Title: Memory Complaints and Depressive Symptoms Over Time: A Construct-Level Replication Analysis

Authors:
Jacqueline Mogle (jam935@psu.edu)
Nikki Hill (nikki.hill@psu.edu)
Sakshi Bhargava (sub229@psu.edu)
Tyler Bell (trb5511@psu.edu)
Iris Bhang (iub22@psu.edu)

Version: 1 Date: 05 Dec 2019

Author’s response to reviews:

We thank the editor for her comments on the reviews for the paper. We have reframed the paper to better emphasize the novelty of the work presented and its contribution to the literature as a whole. We also now include a section on the policy and practice implications. We have not included objective measures of cognitive function for two reasons. First, we have restricted the samples in the current study to only those with intact cognitive performance throughout the follow up period. This reduces the likelihood we will have individuals who are experiencing precipitous declines associated with non-normative cognitive decline but also restricts the variability in cognition. The observed correlations between objective and subjective cognition are typically quite low and this restricted range would further reduce that relationship. Our second reason for not including objective cognition in the analyses is that we believe objective cognitive tests are assessing a different underlying construct related to specific cognitive processes while reports of cognitive problems are assessing more naturalistic experiences of cognitive functioning. This would explain, at least partially, why these measures are not generally significantly correlated.

Fan Zhang (Reviewer 1):

Using an integrative analytic framework, the authors tested the bidirectional relationships between subjective memory complaints and depression with two Longitudinal datasets (NHATS and HRS). By controlling for the same covariates, multilevel linear modeling showed that in both datasets, perceived memory decline was associated with higher level of concurrent and future depressive symptoms, and depressive symptoms predicted more concurrent memory complaints and declines. With the construct-level replication analysis, the study provided evidence to support the generalizability of the bidirectional association between depressive symptoms and
memory complains among older adults. However, there are a few issues that need to be addressed before the manuscript could conform to the publication standard of BMC Geriatrics.

1. It is agreed that the two datasets adopted quite similar measurements for memory ratings and memory declines. However, regarding depressive symptoms, HRS used CES-D, while NHATS used 2 items from PHQ2, with a relatively low reliability (0.57). Would it lead to any problem or differences in the replication analysis in the two datasets?

Thank you for this comment. We have now also included the reliability of both measures that takes into consideration the repeated observations for both of these variables. These reliabilities are higher (NHATS reliability = .80; HRS reliability = .82). We believe that differences in findings across the two studies arising from the depression measure could actually be due to the lower coverage of depressive behaviors in the NHATS measure compared with the HRS measure, and now indicate this in the discussion on page 18.

2. In introduction, the authors mentioned that one of the limitations in existing literature is that the majority only investigated the directional or concurrent relationship between depressive symptoms and memory complaints. In the current study, bidirectional relationships were found, while the different causes and implications in "depression predicting memory complaints" vs. "memory complaints predicting depression" were not clearly addressed in discussion (p14).

Thank you for this comment. We now include a discussion of the potential causes and implications of memory decline specifically predicting increases in future depressive symptoms consistent with our reframing the paper to focus on these results.

3. The author also mentioned that the heterogeneity in the measurement of memory complaints could be problematic, and focused on memory rating and memory declines in the current study. However, the reasons why choosing memory rating and memory declines were not clarified. Meanwhile, what implications the results of memory rating and memory declines may have were also not clear. Moreover, little explanations were provided about why memory declines showed a greater effect than memory rating when predicting depressive symptoms, especially in the dataset of NHATS.

Thank you for this comment. We have clarified on page 10 that the items were selected based on what was available in these secondary datasets. We now include more information on implications, particularly related to policy and clinical implications on pages 17,19. For example, we discuss how the items we’ve identified might have greater sensitivity for picking up on very early cognitive changes or milder deficits that are problematic for older adults but do not rise to the level of a diagnosis. Identifying individuals earlier in the trajectory prior to diagnosis would aid in intervention efforts for depressive symptoms. We have added and clarified in our discussion that the hypothesized mechanism for why perceived declines lead to depressive symptoms (while ratings do not) is potentially due to different reasons for different individuals (pages 16-17). As one example, some older adults may have concerns that perceived declines in memory place them at risk for Alzheimer’s disease or other dementias which could impact emotional well-being.
4. Also, the "so what" question was not addressed sufficiently. Though the bidirectional relationship between memory complaints and depressive symptoms was identified, how would this finding help with diagnosis, treatment, and health care for the people who might develop dementia? This should be added in discussion.

We have now added more information about implications throughout the discussion with a specific discussion of the clinical and policy implications on page 19.

5. Another issue is that the attribution rate across different waves in two datasets should be mentioned, to give the readers a better idea about how many participants stayed in the study, or any selection bias.

We have now mentioned attrition rate for NHATS and HRS datasets on pages 8 and 9. In NHATS, approximately 91% of the individuals who were included at baseline had at least one follow up wave of data, and 60% had at least 4 waves of follow up. In HRS, 86% of individuals who were included at baseline had at least one follow up wave of data, and 50% had at least 4 waves of follow up data.

6. Although the authors mentioned that "participants tended to rate their memory more poorly in HRS compared with NHATS" (P15), no statistic results were provided to compared the memory complaints and depressive symptoms in baseline across the datasets.

Thank you for your comment. We have now removed the above-mentioned sentence from the manuscript.

7. There were many mistakes in the tables and reference. E.g., in Table 1, though both columns were supposed to list n(%), in the column of HRS, it was %(n) in income. In Table 2, the heading of the table should be put above the table not below the table. In the reference no. 23, the title should be "The Patient Health Questionnaire-2: Validity of a two-item depression screener", not "ywo-item". Pls have a thorough check throughout the manuscript.

We thank the reviewer for their patience in reviewing the manuscript. We have identified and removed all of these errors.

Rick Kwan (Reviewer 2): Some unclear messages in the manuscript:

In the background:

1. Why bidirectional relationship is more important/significant to understand compared to the concurrent relationship between depressive symptoms and memory complaints? These relationships have been examined for many years repeatedly. Why is the current understanding on the relationship between depressive symptoms and memory compliant insufficient for clinical/policy decision making?
Thank you for this comment. We have now expanded on the differences between previous analyses and the analyses in the current paper. Specifically, we are examining bidirectional temporal relationships using autoregressive modeling. We expand the previous research by testing models with lagged variables to determine if an individual had higher or lower memory complaints at a previous time point, whether they have higher or lower depressive symptoms in the future. This lagged relationship, which has not been studied extensively to our knowledge, is key to understanding temporal precedence of memory complaints or depressive symptoms which is needed for research and clinical decision making. If memory complaints tend to occur prior to increases in depressive symptoms, then intervening on memory concerns may prevent an older adult from developing clinically significant depressive symptoms. In contrast, if depressive symptoms occur prior to memory complaints, intervening on depressive symptoms would be more appropriate. We have now revised our introduction to make the additions our study makes to the literature clearer and the discussion to highlight the specific new contributions of our analyses.

2. Again, how does your integrative analysis overcome the methodological problems in the previous studies? You have described how you did the analysis without clearly saying how this method is better than the previous methods. This is very important because these relationships have been well-studied. You are advised to justify very clearly how your method produces robust conclusion compared with the previous ones.

Thank you for this comment. As part of our revisions to our introduction, we now make clear how using integrative analysis specifically advances this research area by efficiently accumulating evidence of the temporal relationships that are the innovative contribution of our manuscript.

3. What are the objectives or hypothesis in this manuscript?

We have now added our research questions on page 7 of the manuscript.

Methods:

1. In the HRS data set, you mentioned that you removed 92 Hispanic individuals due to extremely unbalanced sample size. What do you mean by unbalanced sample size, unbalanced race groups? unbalanced sample size between HRS and NHATS? Given your sample size is large, does this 92 subjects make a difference?

In both datasets we removed a small number of participants (HRS: 92; NHATS: 269) who reported their race/ethnicity as something other than White or Black. We did so because we wanted to be able to interpret racial differences in any of the outcomes and including a subgroup of individuals who vary in the racial identity (i.e., Hispanic, Other, or did not report a racial identity) would hinder our interpretation of any race differences. Additionally, this group was very small relative to the other groups being compared which would cause inflated standard
errors and other statistical anomalies that impact model fitting. We have now revised our statements in the methods on pages 8,9 to better reflect why these individuals were removed.

2. Covariates: how do you justify the selection of the covariates?

We now have included justification on page 11. All included variables have been shown to relate to both memory complaints and depressive symptoms, and, thus, would be important to include to prevent potential effects of confounding variables.

3. Statistical analysis: It is clear in terms of what methods you have used. However, it is unclear how your methods are used to answer the research questions that you intended to ask.

We have now clarified our research questions in the introduction and incorporated them in the analytic strategy section to more closely tie our questions to our analytic methods.

Results:

1. In table 3, there are four models (NHATS HRS NHATS HRS), how are they different from one another? Please specify them in the table.

We apologize for not being clear in tables. We have edited the table headers and provided an explanation in the notes for Tables 3 and 4.

2. What is the information about the number of follow-up times, and years between the follow-up times?

We have now added this information on pages 8-9. In NHATS, data were collected annually and up to 5 follow up waves are available for sample from wave 1. For the replenished sample, only one follow up is available. In HRS, data were collected every other year. Current study uses up to 9 waves of follow up data.

3. Your results showed that depressive symptoms are not predictive of future memory complaint. How do you explain/guess that this finding is different from that reported in the previous study that people with depressive symptoms are associated with higher risk of cognitive decline? e.g., Paterniti, S., Verdier-Taillefer, M., Dufouil, C., & Alpérovitch, A. (2002). Depressive symptoms and cognitive decline in elderly people: Longitudinal study. British Journal of Psychiatry, 181(5), 406-410. doi:10.1192/bjp.181.5.406

We would argue that the article cited above is different for several reasons. First, the Paterniti et al. study focused on changes in objective cognitive performance which differs from the focus of the current study which is subjective cognitive functioning (i.e., self-reports of cognitive
problems). Objective cognitive performance (particularly as measured by Paterniti et al. using the Mini-mental State Exam; MMSE) is often poorly correlated with self-reports of cognitive functioning. We believe this is because self-reports of cognitive functioning capture more naturalistic or real-world cognitive functioning, compared with objective cognitive tasks which lack many of the features of naturalistic cognitive demands. Additionally, the MMSE is a global test of cognitive performance as compared with the focus of the current study on memory specifically. Another key difference between our current study and the Paterniti et al. study is their use of a cut off on the CESD to identify individuals who were higher or lower in depressive symptoms. We opted to leave the depressive symptom scores continuous in the current study so as to ensure better representation of changes in depressive symptoms over time. For example, in the Paterniti et al. analyses, changes in depressive symptoms over time would only be detected if an individual’s symptoms changed sufficiently to cause their score to rise above the cut off. Many changes in symptoms could occur that do not meet this strict criterion but are still meaningful to the individual that would go undetected in this type of analysis. Although cut off scores are informative, our goal was to detect change over time and so this approach lends itself to treating scores continuously.

Discussion:

1. What is the "new" that you have found? You repeatedly reported that your findings are consistent with the previous studies. It is no wonder because this is a very well-studied topic.

Thank you for this comment. We feel that the innovative findings come from the autoregressive models clearly pointing toward a temporal sequence for perceived memory decline predicting depressive symptoms. We now highlight these findings specifically in our discussion as the primary result of interest and build from this particular result.

2. Again, you mentioned the strength of the integrative analysis (e.g., consistency across results were not due to differences in covarites/predictors). Then, how is your result different from the previous results after adopting this robust analysis method? If it comes up with a similar result with a more robust method to enhance our confidence, say it so.

We agree with the reviewer that one of the main strengths of this type of analysis is to build a stronger evidence base more efficiently in a given area. We now highlight this as a primary advantage to our analysis strategy in both our introduction to the paper as well as the discussion.

3. Again, there is a lack of implications to clinical practice, research, and policy making from your findings. What changes do you expect to see by knowing what you have found?

We have now added a section on the implications of these findings on page 19. In brief, we think that our findings suggest a potential target for intervening with older adults who have complaints about their memory decline in particular. Older adults are asked to report about their memory and cognition at their annual wellness visits. However, the guidelines for this reporting are vague
and could be enhanced by asking older adults specifically about their memory decline rather than complaints in general based on our findings. We now have evidence to suggest that those older adults who are perceiving a memory decline go on to develop greater levels of depressive symptoms and clinicians could identify this group and potentially intervene to alleviate concerns about changes in memory.