Research Editor
The BMJ

8th of June, 2021

Dear Dr. Islam,

BMJ-2021-065492 entitled "Terminal decline in objective and self-reported measures of motor function over 10-years before death: results from the Whitehall II cohort study"

Thank you for the good news and giving us the opportunity to revise our paper. Please see below our detailed response to comments from the editorial board and the reviewers. The response has been itemized and includes the comment, our response, and changes made to the manuscript (included in the response letter as text in boxes).

We hope that the revised version of our paper has addressed the concerns raised and is now suitable for publication.

We would like to thank you for considering our paper and look forward to hearing from you.

Yours sincerely,

Benjamin Landré, on behalf of all authors

Report from The BMJ's manuscript committee meeting

1. Authors should be very clear as to how this single cohort study advances the field in the context of there being multiple systematic reviews and meta-analyses of nearly a dozen other cohort studies published in BMJ and JAMA.

OUR RESPONSE: We agree that a number of individual studies and meta-analyses have examined the association between motor function and mortality. Our study addresses the following limitations in these studies:

- No previous study has examined both measures of objective and subjective measures of motor function.
- Most studies use time to event analyses where the HR reflects the association of between-person differences in motor function but does not inform how motor function changes in the years preceding death. Within-person processes are less subject to residual confounding, allowing better assessment of the trajectory of motor function before death.
- Some studies include "distance from death" in the analysis of trajectories. We use an innovative approach (backward trajectories) where the repeat measures of motor function in the trajectories are anchored to death. This approach leads to the "distance" from death being the same in all participants who died when they are compared to participants who were alive in the analysis of trajectories.

The revised manuscript highlights these points, with addition of the following text to the introduction and the discussion section of the revised manuscript.

Introduction, line 94 to 100

To address these limitations, the aim of this longitudinal cohort study was to examine multiple measures of motor function for their associations with mortality using time-to-event analyses to capture the importance of between-person differences in motor function and retrospective trajectory analyses to compare within-person change in motor function over 10 years in survivors and deceased participants. Use of this twin analytic strategy allows both between- and within-person differences in motor function to be examined in relation to mortality in the same study, the latter being reflected in the shape of the change in motor function leading to death.

Discussion, line 312 to 318

…Use of the terminal decline framework allows better understanding of the relationship between motor function and mortality due to assessment of within-person changes[9 28 29] in motor function. Time-to-event analysis identifies the relevance of specific motor function measures and the HR estimates reflect between- rather than within-person differences in motor function. The originality of our approach is the use of retrospective trajectories, anchored to the date of death, so that distance to death is the same in those who died in comparisons of motor function with survivors. Increase in heterogeneity in individual trajectories is a hallmark of ageing;[5 6] our analysis shows this heterogeneity to be meaningfully associated with mortality.

2. Of 10,308 recruited at baseline, ~50% were analysed in this study. Could you please elaborate on the potential impact of that on the findings. On this note, could you consider an alternative approach such as multiple imputation.

OUR RESPONSE: Participants in the Whitehall II study were on average 45 years old at recruitment to the study in 1985; measures of motor and (cognitive) function were introduced to the study at later waves when participants were older. As in most longitudinal studies, some participants dropped out immediately after baseline, with little attrition after the introduction of the motor function protocol. The wave to wave response rate in recent waves has been over 80%, suggesting major bias is unlikely. However, we agree that the analyses are based on smaller numbers than the number recruited at baseline. In the revised manuscript we have added analyses using inverse probability weighting (similar to multiple imputation) in the new eTable 3. These results show that these results are broadly consistent with the results reported in the manuscript.

Method, line 220 to 225

Second, to examine the impact of missing data the Cox regression analysis was repeated using inverse probability weighting to reflect the study population at recruitment (1985).[27] This involved calculation of the probability of being included in the present study among those alive using data from baseline on sociodemographic factors and health behaviours as well as data on chronic conditions over the follow-up; then the inverse of these probabilities was used as weights in the Cox regression.

Results, line 266 to 267

The use of inverse probability weighting to account for missing data yielded results similar to those in the main analyses (eTable 3).

3. Multiple measurements on the motor functions is a strength of the study. The Editors think it would be more interesting to examine the association between the 'changes in motor functions' and all-cause mortality. Such analysis will make the best use of longitudinal measurements.

OUR RESPONSE: Thank you for this suggestion. In the revised manuscript we have added analyses on change in motor function between 2007-2009 and 2012-2013 and subsequent mortality in eTable 6. Please see the text box below for details of the results.
We have included these results as supplemental data rather than the main analyses for the following reasons:
1) Examination of terminal decline is not possible in time-to-event analysis due to the requirement of a follow-up period for deaths.
2) The follow-up period in time-to-event analyses varies from zero to the end of follow-up and the hazard ratio, based on the non-proportionality assumption, reflects an averaged effect over the period. However, it is likely that change in motor function in those who die within a few days or months is different from those who die a few years later. In the revised manuscript we have included a section on terminal decline to clarify this point (see below).
3) The time-to-event analyses described below are based only on deaths that occurred after 2012-2013 (316 instead of 484 in the analysis of trajectories). Smaller number of deaths reduces statistical power and the analysis is also based on a more selected population as these participants need to be alive until at least 2013.

Methods, line 230 to 233

Finally, in an alternate approach to assessment of change in motor function we examined the association between change in motor function over the first two measures of motor function and subsequent mortality using Cox regression and the same covariates as in the main analyses drawn from the 2012-2013 assessments.

Results, line 270 to 274

Among the 4,606 participants with motor function data in 2007-2009 and 2012-2013 assessments (eTable 6), decline of one SD in walking speed (HR 1.18, 1.05 to 1.32), grip strength (HR 1.22, 1.04 to 1.42), and PCS score (HR 1.16, 1.03 to 1.29), but not timed 5 chair-rises (HR 0.93, 0.84 to 1.03), was associated with higher risk of mortality. Compared to those with no IADL/ADL limitations at these waves, participants who developed a limitation had a higher risk of mortality (HR 1.37, 1.00 to 1.87).

Discussion, line 360 to 366

Studies with repeat measures of motor function have shown change in walking speed[41] and grip strength in older adults to be associated with mortality in Cox regression.[42 43] In the present study, analysis of change in motor function between 2007-2009 and 2012-2013 found change in both objective (walking speed, grip strength) and self-reported (physical component summary score and limitations in ADL/IADL) motor function to be associated with mortality (eTable 6). However, this approach provides only a mean hazard ratio over the follow-up, which could vary from a few months to several years, rather than change in motor function in the years leading to death.

4. Could you please elaborate on the generalisability of the findings given the study population were predominantly men (n=4106, 73%), white (n=5244, 92.9%), high SES (43.9%), active (57.3%), which limits the generalisability of the findings.

OUR RESPONSE: The generalisability concern is flagged up regularly by reviewers of manuscripts based on the Whitehall data. Our view is that the specificities of the study population is not problematic for risk factor research, such as the present study, two reasons. One, results from the Whitehall II study have been used in a range of collaborative meta-analyses without showing difference in effect estimates of risk factor-outcome association. Two, where possible we have compared the association of risk factors with health outcomes to general population studies and found them to be similar to that in the Whitehall II study. This is because although the distribution (incidence and prevalence) of risk factors and the outcomes is favourable in the Whitehall II study, the association between the two is no different to that in the general population. Thus, data from the study are not suitable for calculation of incidence and prevalence of disease or risk factors but the effect estimates of associations between risk factors and health outcomes are generalizable. We have added the following text to highlight this in the revised manuscript.

Discussion, line 332 to 338

Fourth, data are based on an occupational cohort at recruitment and participants were healthier than the general population, in terms of risk factors levels and incidence of disease. However, this does not necessarily affect risk factor-disease associations.[32] For example, the associations of walking speed with mortality risk factors in Whitehall II, such as smoking, obesity, hypertension and diabetes are comparable to those found in 21 other cohort studies[33 34]and the association between cardiovascular risk factors and CVD incidence in the Whitehall II study is similar to that in general population studies.[33]

References cited

- Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. Int J Epidemiol 2013;42(4):1012-4. doi: 10.1093/ije/dys223 [published Online First: 2013/09/26]

- *Stringhini S, Carmeli C, Jokela M et al. Socioeconomic status, non-communicable disease risk factors, and walking speed in older adults: multi-cohort population based study. BMJ 2018; 360: k1046.*
- *Batty GD, Shipley M, Tabak A, et al. Generalizability of occupational cohort study findings. Epidemiol 2014; 25: 932-3.*

5. How did you measure timings for walking speed? Will this test be representative for patient's 'normal walking speed?' I can imagine some participants will be fine over 8ft but not further. Is 8 ft even long enough to gather momentum for 'normal walking speed.'

OUR RESPONSE: Apologies for unclear description in the original version of the manuscript. We have now clarified the manner in which motor function was measured, please see text box below. The goal of the assessment of walking speed is to measure lower extremity function rather than cardiovascular fitness as for older adults mobility rather than endurance appears to be critical (PMID: 30720555). Walking speed measures, such as ours, are widely used in research settings and have a remarkably robust association with health outcomes and mortality.

Methods, line 127 to 133
Walking speed was measured over an 8-ft (2.44 m) marked course, with no obstructions for an additional 2 feet at either end. Participants wore either low-heeled close-fitting footwear or walked barefoot with instructions to 'walk to the other end of the course at your usual walking pace, just as if you were walking down the street to go to the shops. Walk all the way past the other end of the tape before you stop'. Three tests were conducted and the time taken to complete the test was recorded by a research nurse using a stop-watch; the mean of three trials (meters per second) was used in the analysis. Use of a walking stick, if habitual, was allowed.

Reference cited

- *Mehmet H, Robinson SR, Yang AWH. Assessment of Gait Speed in Older Adults. J Geriatr Phys Ther. 2020;43(1):42-52.*

6. As the reviewer Xu points out, it would be better to identify changes in physical function that predict mortality in individuals, or the patterns that predict a longer survival, for example.

OUR RESPONSE: We have summarised the drawbacks of such analyses in our response to comments 1 and 3 above.

7. How well do the curves in Figure 1 fit – what % of variance is explained?

OUR RESPONSE: The results shown in Figure 1 are from mixed models where the aim is to reflect the observed data after taking covariates into account and % variance explained is not thought to be relevant. In "Response letter results 1" at the end of this response letter, we have included observed mean motor function among survivors and the decedents. These data show similar trajectories to those from the mixed models. In these results we also report log likelihood (for fitted and null models) and R^2 for linear mixed models which range from 0.11 to 0.19. For ADL/IADL limitations, we report quasi-likelihood (for fitted and null model). There does not appear to be a more direct measure of % of variance explained available for these analyses.

8. Re BMI, the dose-response curve is potentially complex - very low BMI is a very strong predictor of mortality, and high BMI to a lesser extent. So BMI should either be in 3-4 groups or quadratic - but certainly not dichotomised.

OUR RESPONSE: Thank you for this suggestion, BMI has now been included in the analyses as a categorical variable: <20, 20-24.9, 25-29.9, ≥30 Kg/m^2.

9. eTable 1 is uninformative.

OUR RESPONSE: Thank you, it has been removed from the revised manuscript.

10. eTable 2 is virtually identical to Table 2 (age vs time in Cox regression). Please consider taking it off.

OUR RESPONSE: Thank you, it has been removed from the revised manuscript.

11. "5-chair rises" looks like raising 5 chairs. Please consider "5 chair-rises" or just "chair rises".

OUR RESPONSE: Thank you, we have modified the term as suggested.

12. Could you please define the 'period leading to death' more precisely ?

OUR RESPONSE: Thank you, we have modified the methods section to clarify that the entire 10-year period preceding death (the deceased) or end of follow-up (the survivors) is considered to be the 'period leading to death'.

Methods, line 275
Retrospective trajectories of motor function over 10 years leading to death

13. Typo: 6chair rise (last sentence).

OUR RESPONSE: Apologies, the typo has been corrected.

14. The phrase "terminal decline" sounds as if it is referring to time immediately preceding deaths. Could you elaborate on this, and make sure it is used properly.

OUR RESPONSE: Thank you, the revised manuscript contains the following clarification:

Discussion, line 303 to 308
…Two, trajectories of motor function over 10 years using a backward time scale showed divergence, or terminal decline, in timed chair rises, physical component summary score (SF-36), and ADL/IADL limitations starting 10, 7, and 4 years before death respectively. Given the definition of terminal decline as accelerated decline in functioning before death,[9] or specifically divergence in trajectories of function, our results suggest important differences in terminal decline as a function of specific measures of motor function.

15. The use of both measured motor activity markers and self-reported markers is quite interesting. Could the authors provide correlation matrices for patients to show the inter-relationship between them?

OUR RESPONSE: Thank you for this suggestion, a correlation matrix has been added to revised manuscript in the supplemental data (eTable 1).

16. Clarify when self-reported measures were measured - was it 3 times, like the motor activity measures, or just once, or annually?

OUR RESPONSE: Apologies for omission. We have added this information to the revised manuscript.

Methods, page 145 Self-reported measures (2007-2009, 2012-2013, and 2015-2016)

Reviewer(s)' Comments to Author

Reviewer: 1

1. This study on age-related decline in motor function, with emphasis on terminal decline, towards the end of life, is an important and timely piece of work. The topic is relevant, as the population ages, and ways of identifying and therefore potentially reducing age-associated morbidity and frailty, are becoming more necessary. The work is clearly presented, and adds new findings to this area of research, as well as emphasising the overall importance of motor function as an indicator of health status.

OUR RESPONSE: Thank you for the positive feedback.

2. Some points to consider: is this population representative of the UK population, making the results generalisable? Could some comments be included on this ?

OUR RESPONSE: Please see our response to comment 4 from the BMJ's manuscript committee meeting.

3. The population is relatively young, from a geriatrician's perspective. Was the rate of death expected for this age group?

OUR RESPONSE: Few of the Whitehall II participants are in the 9th decade of life and none in the 10th decade so the participants are indeed young from a geriatrician's perspective. Mortality, like most health outcomes, was lower in our cohort than in the general population, as in all longitudinal cohort studies where participants tend to be healthier than non-participants. As detailed in our response to

comment 4 from the editors, the fact that mortality rates are lower in the Whitehall II study has little effect on the association between motor function and mortality.

4. It is interesting that walking speed differences did not increase in the period before death. Is there any hypothesis as to why this is the case? Was there any way of capturing whether people became bedbound/unable to walk?

OUR RESPONSE: This is an interesting and intuitive observation. However, as previous studies have not examined this issue we are unable to compare our findings with those from other studies. It is possible that differences in walking speed are established by midlife and track over the course of ageing, particularly as our data are necessarily on people who are not bedbound; terminal decline may be seen in longer distance walking speed which measures endurance rather than lower extremity function(PMID: 30720555).

Reference cited

- *Mehmet H, Robinson SR, Yang AWH. Assessment of Gait Speed in Older Adults. J Geriatr Phys Ther. 2020;43(1):42-52.*

5. Minor point = typo line 56.

OUR RESPONSE: Apologies, the typo has been corrected.

Reviewer: 2

1. RE: "Terminal decline in objective and self-reported measures of motor function over 10-years before death: results from the Whitehall II cohort study" by Landre et al. This study examined motor function in relation to mortality based on the Whitehall II cohort study including 6194 participants with a mean age of 65 at baseline. The results showed that physical function decline is associated with mortality. Following comments may help to improve this study.

OUR RESPONSE: Thank you, please see our response to comments below.

Major concerns

2. Originality. A number of studies have shown that physical function decline or impairment has been associated with higher mortality risk and adverse health outcomes. It would be more interesting to show the patterns of changes in physical function predicting mortality in healthy older adults, and decline in function related to mortality reflects medical conditions, instead of mortality.

OUR RESPONSE: Please see out response to comment 1 from the editors.

3. Study population. 1) The original study participants were 10,308, and only 6,194 were included in the current study. The authors may want to discuss how the dropouts would affect the results and the generalizability. 2) Participants with function impairment/disability or dementia should have been excluded from the study population (i.e., disability-free participants), as these conditions might have

driven the observed associations. 3) The flow chart (eFig 1) is very difficult to understand and seems to show 3 separate populations without showing how many people with all repeated measurements of function. In fact, the flow-chart in this manuscript should not be called a flow-chart, which should show the populations that remained, died, and dropped out at each examination time). The authors should consider improving the Fig.

OUR RESPONSE: Thank you for raising these points, our response is below.

1) Please see our response to comment 2 from the Editors on use of inverse probability weighting to account for missing data and comment 4 from the Editors on generalizability.
2) We undertook two additional analyses (1) excluding participants with dementia or ADL limitations from the analyses and (2) excluding participants with poorest motor performance, using the 2.5 SD cut-off. These exclusion do not affect results substantially ("Response letter results 2" at the end of this letter). We prefer to keep the analyses on all participants in the manuscript to be able to examine the full spectrum of ageing.
3) We have now added the following text to better explain the flow-chart.

Results, page 237 to 241
Assessment of motor function was introduced to the study protocol at the 2007-2009 wave of data collection when the age range of participants was 55 to 79 years, and repeated in 2012-2013 and 2015-2016 leading to smaller numbers in analyses due to drop-out and mortality (eFigure 1). The analyses of motor function trajectories were based on 6,194 of participants with data on at least 1 out of 3 waves of motor function and the covariates.

4. Assessment of physical function. ADL and IADL should not be combined as they measure the different activities of daily living with different scales, and ADL disability indicates a more advanced disability than IADL. Otherwise, they should be merged after standardizations.
The outcome. It is interesting to show function changes related to cause-specific mortality by stratified analysis. The validation of the registry-based ascertainment of death should be reported.

OUR RESPONSE: We agree with the reviewer, ADL and IADL limitations are not the same. It is for this reason that we have included both combined scale (in the main analysis) and separate ADL and IADL analyses in eTable 9. We also agree that cause-specific mortality would be an important outcome but we lack the statistical power to do so, as now noted in the limitations section of the revised manuscript (see below). Please note that mortality data are based on information recorded when deaths are certified and registered, and this information is virtually complete in the UK.

Discussion, line 327 to 329
First, we were not able to examine trajectories of motor function separately by cause of death due to small number of deaths in categories of major causes of death.

5. Statistical analysis. 1) I would be more clinical relevant to identify a cut-off of function decline that may predict mortality 10 years later for people with and with multimorbidity. 2) Stratified analysis in people with and without chronic diseases would show different associations between function decline and mortality. Cognitive function should have been taken into account in the analysis. 3) As data on healthy lifestyle factors are available, the authors may want to identify which lifestyle factors may counteract function decline to prolong survival.

OUR RESPONSE: Thank you for raising these points, please see our response below.

1) We agree with the reviewer that it would be clinically pertinent to have a cut-off of functional decline that predicts mortality. However, our analyses show that the association applies across the range of the motor function measure, leading us to use continuous measures of motor function in the time-to-event analyses (1 SD difference or 1 SD change in motor function).
2) The main analyses are already adjusted for chronic diseases. We undertook additional analyses in those with and without multimorbidity (eTable 4) and the results were fairly similar in both groups. Adding cognitive function to the list of covariates (eTable 5 & 10) also did not modify findings.
3) The focus of our manuscript is on the association of objective and subjective measures of motor function with mortality. The analyses were adjusted for lifestyle factors. While research on the association between lifestyle factors and mortality is interesting it is beyond the scope of the present paper.

Methods, line 225 to 228

…Third, the role of chronic diseases was examined in time-to-event analyses stratified by the status of multimorbidity at the assessment of motor function. Fourth, the possible influence of cognitive function was examined by adding a measure of global cognition (the Mini Mental State Examination) as a covariate to the analyses.

Results, line 267 to 269

The association of motor function with mortality was similar in those with and without multimorbidity (eTable 4). Further adjustment for cognitive function did not alter findings (eTable 5).

Results, line 296 to 297

Adjustment for cognitive function did not alter the main findings (eTable 10).

Discussion, line 391 to 396

Chronic diseases are thought to be important drivers of motor decline; in the present study, adding the multimorbidity score to the analysis attenuated the associations in both time-to-event and backward trajectories analyses. The importance of chronic diseases might be due to processes of chronic inflammation and oxidative stress; these are likely to operate across the lifecourse[52] as demonstrated by diverging motor function trajectories prior to death in early old age in our study. However, in our analyses the association between motor function and mortality was also observed in participants free of multimorbidity.

6. Minor comments.

1. Using BMI as the covariate instead of obesity, as obesity is reversely related to both physical function and mortality among older people.

2. Number of people with dementia or depression might have been underestimated due to the ascertainment of the conditions based on medical records.

OUR RESPONSE: Thank you, please see our response below.

1. In the revised manuscript categorical BMI variable has been included in the analyses as <20, 20-24.9, 25-29.9, ≥30 Kg/m^2.

2. Please note that data on depression come from self-report of longstanding illness, use of antidepressants, and medical records. Use of medical records for conditions such as depression and dementia is not the gold-standard and is likely to miss milder cases. However, it has the advantage of providing data on persons who may not be willing or able to attend a clinical assessment as part of the Whitehall II study to diagnose depression/dementia. Thus, it is possible that the bias introduced by use of medical records is small compared to the advantages linked to availability of data on all participants.

Reviewer: 3

1. Terminal decline in objective and self-reported measures of motor function over 10-years before death: results from the Whitehall II cohort study This study is a helpful addition to the body of evidence on ageing, supporting the case for early interventions to address future decline and I believe the inclusion of self-reported subjective measures is a positive one - it gives participants agency, places their lived experience at the centre of the research and gives a richer understanding of changes to an individual person's function over time. While the authors report no current public involvement in the Whitehall II cohort study I strongly believe there would be considerable value in future public involvement. I would encourage them to consider options to draw in public views that could help identify priorities for further research and recommend strategies to effect change in policy and practice.

OUR RESPONSE: Thank you, we agree with the reviewer that it is important to consider both self-reported and measures of motor function. The results clearly support the importance of both assessment approaches.
We are grateful for the reviewer's advice on public involvement, this is an ongoing area of discussion within the Whitehall II study team. We are keen to involve the study participants as well as the public at large in discussions on future priorities for research. However, only participants of the study have so far been involved in exchange with the research team, particularly when the research questions addressed in the study have changed. These interactions have involved focus groups and feedback after clinical examinations. At the moment the primary mechanism of feedback from the wider public is via exchange at presentations to lay audiences. We recognize the importance of public involvement for instigating change in policy and practice and we are in the process of seeking funding to strengthen this aspect of the study.

2. My comments relate to representativeness of the cohort and the generalisability of the findings to the wider population. Although many of these issues have been raised before, they are still relevant. Changes in the makeup of the working population and changes to the nature of employment in the last 30+ years are significant and should be acknowledged. Such secure employment is now much less common.

OUR RESPONSE: Thank you for raising the issue of generalizability. Whitehall II is a closed cohort, in that it follows the same individuals over time, allowing us to examine changes in motor function in this case and other research on incidence of chronic diseases. As the reviewer notes, this is a group

that mostly experienced stable employment and are healthier than the general population. Given these considerations, we have examined association between risk factors and various chronic diseases and found them to be similar to that in other studies (as in collaborative meta-analyses) and when compared directly to results from general population studies. We have added the following text to the revised manuscript.

Discussion, line 333 to 339

Fourth, data are based on an occupational cohort at recruitment and participants were healthier than the general population, in terms of risk factors levels and incidence of disease. However, this does not necessarily affect risk factor-disease associations.[32] For example, the associations of walking speed with mortality risk factors in Whitehall II, such as smoking, obesity, hypertension and diabetes are comparable to those found in 21 other cohort studies[33 34] and the association between cardiovascular risk factors and CVD incidence in the Whitehall II study is similar to that in general population studies.[33]

3. The Whitehall II study is an occupational cohort of people working in the civil service in London, and women and minority ethnic populations are underrepresented. The ethnicity of the cohort is here identified as white or non-white – the very low percentage of workers identified as non-white makes it incredibly difficult if not impossible to translate findings in order to understand variations between individual ethnic populations e.g. people of South Asian heritage who are known to have higher rates of CVD. Identifying if studies meet the needs of different populations equally is a topic of great urgency and importance and it's important to identify gaps in what we know and to call for further research that can address disparities and gaps in our understanding.

OUR RESPONSE: Approximately 10% of the Whitehall II participants are "non-white", reflecting the ethnic mix of the UK in the 1980s when the cohort was set up. For example, the 2011 census puts the non-white composition at 13%. We agree with the reviewer on the great urgency and importance of understanding health in minority ethnic populations. This would require a different study design whereby specific ethnic groups (e.g. South Asians) are oversampled at recruitment to the study. Clearly, the Whitehall II study set up in 1985 is not well suited for this purpose. We have added this limitation to the revised manuscript.

Discussion, line 3339 to 340

Fifth, the ethnicity distribution in the study reflects the UK population 30 years ago and the study lacks sufficient numbers to allow analyses in specific minority groups.

4. Moreover, the study participants are geographically bound to London and the protective effect of walking to and from work in the work-life patterns of commuting workers in London has already been identified as specific and not representative of wider population across the UK. Recruitment practices have changed significantly since the beginning of this cohort study and equality legislation has been strengthened meaning that the current intake of civil servants are more likely to include people with disabilities. It would also be hugely helpful to know the experiences of people with caring responsibilities -previous research shows that carers have less opportunity to look after their own health.

OUR RESPONSE: Apologies for the confusion. The participants were based in London at recruitment to the study in 1985 but now live all over the country and abroad (Spain), primarily due to retirement. We have added this information to the methods section to better describe data collection.

Methods, line 106 to 107
Since baseline, follow-up clinical examinations have taken place approximately every 4-5 years using home-based assessment for those who choose this option and clinic-based assessments (London and major cities in the UK) for others; each wave takes approximately two years to complete.

We did not look specifically at caring responsibilities but all the analyses were adjusted for the health status of participants using detailed data on chronic diseases.

5. All of these issues may need to be teased out – it would be helpful to explore the differences in the cohort with members of the public who reflect the diverse populations living in and around London in order to identify the relative importance of those differences and what they may imply for findings of this study. I do think these potential limitations should be acknowledged in the discussion.

OUR RESPONSE: As described in our response to the point above, the participants of the Whitehall II study now live all over the country. The clinical examination involves either the research nurse travelling to the participants' homes for data collection (mostly the oldest participants or those who are unwell) or participants travelling to the clinic (in London and major cities in the UK).

6. One minor typo – the link to the participant portal doesn't work in the document https://www.ucl.ac.uk/whitehallII/participants/

OUR RESPONSE: Apologies, this has now been corrected in the revised manuscript. https://www.ucl.ac.uk/epidemiology-health-care/research/epidemiology-and-public-health/research/whitehall-ii/participants-area

Response letter results 1

	Log Likelihood	R squared[a]
Adjusted model[b]	-18,310.3	0.19
Null model[c]	-19,403.9	

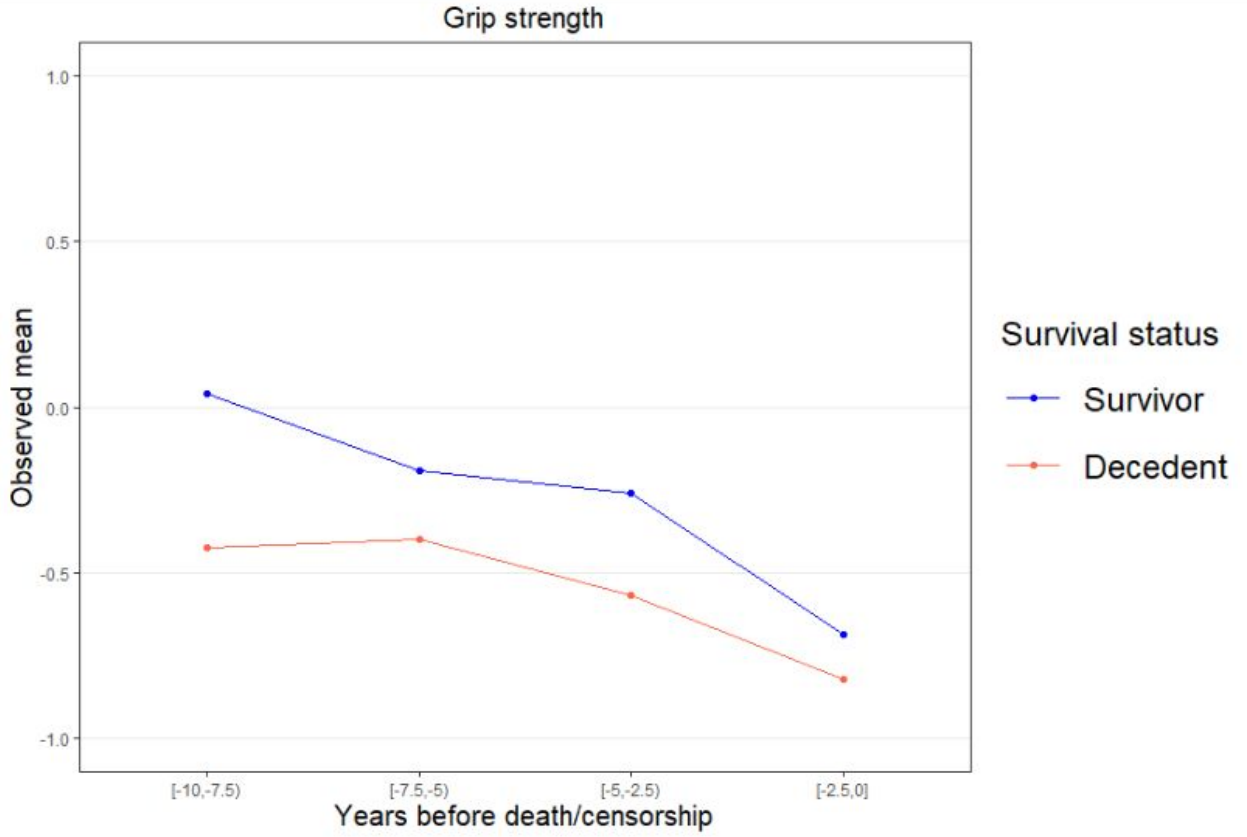
a. Marginal R squared calculated based on Nakagawa et al.(1) formula implemented in MuMIn (version 1.43.17) R package.

b. Model adjusted for age at year 0, sex, ethnicity, marital status, occupational position, vital status, time terms (time & time^2), interactions of these covariates with time terms, and health behaviours, 9-point multimorbidity score and BMI categories at motor function measurement.

c. Intercept only model.

Reference cited

(1) Nakagawa, S., Schielzeth, H. (2013) A general and simple method for obtaining R^2 from Generalized Linear Mixed-effects Models. Methods in Ecology and Evolution 4: 133–142.



	Likelihood	R squared[a]
Adjusted model[b]	-17,902.8	0.19
Null model[c]	-19,818.1	

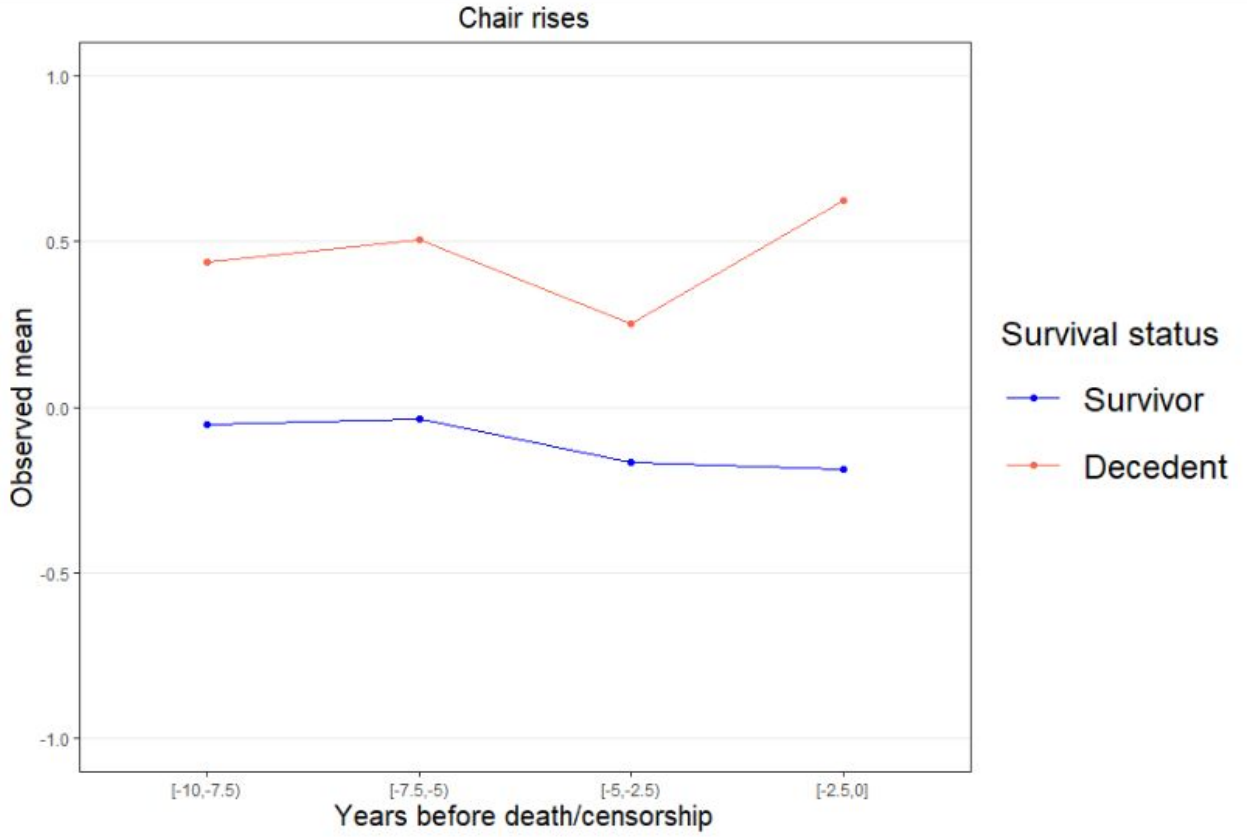
a. Marginal R squared calculated based on Nakagawa et al.(1) formula implemented in MuMIn (version 1.43.17) R package.

b. Model adjusted for age at year 0, sex, ethnicity, marital status, occupational position, vital status, time terms (time & time^2), interactions of these covariates with time terms, and health behaviours, 9-point multimorbidity score and BMI categories at motor function measurement.

c. Intercept only model.

Reference cited

(1) Nakagawa, S., Schielzeth, H. (2013) A general and simple method for obtaining R^2 from Generalized Linear Mixed-effects Models. Methods in Ecology and Evolution 4: 133–142.



	Likelihood	R squared[a]
Adjusted model[b]	-19,826.0	0.11
Null model[c]	-20,363.9	

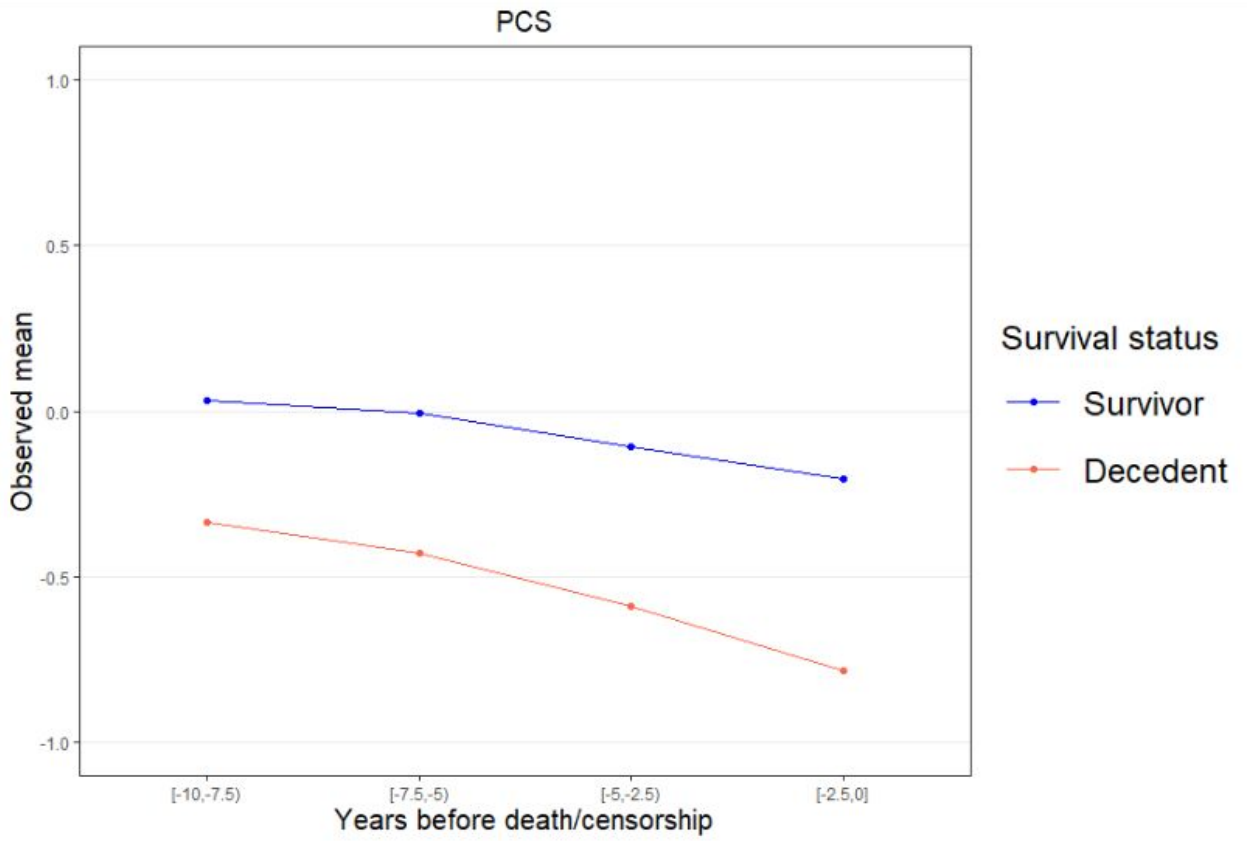
a. Marginal R squared calculated based on Nakagawa et al.(1) formula implemented in MuMIn (version 1.43.17) R package.

b. Model adjusted for age at year 0, sex, ethnicity, marital status, occupational position, vital status, time terms (time & time^2), interactions of these covariates with time terms, and health behaviours, 9-point multimorbidity score and BMI categories at motor function measurement.

c. Intercept only model.

Reference cited

(1) Nakagawa, S., Schielzeth, H. (2013) A general and simple method for obtaining R^2 from Generalized Linear Mixed-effects Models. Methods in Ecology and Evolution 4: 133–142.



	Log Likelihood	R squared[a]
Adjusted model[b]	-18,954.36	0.18
Null model[c]	-20,018.33	

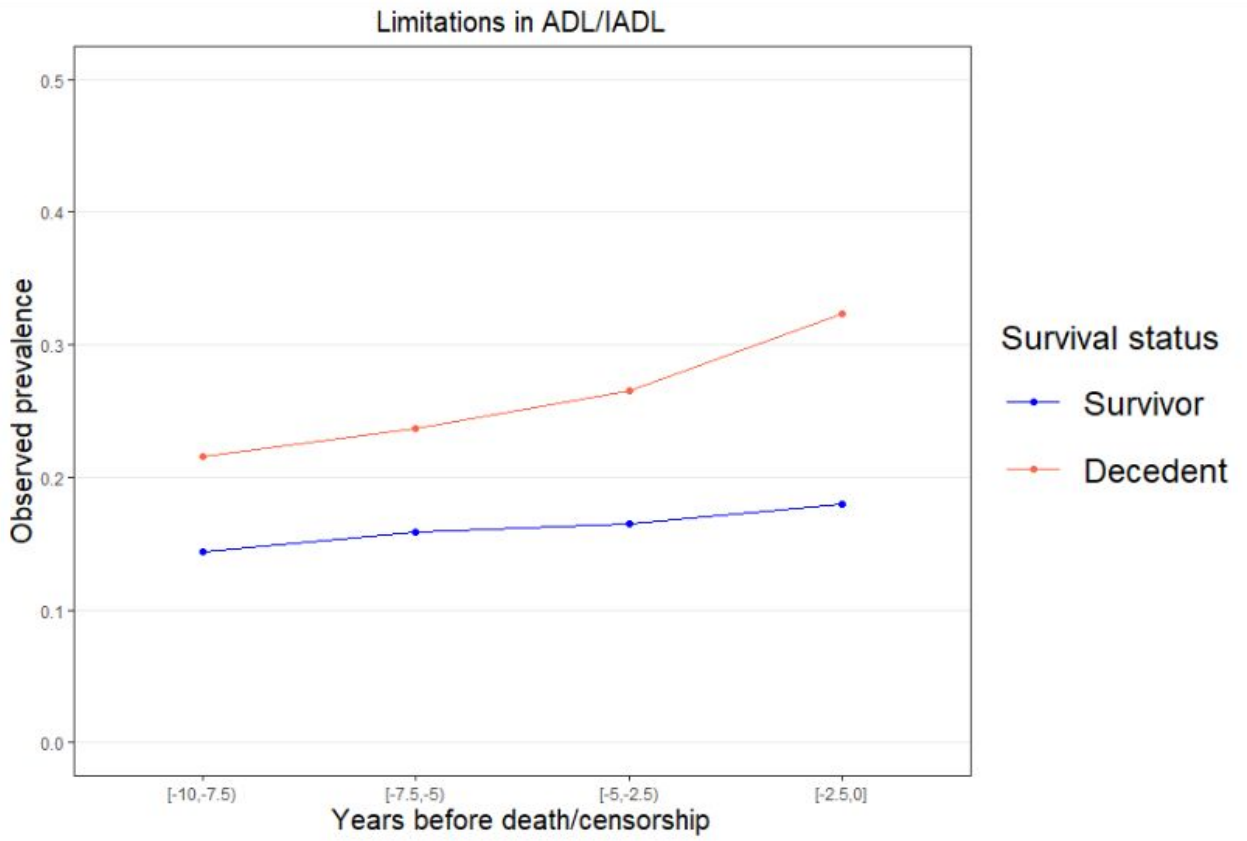
a. Marginal R squared calculated based on Nakagawa et al.(1) formula implemented in MuMIn (version 1.43.17) R package.

b. Model adjusted for age at year 0, sex, ethnicity, marital status, occupational position, vital status, time terms (time & time^2), interactions of these covariates with time terms, and health behaviours, 9-point multimorbidity score and BMI categories at motor function measurement.

c. Intercept only model.

Reference cited

(1) Nakagawa, S., Schielzeth, H. (2013) A general and simple method for obtaining R^2 from Generalized Linear Mixed-effects Models. Methods in Ecology and Evolution 4: 133–142.



	Quasi likelihood
Adjusted model[a]	-6,112.38603
Null model[b]	-6,804.106781

a. Model adjusted for age at year 0, sex, ethnicity, marital status, occupational position, vital status, time terms (time & time²), interactions of these covariates with time terms, and health behaviours, 9-point multimorbidity score and BMI categories at motor function measurement.
b. Intercept only model.

Response letter results 2

Association between standardized measures of motor function[a] and subsequent mortality in "healthy" and all participants (results in the manuscript).

	Excluding participants with poor motor function[b] HR (95% CI)	**Excluding participants with ADL limitations or dementia** HR (95% CI)	**All participants Main results** HR (95% CI)
N death/N total	**d**	**515/5,147**	**610/5,646**
Walking speed	0.82 (0.75 to 0.90)*	0.86 (0.78 to 0.95)*	0.82 (0.75 to 0.90)*
Grip strength	0.89 (0.81 to 0.98)*	0.92 (0.84 to 1.01)	0.87 (0.80 to 0.94)*
Timed 5 chair-rises[c]	1.10 (1.01 to 1.21)*	1.11 (1.02 to 1.22)*	1.14 (1.07 to 1.23)*
SF-36 PCS score	0.87 (0.79 to 0.96)*	0.85 (0.77 to 0.93)*	0.86 (0.79 to 0.92)*
N death/N total	**e**	**274/4,577**	**359/5,083**
Walking speed	0.76 (0.67 to 0.87)*	0.79 (0.68 to 0.92)*	0.73 (0.64 to 0.82)*
Grip strength	0.86 (0.76 to 0.97)*	0.94 (0.83 to 1.07)	0.87 (0.78 to 0.98)*
Timed 5 chair-rises[c]	1.25 (1.11 to 1.42)*	1.23 (1.11 to 1.35)*	1.20 (1.11 to 1.31)*
SF-36 PCS score	0.83 (0.73 to 0.94)*	0.90 (0.79 to 1.01)	0.86 (0.78 to 0.94)*
N death/N total	**f**	**116/3,974**	**150/4,440**
Walking speed	0.68 (0.55 to 0.84)*	0.68 (0.54 to 0.87)*	0.67 (0.56 to 0.80)*
Grip strength	0.82 (0.66 to 1.01)	0.85 (0.70 to 1.04)	0.78 (0.65 to 0.92)*
Timed 5 chair-rises[c]	1.42 (1.18 to 1.72)*	1.26 (1.12 to 1.40)*	1.16 (1.06 to 1.27)*
SF-36 PCS score	0.79 (0.66 to 0.96)*	0.84 (0.71 to 1.00)	0.82 (0.71 to 0.93)*

Abbreviations: ADL: Activities of Daily Living; CI: Confidence interval; SF-36: Short Form 36 General Health Survey; BMI: Body Mass Index.

a. Standardized using mean and SD from 2007-2009, separately in men and women.

b. Defined as measures below 2.5 SD of the distribution of the motor function.

c. Higher values reflected poor motor function.

d. N death/N total: 601/5,618 for walking speed, 587/5,573 for grip strength, 589/5,559 for timed 5 chair-rises, 565/5,480 for SF-36 PCS score.

e. N death/N total: 343/5,044 for walking speed, 348/5,007 for grip strength, 345/5,011 for timed 5 chair-rises, 326/4,893 for F-36 PCS score.

f. N death/N total: 141/4,401 for walking speed, 133/4,282 for grip strength, 141/4,366 for timed 5 chair-rises, 129/4,238 for SF-36 PCS score.

*p<0.05.

Analyses adjusted for age, sex, ethnicity, marital status, occupational position, health behaviours, BMI categories, and 9-point multimorbidity score.