Reply to the Editor and Reviewers
Author's Reply: by Chris Chambers, 19 Dec 2022 10:53 Manuscript: [https://osf.io/ur9v4 | https://osf.io/ur9v4] version v1 Major Revision

I have now received two very detailed and constructive reviews of your submission. As you will see, the evaluations are broadly encouraging while also raising a variety of issues that will need careful attention to achieve Stage 1 in-principle acceptance (IPA). Most of the reviewers' comments seek more detailed clarification and justification of specific design features and analysis choices, but there are also some key points to consider concerning the study rationale.

In my own reading, I noted two additional points:

1. Bias control level

In the section "Existing data", you state: "At the time of this stage-1 protocol, we have access to the baseline anthropometric and medical data and have preprocessed and quality-controlled the imaging data of both time points. We have not gained access to the follow-up anthropometric, medical and cognitive data and have not explored any associations of these measures with WML volume beyond the baseline investigations cited above." In Q8 of the submission checklist you selected Level 3 for the bias-control level: "At least some data/evidence that will be used to the answer the research question has been previously accessed by the authors (e.g. downloaded or otherwise received), but the authors certify that they have not yet observed ANY part of the data/evidence [Level 3]" Level 3 requires that while authors have accessed the data, they have not observed it (in any way). Based on your description from the Method, however, it appears that you have observed some of the data and undertaken preprocessing. Therefore, I believe Level 2 is the most appropriate level: "At least some data/evidence that will be used to answer the research question has been accessed and partially observed by the authors, but the authors certify that they have not yet observed the key variables within the data that will be used to answer the research question AND they have taken additional steps to maximise bias control and rigour (e.g. conservative statistical threshold; recruitment of a blinded analyst; robustness testing, multiverse/specification analysis, or other approach) [Level 2]" At the same time, I also think the risk of additional bias resulting from your prior analyses is minimal, so I would set aside the normal requirement for Level 2 that you take take "additional steps to maximise bias control and rigour" on top of what you have already proposed. If you agree with this assessment then when you submit your revision please alter the checklist entry for Q8 to select Level 2, and please state the assigned level in the "Existing data" section of the Method. If you disagree with my assessment then please include a rebuttal in your response letter.

We agree with this point and have changed the bias control level to Level 2 in the checklist and added it in the “Existing data” section on p. 5.

2. Design table and page numbers Please reduce the font size in the study design table and present it in landscape format so that it is more digestible to readers. Please also ensure that your revised submission includes page numbers.

We adopted the above suggestions for better readability.

Overall, I think your study is promising candidate for IPA and I look forward to receiving your revised submission and point-by-point response to the reviews in due course. Please note that PCI RR is currently closed to all submissions (including revisions) until 3rd January.

Reviews Reviewed by Max Elliott, 16 Dec 2022 23:43

The authors plan to investigate whether blood pressure and waste-to-hip ratio influence the progression of cerebral small vessel disease. Overall, the research question and hypothesis are clear and fit into the literature in the field. The analysis plan is generally thorough, and the power analysis is well-described and justifies the planned statistical tests. Below I describe several questions and limitations I found while evaluating the registered report

Intro

1. In the introduction the authors mention that they will replicate findings and extend findings, however, it is unclear which aims, and hypotheses refer to replications versus extensions. It would be useful to make this explicit. What is a replication and what is an extension or novel aim?

We thank the reviewer for pointing out the need to be more precise. We have changed the introduction to make it more explicit that we replicate previous findings (on blood pressure) with the first three hypotheses and extend previous results (for obesity and gender -differences) with the exploratory analyses.

Aims and Hypotheses

2. H1, H2, and H3 seem to all be replications of previous research, is that right? If the primary goal of this project is to replicate previous results, then this is fine, but the researchers should make it more explicit in their introduction that the primary goal of this project is replication. However, based on the introduction it seems the researchers have more clear goals to ask novel questions about the influence of gender and Waist to hip ratio on longitudinal white matter lesions. If so, if so, I encourage the authors to explicitly state hypotheses about the waist-to-hip ratio and gender analyses.

This is correct, H1 - H3 refer to replication analyses. We have changed the introduction to make this more explicit where it now reads on p. 3: "Here, we therefore aim to replicate previous findings on the relationship of higher blood pressure, more WML progression and worsening of cognitive function in a large cohort of population-dwelling older adults. In exploratory analyses we aim to extend these findings towards abdominal obesity, a risk factor which has been understudied in longitudinal designs. We will explore gender-by-risk factor interactions for WML progression and gender-by-WML progression interaction for cognitive outcomes."

Regarding the effect of baseline waist-to-hip ratio on WML progression, as well as the effects of change in diastolic blood pressure and waist-to-hip ratio, our power analyses showed that we do not have sufficient power to detect the hypothesized effects. Therefore, during the preparation of the registered report, we moved these hypotheses from the confirmatory analyses to the exploratory section. Following the reviewer's comment, we have now explicitly stated the hypotheses E1a – E1c in the exploratory section and added the information that these analyses might be underpowered.
The gender-specific analyses were originally planned to only provide information on regression coefficients for both genders and not test the interaction. This was because we lack appropriate data for conducting power analyses for the interactions. Yet, we have hypotheses about the effects and agree with the reviewer that it would be interesting to test them, even though we may be underpowered. We now added these hypotheses E2a – E3b to the exploratory analyses section.

3. Why did the authors select these covariates for their hypotheses? These covariates seem reasonable, but I encourage the authors to justify each covariate and explain why its inclusion is important to each hypothesis/analysis.

We thank the reviewer for this comment and have added a detailed explanation of all covariates in the statistical analysis section where it now reads on p.15:
Explanation of covariates (M1)

Age_baseline: effect of age at baseline
Age_change: effect of passed time between baseline and follow-up (progression)
DBP_baseline: effect of baseline DBP
DBP_baseline: modifying effect of baseline DBP on progression of WML between baseline and follow-up (effect of interest for H1)
DBP_change: effect of change in DBP between baseline and follow-up on WML progression (effect of interest for E1c)
WHR_baseline effect of baseline WHR
WHR_baseline:Age_change: modifying effect of baseline WHR on progression of WML between baseline and follow-up (effect of interest for E1a)
WHR_change: effect of change in WHR between baseline and follow-up on WML progression (effect of interest for E1b)
Gender: adjust for gender (no power analyses possible for gender/sex interaction, therefore we control for it in confirmatory analyses)
HT_medication: adjust for hypertension medication as this probably influences the effect of DBP on WML progression
TIV: total intracranial volume, trivially linked with WML volume".

Methods

The methods are clear and logical. They are sufficiently detailed for replication and to limit analytical flexibility. I appreciate their thorough review of previous literature to

justify parameters used in their power analyses

4. How did the authors choose p<.033 for establishing statistical significance? Is this a field convention, or based on another justification?

We established this threshold based on the following assumptions: We have one-sided hypotheses but the p-values we obtain come from two-sided tests. We thus adjust the α-level ($\alpha_{TwoSided} = 2 * \alpha_{OneSided}$). We apply Bonferroni-Correction for three confirmatory hypotheses to the α-level ($\alpha_{OneSided} < 0.05/3$). This results in $\alpha_{TwoSided} < 2*0.05/3 < 0.033$ as threshold for the two-sided p-values we obtain from the linear mixed models (see p. 20).

Statistical Analyses

5. In the abstract and intro, sex/gender stratified analyses are discussed as a central focus of this project and these are included in the statistical models. However, I do not see a clear description of how sex/gender inferences will be made or evaluated from these models. Given the introduction, I encourage the authors to thoroughly describe their plans for evaluating and drawing inferences based on sex/gender.

We thank the reviewer for the encouragement. In the first version of the report, we decided to report gender-stratified results and not test the interaction because we could not perform power analyses due to missing references in the literature. Yet, as becomes clear from the introduction, we do have specific hypotheses about these interactions.
We have now added hypotheses E2a, E2b, E2c, E3a and E3b to the exploratory analysis section and investigate potential interactions of gender, vascular risk factors, WML progression and cognition.

Miscellaneous

6. I would highly recommend using fewer acronyms. It is very hard to track what each acronym is, and this makes reading the text burdensome. A few acronyms for the most common and well-known terms are fine, but currently, there is an acronym in nearly every sentence. SBP is not defined as far as I can tell.

We have removed rarely used acronyms and hope to have improved the reading flow. We added the definition of DBP in the beginning of the introduction.

Reviewed by anonymous reviewer, 14 Dec 2022 10:14
In this proposal Beyer et al. have designed a study aimed at assessing the link between blood pressure and progression of white matter hyperintensity load, as well as examining the cognitive implications of white matter hyperintensity progression. Overall, the research questions are pertinent and valid, the methodology is robust, and the transparency of the research plan is satisfactory. I would like to have some clarifications on certain points that are outlined below.

1A. Scientific validity of the research question(s) This proposal uses a longitudinal design to address the prospective effects of blood pressure and abdominal obesity on the progression of white matter hyperintensities. Additionally, it tests for the functional implication of increases in white matter hyperintensity load, since it also examines its associations with longitudinal decline in executive functions and in general cognition. The research question is valid and it is clearly defined. The proposal includes both exploratory and hypothesis-based analyses.

1B. Logic, rationale, and plausibility of the proposed hypotheses:
This proposal states three main hypothesis, which are logic and plausible.

1C. Soundness and feasibility of the methodology and analysis pipeline:
Methodologically speaking, the current proposal is sound and feasible. The authors have performed a sophisticated statistical power analysis where they take into account previous results from the literature and they simulate their predictive power in their dataset, according to their baseline information.

I was missing some clarification regarding certain aspects of the proposal. Namely:

1. Why do the authors focus on systolic blood pressure instead of diastolic blood pressure?
We thank the reviewer for this extremely valuable comment. We used systolic blood pressure assuming that both measures predicted WML progression equally well. Yet, after going through the literature again, we realized that evidence points to a more pronounced effect of diastolic blood pressure, especially in older individuals [1]. Therefore, we changed our analysis to diastolic blood pressure and included a justification in the introduction where it now reads: "While both systolic and diastolic blood pressure (DBP) are important predictors, effects seem to be more pronounced for DBP [1]." We kept the power analyses unchanged because more data was available for systolic blood pressure and effects for DBP should be similar or stronger than for SBP

2. The longitudinal data has already been collected but the authors do not have access yet. From the information that the authors provide, I assume that the follow-up range should be somewhere between 3 and 10 years. Is this correct? Do the authors happen to have a mode accurate description of this?
This is correct. The mean time to follow-up is 6 years (SD=1.9 years), the minimal time between scans is 4, the maximal time 9 years.

3 Regarding the confirmatory analyses, the authors will test 3 regression models that are aligned with the hypotheses. The summary provided in Table 7 highlights what predictors are expected to be significant regarding their hypotheses. I am a bit confused regarding the number of predictors and the highlights. The authors state somewhere earlier that the analyses will be divided by sex/gender. Then, why have they included sex as a predictor variable in the regression models?
In the confirmatory analysis section, we adjusted for gender instead of including an interaction with it because we could not perform a power analysis for gender-specific effects. Therefore, we now added these interaction analyses in the

exploratory section (E2a - E3b). Please also see response to question 5 of reviewer 1.

4 Interactions between variable of interest and age change. Could the authors state a bit clearer what is the rationale for including these interactions?
The coefficient of age change represents the amount of change in the outcome variable between time points in the linear mixed model. Therefore, the interaction of baseline risk factor and age change assesses the modifying effect of the baseline risk factor on the WML progression.

5 Please correct me if I am wrong, in Table 7, M2 and M3 are missing the interaction term hyperintensity load x age change while this is stated in the text.
Models M2 and M3 only model the association of WML progression and cognitive function, without taking into account vascular risk factors. This is to make the model simpler, but ignores additional effects vascular risk factors could have on cognition through different mechanisms than vascular brain lesions.

6. Also, why is it that in M1 the highlighted predictor is the interaction term between systolic blood pressure x age change, while this is not the case for M2 and M3? If I am reading the hypothesis correctly, the predictor of interest in M1 would be systolic blood pressure.
With M1we want to assess the modifying effect of blood pressure on change in WML volume (please also see response to question 4). In M2 and M3 we aim to study whether change in WML volume is directly related to change in cognition when taking into account elapsed time (Age_change) and a potential effect of baseline WML load on the cognitive decline (Age_change:WML_baseline). In these models, we do not specifically investigate the effect of the risk factors on cognition.

7 Regarding the dependent variable specified in models M1-M3, does it reflect the variable collected in the follow-up or is it a change score? Would the authors control for the scores obtained in these very same variables at baseline?
The dependent variables in M1 – M3 include both baseline and follow-up values. This is recommended for linear mixed models which take into account all available data points. To answer the question how baseline WML volume or cognitive function might impact their own change, we would have to include random slopes into the models M1 – M3. Unfortunately, random slopes cannot be estimated in our case because we only have two observations per level of the random effect (subject), which does not allow the estimation of both random intercept and slope.

1D. Replicability The authors have provided detailed methods. I do not have concerns regarding this point.

1E. Consideration of outcome-neutral conditions This part has been addressed satisfactorily. The authors have detailed the thresholds at which they will consider that a certain variable is invalid for each participant.

References

1. Zhang, D., et al., *Age and Diastolic Blood Pressure Play an Important Role in the Progression of White Matter Lesions: A Meta-Analysis.* European Neurology, 2020. **83**(4): p. 351-359.