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)	Adiposity, Physical activity, and risk of Diabetes Mellitus: Prospective Data from the population-based HUNT Study, Norway
AUTHORS	Hjerkind, Kirsti; Stenehjem, Jo; Nilsen, Tom

VERSION 1 - REVIEW

REVIEWER	Dexter Canoy University of Oxford UK
REVIEW RETURNED	29-Jul-2016

GENERAL COMMENTS	Summary: This study investigates the combined influence of physical activity and adiposity in relation to diabetes risk. Using data from HUNT1 and HUNT2 cohort studies, shows the independent associations of various indicators of physical activity level and body mass index with subsequent risk of diabetes. Although the highest risk was found in most sedentary and obese groups, the most active groups who were obese remained to have elevated diabetes risk. The authors show that obesity-related diabetes risk was partly reduced by increased physical activity level particularly in men. This study addresses an important conundrum in diabetes prevention, which involves the relative importance of important risk factors of diabetes, in particular, obesity and sedentary lifestyle. With a long follow-up and relatively a large number of incident cases of diabetes, this prospective study is in a good position to provide evidence to this important research question. Although the manuscript is generally well-written, a number of key issues, which I detail below, have to be addressed to improve clarity and certainty in the findings of this investigation. Comments: In determining the relevant study sample, it is not clear how the 45,000 participants were derived from the 75,000 HUNT1 participants and 66,000 HUNT2 participants. Are HUNT1 and HUNT2 studies two separate and independent cohorts? Or are there overlapping participants in both studies? Or is HUNT2 a follow-up of HUNT1 study participants. The diabetes outcome was defined based on various parameters, including the use of blood-based measurement of glucose level for persons aged 40+. It would be useful to see (and explore) how many cases were found for age >=40 years (ie, excluding cases <40 years) and whether it would impact on their findings if the analysis is limited among participants aged >=40 so that all outcomes are similarly defined and ascertained.

	The authors created a summary score to determine the overall level of physical activity. On what basis was this equation developed? Was there a validation studyy from which the equation was based on? Would changing the weighting for the various components of physical activity in the equation alter the main findings of this study? How was age controlled in the analysis? Age was only mentioned in the interaction between physical activity and age on diabetes outcome, but it was never adjusted in any of the analysis. Although a log-binomial model allows conservative estimation of RRs, I am unfamiliar with the underpinnings of this statistical technique. I assume that this is not a time-to-event analysis since the outcome was assessed at a follow-up questionnaire, and presumably, the follow-up time is similar for most participants. However, is the follow-up time really comparable? (I'm actually unclear when the follow-up questionnaire was sent to obtain information on diabetes outcome.) If the date varies by, say, date when the survey forms were received by the investigators, I don't see why disease rate in person-years could not be calculated. In any case, my concern is that, with diabetes, one would presumably have disease events occurring between study baseline and follow-up survey (e.g. due to death or too unwell to participate in future studies). How did the authors account for this? Is there a way to ensure that the authors are not missing out on participants who develop diabetes but may have been censored prior to the follow-up survey? It would be helpful to clarify these issues. I am somewhat concerned by the number of categories of physical activity summary score (and to an extent, body mass index categories in table 2). Although the overall analytical approach is appropriate, many of the confidence intervals are wide and overlapping, and therefore, findings are not really as convincing as they appear to be (based on the risk estimates provided). My suggestion is for the authors to consider trimming down the number of categories to provide slightly more robust risk estimates for each category. Here are a number of suggestions about the analysis: 1. Table 2: I suggest that both body mass index and the summary score physical activity risk estimates are shown together in this table (and leave Table 3 as a table for the other remaining physical activity indicators). It would also be useful to show two types of adjusted risk estimates where relative risk with and without adjustment for the other factor (ie, body mass index or physical activity summary score) are shown. The authors might consider excluding or combining underweight with normal weight (or simply label categories as specific body mass index cut-offs to avoid categorising participants into definitions based on 'clinical' groups). Further, it is not clear from table 2 how age is adjusted for. 2. Tables 4 and 5: The numbers of events for many cells are really small, and I don't see that this current approach is ideal. For the physical activity summary score, I suggest that the authors reduce this to three categories (similar to the number of categories for body mass index), perhaps by combining <1 week and low levels, as well as medium and high levels. Have the authors calculated any statistical interaction between physical activity and body mass index categories on diabetes risk? As queried above, please adjust for age in the analyses.

	The authors suggest that the effect of physical activity may be able to compensate for the effects of obesity on diabetes risk particularly in men. But the data shown are not really strong enough to support this conclusion. For example, in Table 4, the difference in the risk among obese men between high and no activity levels is about 6%; yet the difference in risk associated with obesity is about 17% in no activity group, and 14% in high activity group (it is similarly high for obesity in women). The impact of physical activity, seems to me, is relatively small. Obviously, the categories of physical activity and body mass index are probably not comparable (ie, it is actually difficult to quantify the equivalent unit for body mass index and physical activity level to make appropriate comparison, so this issue should be considered when interpreting risks based on effect sizes associated with each factor). Moreover, if the number of categories is reduced for physical activity score, a different picture in risk patterns may emerge. Rather than explaining biological plausibility in the discussion section, I wonder if the authors would rather describe the impact of measurement error in their analysis. Although part of this issue is discussed already, the issue I'm referring to is in relation to making comparisons with body mass index which, although an imperfect measure of adiposity, is a good, objectively measured phenotype. I guess the issue lies on whether the errors associated with using physical activity indicators would underestimate the risk estimates for these indicators than for measures of adiposity.

REVIEWER	Joshua Joseph Johns Hopkins University School of Medicine USA
REVIEW RETURNED	13-Aug-2016

GENERAL COMMENTS	Adiposity, physical activity, and risk of diabetes mellitus: the HUNT Study Overall, interesting and important concept to assess the interrelationship of adiposity, physical activity and risk of diabetes mellitus. There are many existing studies examining the association of BMI with incident diabetes and physical activity with incident diabetes, but there is less literature on the interrelationship of BMI and physical activity on diabetes risk. In general, the study is well conceived, clearly described and the manuscript well written. Major Comments: In Table 2, the authors show that a higher BMI (overweight and obesity as compared to normal) is a significant risk factor for incident diabetes in men and women. In Table 3, the authors show that high frequency >= 4x per week, > 60 minutes per exercise, medium to high intensity and High summary score were associated with lower diabetes risk in men. In women there were less graded but significant trends with medium to high intensity physical activity being associated with the lowest diabetes risk.

	The results from Table 2 & Table 3 are consistent with prior literature, but do not add to the current literature. In Table 4, among men there are potentially qualitative differences showing increasing risk with lower levels of physical activity among normal weight, overweight and obese men. It is not clear from the analysis completed if these represent true quantitative differences given the overlapping confidence intervals. For instance, in the Obese category the High, Medium, Low, < 1/week and no physical activity relative risks were 14.3 (5.8, 35.5), 19.2 (8.5, 43.4), 13.6 (5.9, 31.5), 17.8 (8.1, 39.1) and 20.4 (8.9, 46.5). These results show the overlapping confidence intervals and the results do not show a qualitative or quantitative graded increase in diabetes risk. It is possible that the self-reported physical activity, potential underreporting of incident diabetes cases due to reliance on self-reported diabetes and using solely BMI as opposed to other measures of visceral adiposity including waist circumference is confounding the true association. In Table 5, among women it appears physical activity has very little impact on the risk of diabetes in the categories of BMI. For instance, in the Obese category the High, Medium, Low, < 1/week and no physical activity relative risks were 11.3 (5.6-22.9), 10.2 (5.3-19.5), 9.5 (4.9, 18.5), 10.3 (5.4, 19.6) and 13.8 (7.2, 26.3). The authors conclusion that, "Physical activity may to some extent compensate for the adverse effect of adiposity on the risk of diabetes, and seemed to be more evident for men than for women" is thus not fully supported by the data in Tables 4 and 5. Minor Comments: - The authors used a summary score of frequency, duration and intensity. It is unclear how the equation of this frequency score was derived and if this score has been validated. - Analysis were gender specific, did the authors perform a test of interaction prior to stratifying by sex. - The reference given for the validation of the HUNT physical activity questionnaire (Ref 32) notes that, "The HUNT 2 question for "hard" LTPA has acceptable repeatability and appears to be a reasonably valid measure of vigorous activity, as reflected in moderate correlations with several other measures including VO2max, and with corresponding results from IPAQ and ActiReg." In relation to this analysis, the physical activity questionnaire may not be adequately capturing light and moderate leisure time physical activity.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

Comment 1:
In determining the relevant study sample, it is not clear how the 45,000 participants were derived from the 75,000 HUNT1 participants and 66,000 HUNT2 participants. Are HUNT1 and HUNT2 studies two separate and independent cohorts? Or are there overlapping participants in both studies? Or is HUNT2 a follow-up of HUNT1 study participants.

Response 1:
HUNT1 and HUNT2 are two separate cross-sectional cohorts, which both invited all inhabitants of the

'Nord-Trøndelag' county in Norway to participate. There were 45 925 individuals participating in both cohorts, and those were included in the study sample. We have clarified this in the text (manuscript tracked changes page 5).

Comment 2:
The diabetes outcome was defined based on various parameters, including the use of blood-based measurement of glucose level for persons aged 40+. It would be useful to see (and explore) how many cases were found for age >=40 years (i.e., excluding cases <40 years) and whether it would impact on their findings if the analysis is limited among participants aged >=40 so that all outcomes are similarly defined and ascertained.

Response 2:
Diabetes as an outcome variable at HUNT2 was defined as a positive response to the question: "Do you have or have you had diabetes?". No information on blood glucose at follow-up was used to define diabetes. However, diabetes at baseline (HUNT1) was defined by a positive response to the question: "Do you have or have you had diabetes?", and in addition some subjects were excluded due undiagnosed diabetes identified through blood glucose measurements. Thus, follow-up data on diabetes is not differential across age. This has been further clarified in the revised manuscript

Comment 3:
The authors created a summary score to determine the overall level of physical activity. On what basis was this equation developed? Was there a validation study from which the equation was based on? Would changing the weighting for the various components of physical activity in the equation alter the main findings of this study?

Response 3:
This equation was developed on the basis that frequency, duration and intensity should be equally weighted, and the categorization was based on sex-specific medians. The equation has been used in previous studies (citations have been inserted in the revised manuscript) based on HUNT data, however no validation study has been executed. Although changing the weighting for the various components could possibly alter the findings, the results in Table 3 imply that associations between the different physical activity components and diabetes risk are of fairly similar magnitude. Thus, we have chosen to keep the equation and classification in the revised manuscript, and rather commented on possible limitations regarding this in the Discussion section. We hope this is satisfactory.

Comment 4:
How was age controlled in the analysis? Age was only mentioned in the interaction between physical activity and age on diabetes outcome, but it was never adjusted in any of the analysis.

Response 4:
This is unfortunately a typographic error, and age was controlled for in all analysis. We have rectified this in the revised table legends.

Comment 5:
Although a log-binomial model allows conservative estimation of RRs, I am unfamiliar with the underpinnings of this statistical technique. I assume that this is not a time-to-event analysis since the outcome was assessed at a follow-up questionnaire, and presumably, the follow-up time is similar for most participants. However, is the follow-up time really comparable? (I'm actually unclear when the follow-up questionnaire was sent to obtain information on diabetes outcome.) If the date varies by, say, date when the survey forms were received by the investigators, I don't see why disease rate in person-years could not be calculated. In any case, my concern is that, with diabetes, one would presumably have disease events occurring between study baseline and follow-up survey (e.g. due to

death or too unwell to participate in future studies). How did the authors account for this? Is there a way to ensure that the authors are not missing out on participants who develop diabetes but may have been censored prior to the follow-up survey? It would be helpful to clarify these issues.

Response 5:
We appreciate this comment, and have rerun all analyses using a Poisson regression model with robust standard errors (instead of the log-binomial model used in the previous version of the manuscript). As for the log-binomial model, this modified Poisson model estimate risk ratios and has been suggested as a reliable approach for prospective data without person time information (1).

The outcome data (diabetes status) was collected at participation in HUNT2 for all individuals, hence disease rate based on person time could not be calculated. . We are aware that some people were "lost to follow-up" between HUNT1 and HUNT2 due to death, migration, or non-participation. However, for this to bias our estimates of relative risk, loss to follow-up must have been differential between exposure categories. It is conceivable that lean and physically active people to a greater extent were alive and willing to participate in HUNT2 than those who were obese and inactive. This could have underestimated the effects of adiposity and physical activity in our data. We have expanded on this limitation in the Discussion section of the revised manuscript.

Comment 6:
I am somewhat concerned by the number of categories of physical activity summary score (and to an extent, body mass index categories in table 2). Although the overall analytical approach is appropriate, many of the confidence intervals are wide and overlapping, and therefore, findings are not really as convincing as they appear to be (based on the risk estimates provided). My suggestion is for the authors to consider trimming down the number of categories to provide slightly more robust risk estimates for each category. Here are a number of suggestions about the analysis:

1. Table 2: I suggest that both body mass index and the summary score physical activity risk estimates are shown together in this table (and leave Table 3 as a table for the other remaining physical activity indicators). It would also be useful to show two types of adjusted risk estimates where relative risk with and without adjustment for the other factor (ie, body mass index or physical activity summary score) are shown. The authors might consider excluding or combining underweight with normal weight (or simply label categories as specific body mass index cut-offs to avoid categorising participants into definitions based on 'clinical' groups). Further, it is not clear from table 2 how age is adjusted for.

2. Tables 4 and 5: The numbers of events for many cells are really small, and I don't see that this current approach is ideal. For the physical activity summary score, I suggest that the authors reduce this to three categories (similar to the number of categories for body mass index), perhaps by combining <1 week and low levels, as well as medium and high levels. Have the authors calculated any statistical interaction between physical activity and body mass index categories on diabetes risk? As queried above, please adjust for age in the analyses.

Response 6:
We appreciate the reviewer's suggestions and have changed the analyses and presentation of the results accordingly. Both the physical activity summary score and the body mass index variables were trimmed down to three categories; resulting in more robust and precise risk estimates.

1. We included both body mass index and physical activity summary score in Table 2, and kept Table 3 for the remaining physical activity indicators. We also showed risk estimates with and without adjustment for the other factor (physical activity summary score or body mass index) in Table 2. We have adjusted for age in all analyses, and indicated that we did so in the footnotes of the tables.

2. We reduced the physical activity summary score to include only three categories, low, medium and high activity. We calculated statistical interaction between physical activity and body mass index on diabetes risk and included estimates in Tables 4 and 5. We adjusted for age in all analyses, and indicated that we did so in the footnotes of the tables.

The revised statistical approach is described in the 'Methods' section.

Comment 7:
The authors suggest that the effect of physical activity may be able to compensate for the effects of obesity on diabetes risk particularly in men. But the data shown are not really strong enough to support this conclusion. For example, in Table 4, the difference in the risk among obese men between high and no activity levels is about 6%; yet the difference in risk associated with obesity is about 17% in no activity group, and 14% in high activity group (it is similarly high for obesity in women). The impact of physical activity, seems to me, is relatively small. Obviously, the categories of physical activity and body mass index are probably not comparable (ie, it is actually difficult to quantify the equivalent unit for body mass index and physical activity level to make appropriate comparison, so this issue should be considered when interpreting risks based on effect sizes associated with each factor). Moreover, if the number of categories is reduced for physical activity score, a different picture in risk patterns may emerge.

Response 7:
We agree that the estimated relative risks associated with the different exposure variables (physical activity and body mass index) are not directly comparable since the exposures are measured on a completely different scale. We have avoided such comparisons in the revised version of the manuscript. Moreover, there was no strong evidence of statistical interaction (departure from multiplicative effects) between physical activity and body mass index, and although physical activity seems to contribute to a reduction in risk of diabetes at all body mass index levels, we agree that the modifying effect is not strong enough to state that physical activity cancel outs the effect of obesity. We have modified the conclusion of the revised manuscript to reflect this.

Comment 8:
Rather than explaining biological plausibility in the discussion section, I wonder if the authors would rather describe the impact of measurement error in their analysis. Although part of this issue is discussed already, the issue I'm referring to is in relation to making comparisons with body mass index which, although an imperfect measure of adiposity, is a good, objectively measured phenotype. I guess the issue lies on whether the errors associated with using physical activity indicators would underestimate the risk estimates for these indicators than for measures of adiposity.

Response 8:
In the 'Strengths and limitations' section we have discussed measurement error of both body mass index and physical activity. We mentioned that questionnaire based information can overestimate the general physical activity level, while body mass index might both overestimate (in short and heavy individuals with lean body mass) or underestimate (in taller individuals with less lean body mass) adiposity.

Measurement error (overestimation) of physical activity level could result in underestimation of the effect of physical activity on diabetes risk. However, misclassification of body mass index could also be related to physical activity level in the way that 'athletic people with high muscle mass, misclassified as being overweight or obese, would probably report having high levels of physical activity'. The effect of physical activity on overweight and obese individuals could therefore be overestimated. We have discussed these issues in 'Strength and limitations'.

Reviewer 2:

Comment 1:
The authors conclusion that, “Physical activity may to some extent compensate for the adverse effect of adiposity on the risk of diabetes, and seemed to be more evident for men than for women” is thus not fully supported by the data in Tables 4 and 5.

Response 1:
We agree, and have modified our conclusion. See answer 8 to reviewer 1 above.

Comment 2:
The authors used a summary score of frequency, duration and intensity. It is unclear how the equation of this frequency score was derived and if this score has been validated.

Response 2:
See response 3 to Reviewer 1.

Comment 3:
Analysis were gender specific, did the authors perform a test of interaction prior to stratifying by sex.

Response 3:
We did not perform a test of interaction prior to stratifying by sex. Since we had sufficient data to conduct stratified analyses without losing much statistical power we decided that this would give more detailed information, irrespective of possible modifying effects of sex.

Comment 4:
The reference given for the validation of the HUNT physical activity questionnaire (Ref 32) notes that, “The HUNT 2 question for "hard" LTPA has acceptable repeatability and appears to be a reasonably valid measure of vigorous activity, as reflected in moderate correlations with several other measures including VO2max, and with corresponding results from IPAQ and ActiReg.” In relation to this analysis, the physical activity questionnaire may not be adequately capturing light and moderate leisure time physical activity.

Response 4:
Other validation studies (Ref 34 in the revised manuscript) have shown that questionnaires may be useful in classifying people into broad categories of physical activity. Still, questionnaire based information has been known to overestimate the general physical activity level, and this could result in an underestimation of the effect of physical activity on diabetes risk (Ref 36 in the revised manuscript). We have discussed the issue of misclassification of physical activity in the ‘Strengths and limitations’ section.

Reference
1. Zou G. A modified poisson regression approach to prospective studies with binary data. American journal of epidemiology. 2004 Apr 1;159(7):702-6. PubMed PMID: 15033648. Epub 2004/03/23. eng.

VERSION 2 – REVIEW

REVIEWER	Dexter Canoy University of Oxford, UK
REVIEW RETURNED	03-Oct-2016

GENERAL COMMENTS	I have no further comments