PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Temporal changes in predicted risk of type 2 diabetes in Germany: Findings from the German Health Interview and Examination Surveys 1997-1999 and 2008-2011
AUTHORS	Paprott, Rebecca; Mensink, Gert; Schulze, Matthias; Thiele, Silke; Mühlenbruch, Kristin; Scheidt-Nave, Christa; Heidemann, Christin

VERSION 1 - REVIEW

REVIEWER	Lars Vatten Norwegian University of Science and Technology Norway
REVIEW RETURNED	15-Aug-2016

GENERAL COMMENTS	This is an interesting study, but it has one major communicative limitation that needs to be considered carefully. On page 6 - the statistical analyses are incompletely described, and will be nearly impossible to understand for most readers. For example, to reference a procedure in SAS (proc SURVEY) is by no means meaningful for most readers, and the algorithm used in the analyses, derived from GDRS (German Diabetes Risk Score) is not possible to understand, and requires some explicit explanation. Also, the last paragraph of the methods, that includes a formula of cumulative risk - the numbers entered into the formula to give the 5 year risk of diabetes, need to be explained in more detail. As it stands now, most of the statistical approaches will be very confusing to the reader, and should be totally revised. However, if the authors could make it easier for the reader to understand that differences in risk factor prevalence of diabetes over time can be translated into differences in actual diabetes risk (predicted 5 year risk, as estimated here), the paper would be a useful contribution to the literature on diabetes epidemiology.

REVIEWER	Asnawi Abdullah Faculty of Public Health, University Muhammadiyah Aceh, Indonesia
REVIEW RETURNED	28-Nov-2016

GENERAL COMMENTS	The paper presented 5-years risk of type-2 diabetes changes between two cross-sectional surveys by examining the prevalence changes in individual diabetes risk factors. The 5-years type-2 diabetes risk factor was predicted using the formula that has been developed by Muchlenbruch at al., (2014). In the formula, each variable was multiplied by a certain point as has been developed by Schlze et al. (2007). For example, if someone currently smoking with less than 20 cig./day was assigned with the point of 23, or currently

	smoking with more than 20 cig./day was assigned with the point of 77. For some readers, it is interested to know how this number was determined. This is particularly important as the original paper of Schlze et al. (2007) was not in English. It is suggested to elaborate more in this paper. It is also important to describe in method in more detail that for some variables such as smoking status (former and current smoking), family history of diabetes, which are a mutual exclusive variable. The current formula presented in the paper as look as it is not a mutual exclusive variable. It is suggested to rewrite the formula that able to demonstrate which variables mutual and not mutual exclusive. In addition, the readers also interested to see in the paper how score point was developed. In Table 1, why missing GDRS points for family history. This variable has a higher point compared with other variables. Therefore, this variable should not be missing as it affect the total point. In discussion, the readers what to read what are driving factors for declining the point of GDRS. It should be described in early part of discussion section.

REVIEWER	Diego Serraino Unit of Canccer Epidemiology - National Cancer Istitute - Centro di Riferimento Oncologico di Aviano IRCCS; Italy
REVIEW RETURNED	22-Dec-2016

GENERAL COMMENTS	The pairs of columns of the histogram (Fig.1) would be appropriate to separate them for better readability. In my opinion the authors could describe in more detail the concept of selection bias related to the failure history of diabetes (own, siblings and parents) ande the calculation of constants used as weight factors

VERSION 1 – AUTHOR RESPONSE

Reply to Reviewer 1 (Lars Vatten)

1. This is an interesting study, but it has one major communicative limitation that needs to be considered carefully. On page 6 - the statistical analyses are incompletely described, and will be nearly impossible to understand for most readers. For example, to reference a procedure in SAS (proc SURVEY) is by no means meaningful for most readers, and the algorithm used in the analyses, derived from GDRS (German Diabetes Risk Score) is not possible to understand, and requires some explicit explanation.

Response: We rephrased and extended the paragraph on statistical analyses (p. 7). Moreover, we now explain in more detail how the GDRS was derived in EPIC-Potsdam by a separate paragraph ‘Calculation of GDRS points and predicted 5-year type 2 diabetes risk’ (p. 6f).

2. Also, the last paragraph of the methods, that includes a formula of cumulative risk - the numbers entered into the formula to give the 5 year risk of diabetes, need to be explained in more detail. As it stands now, most of the statistical approaches will be very confusing to the reader, and should be

totally revised.

Response: We also rephrased and extended the respective paragraph, which is now summarized under the subheading ‘Calculation of GDRS points and predicted 5-year type 2 diabetes risk’ (p. 6f), to explain how the respective formula was derived in EPIC-Potsdam.

3. However, if the authors could make it easier for the reader to understand that differences in risk factor prevalence of diabetes over time can be translated into differences in actual diabetes risk (predicted 5 year risk, as estimated here), the paper would be a useful contribution to the literature on diabetes epidemiology.

Response: The concept that differences in individual diabetes risk factors may be transferred into a difference in actual diabetes risk is explained in the introduction and discussion section. Moreover, we aimed to improve the paper’s comprehensibility by introducing an extra paragraph on ‘Calculation of GDRS points and predicted 5-year type 2 diabetes risk’ in the method section (p.6f).

Reply to Reviewer 2 (Asnawi Abdullah)

The paper presented 5-years risk of type-2 diabetes changes between two cross-sectional surveys by examining the prevalence changes in individual diabetes risk factors. The 5-years type-2 diabetes risk factor was predicted using the formula that has been developed by Muchlenbruch at al., (2014). In the formula, each variable was multiplied by a certain point as has been developed by Schulze et al. (2007).
1. For example, if someone currently smoking with less than 20 cig./day was assigned with the point of 23, or currently smoking with more than 20 cig./day was assigned with the point of 77. For some readers, it is interested to know how this number was determined. This is particularly important as the original paper of Schulze et al. (2007) was not in English. It is suggested to elaborate more in this paper.

Response: We now provide more information on how the individual points from the score equation were determined in the developmental study in EPIC-Potsdam (p. 6f, paragraph ‘Calculation of GDRS points and predicted 5-year type 2 diabetes risk’). Regarding the original paper of Schulze et al. (2007) (Ref. 5) this is probably a misunderstanding since this paper has been originally published in English. We did not cite other papers from Schulze et al. published in German. However, the reviewer is right about the paper by Mühlenbruch et al. (2014, Ernährungsumschau) which was originally published in German and English. We referred already to the English version. However, we now also provide the link to the electronically available English pdf of the paper (reference 7).

2. It is also important to describe in method in more detail that for some variables such as smoking status (former and current smoking), family history of diabetes, which are a mutual exclusive variable. The current formula presented in the paper as look as it is not a mutual exclusive variable. It is suggested to rewrite the formula that able to demonstrate which variables mutual and not mutual exclusive.

Response: We slightly revised the formula to indicate more clearly which variables are mutual exclusive (p. 6, 2nd paragraph). For example, for parental history of diabetes it is now differentiated between ‘only one parent with diabetes’ and ‘both parents with diabetes’.

3. In addition, the readers also interested to see in the paper how score point was developed.
Response: We rephrased and extended the last paragraph on the calculation of predicted 5-year diabetes risk (p. 6, last paragraph) to explain in more detail how the respective formula was derived in EPIC-Potsdam.

4. In Table 1, why missing GDRS points for family history. This variable has a higher point compared with other variables. Therefore, this variable should not be missing as it affect the total point.

Response: We now provide the GDRS points for family history of diabetes in Table 1.

5. In discussion, the readers what to read what are driving factors for declining the point of GDRS. It should be described in early part of discussion section.

Response: We shifted and adapted the paragraph about the driving factors for the decline in GDRS points (p. 8, 2nd paragraph of 'Discussion').

Reply to Reviewer 3 (Diego Serraino)

1. The pairs of columns of the histogram (Fig.1) would be appropriate to separate them for better readability.

Response: We separated the pairs of columns of Fig. 1 to improve its readability.

2. In my opinion the authors could describe in more detail the concept of selection bias related to the failure history of diabetes (own, siblings and parents) ande the calculation of constants used as weight factors

Response: In DEGS1, missing values for parental history of diabetes were replaced by a constant representing the prevalence of a parental history of diabetes in DEGS1 participants with available information. This approach can be regarded as a type of 'imputation using the overall sample mean' which has been shown to lead to biased results when examining associations (Donders 2006). In GNHIES98, parental history of diabetes was not assessed at all, i.e. information was missing for all study subjects. To still be able to calculate the GDRS, missing values were again replaced by a constant representing the prevalence of a parental history of diabetes in DEGS1. Furthermore, history of diabetes in siblings was neither assessed in GNHIES98 nor in DEGS1. This information was replaced by a constant reflecting the prevalence of a sibling history in EPIC-Potsdam.

Since we are aware about the shortcomings of the applied method, we did not calculate p-values to assess temporal changes in parental history of diabetes and labeled the respective values as 'constant' (please compare Table 1). However, the main finding of the study would only be affected if the prevalence of a family history of diabetes in GNHIES98 was much lower than the prevalence in DEGS1. On p. 10 first sentence we speculate that 'from a genetic point of view we would not expect a considerable change in this component during the relatively short period of about 12 years'. However, the next survey wave is currently planned and by including questions on family history of diabetes will provide detailed information on this issue.

References

Donders, A.R.T., van der Heijden G.J.M.G., Stijnen, T., Moons, K.G.M.: Review: A gentle introduction to imputation of missing values. J Clin Epidemiol. 2006;59(10):1087-91.

VERSION 2 – REVIEW

REVIEWER	Lars Vatten Norwegian University of Science and Technology, Trondheim, Norway
REVIEW RETURNED	03-Feb-2017

GENERAL COMMENTS	I have no further comments to this paper

REVIEWER	Asnawi Abdullah University Muhammadiyah Aceh Indonesia
REVIEW RETURNED	19-Feb-2017

GENERAL COMMENTS	In general, the revised version is much improve. However, some aspects are still need more work. The introduction still poor; beyond the standard (too short). It is required some statistical data to support your argument (i.e. prevalence of changes). The author also argued this paper could help policy makers; it is also required some explanation how it works, what mechanism? Some part of method is still need more description as has been suggested previously. It is also important to describe in detail formula. In more importantly, how the scores or the points in the formula was derived? In the previous version, there was also a table to demonstrate the score; why in this version, the table has been excluded? Results, in the first paragraph, line 44 to 48, the description is not clear enough, it is hard to follow. Suggest rewriting. In Table, particularly variable of coffee consumption, the prevalence was increase; from 2.6 to 3.3; why it is has positive effect for GDRS score? This is interesting as no current meta-analysis study showed coffee significantly positive impact on health.

VERSION 2 – AUTHOR RESPONSE

Reply to Reviewer 2 (Asnawi Abdullah)

Dear Authors, In general, the revised version is much improve. However, some aspects are still need more work.
1. The introduction still poor; beyond the standard (too short). It is required some statistical data to support your argument (i.e. prevalence of changes). The author also argued this paper could help policy makers; it is also required some explanation how it works, what mechanism?

Response: We extended the introduction by information on the pathophysiology and consequences of diabetes, by prevalence changes in risk factors recently observed in Germany (including respective references for those readers interested in detailed prevalence data), and by an explanation on how the monitoring of a summary measure to estimate future diabetes risk could serve as guidance for health policy makers.

2. Some part of method is still need more description as has been suggested previously.

Response: We think that we carefully incorporated all previous suggestions raised by the reviewers regarding the revision of the method section. That was inclusion of more detailed information on how the individual points from the score equation were determined in the developmental study (i.e. EPIC-Potsdam) and on how the formula for predicted 5-year diabetes risk was derived (p. 6f, paragraph 'Calculation of GDRS points and predicted 5-year type 2 diabetes risk'). Besides, we revised the representation of the formula of the GDRS to indicate more clearly which variables are mutual exclusive (p. 7, 1st paragraph). Since this study solely focuses on the application of a previously developed and published GDRS, we think that herewith the method section comprises all necessary

information for the reader (including all respective references on the method of GDRS derivation in a previous study). However, if the reviewer thinks that essential information is still lacking, please state more precisely where changes are required.

3. It is also important to describe in detail formula. In more importantly, how the scores or the points in the formula was derived?

Response: According to the previous comment of the reviewer (from the first revision), we described the formula in more detail and we provided information on how the formula was derived (p. 6f, paragraph 'Calculation of GDRS points and predicted 5-year type 2 diabetes risk'). In addition, we state the respective references in case the reader needs more detailed information.

4. In the previous version, there was also a table to demonstrate the score; why in this version, the table has been excluded?

Response: We included all tables and also the figure from the submitted version in the revised version. However, we noticed that the figure was not additionally included in the pdf version of the revised manuscript showing the tracked changes and therefore might have been overlooked by the reviewer in the first place. For this reason, we now include the figure in the revised manuscript's pdf version.

5. Results, in the first paragraph, line 44 to 48, the description is not clear enough, it is hard to follow. Suggest rewriting.

Response: We rewrote the first paragraph of the results section.

6. In Table, particularly variable of coffee consumption, the prevalence was increase; from 2.6 to 3.3; why it is has positive effect for GDRS score? This is interesting as no current meta-analysis study showed coffee significantly positive impact on health.

Response: In a recent meta-analysis by Jiang et al. it has been shown that the highest level of coffee intake was associated with a reduced risk for type 2 diabetes (Jiang 2014). Therefore, it is in line with the evidence that in the GDRS coffee consumption has a positive effect on diabetes risk.

References:
Jiang X., Zhang D., Jiang W.: Coffee and caffeine intake and incidence of type 2 diabetes mellitus: a meta-analysis of prospective studies. Eur J Nutr. 2014;53(1):25-38.

VERSION 3 – REVIEW

REVIEWER	Asnawi Abdullah University Muhammadiyah Aceh, Indonesia
REVIEW RETURNED	27-Mar-2017

GENERAL COMMENTS	No further comments, the paper now is much better and easy to follow.