PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The association between dementia parental family history and midlife modifiable risk factors for dementia: a cross-sectional study using propensity score matching within the Lifelines cohort
AUTHORS	Vrijsen, Joyce; Abu-Hanna, Ameen; de Rooij, Sophia; Smidt, Nynke

VERSION 1 – REVIEW

REVIEWER	Yokomichi, Hiroshi University of Yamanashi, Department of Health Sciences
REVIEW RETURNED	25-Mar-2021

GENERAL COMMENTS	Paper by Doctor Vrijsen J et al. treated a cross-sectional study of history of parental dementia and midlife risk factors of dementia. I would like to provide comments to improve the manuscript. [Major] 1. Table 3: With vs. without propensity score matching, odds ratios are still different. This would affect interpretation of the study results. Is propensity score matching unnecessary? 2. Table 3: Without propensity score matching, the results in imputed data are slightly different from those in non-imputed data. This may confuse the readers. Is imputation necessary, when the researchers have data of enough number of participants? 3. Details of imputation method need to be disclosed. I mean that the setting and the options of software that the researchers selected need to be described. 4. Main results: I wonder which results were main in this study. I mean that with imputation and propensity score matching, the data may be greatly processed. The process may bend the odds ratios and the interpretation. 5. I wonder how the explanatory variables for propensity score mating were selected. 6. I could not follow the methodology descriptions. Could I ask why propensity score matching and logistic regression were simultaneously adopted for the outcome of parental history of dementia, please? I wonder why the two statistical analysis were needed to investigate the study question. 7. Discussion is needed on the difference of odds ratios between observed data vs. imputed data and data without PSM vs. with PSM. [Minor] 8. Introduction section may be relatively lengthy. As a reader, I would like to read the rationale of the study compactly. 9. Independent variables, Methods section: Representing concrete question and the answer may not necessary. Or the researchers

	could describe it abstractly. For example, following description may not be necessary: (i) 'yes' (1 = having a parent with dementia) and (ii) 140 'no' (0 = not having a parent with dementia). 10. Measurement of independent and dependent variables, Methods section may be lengthy. This part could be shrunk for more readability. Overall, I would like to request the researchers to describe the necessity and appropriateness of methodologies they selected. Discussion on the varied ORs would also be needed.

REVIEWER	Barber, Philip University of Calgary, Neuroscience
REVIEW RETURNED	29-Apr-2021

GENERAL COMMENTS	This large community-based study recruited individuals with parental family history of dementia and established a relationship with increased risk of dementia regardless of genetic risk factors. They found a relationship between parental family history currently established modifiable vascular risk factors among middle-aged individuals using propensity score matching. The results are of interest and the manuscript is succinctly written. The discussion is well-balanced citing the strengths and limitations of the study design from an informed position. One limitation that the authors do not mention is that they do not collect data regarding the age of onset of dementia of the father or mother. This might be important as dementia is a strongly age-related condition and therefore if individual parents were older they were more likely to develop dementia. In contrast, the relationship of early onset of dementia may have a stronger genetic basis. I invite the authors to respond to this comment.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

Point 1: Table 3: With vs. without propensity score matching, odds ratios are still different. This would affect interpretation of the study results. Is propensity score matching unnecessary?

Response 1: The reviewer questions whether propensity score matching is necessary. Perhaps, we have not been very clear in our description, so thanks for the opportunity to clarify and improve the manuscript. In order to investigate the association between having a parental family history of dementia and several modifiable risk factors for dementia, we firstly conducted regression analyses with standard covariate adjustment. However, we found that the two groups (with and without a PFH of dementia) were not similar in several baseline characteristics, especially in age. This means that the model will have to extrapolate to areas that have no proper support in the data. Since age is an important risk factor for dementia, participants with a PFH of dementia were often older and had therefore more often hypertension and high cholesterol levels. By using covariate adjustment, there is the threat that this confounding bias was not tackled sufficiently. By using propensity score matching, a greater proportion of this bias could be eliminated (Austin et al. 2011). Also, we were able to check whether the systematic differences in characteristics of individuals with and without a PFH of dementia were eliminated by assessing the standardized mean differences (SMDs). Since the results without propensity score matching are not yet adjusted for confounding variables, it was actually expected that the estimates would be different before and after propensity score matching.

We have added the following sentences to the Introduction section (Page 6, Lines 111-114):
Since age is an important risk factor for dementia, individuals with a PFH of dementia are often older and could therefore have more often modifiable risk factors for dementia, such as hypertension and high cholesterol levels (9). By using covariate adjustment, there is the threat that this confounding bias is not tackled sufficiently.

Point 2: Table 3: Without propensity score matching, the results in imputed data are slightly different from those in non-imputed data. This may confuse the readers. Is imputation necessary, when the researchers have data of enough number of participants?

Response 2: Although the results may differ only slightly between the observed and imputed data, missing data was imputed to be able to match individuals with a PFH of dementia with individuals without a PFH of dementia based on their propensity score. In addition, multiple imputation does not only contribute to the size of the dataset but also to the reliability of the results, since not all data of participants had to be excluded due to missing values on one or a little number of variables.

Point 3: Details of imputation method need to be disclosed. I mean that the setting and the options of software that the researchers selected need to be described.

Response 3: We thank the reviewer for this comment and have elaborated more on the details of the imputation method.

We have replaced text (1) by text (2):
(1) Five imputed datasets were generated to replace missing values, using Multiple Imputation using Chained Equations (MICE).

(2) Five imputed datasets were generated to replace missing values, using the Multiple Imputation using Chained Equations (MICE) approach. Specifically, we used predictive mean matching (ppm) for continuous data, logistic regression imputation (logreg) for binary data, polytomous regression imputation (polyreg) for unordered categorical data and proportional odds model (polr) for ordered categorical data.

Point 4: Main results: I wonder which results were main in this study. I mean that with imputation and propensity score matching, the data may be greatly processed. The process may bend the odds ratios and the interpretation.

Response 4: In this study, we used multiple imputation and propensity score matching to overcome systematic differences between individuals with and without a PFH of dementia. Therefore, as mentioned in the Results section, we focused on the results of the final model (model 2) in which the propensity score was based on the age, sex, educational level and other modifiable risk factors for dementia. It was within our expectation that the odds ratios in model 2 (and 1) would change after propensity score matching, since systematic differences between individuals with and without a PFH of dementia would be reduced or even eliminated by using this method. In sum, the change in odds ratios testifies for the need of using propensity score matching.

Point 5: I wonder how the explanatory variables for propensity score matching were selected.

Response 5: We thank the reviewer for this comment and have elaborated more on the selection of explanatory variables for the propensity matching score. The propensity score was based on the confounding variables age, sex and educational level (model 1) and other potential confounders (model 2). The other potential confounders in model 2 were carefully a-priori selected per outcome measure in a consensus meeting, in which each potential confounder had to be associated with both the independent (having a parental family history of dementia) and dependent variables (modifiable risk factors for dementia). For example, the most important confounding variable was the age of the participant, as the older the participant the greater the possibility that they have a parent with dementia. Also, the older the participant the higher the chance that they suffer from hypertension, cardiovascular diseases or renal dysfunction etc.

We have added the following sentence to the Method section (Page 12, Lines 287-289): The other potential confounders were a-priori carefully selected per outcome measure in a consensus meeting, in which each potential confounder had to be associated with both the independent and the dependent variables.

Point 6: I could not follow the methodology descriptions. Could I ask why propensity score matching and logistic regression were simultaneously adopted for the outcome of parental history of dementia, please? I wonder why the two statistical analysis were needed to investigate the study question.

Response 6: In order to eliminate the systematic differences between individuals with and without a PFH of dementia, propensity score matching (PSM) was used (in each imputed dataset) before the regression analyses were conducted. First, each individual with a PFH of dementia was matched with one individual without a PFH of dementia based on their propensity score. After PSM the balance in confounding variables, such as age, was improved and regression analyses could be conducted at this stage in order to obtain the odds ratios. In summary, propensity score matching is an adjustment method to eliminate bias and regression analyses were used to calculate the estimates. In this sense, they are not simultaneously used in the meaning of “parallel” but rather in two different stages of the same analysis.

Point 7: Discussion is needed on the difference of odds ratios between observed data vs. imputed data and data without PSM vs. with PSM.

Response 7: We thank the reviewer for this suggestion. In our opinion, it might not be useful to compare the odds ratios with and without PSM, since the odds ratios without PSM were not yet corrected for confounding factors. However, we agree that it might be interesting to elaborate more on the differences between the results after PSM and the sensitivity analyses in supplementary file 4 (covariate adjustment on imputed data).

We have added the following sentences to the Discussion section (Page 16, Lines 376-380):
In comparison to the main analyses, the estimates for physical inactivity and social activity are slightly smaller in the sensitivity results. This could be explained by the smaller sample size in the main results (n=53,644 versus n=89,869). Due to one-to-one matching, a relatively high number of healthy living individuals with a PFH of dementia could not be matched and therefore not included in the main analyses.

[MINOR]
Point 8: Introduction section may be relatively lengthy. As a reader, I would like to read the rationale of the study compactly.

Response 8: We thank the reviewer for this comment and have shortened the rationale of the study.

We have replaced text (1) into text (2):
(1) Since the world's population is ageing, the total number of people with dementia will increase (1). In 2019, around 50 million people were living with dementia worldwide and the number of people with dementia is expected to increase to 152 million by 2050 (2). Dementia affects not only the individual living with dementia, but also their family, caregivers and society as a whole (2).

(2) Since the world's population is ageing, the total number of people with dementia will increase (1). In 2019, around 50 million people were living with dementia worldwide and the number of people with dementia is expected to increase to 152 million by 2050 (2).

(1) The majority of these risk factors were combined in the Lifestyle for Brain Health (LIBRA) score, reflecting someone's potential for dementia risk reduction (DRR) (8,11–12). The predictive accuracy of the LIBRA score was examined and results showed that higher LIBRA scores (presence of more risk factors) were associated to dementia in middle-aged individuals (55–69 years) (HR=1.10, 1.02–1.18)(12), but not in very old individuals (84-102 years) (HR=0.93, 0.83-1.05) (13).

(2) The majority of these risk factors were combined in the Lifestyle for Brain Health (LIBRA) score, reflecting someone's potential for dementia risk reduction (DRR) (8,11–13).

Point 9: Independent variables, Methods section: Representing concrete question and the answer may not necessary. Or the researchers could describe it abstractly. For example, following description may not be necessary: (i) 'yes' (1 = having a parent with dementia) and (ii) 'no' (0 = not having a parent with dementia).

Response 9: We thank the reviewer for this comment and have now described the two categories of the independent variable more abstractly.

We have replaced text (1) into text (2):
(1) Participants could indicate whether their father and/or mother had dementia. This variable was dichotomized into: (i) 'yes' (1 = having a parent with dementia) and (ii) 'no' (0 = not having a parent with dementia).

(2) Participants could indicate whether their father and/or mother had dementia. This variable was dichotomized [yes/no].

Point 10: Measurement of independent and dependent variables, Methods section may be lengthy. This part could be shrunk for more readability.

Response 10: We thank the reviewer for this suggestion to improve the readability and at the same time attempted to be clear about the operationalization of these dependent variables, since this can have influence on the estimates. We have shortened the description of the measurement of a healthy diet.

We have replaced text (1) by text (2):
(1) A quantitative Food Frequency Questionnaire (FFQ) was used to assess dietary intake over the previous month (43,44). The Mediterranean diet was associated with slower cognitive decline (45–

47), however not all food groups of the Mediterranean diet were measured within the Lifelines population on baseline. Therefore, the Lifelines diet score (LLDS) was used to determine adherence to a healthy diet, which includes most food groups of the Mediterranean diet. The LLDS was based on the consumption of nine positive food groups (vegetables, fruit, whole grain products, legumes and nuts, fish, oils and soft margarines, unsweetened dairy, coffee and tea) and three negative food groups (red and processed meat, butter and hard margarines and sugar-sweetened beverages).

(2) A quantitative Food Frequency Questionnaire (FFQ) was used to assess dietary intake over the previous month (43,44). Subsequently, the Lifelines diet score (LLDS) was used to determine adherence to a healthy diet, which is based on the consumption of nine positive food groups (vegetables, fruit, whole grain products, legumes and nuts, fish, oils and soft margarines, unsweetened dairy, coffee and tea) and three negative food groups (red and processed meat, butter and hard margarines and sugar-sweetened beverages).

Reviewer 2

Point 1: This large community-based study recruited individuals with parental family history of dementia and established a relationship with increased risk of dementia regardless of genetic risk factors. They found a relationship between parental family history currently established modifiable vascular risk factors among middle-aged individuals using propensity score matching. The results are of interest and the manuscript is succinctly written. The discussion is well-balanced citing the strengths and limitations of the study design from an informed position.
One limitation that the authors do not mention is that they do not collect data regarding the age of onset of dementia of the father or mother. This might be important as dementia is a strongly age-related condition and therefore if individual parents were older they were more likely to develop dementia. In contrast, the relationship of early onset of dementia may have a stronger genetic basis. I invite the authors to respond to this comment.

Response 1: We thank the reviewer for this useful comment and we agree that the age of onset is an important factor to take into account. Therefore, we have added this in the limitation section of the Discussion.

We have added the following sentences to the Discussion section (Page 17, Lines 418-423): Also, we did not take into account the age of onset of dementia of the parent(s), since the average age of onset of dementia differs between types of dementia (63). However, this might be an important effect modifier as early onset dementia may have a stronger genetic basis. Therefore, these results could be an underestimation of the results for individuals with a parent diagnosed at an older age. Nevertheless, after excluding individuals with a parent diagnosed before the age of 70 years, the results were similar.

VERSION 2 – REVIEW

REVIEWER	Yokomichi, Hiroshi University of Yamanashi, Department of Health Sciences
REVIEW RETURNED	15-Sep-2021

GENERAL COMMENTS	The researchers have addressed all of my concerns. I have no more concern. I appreciate their efforts to report the important results.

REVIEWER	Barber, Philip

	University of Calgary, Neuroscience
REVIEW RETURNED	15-Sep-2021

GENERAL COMMENTS	Thank you for responding to my comment and revising your manuscript.