

## 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

<b>TITLE (PROVISIONAL)</b>	Oral function and cumulative long-term care costs among older Japanese adults: A prospective six-year follow-up study of long care receipt data
<b>AUTHORS</b>	Kojima, Kaori; Saito, Masashige; Miyaguni, Yasuhiro; Okada, Eisaku; Ojima, Toshiyuki

## VERSION 1 – REVIEW

<b>REVIEWER</b>	EPSTEIN, DAVID Universidad de Granada - Campus de Cartuja, Economía Aplicada
<b>REVIEW RETURNED</b>	27-Jul-2022

<b>GENERAL COMMENTS</b>	<p>Overall a nice paper, but please be careful about overclaiming causal relations rather than just association. Also, you need to pay more attention to how deaths are handled in the analysis of costs.</p> <p>Need some background on who pays insurance premiums and how needs are assessed. Physical and mental health criteria? Who is eligible for benefits? People living in their own homes? People in residential or nursing homes? is there a large population ineligible for benefits (eg because they have not paid sufficient social security premiums or are foreign resident etc). Are there special benefits for people who live alone ? and so on..</p> <p>Is there information about who filled in the questionnaire if the person was frail?</p> <p>P6, line 16 "evidence" : are there any RCTs which have compared interventions and can support this statement?</p> <p>"The survey was sent by mail to residents of 12 municipalities between August 2010 and January 2012 and included sex. " This seems to be an incomplete sentence, other variables?</p> <p>"In total, 51302 responses were received (valid response rate: 64.7%) and unknown age was excluded." Unclear what was excluded? Any respondent where age was missing?</p> <p>I understand that you were able to cross-reference the same individual in the claims database with the survey . Please clarify this step. Is there a common ID in both databases that enables this? In that case, in what sense was the ID encrypted if you used the same ID to cross-reference?</p> <p>"Information regarding LTCI costs or deaths was collected from the municipalities." Clairify how you cross referenced with deaths registry.</p>
-------------------------	--

	<p>Cumulative costs will be censored because of death. If you simply add up costs over 6 years you will be underestimating cumulative costs because of the people who die. You must find a way to deal with this in the analysis, as well as the high % of zeros.</p> <p>I would suggest to calculate the annual cost in each of the 6 years, for those that are alive at the end of that year. This would mean to conduct a regression for each year. You can then obtain indirectly an estimate of the mean cumulative cost as the sum of the coefficients over the 6 years. If you did this with bootstrap you could obtain a confidence interval for this cumulative mean cost.</p> <p>Another possible option is to treat the data as a 6 year panel, and include alive/dead status as a covariate that is time-dependent.</p> <p>You dont explain the reason for creating a categorical variable for total cost, rather than analyse it as a continuous variable. Your categories are very unbalanced. The vast majority have no cost and very few are in the highest category</p> <p>The zeros could be handled by a two -part model</p> <p>I think the number of household members is very important and should be a covariate. I understand that all explanatory variables are baseline, including oral status. This needs to be very clearly stated in the abstract and text.</p> <p>Are all respondents living in their home (i.e in the community) ? None in residential homes or nursing homes?</p> <p>"Additionally, according to oral function, the percentages of certification for the need for longterm care and death or displacement" I think this is the first time you have mentioned "certification" and "displacement" . Explain in methods what you are doing here.</p> <p>"The number of days to certification.. " again I think you need to explain this varaible in methods and what does it signify?</p> <p>". If oral function improves, the cumulative cost of longterm care will be reduced by 150.25 million USD per year. " No, this is not a valid statement. This statement assumes direct causality. I think you must be very careful and cautious not to claim that your study shows direct causality between oral function and death or costs. Your study shows an association. Oral function loss may be the proximate cause of future problems , or (as you point out in your paper) may be an indicator of other conditions such as general muscle wasting. Hence patients that display symptoms of oral function loss , or at risk of this, may need a more profound diagnosis and targeted interventions in order to maintain overall health.</p> <p>"respond to both the baseline and follow-up surveys may have consisted of healthy individuals. Measurement bias may have occurred" I disagree that non-response necessarily leads to bias You should check whether</p>
--	---

	the responding individuals are representative of your target population Also, is there any indication of who filled in the questionnaire if the target person could not do this.?
--	--

<b>REVIEWER</b>	Mohammadnezhad, Masoud Fiji National University
-----------------	--

<b>REVIEW RETURNED</b>	23-Oct-2022
------------------------	-------------

<b>GENERAL COMMENTS</b>	<p><b>Abstract</b> Please add “Japan” in the keywords.</p> <p><b>Methods</b> 1. In the Outcome variable, please use reference where you divided the participants to three groups based on cumulative costs. 2. In the Covariate section please use reference where you divided the subjects’ health to good and poor as well as the categorisation you made based on Geriatric Depression Scale for valuating depression.</p> <p><b>Results</b> It is better to have a short explanation for the Table 1 to highlight the main finding for each variable and use this explanation before the Table.</p> <p><b>Discussion</b> 1. There is no need to refer the results to the Table 1 and Table 2. You just can discuss the main finding without referring to tables that presented in the result section 2. I am still not sure why you use Table 3 in the Discussion section without refereeing to its findings. You may just remove Table 3 and explain the main thing as text.</p> <p><b>Conclusion</b> Please provide some recommendations based on the findings of this study in this section.</p>
-------------------------	---

### VERSION 1 – AUTHOR RESPONSE

Reviewer#1:

Dr. DAVID EPSTEIN, Universidad de Granada - Campus de Cartuja.

Overall a nice paper, but please be careful about overclaiming causal relations rather than just association. Also, you need to pay more attention to how deaths are handled in the analysis of costs.

Response: We are very thankful to you for reviewing our manuscript. We have read your Comments carefully and tried our best to address them one by one. We hope that the manuscript has been revised as your expectation. Please note that the responses to your comments have been marked with blue font color in the text.

Comment1: Need some background on who pays insurance premiums and how needs are assessed. Physical and mental health criteria? Who is eligible for benefits? People living in their own homes? People in residential or nursing homes? is there a large population ineligible for benefits (eg because they have not paid sufficient social security premiums or are foreign resident etc). Are there special benefits for people who live alone ? and so on..

Is there information about who filled in the questionnaire if the person was frail?

Response: You have rightly pointed out these shortcomings in our manuscript and we apologize the inadequacy of the description.

As a rule, survey respondents for this study interact with us in person. The payment system of insurance premiums is added in the text in the background explanation of the Japanese long-term care insurance system. The baseline survey is intended for people who are not certified as needing long-term care and therefore live at home. And the delinquency or failure of payment rate for long-term care insurance premiums is not high.

Changes made in the manuscript:

1. -The following has been added in the Introduction section:

Page 5, line 93–111: “Japan's long-term care insurance system was established in 2000 as an insurance system in which society as a whole supports the care of older adults. This insurance system provides benefits to those who need nursing care and supports them to receive appropriate services. It aims to support independence and reduce the burden on family members who provide care. The long-term care insurance system consists of three parties: the insured, the insurer, and the long-term care service provider. Municipalities act as insurers to administer the system, and all citizens aged 40 and over are insured. An insured person is a person who subscribes to long-term care insurance and is eligible to receive long-term care services when he/she is certified as requiring long-term care. If they receive long-term care services, they pay 10% of the cost at the counter if they use the long-term care insurance system. (Depending on income, the co-payment can be up to 30%.) The long-term care service provider that provided long-term care services to the insured person bills the insurer for the cost of the long-term care services, and the insured person receives the cost from the insurer except for the portion paid by the insured person at the counter. The insurer is a municipality. The long-term care insurance system is financed by public funds and long-term care insurance premiums. In total, the municipality pays 50% and the insured pays 50% of the long-term care insurance premiums.”

2. In the Methods section; sub-section, Study sample

Page 7, line 151–153: “The study population was limited to older adults who were not certified as needing long-term care at the time of the survey; it was combined with the actual long-term care insurance benefits held by the government six years later.”

Comment2; P6, line 16 "evidence" :are there any RCTs which have compared interventions and can support this statement?

Response: Our sincere apologies, but we did not properly understand the points raised by the reviewers. Therefore, we have responded to some of the parts of our text that we may point out to you.

We also made corrections where we could.

Changes:

1. Method section; Explanatory variables

– About The Kihon Checklist

Page 9, line 186-195: “Explanatory variables were those related to oral function at the time of the baseline survey. In Japan, to assess whether a person is eligible for nursing care prevention services or LTCL services, use of the Kihon Checklist (KCL) has been recommended by the Ministry of Health, Labour and Welfare.[19] The KCL was created by the Ministry of Health, Labour and Welfare in Japan to help people aged  $\geq 65$  years reflect on their lives and health status and check for any decline in their physical or mental functions [20]. It is used by the local governments, and in community consultations to screen for persons eligible for long-term care prevention programs, and to assess the effectiveness of interventions. The KCL was automatically sent to all individuals  $\geq 65$  years on an annual basis up until 2014, but is now administered at the discretion of each local administration [21].”

2. Reference

“18 Ministry of Health, Labour and Welfare. Guidelines for Comprehensive Projects for Care Prevention and Daily Life Support: Ministry of Health, Labour and Welfare; 2022. chrome-

extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.mhlw.go.jp/file/06-Seisakujouhou-12300000-Roukenkyoku/0000088276.pdf Accessed 6 Nov 2022. (in Japanese).

19 Arai H, Satake S. English translation of the Kihon checklist. *Geriatr Gerontol Int.* 2015;15(4):518–9.

20 Sewo Sampaio PY, Sampaio RA, Yamada M, Ogita M, Arai H. Validation and translation of the Kihon checklist (frailty index) into Brazilian Portuguese. *Geriatr Gerontol Int.* 2014;14(3):561–9.”

– About Oral frailty

There have been cohort studies and systematic scoping review articles in recent years, but currently there are no RCTs. A related Review article has been added to references.

“13 Neelamana, S. K., & Janakiram, C. (2022). Characterizing the Literature on Validity and Assessment Tool of Oral Frailty: A Systematic Scoping Review. *The journal of contemporary dental practice*, 23(6), 659–668.”

Comment3; "The survey was sent by mail to residents of 12 municipalities between August 2010 and January 2012 and included sex. " This seems to be an incomplete sentence, other variables?

Comment4; "In total, 51302 responses were received (valid response rate: 64.7%) and unknown age was excluded." Unclear what was excluded? Any respondent where age was missing?

Response to Comments 3 & 4:

This observation is correct and we apologize for the inadequacy of the text We have now revised the text by adding relevant information. Respondents whose sex and age were missing in the returned questionnaires were not included in the database, because sex and age are essential variables in the analysis.

Changes in the text:

Page7, line 146-151: “The JAGES survey was done by collecting self-administered questionnaires, which were mailed to a random sample of functionally independent individuals aged  $\geq 65$  years, from 12 participating municipalities between August 2010 and January 2012. In total, 51,302 responses were received (valid response rate: 64.7%) and unknown sex and age was excluded.”

Comment5; I understand that you were able to cross-reference the same individual in the claims database with the survey . Please clarify this step. Is there a common ID in both databases that enables this? In that case, in what sense was the ID encrypted if you used the same ID to cross-reference?

Comment6; "Information regarding LTCI costs or deaths was collected from the municipalities."

Clairify how you cross referenced with deaths registry.

Response to comments 5 & 6: Thank you for pointing out this oversight to us, we should have written in detail. I have now revised the text.

Changes in the text:

Page 8, line 155-163: “To ascertain the respondents' subsequent use of LTCIs, information on the actual insurance benefits provided to insurers, data on the certification of long-term care needs held by insurers, deaths, and information on the imposition of LTCI premiums were collected in encrypted forms by the insurers. The provided data and questionnaire survey data were matched on an individual basis by the researcher based on the encrypted IDs to create a cohort data set for analysis. A total of 46,616 individuals (90.9% follow-up rate), excluding untraceable cases including in-migrants and out-migrants were included in the analysis.”

Comment7; Cumulative costs will be censored because of death. If you simply add up costs over 6 years you will be underestimating cumulative costs because of the people who die. You must find a way to deal with this in the analysis, as well as the high % of zeros.

Comment8; I would suggest to calculate the annual cost in each of the 6 years, for those that are alive at the end of that year. This would mean to conduct a regression for each year. You can then obtain indirectly an estimate of the mean cumulative cost as the sum of the coefficients over the 6 years. If you did this with bootstrap you could obtain a confidence interval for this cumulative mean cost.

Response: Thank you for this valuable input.

This survey is conducted every three years. And does not capture annual costs each year. This is the first time that data on cumulative care costs has been attempted. Therefore, we only have six years of cumulative data, and we sincerely apologize for not meeting your expectations. We will definitely keep this point in mind and try it in our future studies when we have enough data.

However, we were able to obtain the mean, standard deviation, and 95% confidence interval of the cumulative costs over the 6-year period, as noted in Table 1. With the data we had, this was the best that we could do.

Comment 9; Another possible option is to treat the data as a 6 year panel, and include alive/dead status as a covariate that is time-dependent.

Response: This is a pertinent suggestion, but unfortunately we are not able to create panel data currently. We will create the panel data sequentially and make it an issue for future research.

Comment10; You dont explain the reason for creating a categorical variable for total cost, rather than analyse it as a continuous variable. Your categories are very unbalanced. The vast majority have no cost and very few are in the highest category. The zeros could be handled by a two -part model

Response: Thank you for pointing this out. This we think is related to comment 8; we have reanalyzed the data with total cost as a continuous variable. I request you to check Table1. We calculated the average for those with  $\geq$ USD 1.

Changes in the text:

Page 11, line 228-231: "Descriptive statistics were sociodemographic variables and mean and percentages of LTCI costs over six years were stratified according to the two stratified groups: zero, and more typology. Percentage comparisons were analyzed using the chi-square test."

Comment11; I think the number of household members is very important and should be a covariate.

Response:

Though this is important, the number of household members could not be ascertained. However, we did have data on the number of people living alone, so we added it to the analysis. As a result, the values in Table 2 were slightly lower. Thank you for this valuable input.

Comment12; I understand that all explanatory variables are baseline, including oral status. This needs to be very clearly stated in the abstract and text.

Response: We agree with your point, and accordingly we have revised the abstract at relevant places.

Changes in the Abstract:

Page 3, lines 54-56: "We adjusted for the presence or absence of oral function problems, age, sex, physical function, and socioeconomic and lifestyle background at the time of the baseline survey."

Changes in the Methods section:

Page 9, line 186: "Explanatory variables were those related to oral function at the time of the baseline survey."

Page 10, lines 211-213: "Sex, age, educational attainment, household equivalized income, and marital status at the time of the baseline survey were used as basic attributes to be considered when examining the association with the use of long-term care services."

Comment13; Are all respondents living in their home (i.e in the community) ? None in residential homes or nursing homes?

Response: At the time of the baseline survey, respondents were not certified as needing long-term care. As such, respondents resided in the community.

Corrections were made, which are shown in the response to Comment 1: 'Changes made in the manuscript - 2. In the Methods section; sub-section, Study sample'.

Comment14; "Additionally, according to oral function, the percentages of certification for the need for longterm care and death or displacement" I think this is the first time you have mentioned "certification" and "displacement" . Explain in methods what you are doing here.

Comment15; "The number of days to certification.. " again I think you need to explain this variable in methods and what does it signify?

Response: This is a very valuable suggestion that you have made, thank you. We have added the description in the Methods section. We have rewritten the Methods and Result accordingly.

Accumulated care costs depend on the length of time care that is needed. Therefore, the Table 3 was created to show the number of days for which care is needed.

Changes in the text: Methods

(In the background, we described the long-term care insurance system.)

Page 11, lines 239-244: "Accumulated care costs depend on the length of time care was needed.

Therefore, the rates of those who were certified as needing care and those who died, and the number of days to get there, were calculated for each group in terms of the number of oral problems.

Information such as the certification of the need for nursing care and the moving out of the country was provided by the insurer."

Comment16; ". If oral function improves, the cumulative cost of longterm care will be reduced by 150.25 million USD per year. " No, this is not a valid statement. This statement assumes direct causality. I think you must be very careful and cautious not to claim that your study shows direct causality between oral function and death or costs. Your study shows an association. Oral function loss may be the proximate cause of future problems , or (as you point out in your paper) may be an indicator of other conditions such as general muscle wasting. Hence patients that display symptoms of oral function loss , or at risk of this, may need a more profound diagnosis and targeted interventions in order to maintain overall health.

Response: We appreciate your valuable suggestion, and we agree with it totally.

Changes in the text:

Page 20, lines 379-380: "The preservation of oral function could lower the individuals cost of care."

Comment17; "respond to both the baseline and follow-up surveys may have consisted of healthy individuals. Measurement bias may have occurred" I disagree that non-Response; necessarily leads to bias You should check whether the responding individuals are representative of your target population

Response; Thank you for this valuable analytical suggestion.

We have shortened this sentence appropriately. (In addition, those who were able to respond to both the baseline and follow-up surveys may have consisted of healthy individuals. Measurement bias may have occurred.)

The data for this survey is a baseline survey combined with administrative data, and

We have not received responses for the 6-year survey. My apologies. The phrase "follow-up survey" is an error on my part.

We have not conducted a follow-up survey, we have only combined the baseline data with data obtained from the administration.

Changes in the text:

Page 21, lines 391-394 (Limitations section): "The study population was limited to older adults who were not certified as needing long-term care at the time of the survey; it was also limited to older adults who could be combined with the actual long-term care insurance benefits held by the government six years later."

Comment18; Also, is there any indication of who filled in the questionnaire if the target person could not do this.?

Response: Forgive me for duplicating my response to comment 1.

We apologize for the inadequacy of the description. As a rule, survey respondents for this study cooperate with us in person. The baseline survey is intended for people who are not certified as needing long-term care and therefore live at home.

Reviewer#2:

Dr. Masoud Mohammadnezhad, Fiji National University

Responses: We appreciate your time and effort to review our manuscript. We have read your Comments carefully and tried our best to address them one by one. We hope that the manuscript has been revised as per your expectation.

Abstract

Comment1 ; Please add "Japan" in the keywords.

Response: Thank you for pointing this out. I have now added Japan to the keywords.

Methods

Comment1

1. In the Outcome variable, please use reference where you divided the participants to three groups based on cumulative costs.

Response: We appreciate your valuable suggestion. I am sorry, but we do not have the related articles to show for these three groups. Previous papers on the cumulative costs of LTC services have often used averages. However, in this study, we wanted to identify cost percentages by severity of oral function, so we divided the patients into three groups.

Comment2

2. In the Covariate section please use reference where you divided the subjects' health to good and poor as well as the categorisation you made based on Geriatric Depression Scale for valuating depression.

Response; We apologize very much for the insufficient references. Papers using these cutoffs have now been added to the reference.

And thank you for pointing out my mistake. Subjective health was removed because it was not used for the adjustment variable. My apologies for not checking before submitting.

Changes in the Reference list:

26 Hoyl MT, Alessi CA, Harker JO, et al. Development and testing of a five-item version of the Geriatric Depression Scale. Journal of the American Geriatrics Society 1999 ; 47 : 873-8.

27 Rinaldi P, Mecocci P, Benedetti C, et al. Validation of the five-item geriatric depression scale in elderly subjects in three different settings. Journal of the American Geriatrics Society 2003 ; 51 (5) : 694-8.

Results

Comment; It is better to have a short explanation for the Table 1 to highlight the main finding for each variable and use this explanation before the Table.



Response: We appreciate your valuable suggestion. We have added description in the results.

Changes in the text:

Page 12, lines 253-259: "The analysis showed that with zero expenses had an average age of 73.0 years, while those using care expenses had an average age of 79.2 years. The minimum cost was 5.00 USD, the maximum cost was 2,35,536.90 USD. Table 1 shows the baseline characteristics of the respondents and the average cumulative LTCI cost. Table 2 shows the Tobit regression differences in cumulative cost of long-term care insurance services by number of oral problems. And Table 3 shows rate of those certified as requiring long-term care, mortality, and days by the oral problem."

## Discussion

Comment1; There is no need to refer the results to the Table 1 and Table 2. You just can discuss the main finding without referring to tables that presented in the result section

Response: Thank you for your valuable feedback. Overall, we have revised the text. In addition, we have reviewed inadequate English expressions. I've changed my writing style to "main finding" instead of considering from Table.

Change1 in the text:

Page 17, lines 305-312: "The cost of care was found to be related to physical function, socioeconomic background, and the care environment. Decline in ADL together with depression, as well as equivalent income, and years of education were also associated with the cost of care. In terms of marital status, costs were higher for those who were not married. Further, women were more likely to be in the higher cost group than men. This is consistent with another report that showed 34.0% of caregivers were male,[1] indicating that women use care services more when they required, which also indicates a problem in the caregiving environment."

Change 2 in the text:

Page 19, lines 348-355: "The cost per beneficiary per year in 2020 for the elderly in Japan was 2.09 million USD.[42] There was a concomitant increase in cost, with each increase in oral function problems. For those with one oral function problem, the cost increase was 19.3% of the total, it was 22.7% for those with two, and 39.5% for those with three. When the cumulative long-term care costs during the follow-up period were analyzed by oral function, it was found to affect the cost of long-term care services. People with good oral function may have a shorter duration of need for long-term care during the follow-up period."

Comment2; I am still not sure why you use Table 3 in the Discussion section without refereeing to its findings. You may just remove Table 3 and explain the main thing as text.

Response; We appreciate your valuable suggestion. We have added description in the method. We have rewritten the Method and the Result to with follows, in accordance with your comments. Table 3 has been discussed with a description of the methods and results added.

## Conclusion

Comment1; Please provide some recommendations based on the findings of this study in this section.

Response: Thank you for your valuable advice. We have reviewed our conclusions in line with your comments.

Changes in text:

Page 21-22, lines 405-421: "The degree of oral function in older people was found to be associated with cumulative LTCI costs. The oral function of older people should be maintained to reduce future accumulated LTCI costs. There was a difference in cumulative LTCI costs between those with preserved oral function and those with declining oral function. Compared to those whose oral function was maintained, those with oral function problems had approximately 4,000 to 8,000 USD higher

cumulative LTCI costs over six years. There was a maximum difference of approximately 82,000 USD in long-term care service costs for those with oral function problems. The more the oral function problems, larger the difference. Maintaining the oral functions of older people may lead to a reduction in future accumulated LTCI costs.”

In addition to the above comments, all spelling and grammatical errors have been meticulously corrected to improve readability.

We thank you again for opportunity provided to us to strengthen our manuscript with your valuable comments and queries. We have worked hard to incorporate your feedback and hope that these revisions are now suitable as per the reviewers’ suggestions. We look forward to hearing from you in due time regarding our submission and to respond to any further questions and comments you may have.