Author’s response to reviews

Title: The association of telomere length with substance use disorders: systematic review and meta-analysis protocol

Authors:

Fernando Navarro-Mateu (Fernando.navarro@carm.es)
María Rubio-Aparicio (maria.rubio1@um.es)
Pedro Cayuela (cayuela.pedro@gmail.com)
Francisco-Javier Álvarez (franciscoj.alvarez@carm.es)
Agustín Roca-Vega (agustin.roca@carm.es)
María Chirlaque (mdolores.chirlaque@carm.es)
María Cayuela (Cayuelamarialuisa@gmail.com)
Mathilde Husky (mathilde.husky@u-bordeaux.fr)
Salvador Martínez (smartinez@umh.es)
Julio Sánchez-Meca (jsmeca@um.es)

Version: 1 Date: 01 Jul 2019

Author’s response to reviews:

Response to the reviewers’ comments

1st July 2019

Dear Dr. Yen-Fu Chen,

Thank you for the reviewers’ comments to our manuscript entitled "The association of telomere length with substance use disorders: systematic review and meta-analysis protocol" (SYSR-D-19-00084).

Please find attached our point-by-point responses to the reviewer’s comments as well as a marked-up copy of the changes made to the submitted version of the article.
We hope this revised version will be acceptable for publication.

Yours sincerely,

Fernando Navarro-Mateu, MD, PhD

Editors Comments to Author:

Reviewer reports:

REVIEWER #1 (R.1):

Major comment

R.1.1) Newcastle-Ottawa Scale will be used to assess the risk of bias of included studies. The two items under "Comparability" requires reviewers to assess whether the most important confounding factor(s) were adjusted for. It is currently unclear in the protocol whether unadjusted or adjusted mean differences in telomere length will be used, and whether any (and if so, which) confounding factors would be considered to be most important when assessing comparability between the groups with and without substance use disorders.

Response: We agree with the Reviewer in that addressing the potential influence of confounding factors on the effect sizes is an important issue. In fact, the effect sizes defined in our meta-analysis are unadjusted SMDs. The reason for using unadjusted SMDs is that using adjusted SMDs in a standard meta-analysis requires that all studies adjust for the same confounding factor/s, and this is practically impossible to find in the literature. To clarify this point, we have added the following text (text added in bold, lines 210-212, marked copy):

“In this meta-analysis unadjusted effect sizes (SMDs) will be used. The potential influence of confounding factors will be assessed as described below.”

The consequence of using unadjusted effect sizes is that additional analyses are needed to assess the potential influence of confounding factors on the unadjusted effect sizes. Thus, we have added information on how we will conduct sensitivity analyses to examine the influence of confounding factors. Regarding this point, please see our response to Reviewer’s 2 point 2.

Minor comments

R.1.2) Page 9, lines 188-194: presumably the effect size index being computed here is what commonly known as Hedge's g, and the correction factor is introduced to mitigate bias
associated with the measure when sample sizes are small? (If so) you may wish to explain these to general readers who are less familiar with these statistics/formulae.

Response: Yes, the effect size defined in our meta-analysis is the standardized mean difference proposed by Hedges (1981). This index includes a correction factor for small sample sizes, which is estimated by means of c(m). To clarify this point we have revised this sentence as follows (line 205, marked copy):

“… with c(m) being a correction factor for small sample sizes defined as: c(m) = 1 - 3/(4N – 9), ...

R.1.3) Page 9, line 205: please explain the reason for using the method proposed by Hartung and Knapp - presumably due to potential/expected departure from normality for the distribution of effect sizes?

Response: To address this comment, we have added the following text (lines 221-227, marked copy):

“Instead of assuming a standard normal distribution, the method by Hartung and Knapp assumes a Student t-distribution with k – 1 degrees of freedom (k being the number of studies) and an improved estimator of the variance of SMD that takes into account the uncertainty in estimating the between-studies variance.”

R.1.4) Page 10, line 218 (and also in the abstract): would 'subgroup analyses and meta-regression' be a more accurate term here (for exploring the sources of heterogeneity) than 'sensitivity analyses', which are mainly conducted to examine the robustness of findings under various assumptions/methodological choices?

Response: We agree with Reviewer 1 in that these terms better describe our statistical analyses. Thus, we have changed the text as follows:

In the Abstract (line 60, marked copy):

“Subgroup analyses and meta-regressions will be conducted to ascertain heterogeneity.”

In the Methods section (line 245, marked copy):

“In cases of moderate-to-large heterogeneity (I2 > 25%), we will seek to identify possible explanations using subgroup analyses and meta-regressions based on the most important characteristics of the studies …”
Major comments

R.2.1.) Although I am not an expert in search term design, the search terms appear to be appropriately sensitive to my eye. The search strategy was designed in collaboration with an author who seems to be an academic reference librarian (ARV), in which case I would recommend explicitly stating this in the manuscript since search term design is a critical component of a high-quality systematic review. I was also pleased that there is a strong plan in place for searching grey literature.

Response: Thank you for your suggestion. The following text has been added to highlight this important point.

Lines 135-136, marked copy: “…and telo*” (see Additional file 2). This search strategy was designed and will be carried out by a librarian expert (ARV).”

R.2.2.) My most important concern is that confounding of the relationship between substance use and telomere length will likely represent a considerable methodological threat to the validity of this meta-analysis’ findings. Yet confounding is currently not discussed at all. Not exhaustively, variables such as socioeconomic status, education, and life stressors likely affect both substance use and telomere length. If the meta-analyzed studies do not adjust for nearly all important confounders of this relationship, or if they are cross-sectional so that confounders and mediators cannot be distinguished (e.g., [8]), then the meta-analytic estimate may be quite biased.

I therefore think it is critical that the authors: (1) plan to conduct subset analyses (or similar) on studies whose designs most effectively minimize confounding (i.e., longitudinal studies controlling for, at minimum baseline telomere length); and (2) plan to conduct some form of sensitivity analysis to assess robustness to residual confounding. The latter will be important if, as I suspect, there are few longitudinal studies. See reference [1] for an approach for doing so that is implemented in R (the package EValue).

Response: We agree with Reviewer 2 in that we need to describe how the potential influence of confounding factors will be addressed. This is particularly important as we do not expect to find longitudinal studies, but mainly case-control and cross-sectional designs. Thus, we have added the following text to describe how potential confounding factors on the effect sizes will be addressed (lines 254-266, marked copy):

“Research on telomere length for persons with SUD must dedicate special attention to the potential influence of confounding factors. In order to address this point, several sensitivity analyses will be conducted. First, the risk of bias items of the NOS will be analyzed by means of subgroup analyses. Second, the comparability of the groups (cases and controls) is a key issue to
assess the potential influence of confounding factors on effect sizes, with age being the most relevant confounding factor. In addition, to apply subgroup analyses to the risk of bias item on comparability of the NOS, simple meta-regressions will also be conducted using the SMDs as a dependent variable and the mean difference in age between cases and controls, the difference in age SDs, the difference in the proportion of males and of Caucasians, the difference in education levels, and the difference in life stressors as predictors. Third, when studies report all the potential confounding factors, or the majority of them, a multiple regression model will be applied in order to determine predicted effect sizes once adjusted for confounding factors.”

R.2.3.) The authors refer to meta-regressively examining moderators of the relationship between substance use and telomere length, but the moderators are not specified. Lines 117-118 briefly mention different types of substances and refer without specificity to other "factors implicated in this heterogeneity". What moderators will be examined? Or will this decision be made post hoc?

Response: Lines 167-178 (marked copy) describe the characteristics that will be extracted from the studies. Some relate to sample sociodemographic characteristics (e.g., mean age, gender distribution, etc.); others relate to methodological characteristics of the studies (e.g., study design, diagnostic tools used to classify participants, etc.), and others refer to risk of bias items (e.g., using the Newcastle-Ottawa Scale, described in lines 192-197, marked copy). The purpose of extracting all of these characteristics was to have the opportunity to link them with effect sizes. Therefore, subgroup analyses and meta-regressions will be conducted for all of the characteristics extracted from the studies in order to examine their potential association with effect sizes. To clarify this point, we have added the following text in the aims of the study (lines 119-124, marked copy):

“iii) if heterogeneity is confirmed, what are the substantive and methodological factors implicated in this heterogeneity. In particular, we will investigate the potential association of individual sociodemographic characteristics of the participants with the effect sizes. In addition, the potential confounding effects on the effect sizes will be investigated by analyzing the influence of methodological characteristics of the studies on the effect sizes, such as the design type and risk of bias items.”

In addition, we have also added the following text in the ‘Data extraction’ section (lines 176-178, marked copy):

“The purpose of extracting all of these characteristics is to have the opportunity of examining their association with the effect sizes.”

We have also added ‘education level’ as another characteristic to be extracted from the studies (line 173, marked copy).
R.2.4.) Also regarding heterogeneity, the statistics $I^2$ and $Q$ (while useful to report) characterize heterogeneity relative to the total variation in point estimates and are a function, for example, of the meta-analyzed studies' sample sizes. For the authors' intended purposes, I think these measures should be supplemented with an absolute measure of heterogeneity (tau or tau$^2$), and that decisions about the amount of heterogeneity should consider tau and its confidence interval as well [2].

Response: We agree with the Reviewer in that tau$^2$ is an important measure in meta-analyses. In fact, in p. 10 we had specified that the between-studies variance will be estimated by restricted maximum likelihood. The problem of tau$^2$ is that it is a statistic dependent on the metric scale of the effect size index. This circumstance complicates its use to make decisions on when there is enough heterogeneity among individual effect sizes to search for moderator variables. This is the reason for not mentioning tau$^2$ for making decisions on whether moderator variables should be analyzed. However, following the reviewer recommendation, we have added the following text (lines 238-239, marked copy):

“In addition, heterogeneity will be assessed with the between-studies variance and 95% confidence interval.”

Additionally, since heterogeneity seems to be a central focus of this protocol, the authors should consider reporting statistical metrics that, unlike the pooled point estimate alone, directly describe the heterogeneous distribution of effect sizes. Two options that complement each other well are a prediction interval for the effect size in a new study [2-3] and the estimated proportion of true effect sizes that are of scientifically meaningful size [4].

Response: Following these recommendations, we have added the following text (lines 225-227, marked copy):

“In addition, a 95% prediction interval around the average effect size will be calculated, in order to provide a prediction of the expected true effects if a new study is conducted. [47, 50]”

And in lines 238-243, marked copy:

“Finally, following Mathur and VanderWeele’s (2019) proposal, the estimated proportion (and 95% confidence interval) of true effect sizes exceeding a scientifically meaningful threshold will be calculated. In terms of standardized mean difference, we will consider -0.20 the threshold effect size for these calculations. [60]”

R.2.5.) I appreciate the attention to publication bias. However, funnel plots and their corresponding statistical methods (trim & fill and Egger's test) have some important limitations.
For example, they assume that publication bias operates on effect sizes rather than p-values and that publication bias is the only source of correlation between effect sizes and standard errors) [5]. For this reason, I would recommend: (1) using the funnel plot-based methods only with strong caveats about their limitations; and (2) supplementing or replacing these measures with one of various likelihood-based "selection model" approaches [5-6]. For example, reference [7] has a nice approach that also has an R package, weightr.

Response: We agree with the Reviewer in that publication bias is an important issue for any meta-analysis to address. Simulation studies have indeed revealed that Egger’s test and especially trim-and-fill methods do not perform well under all conditions. As mentioned by the Reviewer, new methods have been developed, though none of the existing methods devised to assess or correct publication bias perform well under all conditions. For example, Yin et al. (2015, p. 357) concluded:

“But considered realistically, we had to face the fact that the available evidence in meta-analysis data about the amount of study selection is extremely limited, and so no statistical method is going to solve the publication bias problem without making strong and essentially unverifiable assumptions”.

As there is no ‘gold standard’ to assess and correct publication bias, the best strategy may be to apply several methods and to compare their results. Approaches based on weighted selection models have yielded a reasonably good performance under many conditions. The problem with these models is that they require strength assumptions on how publication bias has occurred, as McShane et al. (2016) acknowledged: “Idealistic model assumptions underlie even the most advanced selection methods, and population average effect size estimates can be highly sensitive to these assumptions” (p. 732, Table 1, point 7). In contrast, the PET-PEESE (Precision-Effect Test–Precision-Effect Estimate with Standard Error) approach proposed by Stanley and Doucouliagos (2014) does not posit assumptions on a data model and a selection model. In addition, a recent simulation study (Carter et al., 2019) has demonstrated that the PET-PEESE approach has a performance that is very similar to that of selection models. Therefore, we have decided to complement the methods currently included in our protocol (Egger’s test and trim-and-fill) with the PET-PEESE. We have added the following text (lines 268-270, marked copy):

“The presence of publication bias will be examined using the “funnel plot” method using Duval and Tweedie’s trim-and-fill method,[55] the Egger test[56], and the Precision-Effect Test–Precision-Effect Estimate with Standard Error (PET-PEESE) method.[61]”

R.2.6.) Otherwise, the statistical methods look good. The choice of REML with Knapp-Hartung SE adjustments is a good one.

Response: Thank you for your comment.
R.2.7.) I commend the section stating that the authors will explicitly disclose deviations from the preregistered protocol.

Response: Thank you for your recognition.

R.2.8.) Will data and code be made publicly available upon publication?

Response: All relevant data and code is expected to be rendered publicly available in future publications.

Minor comments

R.2.9.) Why restrict the search to studies published before 12/18 (per Abstract)? Has the search already been conducted? If so, was the protocol written prior to conducting any analyses? (I believe this is also journal policy.)

Response: A pre-registered protocol was published in PROSPERO, the international prospective register of systematic review (PROSPERO 2019 CRD42019119785), after the protocol was completed and before the formal screening of search results was started. We confirm the protocol was written prior to any data extraction and no analyses has been yet conducted at the time of submission.

R.2.10.) Lines 173-177: If an article reports on multiple, separate samples, will the largest still be used? Or will all point estimates be included with some modification to the meta-analysis model to account for possible non-independence due, e.g., to similarity of methods or sample characteristics within a paper?

Response: To clarify this point, we have added the following text (lines 184-188, marked copy):

““If an article reports on two or more studies with independent samples, then each independent study will be included as an analysis unit in the meta-analysis. As a large number of moderator variables will be extracted from the studies, the potential non-independence of these studies due to similarity of methods or sample characteristics will be controlled for in subsequent subgroup analyses.””

References cited in the Response Letter


