

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Residential status and health in middle and late life. A population-based study with new data from Spain
AUTHORS	Requena, Miguel; Reher, David

VERSION 1 - REVIEW

REVIEWER	Carlo La Vecchia Università degli Studi di Milano, Italy
REVIEW RETURNED	07-Aug-2019

GENERAL COMMENTS	<p>This is an interesting analysis of mortality according to "residential status" on a sample of 10% of Spanish population. This should be made clear, including in the Abstract, and also the number of deaths should be given together with the death rates and the RRs in the text, since these are relatively small in several strata.</p> <p>This paper has two messages. The first one, i.e. living with a partner is associated to lower mortality is well recognized, and this work essentially contributes to quantification.</p> <p>The second one, i.e. living alone is associated to lower mortality than living with others (mainly in the elderly) is due to the fact that to live alone one cannot be (very) sick. This message is not clear to the reader in the present text..</p>
-------------------------	---

REVIEWER	Kaarina Korhonen University of Helsinki, Finland
REVIEW RETURNED	26-Aug-2019

GENERAL COMMENTS	<p>Thank you for the invitation to review the manuscript entitled "Residential status and health in middle and late life. A population-based study with new data from Spain". The purpose of the study is to examine how living arrangements affect the survival of older adults (p. 4, lines 4–7). The study was performed on a 10% sample of household-dwelling individuals present in the Spanish 2011 Census. Data on mortality in 2012 was obtained from linked vital registers. Residential status in 2011 was classified as living with a partner, living with others, and living alone. Mortality from infectious, endocrine and chronic diseases, cancer, diseases of the circulatory system, and mental disorders, suicides, accidents</p>
-------------------------	--

	<p>and other external causes were analysed separately to tap into the potential mechanisms at play. Analyses were further stratified by sex and age to observe any sex-specific associations and age-patterns.</p> <p>The topic of is very important and timely in the sense that shifting patterns in the way care provision for the elderly is arranged may indeed have implications for the health and survival of older people. However, I have serious concerns about the manuscript.</p> <p>First, the manuscript lacks a coherent conceptual framework. This is shown in, for example, the mismatch between the study aim and the analyses. With the current study design the research questions cannot be answered as the causality of any observed associations cannot be addressed. It is very likely that what the analyses mainly capture is health-related selection into differential living arrangements (one-year mortality follow-up), and not the effect of living arrangement on mortality. The authors also make rather bold causal interpretations based on the cause-of-death analyses. For example, it is true that people living with a partner may have better access to help at myocardial infraction or stroke incidence (p.3, lines 36–43), but it is also possible that these people have better cardiovascular health and thus have lower risk of such an event. The exact mechanisms cannot be known with the available data but these limitations are not discussed in the manuscript in any way. A clearer framework would help to conceptualize the research aims more appropriately. Furthermore, as it is now, the contribution of these analyses to what we already know on the topic remains unclear. A more thorough literature review might assist the authors to better identify the gaps in the literature.</p> <p>Overall, the manuscript appears to be unfinished as it lacks important sections such as methodological considerations, the discussion section is very limited, and the interpretation of the results is very cursory and it seems that some of the results are quoted incorrectly (e.g. p. 9, lines 42–46).</p>
--	--

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Carlo La Vecchia

Institution and Country: Università degli Studi di Milano, Italy

Please state any competing interests or state 'None declared': None declared

This is an interesting analysis of mortality according to "residential status" on a sample of 10% of Spanish population. This should be made clear, including in the Abstract [Clarified, these data have been included in the abstract], and also the number of deaths should be given together with the death rates and the RRs in the text, since these are relatively small in several strata. [The number of deaths by cause has been included as a new panel in Table 3; the number of deaths for sex by age group by living arrangement was already included in Table 2. The final text is updated to underscore the analytical limitations derived from the use of relatively small numbers (see: Article summary, page 5 and the new 'Limitations' section now included at the end of the paper)].

This paper has two messages.

The first one, i.e. living with a partner is associated to lower mortality is well recognized, and this work essentially contributes to quantification.

The second one, i.e. living alone is associated to lower mortality than living with others (mainly in the elderly) is due to the fact that to live alone one cannot be (very) sick. This message is not clear to the reader in the present text. [We have clarified this message in the text, p. 3, p.10 (in the first paragraph of the Discussion section)].

Reviewer: 2

Reviewer Name: Kaarina Korhonen

Institution and Country: University of Helsinki, Finland

Please state any competing interests or state 'None declared': None declared

Thank you for the invitation to review the manuscript entitled "Residential status and health in middle and late life. A population-based study with new data from Spain". The purpose of the study is to examine how living arrangements affect the survival of older adults (p. 4, lines 4–7). The study was performed on a 10% sample of household-dwelling individuals present in the Spanish 2011 Census. Data on mortality in 2012 was obtained from linked vital registers. Residential status in 2011 was classified as living with a partner, living with others, and living alone. Mortality from infectious, endocrine and chronic diseases, cancer, diseases of the circulatory system, and mental disorders, suicides, accidents and other external causes were analysed separately to tap into the potential mechanisms at play. Analyses were further stratified by sex and age to observe any sex-specific associations and age-patterns.

The topic of is very important and timely in the sense that shifting patterns in the way care provision for the elderly is arranged may indeed have implications for the health and survival of older people. However, I have serious concerns about the manuscript.

First, the manuscript lacks a coherent conceptual framework. This is shown in, for example, the mismatch between the study aim and the analyses. [Unfortunately, we disagree with Reviewer #2 on this point. As we explained in the Introduction, our point of departure is that 'implications of residential choices for the health and well-being among mature adults and the elderly are far from clear'. The study aim is to shed light on this issue using an ample and recent sample of the Spanish population. Our basic premise is that one of the main reasons for this lack of clarity is due to the complex relationship between living arrangements and mortality. On the one hand, certain living arrangements can have a direct effect on the risk of death—for example, living with a partner decreases it while living alone tends to increase it. On the other hand, the risk of death can also affect residential options because health-related selection processes exist that may facilitate certain living arrangements for some people but prevent it for others. In this sense, there would be a baseline health component affecting certain living arrangements that would in turn impact the likelihood of death. Thus, the residential arrangement itself would, at least in part, act as an intermediate variable. This is the reason why people living alone in good health may be able to survive longer in their solo households than those living with others provided they moved in with others because they had relatively poor health. Should this be the case, higher mortality will be observed among those not living alone. Both kinds of effects—the direct effect (whether protective or damaging), and the selection by health effect—have been documented in the pertinent literature (see references in the paper). The empirical analysis of residential gradients in mortality enables us to clarify this complex interrelation between residential options and mortality by means of a population-based study. Our findings corroborate the existence of both types of effect in Spain. In this sense, we believe that there is no mismatch at all between the study aim and the analyses. In the revised version of this paper, we have made an effort

to make this dual argument clear to the reader. These adjustments can be seen on page 3-4 (Introduction) and on page 10 (Discussion) and on page 11].

With the current study design the research questions cannot be answered as the causality of any observed associations cannot be addressed. [We agree with the reviewer that with observational population-based studies it is extremely difficult or even impossible to completely prove any causal explanation. However, that is not a reason to turn our backs on the valuable information contained in big databases. The issue of causality and the role of experimental methods has always been a hotly contested one (see, for instance, recent approaches in Deaton AS. Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development. Working Paper No. 14690, National Bureau of Economic Research, 2009; Imbens GW. Better LATE Than Nothing. Some Comments on Deaton. *Journal of Economic Literature*, 2010; 48: 399-423; Imbens GW, Rubin DB. *Causal inference for statistics, social, and biomedical sciences: an introduction*. Cambridge: Cambridge University Press, 2015). In a strict sense, proving that two variables are causally related (explanation) requires a randomized experiment because, in principle, 'correlation is not causation'. But very often such experiments with human beings are not feasible or are ethically unacceptable. Therefore, even strong proponents of experimental methods like Imbens argue that "it would be regrettable if this trend led researchers to avoid questions that cannot be answered through randomized or natural experiments" (Imbens 2010: 420). If the use of observational data is discarded, it would make it almost impossible to assess adequately the impact of many social processes on health. In practice, this means that we can only (try to) control for what are arguably relevant confounders and make sure that potential cofounders are observed at a point in time prior to the outcome being analyzed (death). In our paper we resorted to stratification and adjustment for potential confounders. Still, even when stratifying the sample (by sex and age) and controlling for several possible confounders (educational attainment, size of municipality and migratory status in our case), the possibility always exists that there are unmeasured exposures that would influence the outcome if it were possible to include them in the model. It is true that, for the sake of simplicity, we have used the term 'effects' several times in the text, but the methodological caveats discussed above regarding observational studies still apply. We believe the reader of *BMJ Open* will know, understand and admit these otherwise well-known methodological limitations of any observational study. Many relevant contributions to the study of health and aging are based on observational data that deal with causality only in an indirect way. The paper has been revised to specify these limitations as clearly as possible (most importantly in the Article summary and in a new 'Limitations' section)].

It is very likely that what the analyses mainly capture is health-related selection into differential living arrangements (one-year mortality follow-up), and not the effect of living arrangement on mortality. [That is precisely a key aspect of this paper and is one of our main conclusions. On this point we entirely concur with Reviewer #2, and these arguments are clearly stated in the Abstract, the Introduction and in the Discussion section of this text. In the Abstract, please see : 'These findings point to the existence of powerful health-related selection mechanisms influencing residential choices as people age'; and page 10 please see: 'this paper emphasizes the positive selection effect of good health among those living alone at older ages'; 'This dual pattern associated with age points to the existence of important selection effects whereby people with no spouse who have better health tend to be selected into living alone and those with worse health tend to be selected into living with others'. See also the new 'Conclusions' at the end of the paper].

The authors also make rather bold causal interpretations based on the cause-of-death analyses. [Here we disagree once again. We cannot make 'bold causal interpretations' because of the limitations imposed by the research design and the available data. Once again, we would like to make clear that we are perfectly aware of the methodological limitations of observational designs. This was made clear at the very outset of the paper. See, for example, 'The use of cause of death data helps illustrate the role of specific mechanisms behind differential mortality by residential status' (emphasis added). And in the discussion section, we have indicated that 'cause of death data illustrate the basic

mechanisms involved...'. We are confident that the verb 'to illustrate' does not suggest the existence of 'rather bold causal interpretations', as Reviewer #2 says. In any case, we have added a new clause in the 'Strengths and limitations of this study' section stating that 'Given the observational research design used, only statistical associations are observed and hence strong causal claims are not strictly proven' and a new section on 'Limitations' has been added to the paper].

For example, it is true that people living with a partner may have better access to help at myocardial infarction or stroke incidence (p.3, lines 36–43), but it is also possible that these people have better cardiovascular health and thus have lower risk of such an event. [In fact, it is likely that both prior health-selection effects for living with a partner and the ability of a partner to intervene actively in these catastrophic events are important. With these types of observational data, however, it is impossible to assess the relatively weight of these two factors. So we agree with reviewer #2 on this point that constitutes, in fact, yet another instance of both the positive health selection into certain marital statuses or certain living arrangements, and the specific importance of residing with another person (partner) when a myocardial infarction takes place. This specific issue has been addressed in the introductory section of the paper where several bibliographical references were cited on the importance of a good-health selection process into marriage. (See also references 34, 41-44)].

The exact mechanisms cannot be known with the available data but these limitations are not discussed in the manuscript in any way [We agree with Reviewer #2. In the revised version, we state that our analysis of cause of death is mainly exploratory (page 5 and the new 'Limitations' section). In the revised version of this text, we have provided an example of the mechanisms involved — specifically those related to acute myocardial infarctions. See the text inserted at the start of page 4].

A clearer framework would help to conceptualize the research aims more appropriately. Furthermore, as it is now, the contribution of these analyses to what we already know on the topic remains unclear. A more thorough literature review might assist the authors to better identify the gaps in the literature. [As stated above, we are convinced that the conceptual framework and the research goal of the paper are clearly formulated in its introductory section. This paper contributes to clarifying the complex interrelation between health status, residential options and mortality with a large database based on the Spanish population. Its value added is derived from the statistical analysis of a large national data set that points to: (1) the existence of both direct effects of living arrangements on mortality and health-related selection effects on differential living arrangements; (2) the changing importance of both these effects over the life-cycle, with direct impacts of living arrangements on mortality prevailing among mature adults, and health-related selection effects prevailing among the elderly; and (3) differential strengths of these effects in different national contexts (Spain or Belgium compared with Nordic Countries). These three outputs provide a crucial framework that helps explain the often ambiguous and unclear results that the pertinent literature has often shown when looking at the implications of residential gradients for mortality (see the new 'Conclusions' section at the end of the amended version of the paper). In this sense, we believe that most of the relevant literature on living arrangements and mortality is cited in the paper. Should this not be the case, we would be grateful for any suggestions from Reviewer #2 about neglected references].

Overall, the manuscript appears to be unfinished as it lacks important sections such as methodological considerations, the discussion section is very limited, and the interpretation of the results is very cursory and it seems that some of the results are quoted incorrectly (e.g. p. 9, lines 42–46). [The paper has a complete Methodology section, the limitations of our study are explicitly mentioned throughout the study and especially in a new 'Limitations' section and, after carefully reviewing lines 42–46, we have not found any inexact or incorrect quotation of data].