PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	An observational study of the differential impact of time-varying depressive symptoms on all-cause and cause-specific mortality by health status in community dwelling adults: The REGARDS study
AUTHORS	Moise, Nathalie; Khodneva, Yulia; Jannat-Khah, Deanna; Richman, Joshua; Davidson, Karina; Kronish, Ian M.; Shaffer, Jonathan; Safford, Monika

VERSION 1 – REVIEW

REVIEWER	Emmanuel Wiernik Inserm U894, Paris, France Université Paris Descartes, Sorbonne Paris Cité, Faculté de Médecine, Paris, France
REVIEW RETURNED	29-May-2017

GENERAL COMMENTS	This interesting and clear paper shows surprising results (more important effect of depression among participants with good self-rated health). The discussion mentions some possible explanations but this new finding should be further investigated. p.2: Abstract: the first sentence (objective) is not finished. In the last sentence of results, I suggest to be more clearer because mortality is not considered at baseline. p.8: For Cox regressions, some assumptions should be checked (in particular no time-dependent effect). p.8: what is the lapse time between the first measure of CES-D (initial telephone call) and date of inclusion (in-home visit)? Is it enough short to not be an issue? p.9: you should give some data about missing covariates (how many for each covariate and how many participants with no missing covariates?). Furthermore, are results without imputation similar? p.12: "short-term relationship between elevated depressive symptoms and mortality" this sentence should be moderated because depressive symptoms at baseline could be long-lasting depressive symptoms (especially for participants with depressive symptoms at follow-up) p.12: the exclusion of participants with active cancer diagnosis should be mentionned in methods section.

	Furthemore, why are excluded participants with active cancer and not participants with history of CVD? A sensibility analysis without participant with CVD could be relevant. table 1A & 1B: some legends are missing. In addition, the N for overall and each CES-D subgroup seems to mean that no data is missing. table 2: "Model 5 + intervening non-fatal event" is not clear. Furthermore, this model is not described in methods section table 3: why was the model 5 not computed? Figure 2: This figure could be problematic: if K-M curves are plotted, depressive symptoms are probably not time-dependant. Other methods (for example Simon and Makuch) could be more appropriate. In addition, authors should indicate the variation of the number of participants at risk along the study period as well as the confidence interval for the survival function.

REVIEWER	Aurelie Lasserre Department of Psychiatry Lausanne University Hospital 1011 Lausanne Switzerland
REVIEW RETURNED	22-Jun-2017

GENERAL COMMENTS	This well-written and easy-to-read article aimed to assess the association between time varying depressive symptoms with all-cause and specific mortality, among a cohort of about 30'000 participants who have been followed-up for 6.5 years. This subject is of major public health importance regarding the prevalence of depression worldwide. This study has major stenghs: the size of the population, the multiple measured physical, lifestyle and socio-demographic covariates, the three time-point of symptoms assessment and the assessment of the self-report health. Some limitations could also be raised: the assessment of depressive symptoms was only a 4-questions questionnaire, which is a rough indicator of clinical depression. Psychiatric comorbidities were not assessed (eg anxious disorders). For a mortality study, the length of the follow-up seems relatively short. I have some minor comments: - In the introduction, no studies about time varying association are reported. However, to my knowledge, some studies have addressed this question: (Bruce et al.,1994, Am J Psychiatry, Surtees et al., 2008, Am J Psychiatry and Lasserre et al., 2016, J Aff Dis). - In the methods section, more could be said about the main exposure: time-varying depressive symptoms. How this variable was defined as elevated or not for example. - More could be said about the imputation of missing variables in the statistical section.

	- In the discussion, page 22, the authors use the term "short-term relationship". Is it a short-term comparing to other studies or did they compare short and long-term in their follow-up? This needs to be clarified. - In the limitations, the authors mention the use of the CES-D. They could mention the review from Schulz et al, 2002 that showed variance between studies using scales and those using interviews.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
"This interesting and clear paper shows surprising results (more important effect of depression among participants with good self-rated health). The discussion mentions some possible explanations but this new finding should be further investigated"

• Thank you for your thoughtful and helpful comments. We agree this new finding should be further investigated and have stated as such in our discussion (pages 13, 15).

"p.2: Abstract: the first sentence (objective) is not finished. In the last sentence of results, I suggest to be more clearer because mortality is not considered at baseline."

• Thank you for identifying this typographical error. We have now completed the abstract. We also now state "The analyses of the association of one measure of baseline depressive symptoms and mortality analyses yielded similar results" (page 3).

"p.8: For Cox regressions, some assumptions should be checked (in particular no time-dependent effect)."

• We now clarify that relevant assumptions were checked (page 9).

"p.8: what is the lapse time between the first measure of CES-D (initial telephone call) and date of inclusion (in-home visit)? Is it enough short to not be an issue?"

• The time lapse between the first measure of CES-D and the date of inclusion was 28.0 (interquartile range = 21.0) days. We now include this data in our methodology and believe that this period was short enough to not be an issue (page 6).

"p.9: you should give some data about missing covariates (how many for each covariate and how many participants with no missing covariates?). Furthermore, are results without imputation similar?"

• We have now included detailed information on missing data for each covariate on page 10. Results were similar without imputation and we now state so and within two decimal places (page 12).

"p.12: "short-term relationship between elevated depressive symptoms and mortality" this sentence should be moderated because depressive symptoms at baseline could be long-lasting depressive symptoms (especially for participants with depressive symptoms at follow-up)"

• Thank you for this astute comment. We clarify that we used a broken up follow up time, with the average follow up time between baseline and second measures being 5 years, and between second and third measures 2 years.

Nonetheless, we agree we did not specifically assess long-term vs. short-term depressive symptoms and have moderated our wording to state “time-varying relationship” in lieu of “short-term relationship” (page 12). We also state these concerns in the limitations section (page 15).

"p.12: the exclusion of participants with active cancer diagnosis should be mentioned in methods section. Furthermore, why are excluded participants with active cancer and not participants with history of CVD? A sensibility analysis without participant with CVD could be relevant."

• Thank you. We have explained the exclusion of patients with active cancer in the methods section (page 5). Of note, those with active cancer were excluded as part of the overall REGARDS study, the rationale of which has previously been described (Neuroepidemiology. 2005;25:135-143). However, a history of malignancy was not a specific exclusion criteria. As such, we include those with history of CVD. This is in line with most prior analyses on depression and mortality, which included those with a history of CVD and/or cancer. Our prior analyses of the REGARDS data excluded those with a history of CVD, and yielded a similar relationship between time varying depressive symptoms and CVD death (Journal of the American Heart Association. 2016;5: e003767). We now describe this rationale in our discussion and reference our prior results (page 14-15).

"Table 1A & 1B: some legends are missing. In addition, the N for overall and each CES-D subgroup seems to mean that no data is missing."

• Thank you. We have revised the legends of both tables. Of note, only those with available baseline CES-D data were included in our analyses and as such no baseline CES-D measures were missing.

"table 2: "Model 5 + intervening non-fatal event" is not clear. Furthermore, this model is not described in methods section"

• On page 9, we provide further detail on Model 5, which considers intervening first non-fatal stroke and/or myocardial infarction as a time-dependent covariate.

"table 3: why was the model 5 not computed?"

• Table 3 includes the association between 1 baseline measure of depression and mortality (similar to what has been done in prior analyses), and as such we did not control for time varying CVD (see above)

"Figure 2: This figure could be problematic: if K-M curves are plotted, depressive symptoms are probably not time-dependant. Other methods (for example Simon and Makuch) could be more appropriate."

• We now describe that analyses met proportionality assumptions for time dependency (page 9). We acknowledge that there are inherent issues with using K-M curves in time-varying analyses and have excluded these from our resubmission. We thank the reviewers for their suggestion, which enhanced our methodology. In response, we constructed Simon and Makuch plots, which were largely confirmatory and consistent with our overall finding of a relationship between depressive symptoms and mortality. However, Simon and Makuch methods are not widely used in clinical research, and there are inherent concerns in interpretability by clinical audiences as well as the fact that the results depend on the selection of arbitrary landmark times.1 As such, we include these plots in our supplementary material (eFigure 3).

"In addition, authors should indicate the variation of the number of participants at risk along the study period as well as the confidence interval for the survival function."

• We excluded any plots from the main paper given inherent difficulty in interpretability (see above).

Reviewer: 2
Please leave your comments for the authors below

"This well-written and easy-to-read article aimed to assess the association between time varying depressive symptoms with all-cause and specific mortality, among a cohort of about 30'000 participants who have been followed-up for 6.5 years. This subject is of major public health importance regarding the prevalence of depression worldwide."

• Thank you for our positive review of our manuscript.

"This study has major strengths: the size of the population, the multiple measured physical, lifestyle and socio-demographic covariates, the three time-point of symptoms assessment and the assessment of the self-report health."

• Thank you again for highlighting the strengths of our paper.

"Some limitations could also be raised: the assessment of depressive symptoms was only a 4-questions questionnaire, which is a rough indicator of clinical depression. Psychiatric comorbidities were not assessed (e.g., anxious disorders). For a mortality study, the length of the follow-up seems relatively short."

• You have identified important limitations to our study, and as such we further expand upon these in the discussion (page 15), including the short follow-time and lack of psychiatric comorbidity assessment and short follow up time. Of note, few prior depression to mortality studies accounted for comorbid psychiatric illnesses, a noted weakness in the literature. We believe that the short length of follow up lends further support and urgency to the acute effect of depressive symptoms on mortality. A prior meta-analysis (Pinquart et al.) also suggested that shorter follow up time moderated the relationship between depressive symptoms and mortality and we expand upon this point in the discussion (page 14).

I have some minor comments:

"In the introduction, no studies about time varying association are reported. However, to my knowledge, some studies have addressed this question: (Bruce et al.,1994, Am J Psychiatry, Surtees et al., 2008, Am J Psychiatry and Lasserre et al., 2016, J Aff Dis)."

• Thank you for including these helpful studies. On our review, none of these studies specifically assessed time varying depressive symptoms and mortality (page 4). We do now include them in our introduction and further specify the knowledge gap bridged by our current study. Surtees et al. (2008) cross-sectionally assessed depressive symptoms during a 2-year baseline and incident CVD mortality only, and did not include time-varying analyses nor assess nonCVD mortality. Bruce et al (1994) again assessed baseline depressive symptoms to 9-year follow-up mortality.

The most recent trial, Lasserre et al., 2016, J Aff Dis, found that current but not remitted depression was a strong predictor of all-cause mortality, but used an interview to ascertain a history of depression and current depression at one time point without multiple measures of depression in real time and also only assessed all-cause and not cause-specific mortality. None of these studies assessed the moderating effect of self-reported health.

"In the methods section, more could be said about the main exposure: time-varying depressive symptoms. How this variable was defined as elevated or not for example."

• This variable was defined as CES-D ≥4 vs. <4 at each time period during which depressive symptoms were measured (baseline, on average 5 years later, on average 7 years later). We expand upon this point in our methods section (page 8-9): “we considered depressive symptoms (CES-D≥4 vs. <4) as a time-varying exposure, with updates of exposure at 5-year and 7-year follow-up. Therefore, each participant contributed up to 3 measures of CES-D (≥4 vs. <4) with up to 3 broken-up follow up times. Follow-up time for each participant was calculated from the date of the in-home visit to the date of the earliest of: death, last telephone follow-up, end of follow-up or next CES-D measure”

"More could be said about the imputation of missing variables in the statistical section."

• We have included further information about missing variables (page 10).

"In the discussion, page 22, the authors use the term "short-term relationship". Is it a short-term comparing to other studies or did they compare short and long-term in their follow-up? This needs to be clarified."

• We now use the terminology “time-varying” in lieu of “short term” throughout. Per the editor’s request, we have removed the conclusions box on page 22. While our article is in fact short-term due to the overall length of the study compared to others and even shorter length of time between CES-D measures, we did not specifically compare short and long-term follow up, and state so in our limitations (page 14).

"In the limitations, the authors mention the use of the CES-D. They could mention the review from Schulz et al, 2002 that showed variance between studies using scales and those using interviews."

• Thank you for sharing this citation. We now include it in our limitations section. We also cite a meta-analysis showing stronger relationships between clinical diagnoses of depression and mortality compared to continuous measures, suggesting we may have underestimated our already striking mortality findings.

References

1. Anderson JR, Cain KC and Gelber RD. Analysis of survival by tumor response. J Clin Oncol. 1983;1:710-9.
2.

VERSION 2 – REVIEW

REVIEWER	Emmanuel Wiernik, postdoctoral researcher Inserm, U894 France
REVIEW RETURNED	11-Sep-2017

GENERAL COMMENTS	I thank the authors for having taken into account my suggestions. The most of my previous comments have been addressed satisfactory. Nonetheless, some suggestions or questions still remain. The major concern is about the proportionality assumption and the interest of a time-dependent model (p.9.). You suggest that time-varying covariates were appropriate because proportional hazard hypothesis was violated. However, all covariates, except depressive symptoms (and CV events in model 5), are measured at baseline. Furthemore, in the following sentence, the proportionnality assumption was satisfied for depressive symptoms. Theses sentences should be modified in order to understand if you used a model with time-dependent covariates (as suggested in the rest of the paper) or a model with time-dependent coefficients (as usual when proportional hazard hypothesis is violated). Finally, you could compare the results with time-dependent depressive symptoms and the results of your sensitivity analysis in order to show if the time-dependent model represents a real added value. Furthemore, I have some very minor comments: P.9. "We also conducted a formal test for interaction between depressive symptoms and self-reported health (defined as excellent or very good vs. good, fair or poor health) in the fully-adjusted models": It could be clearer to specify the fully-adjusted model (i.e. model 4 and not model 5). P.17. "Results were similar without multiple imputations within 2 decimal places (Table 3)". This sentence suggests that the analyzes without imputations have been only conducted for sensitivty analyses whereas they are not the main results. Table 1A&B: The n in these tables are not totally clear. How are considered missing data for each variable? For example, the n for income in table 1a seems to mean that no data is missing. You could just specify that 29 491, 26 817 etc. are the maximal n when a variable has no missing data. Table 2. Model 5: "+ intervening non-fatal CVD even" for "NonCVD Death" and "Cancer death" could be retired (as for "CVD death").

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
Please leave your comments for the authors below I thank the authors for having taken into account my suggestions. The most of my previous comments have been addressed satisfactory. Nonetheless, some suggestions or questions still remain.

Thank you for your careful and helpful review of our manuscript.

The major concern is about the proportionality assumption and the interest of a time-dependent model (p.9.). You suggest that time-varying covariates were appropriate because proportional hazard hypothesis was violated. However, all covariates, except depressive symptoms (and CV events in model 5), are measured at baseline. Furthermore, in the following sentence, the proportionality assumption was satisfied for depressive symptoms. Theses sentences should be modified in order to understand if you used a model with time-dependent covariates (as suggested in the rest of the paper) or a model with time-dependent coefficients (as usual when proportional hazard hypothesis is violated).

Thank you for this comment. We now state that the proportionality hazards hypothesis was violated for depressive symptoms as opposed to baseline covariates (page 9). We also specify that covariates were only measured at baseline in the methods section (pages 7, 9). Finally, we clarify that our model did in fact use time-dependent covariates throughout. According to Harrell, tests of association can be performed with time dependent variables (see page 457; Harrell FE, E. F. Regression Modeling Strategies : With Applications to Linear Models, Logistic Regression, and Survival Analysis. Springer; 2001. http://dl.acm.org/citation.cfm?id=1196963. Accessed September 22, 2017).

Finally, you could compare the results with time-dependent depressive symptoms and the results of your sensitivity analysis in order to show if the time-dependent model represents a real added value.

Because of the violation of proportionality hazards, we felt it was more accurate to use time-dependent models regardless. However, we now compare the two results in our discussion (page 20). While overall results were similar, we suggest that baseline data amongst those with excellent health remained significant even in fully adjusted models for cancer mortality because baseline (as opposed to time varying) analyses would bias away from the null. Adjusting for time-dependent depressive symptoms likely allowed for more rigorous, realistic results, particularly for cancer mortality, where we would not expect to see a significant interaction of health status by depressive symptoms in a cohort that used active malignancy as an exclusion criterion.

Furthemore, I have some very minor comments:

P.9. "We also conducted a formal test for interaction between depressive symptoms and self-reported health (defined as excellent or very good vs. good, fair or poor health) in the fully-adjusted models": It could be clearer to specify the fully-adjusted model (i.e. model 4 and not model 5).

Thank you. We now specify that formal tests for interaction were conducted in model 4 (page 9).

P.17. "Results were similar without multiple imputations within 2 decimal places (Table 3)". This sentence suggests that the analyzes without imputations have been only conducted for sensitivty analyses whereas they are not the main results.

Thank you for this helpful suggestion. We moved the sensitivity analyses section to the end of the methods section to avoid confusion (pages 9-10).

Table 1A&B: The n in these tables are not totally clear. How are considered missing data for each variable? For example, the n for income in table 1a seems to mean that no data is missing. You could just specify that 29 491, 26 817 etc. are the maximal n when a variable has no missing data.

We now highlight that n's are the total number of participants assuming no missing data to Tables 1A and 1B (pages 13-15). Under the methods, we also quantify our missing data for each variable (page 9).

Table 2. Model 5: "+ intervening non-fatal CVD even" for "NonCVD Death" and "Cancer death" could be retired (as for "CVD death").

Thank you for this helpful suggestion. We made changes accordingly.

VERSION 3 – REVIEW

REVIEWER	Emmanuel Wiernik, postdoctoral researcher Inserm U894, Paris, France Université Paris Descartes, Sorbonne Paris Cité, Faculté de Médecine, Paris, France
REVIEW RETURNED	13-Oct-2017

GENERAL COMMENTS	I thank the authors for considering my suggestions. The most of issues have been resolved. However, some minor clarifications could further improve the paper. With regard to the interest of a time-dependent model (p.9.), it seems to me that the justification is not correct. Indeed, the text suggests that time-varying depression was appropriate because proportional hazard hypothesis was violated. I think that would be true if the coefficient for depression was time-dependent but not, as in the present case, if the coefficient is constant over time (only the value of the variable changes). Furthemore, the issue of proportional hazard hypothesis is not clear because you mentionned in a sentence a violation whereas in the following sentence "the proportionnality assumption was satisfied". The second point concerns the analyses without multiple imputations. Did you perform these analyses for comparing the results with the associations of time-variant depressive symptoms (table 2)? The text as it is suggests that analyses without multiple imputations have been only conducted for verifying results of sensitivity analyses (table 3).

VERSION 3 – AUTHOR RESPONSE

Comment: "I thank the authors for considering my suggestions. The most of issues have been resolved. However, some minor clarifications could further improve the paper."

Response: Thank you for your useful input. It has been integral to improving this manuscript.

Comment: “With regard to the interest of a time-dependent model (p.9.), it seems to me that the justification is not correct. Indeed, the text suggests that time-varying depression was appropriate because proportional hazard hypothesis was violated. I think that would be true if the coefficient for depression was time-dependent but not, as in the present case, if the coefficient is constant over time (only the value of the variable changes). Furthemore, the issue of proportional hazard hypothesis is not clear because you mentionned in a sentence a violation whereas in the following sentence "the proportionnality assumption was satisfied".

Response: We thank the reviewer for pointing out the lack of clarity in our description of evaluating the possibility of non-proportional hazards and we agree that it was confusing. We have re-written that passage as follows:

“To evaluate the possibility of non-proportional hazards, we graphically inspected the log-log survival plots for depressive symptoms. We tested the Schoenfeld residuals for each model for a non-zero slope and all p values were greater than 0.05, indicating compatibility with the proportional hazards assumption.” (page 9)

Comment: "The second point concerns the analyses without multiple imputations. Did you perform these analyses for comparing the results with the associations of time-variant depressive symptoms (table 2)? The text as it is suggests that analyses without multiple imputations have been only conducted for verifying results of sensitivity analyses (table 3)."

Response: Thank you for this important input. All the results presented in this manuscript are from models run on the datasets created from multiple imputation. We specify within table 2 and 3 footnotes that results were from models ran on multiple imputed datasets. We also now place the wording about multiple imputations at the end of the statistical section and state that imputations were used for all analyses (pages 9-10).

VERSION 4 – REVIEW

REVIEWER	Emmanuel Wiernik, postdoctoral researcher Inserm U894, Paris, France Université Paris Descartes, Sorbonne Paris Cité, Faculté de Médecine, Paris, France
REVIEW RETURNED	16-Nov-2017

GENERAL COMMENTS	The authors responded satisfactorily to the first comment. For the second point, they could just add this sentence: "Results were similar without multiple imputations within 2 decimal places" (p. 15, just before "Table 2") as they did for "Table 3" (if that is the case of course). Otherwise, it seems that the analyses without missing data were not performed for the main analyses but only for the sensitivity analyses.

VERSION 4 – AUTHOR RESPONSE

Comment: -"For the second point, they could just add this sentence: "Results were similar without multiple imputations within 2 decimal places" (p. 15, just before "Table 2") as they did for "Table 3" (if that is the case of course). Otherwise, it seems that the analyses without missing data were not performed for the main analyses but only for the sensitivity analyses"

Response: Thank you to the reviewer for this very helpful comment. All analyses were first performed without multiple imputations. We then performed multiple imputation and re-did all of the analyses. All of the results presented in this manuscript are from the multiply imputed dataset- not just for the sensitivity analysis. As recommended we now state:
“Results were similar without multiple imputations within 2 decimal places" (page 15)