PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Cohort profile: The Obesity and Disease Development Sweden (ODDS) study, a pooled cohort
AUTHORS	Stocks, Tanja; da Silva, Marisa; Fritz, Josef; Mboya, Innocent; Sun, Ming; Wahlström, Jens; van Guelpen, Bethany; Michaëlsson, Karl; Magnusson, Patrik; Melander, Olle; Sandin, Sven; Yin, Weiyao; Trolle Lagerros, Ylva; Nwaru, Bright; Leppert, Jerzy; Chabok, Abbas; Pedersen, Nancy; Elmståhl, Sölve; Isaksson, Karolin; Ingvar, Christian; Hedman, Linnea; Backman, Helena; Häggström, Christel

VERSION 1 – REVIEW

REVIEWER	Tutunchi, Helda Tehran University of Medical Sciences
REVIEW RETURNED	21-Feb-2024

GENERAL COMMENTS	This pooled cohort is so important, and the results support the authors' conclusions. It is well-written and structured. I accept this manuscript for publication.

REVIEWER	Oshakbayev, Kuat University medical center
REVIEW RETURNED	28-Feb-2024

GENERAL COMMENTS	Marisa da Silva et al. designed to create a large cohort to study body mass index (BMI), waist circumference (WC), and changes in weight and WC, in relation to morbidity and mortality within the Obesity and Disease Development Sweden (ODDS) study. The authors concluded that the large sample size and long follow-up of the ODDS study will provide robust results on anthropometric measures in relation to risk of common diseases and causes of deaths, and novel findings in subgroups and rarer outcomes. Several issues should be addressed before the paper is suitable for publication. Though the topic is relevant, some flaws need to be acknowledged: 1. There is no conclusion in the paper, neither in Abstract nor in the text. 2. The aim of the study is formulated vaguely. The aim should be clear and understandable. 3. The aim in Abstract and the text are not the same. 4. The authors should rewrite the paper using the IMRAD principle. For instance, it was not clear where Method, Results, or Conclusion sections began or distinction between them. 5. Inclusion and exclusion criteria for selecting study cohorts from various databases should be explicitly stated in the methods section. 6. The variables being compared must be homogeneous. For example, the years covered in ODDS were different, so it is not

reliable to compare between each other, in all probability. The study cohorts and study population have large heterogeneity, and some important information is missing from the tables (see below).

7. Page 8, lines 39-43: how was it practically fulfilled to avoid the Cohort overlap?

8. Some info in the paper is redundant, in my opinion. For instance, info in lines 15-52, page 9.

9. P.6, lines 39-44 and P.7, lines 53-60 are controversial between each other, because in the first the authors write "…Cohort overlap,… is not considered in the table…", but in the second they write "…includes overlap with…" in the table.

10. Tabl. 2: difficult to understand the first time due to it has very much info (Swedish register, Description, Type of information retrieved, Years covered in ODDS, Coverage of target population). But, it provides little information on all-cause and cause-specific mortality across different pathologies.

11. Of course, the authors showed a lot of interesting statistical information, but I expected to know scientific and analytic info related to an influence of the risk factors (overweight, smoking) on morbidity/mortality depending on age, region, gender, etc.

12. The cohort data (with information on weight, height, and WC more between 1963 and 2020) is a very good perspective to retrieve a correlation/regression between morbidity/mortality and weight, BMI/WC, etc., or impact of overweight in development of morbidity/mortality.

13. Also there was not clear about smoking habit as risk factor for the prevalence of diseases, morbidity/mortality, or the relation to BMI/WC.

14. Tabl. 3: line 19: there is no info related to Age (years).

15. In Tabl. 3 presented a lot of info, but there is no scientific analysis. For example, there are presented data "Objectively measured", "Self-reported current weight", "Self-reported recalled weight", but how to analyze it?

16. In Tabl. 3, another example: there was very interesting dynamic of smoking frequency during >20 years, but there is no analysis about it.

17. In Tabl. 3, other example: dependence of weight on marital status, or vice versa.

18. In Tabl. 3, other example: dependence of weight/WC on urban or rural residents, or vice versa.

19. Page 12, lines 29-41: the info should present in a table.

20. Fig. 1: I do not exactly understand the purpose of the figure, for what for practice or science?

REVIEWER	Herrán, Oscar
	Universidad Industrial de Santander, Escuela de Nutrición y Dietética
REVIEW RETURNED	12-Mar-2024

GENERAL COMMENTS	The text is well written. However, it is dense, and a map is needed to consolidate the results, a "graphical summary." This text is not research in itself; it is the preparation of a database that will be used for very high-level research due to the characteristics of the database. The richness of the data and the procedures should be known since the cohort assembly is the most significant work when establishing the relationships between morbidity and mortality.

REVIEWER	Mehlig, Kirsten
	University of Gothenburg

REVIEW RETURNED	19-Mar-2024

GENERAL COMMENTS	I am grateful for the opportunity to review this cohort profile. as the manuscript describes proposed research the questions above do not entirely apply. please find specific comments below.
	Summary: This cohort profile describes the combination of different smaller cohort studies to a nationwide cohort of adults living in Sweden with aim to investigate the risk for disease related to body weight and waist circumference and their longitudinal change. The large number of observations and the long register follow-up are a strength but the lack of important confounders such as smoking status is a limitation.
	Overall, the approach looks promising but the cohort profile lacks some important information.
	Questions of major importance:
	1. Novelty: which are the research questions that this cohort study aims to answer that have not been examined before?
	Overweight/obesity or central obesity measured at a single time point are now well established risk factors for many outcomes including different CVD and cancer endpoints as well as total mortality, and it is questionable whether this study could add new insight. The authors claim that change in weight status was less well investigated (p.5 lines 27-30) but there are a lot more studies than cited here (refs 5+7), even regarding change in waist circumference that is less often measured than body weight:
	de Hollander EL et al., Associations between changes in anthropometric measures and mortality in old age: a role for mid-upper arm circumference? J Am Med Dir Assoc 2013;14:187-193
	Berentzen TL et al., Changes in waist circumference and mortality in middle-aged men and women. PloS One 2010;5:e13097
	Karahalios A et al., Change in body size and mortality: results from the Melbourne collaborative cohort study. PloS One 2014;9:e99672
	Klingberg S et al., Increase in waist circumference over 6 years predicts subsequent cardiovascular disease and total mortality in nordic women, Obesity 23 (10), 2123-2130 (2015)
	Mulligan A et al., Changes in waist circumference and risk of all-cause and CVD mortality: results from the European Prospective Investigation into Cancer in Norfolk (EPIC-Norfolk) cohort study, BMC Cardiovascular disorders 19:238 (2019)
	Norman et al., The impact of weight change on cardiovascular disease risk factors in young black and white adults: the CARDIA study, IJO 27: 369-376 (2003)
	Neiberg et al., Patterns of Weight change Associated With Long-term Weight change and cardiovascular Disease Risk Factors in the Look AHEAD Study, Obesity 20:2048-2056 (2012).
	The authors are encouraged to 1.) update their reference list and 2.) formulate specific research questions that are new and 3.) show that these can be answered with the data set at hand.
	2. Aim of the study: the only place where research questions are formulated are in lines 30-33 in the Introduction (p.5); here, the authors hypothesize that their data could allow to determine the age at which weight status per se or a change in weight status have the largest impact on incident disease; alternatively, they want to examine whether the cumulative exposure to obesity poses a larger risk factor than obesity at a particular time point (age).
	These hypotheses need to be expanded on and supported by references: which are the underlying mechanisms? why should there

be a certain age range, when risk of subsequent disease is largest? for instance, individuals who are overweight early in life might have adapted to some extent and be at lesser risk for CVD that individuals with large weight gain later in life. Also, body weight doesn't change overnight, and the respective age interval could be quite large. In addition, the aspect of weight cycling should be discussed, cf. Zou et al., Body-Weight Fluctuation Was Associated With Increased Risk for Cardiovascular Disease, All-Cause and Cardiovascular Mortality: A Systematic Review and Meta-Analysis (2019).

Just an idea: the association between weight status early in life, or change in weight status in relation to prostate cancer is still an open question, cf. which this cohort could address, in particular as linkage to the prostate cancer register is planned.

3. How were the individual data sets selected? I also miss an ethics statement. 'The ODDS population is large and has a nationwide coverage of Swedish cohorts, which generally have a high representativeness of their background population (p.10, lines 50-53)' – so overall, which population is the ODDS representative of? Also, which studies and how many participants have repeated measurements of weight or waist circumference? The conscript register presents the largest contribution to the male population in this study but conscription is done only once (Table 1) – it appears that there are rather few men with repeated anthropometric measurements in this study (Table 3), and this does not seem to allow to match the aim of the study, at least not using the large sample size stated in the abstract.

4. self-reported or self-assessed? Did people weigh themselves or were they asked to report the values assessed by medical personal? How valid would self-assessed waist circumference be? Why were self-reported values used at all, as overweight is often misreported given the social stigma related to it? Please discuss!

Specific questions:
Abstract – participants: since change in anthropometric measures is crucial for ODDS study, it would be interesting to know which is the average time interval between measures of weight and waist circumference (WC) + range (IQR is less informative). Please also add the number of studies included.
Abstract – results: sex-specific results are not comparable as women were on average 11 years older than men and examined about 16 calendar years later (i.e. at a later stage of the obesity epidemic), resulting in higher BMI, as authors state on p.9 lines 53-56. Why is the number of deaths only half for women compared to men? please add incidence rates that account for length of follow-up!
Overall, the manuscript needs language editing, cf. 'enormous' (p.5 line 12), 'full health' (p.5 line 14) and so on. Often, the authors use the comparative without giving a reference, e.g. 'larger population size', 'longer follow-up' (p.5 line 22), 'fewer and smaller' (p.5 line 29), 'higher priority' (p.10 line 29), 'rarer outcomes' (p.11 line 10) – compared to what? 'long history' (p.5 line36) – please specify.
Introduction: inclusion of individuals should be moved to the Methods section, and discussed in detail: how are measurements from individuals treated that are part of more than one cohort? Have they been measured at identical dates, or do the dates differ, and the methods of assessment?
Cohort description (p.6): in which way could height be derived from repeated measurements, if many people were only examined once? What about shrinking and misclassification of BMI?
The description of weight calibration in the medical birth registry is

not clear (refs are not enough, it should be made clear here in the cohort profile, p.7). 'We first regress…' – is this done for the years after 1989, such that the trend could be extrapolated to 1982-89? Please explain, and distinguish between time trends obtained from unrelated individuals and longitudinal change of weight within individuals.

p.6 'weight assessment and completeness of the two largest cohorts' – what is meant to be complete, the weight measurement or the cohort per se (participation)?

Please add a statistical methods section, where you explain how you will account for the correlation of data from the same sub-study, especially if data come from more than one study for some individuals.

p.8 line 9: 'in the pooled ODDS population, recalled weight is counted as a distinct weight assessment' – please explain smoking status (p.8): it seems very useful to enrich information on smoking status by combining the information from different studies or time points, but this is not 'imputation'. What is meant by 'more advanced imputation methods' – use the correlation with other variables to predict smoking status? Please specify these variables, would this be enough information to impute smoking status with sufficient accuracy?

Register linkage: date of birth was only known for each quarter of the year but the precise date was missing 'resulting in a recorded birth date of the 15th in the middle month per quarter of the year' – please reformulate

Findings to date (p.9) – 'preliminary results'?

Please don't repeat figures from the tables in the text, but try to emphasize trends, e.g. women were on average 11 years older the men etc.

p.11 first sentence – unclear, please reformulate.

p.11 line 17: how does heterogeneity hamper stratified analyses?

Table 1: please add the number of participants with at least one follow-up measure of weight or WC. Does the number of people with WC refer to objectively measured WC?

Why is the upper age limit for conscripts so high (58), and for construction workers > age of retirement?

Do some of the twin studies overlap, e.g. the Swedish Twin registry and TwinGene?

LifeGene: what does 44% mean, that 56% were recruited through volunteering? How was this done?

Footnote b should be moved to methods.

Table 3: please give the range instead of IQR, which is more informative here. Please also give the number and percentage of participants with at least 2 measurements of weight or WC among all participants.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Helda Tutunchi, Tehran University of Medical Sciences
Comments to the Author:
This pooled cohort is so important, and the results support the authors' conclusions. It is well-written and structured. I accept this manuscript for publication.

Re: We thank the reviewer for the positive feedback.

Reviewer: 2

Dr. Kuat Oshakbayev, University medical center
Comments to the Author:
Manuscript ID: bmjopen-2023-084836
Marisa da Silva et al. designed to create a large cohort to study body mass index (BMI), waist circumference (WC), and changes in weight and WC, in relation to morbidity and mortality within the Obesity and Disease Development Sweden (ODDS) study.
The authors concluded that the large sample size and long follow-up of the ODDS study will provide robust results on anthropometric measures in relation to risk of common diseases and causes of deaths, and novel findings in subgroups and rarer outcomes.
Several issues should be addressed before the paper is suitable for publication. Though the topic is relevant, some flaws need to be acknowledged:
1. There is no conclusion in the paper, neither in Abstract nor in the text.

Re: A conclusion section is not included in BMJ Open's guidelines for cohort profiles.

2. The aim of the study is formulated vaguely. The aim should be clear and understandable.
3. The aim in Abstract and the text are not the same.

Re: According to the guidelines, a "Purpose – describe why the cohort was set up", but no specific aims, should be included in a cohort profile. The purpose of ODDS is formulated broadly to cover a range of research questions that can be addressed in ODDS. The new section "Ongoing and planned studies" (page 11), further specifies ongoing work and plans.

4. The authors should rewrite the paper using the IMRAD principle. For instance, it was not clear where Method, Results, or Conclusion sections began or distinction between them.

Re: Our cohort profile follows BMJ Open's guidelines for cohort profiles.

5. Inclusion and exclusion criteria for selecting study cohorts from various databases should be explicitly stated in the methods section.

Re: The first sentence of "Study population", page 6, reads: "*Swedish cohorts for research and national registers with individual-level information on weight and height were identified for inclusion in ODDS.*" Other than identifying these cohorts, there was no other selection process of cohorts. Also, there should be no "methods section" according BMJ Open's cohort profile guidelines.

6. The variables being compared must be homogeneous. For example, the years covered in ODDS were different, so it is not reliable to compare between each other, in all probability. The study cohorts and study population have large heterogeneity, and some important information is missing from the tables (see below).

Re: We would like to argue that whether heterogenous information between cohorts is an issue depends on the specific research question and will be addressed according to need in each original study of ODDS. For the broad research purpose of ODDS with anthropometric information as exposures, the different calendar years and ages at baseline and during follow-up, between and within cohorts, will be adjusted for using standard methods such as by model adjustment and stratification. For some research questions, such as in our submitted study of time trends of the BMI-mortality association, a long calendar time span of exposures is not a problem but rather a necessity for the investigation.

7. Page 8, lines 39-43: how was it practically fulfilled to avoid the Cohort overlap?

Re: The reviewer refers to the formulation referring to Table 1: "*Cohort overlap, i.e. individuals participating in more than one cohort, is not considered in the table.*" One of the key points of our cohort profile, as indicated in the bullet points of "Strengths and limitations of this study", is the use of the unique personal identity number assigned to all residents of Sweden, which enables tracking of all cohort assessments of individuals in ODDS, including assessments originating from different subcohorts for an individual. We believe that the terminology "Cohort overlap" may be confusing. We have therefore rephrased the formulation referring to Table 1 to "*Each cohort is presented separately so individuals participating in more than one cohort are counted in each of these cohorts.*" (page 6).

8. Some info in the paper is redundant, in my opinion. For instance, info in lines 15-52, page 9.

Re: The reviewer refers to the information about the Medical Birth Register. Cohort profiles of single cohorts include much more detailed information than what we could provide for the many cohorts included in our pooling. However, because the Medical Birth Register and the Military Conscription Register are by far largest in our pooling, we suggest that more information than that provided for all other cohorts in Table 1 is warranted for these two cohorts. We therefore suggest keeping this information. In the section about the Medical Birth Register, we have also replaced the statement that no appreciable pregnancy-related weight gain has occurred at gestational week 8-10, when weight is generally recorded, with more precise information (page 7): "*A large population-based study in Sweden showed that weight gain in pregnancy was minimal during the first 15 weeks of gestation, for example with a median weight gain in normal weight women of 0.7 kg between gestational week six and ten[43].*"

9. P.6, lines 39-44 and P.7, lines 53-60 are controversial between each other, because in the first the authors write "…Cohort overlap,… is not considered in the table…", but in the second they write "…includes overlap with…" in the table.

Re: Regarding "Cohort overlap" on lines 38-44, please see our response to comment no. 7 relating to Table 1. Because individuals included in more than one cohort are counted in each of the cohorts presented in Table 1 and not only once, we thought that when we later report the number of individuals in all cohorts combined other than the two largest, an explanation was needed *"..(includes overlap with the military conscription and the Medical Birth Register)..*". However, to avoid confusion and unnecessary details, we have now excluded this information.

10. Tabl. 2: difficult to understand the first time due to it has very much info (Swedish register, Description, Type of information retrieved, Years covered in ODDS, Coverage of target population). But, it provides little information on all-cause and cause-specific mortality across different pathologies.

Re: Table 2 is an overview of the type of information retrieved from each national register, similar to the tables presented in two of the reference cohort profiles (Table 1 in Selin D, et al. BMJ Open 2022;12:e059877. doi:10.1136/bmjopen-2021-059877, and Table 6 in Häggström C, et al. BMJ Open 2022;12:e064898. doi:10.1136/bmjopen-2022-064898). The information in the table enables the reader to understand the type of information available in ODDS, it is detailed nowhere else and will therefore be very useful to refer to in original articles of ODDS.

We are not quite sure what type of "different pathologies" the reviewer refers to, perhaps specific causes of death. We have reported the information retrieved from each register, e.g. ICD codes of cause of death, and ICD codes and hisopathological information for cancer. Detailing this further will result in a very long list of potential outcomes.

11. Of course, the authors showed a lot of interesting statistical information, but I expected to know scientific and analytic info related to an influence of the risk factors (overweight, smoking) on morbidity/mortality depending on age, region, gender, etc.

Re: The suggested will be the purpose of original articles but should not be included in a cohort profile according to guidelines.

12. The cohort data (with information on weight, height, and WC more between 1963 and 2020) is a very good perspective to retrieve a correlation/regression between morbidity/mortality and weight, BMI/WC, etc., or impact of overweight in development of morbidity/mortality.
13. Also there was not clear about smoking habit as risk factor for the prevalence of diseases, morbidity/mortality, or the relation to BMI/WC.

Re: It is not the purpose of the cohort profile to investigate exposure-outcome associations, or to investigate the association of risk factors with potential confounders.

14. Tabl. 3: line 19: there is no info related to Age (years).

Re: In Table 3, Age is the heading for the next rows, which show 1) the median (IQR) age, and 2) n (%) individuals in categories of age. The information for e.g. year is set up the same way.

15. In Tabl. 3 presented a lot of info, but there is no scientific analysis. For example, there are presented data "Objectively measured", "Self-reported current weight", "Self-reported recalled weight", but how to analyze it?

Re: Table 3 presents all observations of the ODDS database. An analysis plan or statistical analysis section, if that is what the reviewer suggests, should not be included in a cohort profile according to guidelines. However, in relation to the form of weight/WC assessment, we did add to the new heading "Characteristics of repeated weight and waist circumference assessments", page 10, that "*For each assessment, age, form of assessment (objectively measured/self-reported current/self-reported recalled) and, in women, a pregnancy indicator (Medical Birth Register, yes/no) will be included as predictor variables to model weight and WC trajectories in mixed effects models[10]*".

Please see also our response to reviewer no. 4 regarding "4. self-reported or self-assessed?"

16. In Tabl. 3, another example: there was very interesting dynamic of smoking frequency during >20 years, but there is no analysis about it.

Re: We do not understand the question about smoking frequency as we have not reported such information. Please see our response to the previous question regarding analysis plan.

17. In Tabl. 3, other example: dependence of weight on marital status, or vice versa.

Re: Please see our response to previous questions, especially no. 13.

18. In Tabl. 3, other example: dependence of weight/WC on urban or rural residents, or vice versa.

Re: Please see our response to previous questions, especially no. 13.

19. Page 12, lines 29-41: the info should present in a table.

Re: The reviewer refers to a paragraph with information on number of incident cancers. This paragraph and the previous one on the number of deaths includes many numbers which are heavy to read in text form. Therefore, as suggested, we have now instead included this information in a new table, Table 4, which shows the number of individuals and events in the full population, with ≥2 weight assessments, with WC information, and with ≥2 WC assessments.

20. Fig. 1: I do not exactly understand the purpose of the figure, for what for practice or science?

Re: The figure illustrates the cumulative number of men and women in ODDS across calendar years. The figure is not essential, but compared to the tables in our manuscript, the figure allows the reader to more quickly grasp some of the basics of our data, so we suggest keeping it.

Reviewer: 3
Prof. Oscar Herrán, Universidad Industrial de Santander
Comments to the Author:
The text is well written. However, it is dense, and a map is needed to consolidate the results, a "graphical summary." This text is not research in itself; it is the preparation of a database that will be used for very high-level research due to the characteristics of the database. The richness of the data and the procedures should be known since the cohort assembly is the most significant work when establishing the relationships between morbidity and mortality.

Re: We thank the reviewer for the positive feedback. We did not come up with a "map" or figure to illustrate the cohort with registers in a way that complements information in the present tables. We do believe that the revision has resulted in a clearer cohort profile, and hope that the reviewer agrees with this.

Reviewer: 4
Dr. Kirsten Mehlig, University of Gothenburg

Comments to the Author:
I am grateful for the opportunity to review this cohort profile. as the manuscript describes proposed research the questions above do not entirely apply. please find specific comments below.

Summary: This cohort profile describes the combination of different smaller cohort studies to a nationwide cohort of adults living in Sweden with aim to investigate the risk for disease related to body weight and waist circumference and their longitudinal change. The large number of observations and the long register follow-up are a strength but the lack of important confounders such as smoking status is a limitation.

Overall, the approach looks promising but the cohort profile lacks some important information.

Questions of major importance:

1. Novelty: which are the research questions that this cohort study aims to answer that have not been examined before? Overweight/obesity or central obesity measured at a single time point are now well established risk factors for many outcomes including different CVD and cancer endpoints as well as total mortality, and it is questionable whether this study could add new insight. The authors claim that change in weight status was less well investigated (p.5 lines 27-30) but there are a lot more studies than cited here (refs 5+7), even regarding change in waist circumference that is less often measured than body weight:

de Hollander EL et al., Associations between changes in anthropometric measures and mortality in old age: a role for mid-upper arm circumference? J Am Med Dir Assoc 2013;14:187-193 PubMed

Berentzen TL et al., Changes in waist circumference and mortality in middle-aged men and women. PloS One 2010;5:e13097

Karahalios A et al., Change in body size and mortality: results from the Melbourne collaborative cohort study. PloS One 2014;9:e99672

Klingberg S et al., Increase in waist circumference over 6 years predicts subsequent cardiovascular disease and total mortality in nordic women, Obesity 23 (10), 2123-2130 (2015)

Mulligan A et al., Changes in waist circumference and risk of all-cause and CVD mortality: results from the European Prospective Investigation into Cancer in Norfolk (EPIC-Norfolk) cohort study, BMC Cardiovascular disorders 19:238 (2019)

Norman et al., The impact of weight change on cardiovascular disease risk factors in young black and white adults: the CARDIA study, IJO 27: 369-376 (2003)

Neiberg et al., Patterns of Weight change Associated With Long-term Weight change and cardiovascular Disease Risk Factors in the Look AHEAD Study, Obesity 20:2048-2056 (2012).

The authors are encouraged to 1.) update their reference list and 2.) formulate specific research questions that are new and 3.) show that these can be answered with the data set at hand.

Re: Our understanding is that the primary aim of a cohort profile article is to present the cohort and the data to be used in original studies of the cohort. While "the rationale for a cohort's creation" and "any specific research questions that motivated the project" should be included according to BMJ Open's cohort profile guidelines, we do not expect that a full description of the literature or strong arguments for the novelty is warranted, which is more of the job of an original article. Also, in response to that "Overweight/obesity or central obesity measured at a single time point…. it is questionable whether this study could add new insight..", we can share with the reviewer that our submitted ODDS study of BMI and cancer risk has >50% more cases than the largest studies to date and includes many more and rarer cancers than in previous studies. Since 15 April it is under consideration for publication in the Lancet Regional Health – Europe, and on 15 May we were invited to submit a (minor) revision (preprint: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4796038). This illustrates the impact that the large sample size of ODDS can have, including for once-measured body size.

With regards to 1) and 3): We have added more information to the Introduction (page 5). Due to many references already (52 in the first submitted manuscript, 62 in the revised version) and our above raised points regarding our understanding of the purpose of cohort profiles, we have limited the number of additional references by giving priority to meta-analyses/systematic reviews. We do not provide a full background of the literature for all possible body size measures and outcomes to be studied, but we exemplify with some of the research conducted on mortality including the number of events. These can be compared to the numbers in ODDS in our new Table 4, which further highlights the main strength (and novelty) of our study: the large sample size.

With regards to 2): We expect more ideas to come up while working on the initial studies, and therefore, to not outdate the purpose, we want to keep the broad purpose as is "*a large population to investigate the association of anthropometric measures (BMI, WC, and changes in weight and WC) with the risk of morbidity and mortality*" (last in Introduction, page 6). However, we have added a section "Ongoing and planned studies" (page 11), which further specifies the research questions to be addressed in the first studies.

2. Aim of the study: the only place where research questions are formulated are in lines 30-33 in the Introduction (p.5); here, the authors hypothesize that their data could allow to determine the age at which weight status per se or a change in weight status have the largest impact on incident disease; alternatively, they want to examine whether the cumulative exposure to obesity poses a larger risk factor than obesity at a particular time point (age).

These hypotheses need to be expanded on and supported by references: which are the underlying mechanisms? why should there be a certain age range, when risk of subsequent disease is largest? for instance, individuals who are overweight early in life might have adapted to some extent and be at lesser risk for CVD that individuals with large weight gain later in life. Also, body weight doesn't change overnight, and the respective age interval could be quite large. In addition, the aspect of weight cycling should be discussed, cf. Zou et al., Body-Weight Fluctuation Was Associated With Increased Risk for Cardiovascular Disease, All-Cause and Cardiovascular Mortality: A Systematic Review and Meta-Analysis (2019).

Just an idea: the association between weight status early in life, or change in weight status in relation to prostate cancer is still an open question, cf. which this cohort could address, in particular as linkage to the prostate cancer register is planned.

Re: Please see our revised and longer Introduction on page 5, which now addresses weight changes in different ages more, and also includes the aspect of weight fluctuation with reference to the suggested meta-analysis. The current cohort profile is aimed to describe the rationale for the study and the data to be used. In original studies, we will go into more detail with the underlying mechanisms and more detailed hypotheses. For the reviewer's information, we are currently finalising a study on weight changes and prostate cancer incidence and mortality. Please see also our response to the previous comment.

3. How were the individual data sets selected? I also miss an ethics statement. 'The ODDS population is large and has a nationwide coverage of Swedish cohorts, which generally have a high representativeness of their background population (p.10, lines 50-53)' – so overall, which populatiois the ODDS representative of? Also, which studies and how many participants have repeated measurements of weight or waist circumference? The conscript register presents the largest contribution to the male population in this study but conscription is done only once (Table 1) – it appears that there are rather few men with repeated anthropometric measurements in this study (Table 3), and this does not seem to allow to match the aim of the study, at least not using the large sample size stated in the abstract.

Re: The first sentence of "Study population", page 6, reads: "*Swedish cohorts for research and national registers with individual-level information on weight and height were identified for inclusion in ODDS.*" Other than identifying these cohorts, there was no other selection process of cohorts.

In line with other cohort profiles of BMJ open, an "Ethics approval" section is included (page 13).

The high representativeness of the included cohorts is evident from the relatively high attendance rate of each cohort, as reported in Table 1, and in more detail in the text (page 7) for the two largest cohorts which have particularly high representativeness. This means that the ODDS population is representative of the background population from which the cohorts originated, i.e. military conscripts, women giving birth, construction workers, etc. It is difficult to explain this clearly in a different way than in general terms, but for clarification, we have added to the sentence **in bold**: "*..which, **although not completely population-based,** generally have a high representativeness of their background population.*" (page 12).

In response to the comment on repeated measurements, the information in Table 1 of the number of individuals and assessments in each cohort indicates whether a cohort has few or many repeated measurements. To further describe this part, we have now divided "Findings to date" in four sections, one being "Characteristics of repeated weight and waist circumference assessments" (page 10). In

that section, we report the cohorts contributing the largest number of assessments in men and women with two or more assessments. Please see the paragraph for a full read.

4. self-reported or self-assessed? Did people weigh themselves or were they asked to report the values assessed by medical personal? How valid would self-assessed waist circumference be? Why were self-reported values used at all, as overweight is often misreported given the social stigma related to it? Please discuss!

Re: Self-reported weight has very high validity, and self-reported WC also has high validity. We have added more information on this in "Weight and waist circumference assessments in other cohorts", page 8: "…*Self-reported current weight has shown very high validity in western populations with a correlation coefficient with objectively measured weight of over 0.95[46 47]. High validity has also been shown for self-reported WC in western populations, showing correlation coefficients with objectively measured WC of over 0.85 when instructions were provided on how to measure WC [46 47], and only slightly lower when no instructions were provided [48]. In ODDS, the SIMPLER cohorts and the National March Cohort collected self-reported WC information from questionnaires without further instructions, and in Women's Lifestyle and Health, a picture indicated where to place the tape measure. WC in the other cohorts were objectively measured.*"

Specific questions:
Abstract – participants: since change in anthropometric measures is crucial for ODDS study, it would be interesting to know which is the average time interval between measures of weight and waist circumference (WC) + range (IQR is less informative). Please also add the number of studies included.

Re: We agree that the suggested is relevant information but after including what we think is the most important information, we have already reached the Abstract word limit of 300 words (this is not included in the instructions but shows in the submission system). Another reason to not report a measure for the mean/median time between assessments is that this varies largely between and within cohorts, which is why we will use methods accounting for this variation when modelling weight and WC trajectories. According to the cohort profile guidelines, statistical analyses should not be reported. However, we think that this brief information under the new subheading "Characteristics of repeated weight and waist circumference assessments" on page 10, will be helpful to picture the design of studies of weight/WC changes: "*Individuals with two or more assessments of weight or WC, respectively, will be eligible for inclusion in studies of change of these factors, on outcomes. In these studies, we will utilise all assessments in an individual, i.e. allowing assessments to take place in different cohorts. For each assessment, age, form of assessment (objectively measured/self-reported current/self-reported recalled) and, in women, a pregnancy indicator (Medical Birth Register, yes/no) will be included as predictor variables to model weight and WC trajectories in mixed effects models[10]*". The reference is a study using the method that we are currently using in our first two studies of weight changes (on prostate cancer and on mortality). The paragraph also includes information on the characteristics of repeated weight and WC assessments.

The number of studies included in ODDS is not a straightforward count because even the categorisation into cohorts and subcohorts in Table 1 is a simplification. For example, the NSHDS and the Malmö cohorts include the six listed distinct subcohorts, which have both similarities and differences, and the same goes for Q63, 67 and 70, which are grouped together as one subcohort of the Swedish Twin Registry. To not confuse the reader, we suggest to not indicate the number of cohorts and leave this information to be observed in Table 1.

Abstract – results: sex-specific results are not comparable as women were on average 11 years older than men and examined about 16 calendar years later (i.e. at a later stage of the obesity epidemic), resulting in higher BMI, as authors state on p.9 lines 53-56. Why is the number of deaths only half for women compared to men? please add incidence rates that account for length of follow-up!

Re: We agree with the reviewer that the differences in age and year of weight assessments between men and women make the descriptive information of men and women incomparable. In addition, women have longer life expectancy than men. The reviewer suggests adding incidence rates (likely meant to be mortality rates) that account for length of follow-up. Because of the declining mortality rate in Sweden, i.e. a higher life expectancy over time, the correct analysis to compare men and women would be to calculate age-standardized mortality rates across calendar years. However, this is

beyond the scope of our cohort profile article and would be more suitable for an original article. In the referent cohort articles, only one of six reported rates. The five others, like us, reported estimated or real numbers of events during follow-up.

Overall, the manuscript needs language editing, cf. 'enormous' (p.5 line 12), 'full health' (p.5 line 14) and so on.

Re: The reviewer refers to the expression "*The negative impact of obesity on public health is enormous*", we have changed "enormous" to "large".

The reviewer further refers to the expression *"…for example with disability-adjusted life years, i.e. years lost in full health"*. The expression "full health" is the WHO's definition of disability-adjusted life years, which we suggest should not be changed (https://www.who.int/data/gho/indicator-metadata-registry/imr-details/158).

Two of the co-authors are native English speakers and had read through and agreed to the submitted version. We have read through the revised manuscript and have made minor corrections to the language, and we also had it read and approved by a third native English speaker.

Often, the authors use the comparative without giving a reference, e.g. 'larger population size', 'longer follow-up' (p.5 line 22), 'fewer and smaller' (p.5 line 29), 'higher priority' (p.10 line 29), 'rarer outcomes' (p.11 line 10) – compared to what? 'long history' (p.5 line36) – please specify.

Re: The reviewer refers to: "*Less is known about the association of obesity with rarer diseases, which, in a prospective cohort setting, requires a larger population size and/or a longer follow-up for investigation.*" This sentence is not meant to be specific, it is stating the fact that a larger population and/or longer follow-up will result in more cases, which enables to study rarer diseases. There is no specific number to quantify this, or to compare with.

The reviewer refers to: "*Studies based on measures of central obesity, such as waist circumference (WC), and of changes in body size across life on the risk of morbidity and mortality, have been much fewer and smaller.*" In response to this comment and the reviewer's comments no. 1-2, we have added more information and number of events in previous research to the Introduction, which quantifies "fewer and smaller".

The reviewer refers to *"…, cancer has a higher priority in the ODDS working group.*" This information is provided for the reader to understand why we report numbers of incident cancers, but not of other diseases. There is no relevant measure of how we could quantify how or why cancer has a higher priority. However, we have changed the wording to "*cancer is of main focus*".

The reviewer refers to "*Lastly, the long follow-up time together with the large sample size has accumulated many events resulting in overall high statistical power and the possibility to investigate subgroups and rarer outcomes.*" Please see our response to the first expression in this comment. Also, we have exemplified "rarer outcomes" with the number of incident cases of some rare cancers (previously in text, now in the new Table 4).

Lastly, the reviewer refers to that "*The long history, … of many Swedish nationwide registers..*" is unspecific. The "long history" is exemplified with the starting year of the Swedish Cause of Death Register and the Cancer Register in the same paragraph, and with the starting year of each register in Table 2. It would be inappropriate to point out specific countries with a shorter history of these registers.

Introduction: inclusion of individuals should be moved to the Methods section, and discussed in detail: how are measurements from individuals treated that are part of more than one cohort? Have they been measured at identical dates, or do the dates differ, and the methods of assessment?

Re: There should be no Methods section in cohort profiles according to the guidelines. Like the reference cohort profiles, in our originally submitted version, we had placed the information about the study population first in Cohort description (page 6). We had stressed the tracking of individuals by use of the unique personal identification number throughout the manuscript, for example in the key points of Strengths and limitations, and in the Introduction: "*The personal identity number further facilitates cross-identification of individuals between different cohorts in pooled studies…*". Our

revised sentence in relation to Table 1 (in bold below) may further clarify the unique individual tracking of assessments also between cohorts (including retrieving information of the date of assessment in each cohort): "*The characteristics of the included cohorts forming the ODDS study population of 4,295,859 **unique** individuals with 7,733,901 weight assessments are described in Table 1. **Each cohort is presented separately so individuals participating in more than one cohort are counted in each of these cohorts**".

Cohort description (p.6): in which way could height be derived from repeated measurements, if many people were only examined once? What about shrinking and misclassification of BMI?

Re: Information on height was very complete and 68% of the population had two or more assessments. We have changed the sentence that we think that the reviewer refers to, to (added word in **bold**): "*The availability of height was not used as an inclusion criterion for each individual assessment as it could **often** be derived from repeated assessments, and because other factors, such as sociodemographic factors, were regarded to be equally important as height in studies of changes in weight and WC".*

Shrinking in older age affects BMI, as does a changing body composition across age. One way of handling this is through model adjustment or subgroup analysis by age. We suggest that these are detailed methodological questions to consider in the original articles of ODDS.

The description of weight calibration in the medical birth registry is not clear (refs are not enough, it should be made clear here in the cohort profile, p.7). 'We first regress…' – is this done for the years after 1989, such that the trend could be extrapolated to 1982-89? Please explain, and distinguish between time trends obtained from unrelated individuals and longitudinal change of weight within individuals.

Re: For clarity, we have added more information on page 7, indicating the linear regression mode of weight and calendar year at baseline examination (added text in **bold**): "*Therefore, we corrected weight assessments in 1982-89 in the Medical Birth Register **using a linear regression model of weight and calendar year from 1992 onwards**."*

p.6 'weight assessment and completeness of the two largest cohorts' – what is meant to be complete, the weight measurement or the cohort per se (participation)?

Re: "*Weight assessment and completeness of the two largest cohorts*" is a heading and includes information on how and when weight was assessed, completeness of weight assessments, and the coverage of the background population, of the two largest cohorts, the Military Conscription Register and the Medical Birth Register. We have now changed the title to "*Weight assessments, completeness, and population coverage of the two largest cohorts*".

Please add a statistical methods section, where you explain how you will account for the correlation of data from the same sub-study, especially if data come from more than one study for some individuals.

Re: There should not be a statistical methods section in a cohort profile article. The information provided in the manuscript describes that the unique personal identity number will be used to track all assessments for an individual originating also from different cohorts, for example in the key points, and in the Introduction "*It also enables the use of repeated individual assessments between cohorts to enlarge the sample size for the investigation of factors such as longitudinal changes of anthropometrics.*". Using repeated assessments also between cohorts is possible because weight and WC are quite standardised measures after accounting for whether the assessment was objectively measured or self-reported. Please see also our response to the comment "Specific questions – Abstract", and the new information added in "Characteristics of repeated weight and waist circumference assessments" on page 10.

p.8 line 9: 'in the pooled ODDS population, recalled weight is counted as a distinct weight assessment' – please explain

Re: We have added (**in bold**): "*In the pooled ODDS study population, recalled weight is counted as a distinct weight assessment **with the date set to that of the age for the recalled weight**." (page 8).

smoking status (p.8): it seems very useful to enrich information on smoking status by combining the information from different studies or time points, but this is not 'imputation'. What is meant by 'more advanced imputation methods' – use the correlation with other variables to predict smoking status? Please specify these variables, would this be enough information to impute smoking status with sufficient accuracy?

Re: Carrying information forwards and backwards is one of the methods used for "single" imputation[1]. We have now specified that "…*we performed* **single** *imputation by carrying information forwards or backwards…*", (page 8).

With regards to the "more advanced imputation methods", we have rephrased the sentence to (page 9): "*To increase the coverage of smoking information, multiple imputation[49] based on all available data may be considered in single ODDS studies.*" We have previously used such method[2], suitable for when the information is missing at random (which is the majority case in ODDS). We do not think that more information on this should be included in the cohort profile, but rather in any original studies where multiple imputation is used.

1. Jakobsen J C, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Medical Research Methodology.* 2017;17(1):162.

2. Sun M, Fritz J, Häggström C, Bjørge T, Nagel G, Manjer J, Engeland A, Zitt E, van Guelpen B, Stattin P, Ulmer H, Stocks T. Metabolically (un)healthy obesity and risk of obesity-related cancers: a pooled study. *J Natl Cancer Inst.* 2023;115(4):456-467.

Register linkage: date of birth was only known for each quarter of the year but the precise date was missing 'resulting in a recorded birth date of the 15th in the middle month per quarter of the year' – please reformulate

Re: We have added **in bold**: "*…date of birth was obtained only for the year and quarter of a year, resulting in a recorded birth date with day 15 and the middle month of the quarter,* **e.g. February for the first quarter,** *in addition to the birth year.*".

Findings to date (p.9) – 'preliminary results'?
Please don't repeat figures from the tables in the text, but try to emphasize trends, e.g. women were on average 11 years older the men etc.

Re: We agree to add descriptive information to the numbers, which, however, we have kept because the absolute year and age is also important. The sentence now reads "*The median (IQR) year and age of weight assessment was lower in men than in women; year 1985 (1977-94) vs. 2001 (1991-2010), and 19 (18-40) vs. 30 (26-36) years of age.*"

p.11 first sentence – unclear, please reformulate.

Re: The reviewer refers to the formulation "*The combination of anthropometric measures in the pooled population and register information further enables investigation of the combined exposure of anthropometric measures and, for example, cardiometabolic diseases and medications on outcomes.*". We have reformulated to "*The combination of anthropometric measures in the pooled population with, for example, cardiometabolic diseases and medications retrieved from national registers, enables the investigation of various risk factors jointly with anthropometrics.*".

p.11 line 17: how does heterogeneity hamper stratified analyses?

Re: We agree that the formulation is misleading and have reformulated to "*The large heterogeneity of the study population in terms of geographical region, age, and year of weight assessment, enables subgroup investigation; however, certain subgroups, for example older age individuals, will be too small for a separate investigation.*".

Table 1: please add the number of participants with at least one follow-up measure of weight or WC. Does the number of people with WC refer to objectively measured WC?

Re: We agree with the reviewer that more information would be of interest; however, the table is already extensive with nine columns, and the information only just fits in one row with Arial size 8. The table includes the number of individuals and assessments for each cohort, which indicates the cohorts with the most repeated assessments. Instead of providing further detailed information on all cohorts in Table 1, most of which contribute <10% of the total population, we have added a separate subtitle on repeated assessments in "Findings to date", "Characteristics of repeated weight and waist circumference assessments" (page 10), where we report which cohorts that contribute with the most assessments in individuals with two or more assessments of weight and WC.

The number of individuals with weight and WC assessments in Table 1 refer to all assessments. The proportion with objectively measured and self-reported weight and WC are described in text.

Why is the upper age limit for conscripts so high (58), and for construction workers > age of retirement?

Re: The age at health examination for some conscripts and construction workers is high but not unreasonable as further indicated by that these observations are not observed only in a few individuals. The 95th percentile of age among the conscripts is 22 years and it is 64 years among the construction workers, further showing that the high ages are exceptions from the standard. The cohort profile of the Conscription Register reports that 2% are older than 22 years but it does not provide further detail. Contact with a person at the Conscription Register informed us that individuals with high ages are volunteers, i.e. they were not called through mandatory recruitment. We have also been in contact with a key person of the Construction Workers Cohort (Bengt Järvholm) who informs that older individuals in the cohort may have been in temporary work and therefore been invited to a health examination, and retired individuals may have requested a health check-up in follow-up of examinations during previous working years, which then may have been accepted.

Whilst this information is good to know, this detailed information is not included in the "cohort profile" article of the Conscription Register or in any article of the Construction Workers Cohort, so we think it is too much detail for our cohort profile.

Do some of the twin studies overlap, e.g. the Swedish Twin registry and TwinGene?

Re: The Swedish Twin Registry (STR) includes all the subcohorts in Table 1 listed under the STR with an indent, TwinGene being one of them. There is overlap between some of the STR cohorts as well as between several other cohorts of ODDS, especially between those from the same region as described in cohort profiles and original articles of those cohorts.

LifeGene: what does 44% mean, that 56% were recruited through volunteering? How was this done?

Re: We have now changed the short description of LifeGene in Table 1 to "*Residents across Sweden were invited through randomization (~44%) or in other ways, e.g. shared household with the one invited, or volunteering, to a health examination and a web-based survey for research purposes*".

Footnote b should be moved to methods.

Re: We have now omitted footnote b in Table 1 and included the information in "Smoking information", page 8.

Table 3: please give the range instead of IQR, which is more informative here. Please also give the number and percentage of participants with at least 2 measurements of weight or WC among all participants.

Re: We believe that reporting the range or the IQR is a matter of taste and there is no specific suggestion in BMJ Open's instructions to cohort profiles. Range is reported in Table 1 for the individual cohorts so for age and year, the information in Table 3 would be repetitive for range but not for IQR. Additionally, whilst IQR is independent of sample size, range is dependent on sample size – the larger the samples size, the more extreme values and the larger the range tends to be. We therefore suggest keeping IQR in Table 3.

VERSION 2 – REVIEW

REVIEWER	Oshakbayev, Kuat University medical center
REVIEW RETURNED	03-Jun-2024

GENERAL COMMENTS	My comments are the same, but the decision remains with the editor of BMJ Open.

REVIEWER	Mehlig, Kirsten University of Gothenburg
REVIEW RETURNED	07-Jun-2024

GENERAL COMMENTS	Thank you for the explanations and changes made in the manuscript, which are overall very good. Regarding the number of repeated measures in ODDS I still think that the presentation is a bit too positive and hides the fact that the main contribution to the male part, the conscript registry, does not have repeated measurement of weight. This should be made clear in at least 2 places: in the new paragraph on "Characteristics of repeated weight and WC assessments', and in the "limitations", e.g. "because repeated weight measurements were not performed in the conscript study, which is also the largest male study in ODDS, only 22% of men had repeated measurements on weight. In contrast, almost 60% of women have two or more weight measurements." Anthropometric change is a main exposure in this cohort study, and it should be made very clear how much data are available. Some minor comments: Register linkage: " … date of birth was obtained only for the year and quarter of year …" should be replaced by e.g. " if date of birth was only known for the quarter of the year, we imputed the missing date by the mean value of possible dates, e.g. Feb 15 for the first quarter of the year." p.11 line 17: how does heterogeneity hamper stratified analyses? Re: We agree that the formulation is misleading and have reformulated to "The large heterogeneity of the study population in terms of geographical region, age, and year of weight assessment, enables subgroup investigation; however, certain subgroups, for example older age individuals, will be too small for a separate investigation.". I would suggest the following formulation: "The large heterogeneity of the study population in terms of geographical region, age, and year of weight assessment, enforces subgroup investigation, to examine whether associations differ between subgroups…"

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Dr. Kuat Oshakbayev, University medical center
Comments to the Author:
My comments are the same, but the decision remains with the editor of BMJ Open.
****editor's note: this is fine as we believe there was a misunderstanding about the article type 'cohort profile' with BMJ Open. No further revisions/responses are required to this reviewer's comments.
Reviewer: 4
Dr. Kirsten Mehlig, University of Gothenburg
Comments to the Author:
Thank you for the explanations and changes made in the manuscript, which are overall very good.
Regarding the number of repeated measures in ODDS I still think that the presentation is a bit too positive and hides the fact that the main contribution to the male part, the conscript registry, does not have repeated measurement of weight. This should be made clear in at least 2 places: in the new paragraph on "Characteristics of repeated weight and WC assessments', and in the "limitations", e.g. "because repeated weight measurements were not performed in the conscript study, which is also the largest male study in ODDS, only 22% of men had repeated measurements on weight. In contrast, almost 60% of women have two or more weight measurements." Anthropometric change is a main exposure in this cohort study, and it should be made very clear how much data are available.
Response: We have added to "Characteristics of repeated weight and waist circumference assessments" on page 10-11: "The much smaller number of assessments in men than in women is due to the lack of repeated weight assessments in the Military Conscription Register as opposed to the many assessments in the Medical Birth Register."
As further suggested, we have added to limitations, page 12: "Because repeated weight measurements were not performed in military conscripts, which make up the largest male cohort in ODDS, only 22% of men have repeated measurements on weight. In contrast, almost 60% of women have two or more weight measurements."
Some minor comments:
Register linkage: " … date of birth was obtained only for the year and quarter of year …" should be replaced by e.g. " if date of birth was only known for the quarter of the year, we imputed the missing date by the mean value of possible dates, e.g. Feb 15 for the first quarter of the year."
Response: Date of birth was retrieved from the Total Population Register as year and quarter of the year for all individuals. We agree with the suggested by the reviewer and have reformulated as follows (page 9): "…date of birth was obtained only for the year and quarter of the year, so we imputed the missing date by the mean value of possible dates, e.g. Feb 15 for the first quarter of the year."
p.11 line 17: how does heterogeneity hamper stratified analyses?
Re: We agree that the formulation is misleading and have reformulated to "The large heterogeneity of the study population in terms of geographical region, age, and year of weight assessment, enables subgroup investigation; however, certain subgroups, for example older age individuals, will be too small for a separate investigation.".
I would suggest the following formulation:
"The large heterogeneity of the study population in terms of geographical region, age, and year of weight assessment, enforces subgroup investigation, to examine whether associations differ between subgroups…"
Response: We have reformulated as suggested (page 12).