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)	Effects of Reproductive Period Duration and Number of Pregnancies on Mid-Life Electrocardiographic Indices: A Secondary Analysis from the Women's Health Initiative Clinical Trial
AUTHORS	Parikh, Nisha; Kapphahn, Kristopher; Hedlin, Haley; Olgin, Jeffrey; Allison, Matthew; Magnani, Jared; Ryckman, KK; Waring, Molly; Perez, Marco; Howard, Barbara

VERSION 1 – REVIEW

REVIEWER	Angela H.E.M.Maas, MD PhD Radboud University Medical Center, Nijmegen, the Netherlands
REVIEW RETURNED	27-Oct-2017

GENERAL COMMENTS	Large dataset of 40 K women from the Women's Health Initiative trial, in which a high number of live births and RD was found to be related to atrial and ventricular conduction ECG parameters. Interesting study, some issues need more clarification. 1. Please explain why 15.543 women with a history CVD were excluded. Would a subanalysis of these patients provide relevant clinical information related to the measured ECG characteristics? 2. What is meant by hormone therapy? Premenopausal oral contraceptives and/or postmenopausal hormone therapy?? Duration of OC and HT treatment may also be relevant, but is not further explored. 3. How reliable is age at menopause in women using oral contraceptives or having an IUD? Explain. 4. What is the association of menstrual irregularities/fertility disorders/endometriosis on studied ECG parameters? We know that these may convey an increased future CVD risk. 5. Data on the type/duration of antidepressants used may be important. 6. what may be the role of over-the-counter vitamins etc, often used by females, on these ECG parameters?

REVIEWER	Professor Michael Lewis Swansea University, UK
REVIEW RETURNED	11-Dec-2017

GENERAL COMMENTS	I have highlighted and annotated some specific sections of the attached files (one has comments on the text, the other has comments on the tables). These notes in the main highlight minor errors, omissions, lack of clarity etc. My two main concerns with the paper are 1) the statistical analysis (in as much as I have some queries about the procedures used,

	which were not explicit - please see my comment below), and 2) the quality of the Discussion: I felt that the discussion was broadly speculative and did not explore an evidence-based mechanistic explanation for the results, and it did not present the reader with an adequately interpretation of the clinical relevance of the work. Note on statistical analysis: The concern here relates to the methods used to perform regression analysis: 1) Number of Pregnancies in the study is a categorical variable - how was this handled in the statistical analysis? 2) RD has presumably been employed as a continuous variable (although this was not explained)? However the RD has a very consistent median and narrow IQR across all study participants (36[8]). What is the implication of this narrow range for the RD predictor value in the regression analyses for predicted QTc and P/PR? Might RD have been better interpreted as a categorical variable? The authors should explain their choices and procedures so that these considerations are explicit.

REVIEWER	Ewoud Schuit Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, University Utrecht, Utrecht, The Netherlands
REVIEW RETURNED	20-Dec-2017

GENERAL COMMENTS	Effects of Reproductive Period Duration and Number of Pregnancies on Mid-Life Electrocardiographic Indices: A Secondary Analysis from the Women's Health Initiative Clinical Trial This study is a secondary analysis of the WHI Trial which studied the association between number of pregnancies and reproductive period duration (RD, time from menarche to menopause) with electrocardiographic intervals. This is relevant as up till now only relatively small studies have suggested that there are small but measureable changes in electrocardiographic intervals during pregnancy. The current study confirms these changes with an increased PR interval and longer QTc interval for those with 5+ livebirths compared to 0 prior pregnancies. Reproductive period duration (RD) was associated with longer PR interval and maximum P wave duration among never users of HT. For every year increase in reproductive period, QTc decreased. The authors conclude that an increasing number of live births are related to increased ventricular repolarization time whereas RD is related to decreased ventricular repolarization time. Both longer RD and grandmultiparity are related to increased atrial conduction time. The premenopausal hormonal milieu appears to have effects on midlife cardiac electrical conduction system remodeling in women that may modestly influence CVD risk in later life. In general, I think this is a well done and nicely presented study. I do have some comments/remarks, mostly related to things that weren't entirely clear to me. Major comments 1. Abstract. Conclusion: I think the conclusion needs some rewriting and I actually like the conclusion at the end of discussion better. E.g. it is stated that "An increasing number of live births are (IS?) related to increased ventricular repolarization time.", but the QTc interval varies from 0.66 to 0.15, to 0.25 to 1.15, and statistical significance was only found for 5+ livebirths. Additionally, "hormonal milieu" is not properly introduced in the abstract, which makes that this statement seems to come out of nowhere. The same is true for "CVD risk in later life". It would be helpful if the authors could add some

	information to the objective section to make that the conclusion better links to the rest of the abstract and to make sure that the conclusion easily follows from the results presented. 2. Methods. Study sample/missing data. Over 10,000 (roughly 20% of participants) were excluded because of missing data. Did the authors consider to use statistical techniques (e.g. multiple imputation) to account for these missing data? Not accounting for missing data may potentially lead to bias when the missingness of the data related to the exposures or outcomes of interest. Later on in the methods section the authors do indicate that they performed multiple imputation, however, as reported now, it is not clear to me whether the imputation relates to missing data in the 40,687 women or missings in the initial cohort from which 10,000+ women were excluded? If imputation was used anyway, why not impute data for the 10,000+ women that were excluded for having missing data? 3. Methods. Statistical methods. Can the authors explain why they did not fit models in which both number of pregnancies and RD were included? I understand that the main interest is in these two exposures, but could there be a possibility that some of the results found for these exposures may be explained by the other exposure? 4. Discussion. I miss two things in the discussion: a. What do the results of this study mean for clinical practice? Especially considering that 5+ live births seems to have the biggest association, and 5+ live births becomes less and less common these days. I know some of this has been described in the summary, but I think it would be helpful if it could be added to the discussion as well. b. what questions weren't answered by this study and what would be future research? Minor comments Abstract 5. Methods. Please indicate which outcomes were primary and which were secondary outcomes. 6. Results: “HT” has not been defined before, please replace by “hormone therapy”. 7. Limitations: Recall bias of exposure is mentioned as a limitation. Please specificy that this relates to the reproductive period duration and not number of pregnancies. Methods 8. Study sample, line 145. I assume “HT” stands for "hormone therapy", but it has been introduced earlier in the introduction as "hormone trial". 9. Can the authors provide some basic information about the women in the WHI, e.g. some broad in- and exclusion criteria, e.g. that they all reached menopause? 10. The authors state in lines 148-150 that “This analysis drew from the cohort of women enrolled in the clinical trial. Figure 1 shows the creation of the study sample. Of 68,132 women in WHI Studies (hormone therapy, diet and calcium/vitamin D and observational studies),...” There are some inconsistencies in this section: c. Earlier was the WHI introduced as “a set of clinical trials”, but the now the authors state that the “analysis drew from the cohort of women enrolled in the clinical trial”, not “trialS”. I assume this is a typo. The same holds for observational study (in description of WHI) vs. observational studIES in the quoted statement. d. The statement that “analysis drew from the cohort of women enrolled in the clinical trial” suggests that the data from the observational study/studies was not used, but in the quoted statement it seems that that data was used. Please clarify.

	11. Electrocardiographic parameters, line 178. Typo, “additional” should be “addition”. 12. Line 210. A sentence seems to be missing (“Further,.”). 13. Multiple Imputation Analysis. I would suggest to use more imputations than the 5 currently used. Results 14. QTc. The authors state that “For each additional year in reproductive period duration, there was a 0.4 ms shorter QTc (Table 3).” Please explain what “additional year” means here. Additional to what? I assume the lowest value of RD in that data. What was that value? Another way to deal with this would be to standardize RD (e.g. by subtracting the lowest value or the mean RD from all RD value) and refit the model. Tables & Figures 15. Table 1. I think it would be helpful for the readers to know the baseline characteristics of the population as a whole, to allow comparison to other cohorts. Can the authors add an additional column to the table presenting these study characteristics?

REVIEWER	Dr Victoria Allgar University of York, England
REVIEW RETURNED	08-Jan-2018

GENERAL COMMENTS	The aim is to determine if there is a positive or negative association between number of pregnancies and reproductive period duration with mid-life electrocardiogram intervals (PR interval and QTc) and p wave parameters (p wave 139 maximum duration, dispersion and index). The introduction should include definitions of the dependant variables: PR interval, P wave indices (duration and dispersion) and QTc from enrollment electrocardiogram. e.g. The PR interval is the time from the onset of the P wave to the start of the QRS complex, Data from the Women’s Health Initiative Hormone Trial was analysed for the study and a consort flow diagram is included. However in the flow chart it stats that 6,685 were further excluded for having missing covariate data, leaving a final sample of 40,687. Yet in the analysis plan " multiple imputation techniques to impute missing covariates", so i am unclear with regards to the final numbers included and this needs to be clarified. There is no summary statistics for PR interval, QTc, P wave duration and dispersion e.g. mean (sd) or any test of normality. For the regression models were the residual plots investigated so you can trust the results and what are the measures used for the goodness-of-fit and R-squared? In Table 2 (incorrectly labelled Table 1), Table 3 and Table 4 it would be useful to include summary statistics for PR interval, QTc and P wave duration and dispersion for the categories for number of live births and reproductive period duration. For PR interval and QTc the p-values should be included in the text or table. For P wave duration and dispersion Table 5 is confusing - were individual models use for each subgroup or is this the same as for

	PR interval and QTc?

VERSION 1 – AUTHOR RESPONSE

Reviewers' Comments to Author:

Reviewer 1 (R1):
Reviewer Name: Angela H.E.M.Maas, MD PhD
Institution and Country: Radboud University Medical Center, Nijmegen, the Netherlands
Competing Interests: none declared

R1 Comment 1: Please explain why 15.543 women with a history CVD were excluded. Would a subanalysis of these patients provide relevant clinical information related to the measured ECG characteristics?

> Author Response R1 Comment 1: We thank Dr. Maas for her thoughtful review of our study. Because number of pregnancies and reproductive period (in particular age at menopause) are known to be associated with later CVD and a history of CVD is related strongly with ECG changes including QTc and certainly increased PR, we sought to exclude women with a history of CVD in order to assess associations between reproductive period duration and number of pregnancies that were not directly mediated through CVD. We have clarified this point in the methods section beginning on line 144 through 151 in the revised manuscript.

R1 Comment 2: What is meant by hormone therapy? Premenopausal oral contraceptives and/or postmenopausal hormone therapy?? Duration of OC and HT treatment may also be relevant, but is not further explored. *

> Author Response R1 Comment 2: We now clarify that hormone therapy as defined in our paper refers to postmenopausal hormone therapy and does not refer to oral contraceptive use on line164. We now additionally adjust for oral contraceptive usage and duration of oral contraceptive use and summarize the results in the text section of manuscript. We did not find that these materially affected our results.

R1 Comment 3: How reliable is age at menopause in women using oral contraceptives or having an IUD? Explain.

> Author Response R1 Comment 3: Although we do not have any reason to think that age at menopause is unreliable in women with oral contraceptives, one prior study has demonstrated that use of OCP's was associated with a later age at menopause.(1) Accordingly, we now adjust for OCP's in a secondary analysis and have verified that it does not change our results. We mention this secondary analysis in the methods and results sections of the revised manuscript.

R1 Comment 4: What is the association of menstrual irregularities/fertility disorders/endometriosis on studied ECG parameters? We know that these may convey an increased future CVD risk.

> Author Response R1 Comment 4: Thank you for raising this important point. To our knowledge, these factors have not been well studied in relation to ECG parameters. Therefore, we did assess their associations with ECG parameters and have not found any

significant associations (data not shown). We mention this secondary analysis in the methods and results section of the revised paper.

R1 Comment 5: Data on the type/duration of antidepressants used may be important.

Author Response R1 Comment 5: Thank you for this suggestion. WHI participants did report their current medications at the baseline visit but data on duration was not well ascertained. Participant medications were classified using the National Drug Classification system. We further categorized the medications into the following categories ('Anti-anxiety, Hypnotics', 'Non-SSRI Antidepressants' and 'SSRI Antidepressants') and adjusted for their use in our multivariable models. Entering these variables into the models did not affect the association between reproductive measures and ECG parameters. We have added to the methods section and results (both under secondary analysis).

R1 Comment 6: What may be the role of over-the-counter vitamins etc., often used by females, on these ECG parameters?

Author Response R1 Comment 6: We studied the effects of Ca-D on ECG parameters and did not appreciate any significant associations. We add this to the methods section and results (both under secondary analysis).

Reviewer 2 (R2)

Reviewer Name: Professor Michael Lewis
Institution and Country: Swansea University, UK
Competing Interests: None declared

Please see the files attached.

I have highlighted and annotated some specific sections of the attached files (one has comments on the text, the other has comments on the tables). These notes in the main highlight minor errors, omissions, lack of clarity etc.

R2 Comment 1: My two main concerns with the paper are 1) the statistical analysis (in as much as I have some queries about the procedures used, which were not explicit - please see my comment below), and 2) the quality of the Discussion: I felt that the discussion was broadly speculative and did not explore an evidence-based mechanistic explanation for the results, and it did not present the reader with an adequately interpretation of the clinical relevance of the work.

Author Response R2 Comment 1: We thank Dr. Lewis for his thoughtful critique. We agree about the lack of evidence-based mechanistic studies to support the discussion section. However, this topic has not been fully studied and thus we have included studies which are relevant but we feel that further research on this topic should be undertaken to explore mechanisms underlying the observations that we found. We have added a directions for future research section as follows on line 351:

"Directions for future research:
Future studies that disentangle specific hormonal and molecular mechanisms that underlie the association demonstrated in our study will help us better understand our study findings. Understanding which specific fertility factors alter electrical remodeling in women is an important direction for future research"

R2 Comment 2: Note on statistical analysis: The concern here relates to the methods used to perform regression analysis: 1) Number of Pregnancies in the study is a categorical variable - how was this handled in the statistical analysis?

> Author Response R2 Comment 2: Thank you for this comment. We made the decision to categorize this variable based on its distribution and for ease of interpretation. There were relatively small numbers of women reporting 6,7 and 8+ pregnancies/ livebirths. We therefore analyzed the variable with the category "never pregnant" as the referent and provide a p value for linear trend across categories.

R2 Comment 3: RD has presumably been employed as a continuous variable (although this was not explained)? Yes it was continuous. However the RD has a very consistent median and narrow IQR across all study participants (36[8]). What is the implication of this narrow range for the RD predictor value in the regression analyses for predicted QTc and P/PR? Might RD have been better interpreted as a categorical variable? The authors should explain their choices and procedures so that these considerations are explicit.

> Author Response R2 Comment 3: Thank you for this comment. Given that in a linear regression analysis, the model can be used with an exposure even if it has a narrow range, we felt it was appropriate in this context to consider RD as a continuous variable. When we considered RD's distribution prior to running analysis, there was no reason to suggest that we should categorize.

Reviewer 3 (R3)

Reviewer Name: Ewoud Schuit
Institution and Country: Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, University Utrecht, Utrecht, The Netherlands
Competing Interests: none declared

This study is a secondary analysis of the WHI Trial which studied the association between number of pregnancies and reproductive period duration (RD, time from menarche to menopause) with electrocardiographic intervals. This is relevant as up till now only relatively small studies have suggested that there are small but measureable changes in electrocardiographic intervals during pregnancy. The current study confirms these changes with an increased PR interval and longer QTc interval for those with 5+ livebirths compared to 0 prior pregnancies. Reproductive period duration (RD) was associated with longer PR interval and maximum P wave duration among never users of HT. For every year increase in reproductive period, QTc decreased. The authors conclude that an increasing number of live births are related to increased ventricular repolarization time whereas RD is related to decreased ventricular repolarization time. Both longer RD and grandmultiparity are related to increased atrial conduction time. The premenopausal hormonal milieu appears to have effects on midlife cardiac electrical conduction system remodeling in women that may modestly influence CVD risk in later life.
In general, I think this is a well done and nicely presented study. I do have some comments/remarks, mostly related to things that weren't entirely clear to me.

Major comments

R3 Comment 1: Abstract. Conclusion: I think the conclusion needs some rewriting and I actually like the conclusion at the end of discussion better. E.g. it is stated that "An increasing number of live births are (IS?) related to increased ventricular repolarization time.", but the QTc interval varies from 0.66 to 0.15, to 0.25 to 1.15, and statistical significance was only found for 5+ livebirths.

Author Response R3 Comment 1: We thank Dr. Schuit for his insightful review. We have rewritten the conclusion to reflect a more accurate interpretation of our findings as suggested:

"Conclusions

We found that having an increasing number of five or more pregnancies compared to none is related to small but significant changes in atrial conduction time and ventricular repolarization time. A longer reproductive period duration in women not exposed to exogenous hormone therapy is related to a modest increase in atrial conduction time and to a modest decrease in ventricular repolarization. Reproductive health factors reflective of endogenous sex hormone exposure may be significant determinants of cardiac electrical remodeling in mid-life."

R3 Comment 2: Additionally, "hormonal milieu" is not properly introduced in the abstract, which makes that this statement seems to come out of nowhere. The same is true for "CVD risk in later life". It would be helpful if the authors could add some information to the objective section to make that the conclusion better links to the rest of the abstract and to make sure that the conclusion easily follows from the results presented.

Author Response R3 Comment 2: Thank you for these suggestions. We have changed the first sentence of the abstract to better reflect our study aims.

"Objective: Pregnancy, menses and menopause are related to fluctuations in endogenous sex hormones in women, which cumulatively, may alter cardiac electrical conduction. Therefore, we sought to study the association between number of pregnancies and reproductive period duration (RD, time from menarche to menopause) with electrocardiographic intervals in the Women's Health Initiative Clinical Trials.

R3 Comment 3: Methods. Study sample/missing data. Over 10,000 (roughly 20% of participants) were excluded because of missing data. Did the authors consider to use statistical techniques (e.g. multiple imputation) to account for these missing data? Not accounting for missing data may potentially lead to bias when the missingness of the data related to the exposures or outcomes of interest. Later on in the methods section the authors do indicate that they performed multiple imputation, however, as reported now, it is not clear to me whether the imputation relates to missing data in the 40,687 women or missings in the initial cohort from which 10,000+ women were excluded? If imputation was used anyway, why not impute data for the 10,000+ women that were excluded for having missing data?

Author Response R3 Comment 3: Thank you for this comment, we did impute data for the women with missing covariates (n=6,685).We now better clarify this in the methods section and have revised our study sample creation figure to reflect this as well and place the excluded women in revised Table 1. We did not want to include women with prevalent CVD or with missing ECG data and did want to impute the exposure or dependent variables.

R3 Comment 4: Methods. Statistical methods. Can the authors explain why they did not fit models in which both number of pregnancies and RD were included? I understand that the main interest is in these two exposures, but could there be a possibility that some of the results found for these exposures may be explained by the other exposure?

Author Response R3 Comment 4: Thank you for this comment and suggestion. We have run analyses where both variables are in the same model and have found that this did not affect our results. We now mention this in the secondary analysis and results sections respectively.

R3 Comment 5a: Discussion. I miss two things in the discussion. What do the results of this study mean for clinical practice? Especially considering that 5+ live births seems to have the biggest association, and 5+ live births becomes less and less common these days. I know some of this has

been described in the summary, but I think it would be helpful if it could be added to the discussion as well.

Author Response R3 Comment 5a: Thank you for this suggestion. We now add a section of about women reporting 5+ births in the context of other scientific data/ literature and declining prevalence of grandmultiparity.

"Cardiac electrical remodeling often reflects myocardial remodeling. We previously demonstrated that an increasing number of pregnancies were related to left ventricular volume increase and increase in left ventricular mass in a multiethnic cohort of women.(9) The increase in cardiac volume and mass were more marked in grandmultipara's or women who had 5 or more pregnancies leading to livebirths.(9) It is important to note that grandmultiparity is less common with declining parity levels in the United States."

R3 Comment 5b: what questions weren't answered by this study and what would be future research?

Author Response R3 Comment 5b: Thank you for this suggestion. We have now added a future research directions section. Please see our response to R2 Comment 1.

R3 Minor comments

Abstract

R3 Comment 6: Methods. Please indicate which outcomes were primary and which were secondary outcomes.

Author Response R3 Comment 6: We now clarify that all ECG parameters were the dependent variables in our analysis.

R3 Comment 7: Results: "HT" has not been defined before, please replace by "hormone therapy".

Author Response R3 Comment 7: We have made this change.

R3 Comment 8: Limitations: Recall bias of exposure is mentioned as a limitation. Please specify that this relates to the reproductive period duration and not number of pregnancies.

Author ResponseR3 Comment 8: Thank you for this suggestion we have made the following change in the abstract on line 63 and in the limitations section on line 342 as well :

***"Limitations:** Potential misclassification of RD due to participant recall"*

Methods.

R3 Comment 9: Study sample, line 145. I assume "HT" stands for "hormone therapy", but it has been introduced earlier in the introduction as "hormone trial".

Author Response R3 Comment 9: Thank you for this request for clarification. We have clarified this as follows:

"The WHI recruitment began in 1991 and consisted of a set of clinical trials/ and an observational study on hormone therapy, dietary modification and calcium/ vitamin D supplementation on cardiovascular disease, cancer and fractures.(2)"

R3 Comment 10: Can the authors provide some basic information about the women in the WHI, e.g. some broad in- and exclusion criteria, e.g. that they all reached menopause?

Author Response R3 Comment 10: We now added the following sentence and reference in the study sample section of the methods section in the manuscript:

"At the time of enrollment, all women enrolled in the WHI were required to be between 50 and 79 years old, postmenopausal, and intending to reside in the area for at least 3 years. Other enrollment criteria have been previously described.(3)"

R3 Comment 11: The authors state in lines 148-150 that "This analysis drew from the cohort of women enrolled in the clinical trial. Figure 1 shows the creation of the study sample. Of 68,132 women in WHI Studies (hormone therapy, diet and calcium/vitamin D and observational studies), " There are some inconsistencies in this section: ask for clarification of #s if needed. Earlier was the WHI introduced as "a set of clinical trials", but the now the authors state that the "analysis drew from the cohort of women enrolled in the clinical trial", not "trialS". I assume this is a typo. The same holds for observational study (in description of WHI) vs. observational studies in the quoted statement.

Author Response R3 Comment 11: Thank you for these comments. The WHI was indeed a set of 3 clinical trials and one observational study. ECGs were only done on women in the clinical trials cohort so that is where we drew our study sample. We now clarify this important point in the methods section starting on line 138 in the revised manuscript.

R3 Comment 12: The statement that "analysis drew from the cohort of women enrolled in the clinical trial" suggests that the data from the observational study/studies was not used, but in the quoted statement it seems that that data was used. Please clarify.

Author Response R3 Comment 12: This analysis drew from the cohort of women enrolled in the WHI clinical trials (and not observational study), as they had ECG's done for clinical trial participants. We have now clarified in this the methods section of the study.

R3 Comment 13: Electrocardiographic parameters, line 178. Typo, "additional" should be "addition".

Author Response R3 Comment 13: Thank you we have fixed this.

R3 Comment 14: Line 210. A sentence seems to be missing ("Further,.").

Author Response R3 Comment 14: Thank you we removed this typo.

R3 Comment 15: Multiple Imputation Analysis. I would suggest to use more imputations than the 5 currently used. Run this for 20 imputations.

Author Response R3 Comment 15: Thank you – we ran 20 imputations in a secondary analysis and it did not affect our findings therefore we clarify this in the methods section.

Results:

R3 Comment 16: QTc. The authors state that "For each additional year in reproductive period duration, there was a 0.4 ms shorter QTc (Table 3)." Please explain what "additional year" means

here. Additional to what? I assume the lowest value of RD in that data. What was that value? Another way to deal with this would be to standardize RD (e.g. by subtracting the lowest value or the mean RD from all RD value) and refit the model. (no matter what one compares this to).

> Author Response R3 Comment 16: There is no referent group since RD is continuous. The mean reproductive duration and the standard deviations are presented therefore readers will be able to contextualize these results. We do not want to standardize RD because then we would not be able to interpret this in terms of years, which is reachable to most audiences.

Tables & Figures

Reviewer 3 Comment 17: Table 1. I think it would be helpful for the readers to know the baseline characteristics of the population as a whole, to allow comparison to other cohorts. Can the authors add an additional column to the table presenting these study characteristics? Add a column with everyone.

> Author Response R3 Comment 17: Thank you for this suggestion- we now include everyone in table 1- excluded and included.

Reviewer: 4

Reviewer Name: Dr Victoria Allgar
Institution and Country: University of York, England
Competing Interests: None declared

The aim is to determine if there is a positive or negative association between number of pregnancies and reproductive period duration with mid-life electrocardiogram intervals (PR interval and QTc) and p wave parameters (p wave 139 maximum duration, dispersion and index).

Reviewer 4 Comment 1: The introduction should include definitions of the dependent variables: PR interval, P wave indices (duration and dispersion) and QTc from enrollment electrocardiogram. e.g. The PR interval is the time from the onset of the P wave to the start of the QRS complex,

> Author Response R4 Comment 1: Thank you for this suggestion-we now include these definitions in the introduction as you have suggested starting on line 87.

Reviewer 4 Comment 2: Data from the Women's Health Initiative Hormone Trial was analyzed for the study and a consort flow diagram is included. However in the flow chart it stats that 6,685 were further excluded for having missing covariate data, leaving a final sample of 40,687. Yet in the analysis plan "multiple imputation techniques to impute missing covariates", so i am unclear with regards to the final numbers included and this needs to be clarified.

Author Response R4 Comment 2: Thank you for this comment. We now further clarify this point and have redrawn our flow diagram. We now include table 1 with the total n of women included and excluded due to missing covariates in order to further clarify. The multiple imputation analysis was a sensitivity analysis and therefore the n=40,687 was the n for the primary analysis and 54,057 was the n for the sensitivity analysis. Our revised figure better clarifies this.

Reviewer 4 Comment 3: There is no summary statistics for PR interval, QTc, P wave duration and dispersion e.g. mean (sd) or any test of normality. For the regression models were the residual plots investigated so you can trust the results and what are the measures used for the goodness-of-fit and R-squared?

> Author Response Reviewer 4 Comment 3: Thank you for these comments. We now place descriptive statistics for the ECG indices in Table 1 at the bottom of the table. In terms of normality and distribution of exposure variables, we did look at the distributions of the

variables and because of the large size of our cohort- in our analyses, we will most always reject normality. The exposure variable distributions did not diverge greatly from normality despite this and therefore our models were robust and appropriate for this data.

Reviewer 4 Comment 4: In Table 2 (incorrectly labelled Table 1), Table 3 and Table 4 it would be useful to include summary statistics for PR interval, QTc and P wave duration and dispersion for the categories for number of live births and reproductive period duration.

Author Response Reviewer 4 Comment 4: Thank you for this suggestion- we now display these in table 1.

Reviewer 4 Comment 5: For PR interval and QTc the p-values should be included in the text or table.

Author Response Reviewer 4 Comment 5: We have included the p values for trend in each table including Table 2 and 3 (for PR and QTc respectively).

Reviewer 4 Comment 6: For P wave duration and dispersion Table 5 is confusing - were individual models use for each subgroup or is this the same as for PR interval and QTc?

Author Response Reviewer 4 Comment 6: Thank you for this comment. I have now clarified this by renaming this table to clarify and also standardizing dependent variable names with preceding tables.

1. Gold EB, Crawford SL, Avis NE, Crandall CJ, Matthews KA, Waetjen LE, Lee JS, Thurston R, Vuga M, Harlow SD. Factors related to age at natural menopause: longitudinal analyses from SWAN. American journal of epidemiology. 2013;178(1):70-83. Epub 2013/06/22. doi: 10.1093/aje/kws421. PubMed PMID: 23788671; PMCID: PMC3698989.
2. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. Controlled clinical trials. 1998;19(1):61-109. Epub 1998/03/11. PubMed PMID: 9492970.
3. Hays J, Hunt JR, Hubbell FA, Anderson GL, Limacher M, Allen C, Rossouw JE. The Women's Health Initiative recruitment methods and results. Annals of epidemiology. 2003;13(9 Suppl):S18-77. Epub 2003/10/25. PubMed PMID: 14575939.

VERSION 2 – REVIEW

REVIEWER	Professor Michael Lewis Swansea University, UK
REVIEW RETURNED	23-Apr-2018

GENERAL COMMENTS	No further comments.

REVIEWER	Prof dr Angela Maas Radboud University Medical center, Nijmegen, the Netherlands
REVIEW RETURNED	27-Apr-2018

GENERAL COMMENTS	excellent revision

REVIEWER	Dr Victoria Allgar University of York, England
REVIEW RETURNED	01-May-2018

GENERAL COMMENTS	In terms of the number of pregnancies, can you clarify why the categories were chosen e.g. Never pregnant, none, 1, 2-4 and 5+. The 2-4 group represents 26599 women, whereas the other groups are much smaller. I was a little unclear of definition - Never pregnancy vs none or live births. I presume 'none' are those who did get pregnant but had no live birth (miscarriage, abortion, still birth) - but these could be multiple pregnancies which may explain some of the findings. I would like to have seen summary statistics for the dependant variables (e.g. PR) split by number of births and hormone use etc. This would allow the reader to see the magnitude of differences and whether clinically significant. The statement "5+ live births versus 0 prior pregnancies was associated with a 1.32 ms increase in PR interval [95% CI (0.25, 2.38)]." Did the authors consider other comparisons here e.g. 2-4 vs 5? 'None' was associated with a 1.15 ms increase in PR interval versus 0 prior pregnancies, not significant but worth discussion.e.g. there was no graded association with PR, nor with the QTc interval (as stated in the text) - which could be due to the classifications discussed above with regards to the 'none' group and grouping 2-4 together. In Table 2 the Reproductive period duration is stated as a continuous variable in years but then categorised with no reference group. Can this be clarified? The discussion should set the findings in context e.g. is 1.32 ms increase in PR interval clinically significant or too small a difference to be important - The sample sizes are very large, so small differences will be statistically significant but may not be clinically significant.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 4
Reviewer Name: Dr Victoria Allgar
Institution and Country: University of York, England
Competing Interests: None declared

Reviewer 4 Comment 1: In terms of the number of pregnancies, can you clarify why the categories were chosen e.g. Never pregnant, none, 1, 2-4 and 5+. The 2-4 group represents 26599 women, whereas the other groups are much smaller. I was a little unclear of definition - Never pregnancy vs none or live births. I presume 'none' are those who did get pregnant but had no live birth (miscarriage, abortion, still birth) - but these could be multiple pregnancies which may explain some of the findings.

Author response R4C1: Thank you for these important comments which we agree need to be clarified. In the text and tables we now refer to the “none” category as follows: none (pregnant, no livebirths). We have now added the following section to the methods section which we hope will better clarify the choice of categories (beginning on line 166):

“In order to be able to also study women who had not experienced pregnancy and/or childbirth and in an effort to make our study as representative as possible, we separately categorized women who had had no prior pregnancies and women who had experienced a pregnancy but no livebirths (i.e. due to miscarriage, stillbirth, or abortion) as separate categories. We further categorized women based on our prior work demonstrating that having 5 or more pregnancies was associated with greater cardiac remodeling.(9) Due to small cell sizes we combined women with 5 or more pregnancies leading to livebirths into one category. Preliminary data analysis reflected that 2-4 had similar effects sizes for PR and QTc and thus these categories were collapsed into a single category for ease of interpretation. Therefore the exposure categories for number of pregnancies leading to livebirths were as follows: no pregnancies (referent), none (pregnant, no livebirths), 1, 2-4, 5 or more.”

R4 Comment 2: I would like to have seen summary statistics for the dependant variables (e.g. PR) split by number of births and hormone use etc. This would allow the reader to see the magnitude of differences and whether clinically significant.

Author Response R4C2: Thank you for this comment. We recognize that these tables would give readers an idea about unadjusted differences between groups. However, magnitude of differences for unadjusted comparisons aren't necessarily informative when the relationship between outcome and predictors are confounded, and thus the inclusion of these tables might encourage readers to draw misleading conclusions. We do add a section to the discussion regarding clinical significance to better convey this to readers which highlights the modest nature of our findings with respects to clinical significance (please see our response to Comment 5).

R4 Comment 3: The statement "5+ live births versus 0 prior pregnancies was associated with a 1.32 ms increase in PR interval [95% CI (0.25, 2.38)]." Did the authors consider other comparisons here e.g. 2-4 vs 5? 'None' was associated with a 1.15 ms increase in PR interval versus 0 prior pregnancies, not significant but worth discussion.e.g. there was no graded association with PR, nor with the QTc interval (as stated in the text) - which could be due to the classifications discussed above with regards to the 'none' group and grouping 2-4 together.

Author Response R4C3: Thank you- we appreciate this feedback aimed at making our paper more interpretable and clear to the audience/ readership. We did not chose to discuss the non-significant finding for the “None” category as there were several tests that we conducted in our study and this represented a non-significant finding. We now clarify our choice of categories in the methods section (please see response to R4C1).

R4 Comment 4: In Table 2 the Reproductive period duration is stated as a continuous variable in years but then categorised with no reference group. Can this be clarified?

Author Response R4C4: Thank-you for this comment. We now appreciate how Table 2 can appear confusing the reader. Essentially the reproductive period duration estimates are given by strata of HT since there was statistically significant effect modification of the association between reproductive duration and PR interval by HT use. Therefore it was most methodologically appropriate not to present the pooled estimate but rather estimates within each strata of HT use. We have added a row to Table 2 that clarifies this and also added a similar row to table 5 that clarifies the displaying of data in HT use strata.

R4 Comment 5: The discussion should set the findings in context e.g. is 1.32 ms increase in PR interval clinically significant or too small a difference to be important - The sample sizes are very large, so small differences will be statistically significant but may not be clinically significant.

Author Response R4C5: Thank you for raising this important point/ suggestion for additional discussion. We have added a section to the discussion section which specifically discussed potential clinical relevance and provide it here for your review (beginning on line 354):

“Clinical relevance of our findings.

The PR interval normally ranges from 120 to 200 ms in duration. Therefore our finding that having 5 or more livebirths versus never having been pregnant was associated with an adjusted increase in PR interval of 1.32 ms, has modest clinical significance. For an individual with a PR interval at the upper limits of normal, 1.32 ms may be more clinically relevant in terms of the increased risks of later cardiovascular diseases with PR >200 ms.(1) The association of number of pregnancies leading to livebirths with QTc (with 5 or more pregnancies leading to livebirths having a 1.15 ms increase in QTc compared to nulligravid women) is similarly modest with a normal QTc ranging from ~350 to 460 ms in women. The effect sizes for reproductive duration were even more modest in size than those for P wave indices and therefore likely have more relevance in terms of uncovering novel biologic mechanisms related to cardiac electrical remodeling rather than reflecting clinically significant differences among individuals.”