hypothesis", men and women are differently exposed to neighborhood unemployment rate.
Among men, the detected regional variation in the prevalence of type 2 diabetes was statistically explained by controlling the analysis for age, social class, employment status, marital status and neighborhood unemployment rate. In contrast, a large fraction of variance in the prevalence of type 2 diabetes remained statistically unexplained on the level of neighborhoods and regions in women. We reconsidered our previous conclusions. In general, it is not possible to conclude that the neighborhood context is more important for women than men. The source of variation could be as well on the individual level. Thus, we concluded that there were characteristics on the individual, neighborhood and regional level that determine the presence of type 2 diabetes in women and which were not considered in our analysis. We revised the paragraph in the discussion section on page 12-13.

31. A more complete discussion around the effects of neighbourhoods on T2D is warranted. For example, see Krishnan S, Cozier YC, Rosenberg L, Palmer JR. Socioeconomic status and incidence of type 2 diabetes: results from the Black Women's Health Study. Am J Epidemiol 2010;171(5):564e70.

Response: We added a paragraph on the potential mechanisms in the association between neighborhood unemployment and type 2 diabetes to the discussion on page 12. For this purpose, we added two more references to the manuscript. In the literature, we found a number of potential mechanisms that operate in the relationship between neighborhood socio-economic status and type 2 diabetes. That includes neighborhood resources such as the availability of grocery stores offering healthy food and recreational facilities (Auchincloss et al. 2008), the adoption and maintenance of risky health behavior but also psychosocial factors such as chronic stress (Diez-Roux et al. 2002; Krishnan et al. 2010).

References:
Diez-Roux AV, Jacobs DR, Kiefe CI. Neighborhood characteristics and components of the insulin resistance syndrome in young adults. DIABETES CARE 2002;25(11):1976-82.
Auchincloss AH, Diez Roux AV, Brown DG, Erdmann CA, Bertoni AG. Neighborhood resources for physical activity and healthy foods and their association with insulin resistance. Epidemiology 2008;19(1):146-57.
Krishnan S, Cozier YC, Rosenberg L, Palmer JR. Socioeconomic Status and Incidence of Type 2 Diabetes: Results From the Black Women's Health Study. American Journal of Epidemiology 2010;171(5):564-70.

Krishnan et al. 2010 and Auchincloss et al. 2008 were added as references to the manuscript.

32. p. 13, lines 36-40 The statement around "no prior studies…" as suggested in the last comment is not appropriate.

Response: We revised this statement on page 14.

VERSION 2 – REVIEW

REVIEWER	Brendan T. Smith, PhD Candidate Dalla Lana School of Public Health Canada I have no competing interests to declare.
REVIEW RETURNED	17-Apr-2013

GENERAL COMMENTS	The authors should be commended on the additional work on the revisions to the manuscript. It has undergone significant changes and it is my opinion that it be accepted for publication with a few minor editorial changes. Introduction 1. An overall statement(s) regarding the state of the literature on the gender affects of the association being studied. The review of the many studies leaves the reader without a solid understanding of research to date. Methods 2. When discussing neighbourhood unemployment rate, caution should be used in the description of the tertiles as high and low. The neighbourhood unemployment rates included in the lowest tertile group may vary substantially between studies. For example, KORA and HNR have substantial lower rates than CARLA. Therefore saying a neighbourhood has a low unemployment rate is a relative term compared to the study from which the neighbourhood is derived. 3. P. 8 line11-13. More detail should be provided on the question used to measure exercise frequency and duration and period of time for which exercise was assessed (e.g. in the past month, year?)? Results 4. P. 11 lines 13-21. A more clear interpretation of model 4 should be provided here. Introducing risk factors into the analysis is not testing effect modification. 5. P. 11 lines 23. A description has been provided of the between-study and between neighbourhood variation in the prevalence of T2DM. What happened to the table presenting these results (was table 4 in previous draft)? This should at minimum be presented as a web table.

VERSION 2 – AUTHOR RESPONSE

Reviewer
Introduction
1. An overall statement(s) regarding the state of the literature on the gender affects of the association being studied. The review of the many studies leaves the reader without a solid understanding of research to date.
Respond: We edited the section on page 5 to give an overall statement on the state of research.

Methods
2. When discussing neighbourhood unemployment rate, caution should be used in the description of

the tertiles as high and low. The neighbourhood unemployment rates included in the lowest tertile group may vary substantially between studies. For example, KORA and HNR have substantial lower rates than CARLA. Therefore saying a neighbourhood has a low unemployment rate is a relative term compared to the study from which the neighbourhood is derived.
Respond: We added a sentence to the methods section on page 7 explaining that low, medium and high levels of unemployment rate is spoken in relative terms and that the level of unemployment is considerably different across studies.

3. P. 8 line11-13. More detail should be provided on the question used to measure exercise frequency and duration and period of time for which exercise was assessed (e.g. in the past month, year?)?
Respond: We gave a more detailed description on the assessment of physical exercise on page 7/8.

Results
4. P. 11 lines 13-21. A more clear interpretation of model 4 should be provided here. Introducing risk factors into the analysis is not testing effect modification.
Respond: We intended to check for mediation through life style variables in the relationship between T2DM and social variables. Testing for effect modifications was not intended. We formulated that more precise on page 10-11.

5. P. 11 lines 23. A description has been provided of the between-study and between neighbourhood variation in the prevalence of T2DM. What happened to the table presenting these results (was table 4 in previous draft)? This should at minimum be presented as a web table.
Respond: We added a table displaying the estimates for between-study and between neighbourhood variations as online supplemental material.