Dear editor,

We would like to thank you and the three reviewers for the opportunity to revise and improve our manuscript. We have taken the very valuable comments of the reviewers very seriously and have revised the manuscript accordingly. In the attached table we have listed the reviewers’ comments and suggestions and how we have improved the article on this basis.

Many thanks and best regards

|  |  |  |  |
| --- | --- | --- | --- |
| **REVIEWER 1** | | | |
|  | **Short comment** | **Full comment** | **Response** |
| 1 | More “Content” in highlights | Regarding the 'highlights' (p2): those are rather abstract; please provide more content (e.g. what is the paper about, what are the interesting outcomes?) | We changed the highlights into full sentences and included as an additional highlight the labels of the six system types. |
| 2 | Elaborate on Esping Andersen typology | I'd suggest to elaborate a bit further on the typology of Esping Andersen in the introduction (and come back to this at the end) | Thank you for this suggestion. We included Esping-Andersen’s typology in the introduction as this typology is still used in welfare state research and also guides case selection in comparative LTC studies. Instead of further elaborating on Esping-Andersen’s typology we refer to additional literature on this discussion. |
| 3 | Imbed the four dimensions better in literature | On page 5, the authors identify 4 dimensions (i.e. supply; public-private mix; access regulation; performance); it would be good to provide some literature references that justify these categories (lines 7-14)/ better imbed them in the literature. | As suggested by the reviewer, we added further references to the description of the dimensions. |
| 4 | Qualitative indicators? | The authors claim they have used both quantitative and qualitative data; however, what is the qualitative data set they rely upon? I think their study is a quantitative study, and it would be good to use some qualitative methods to deepen their insights (see below), perhaps expert interviews and/or policy document analysis? | As emphasized by the reviewer, in our study we use quantitative data. The term “qualitative data” is therefore misleading. We changed the term into “institutional” (page 4). |
| 5 | 6 vs 9 cluster | P12 the authors identify 6 or 9 (?) different systems; I found this paragraph (lines 47-60) rather confusing: how does the 6 versus 9 systems relate in their analysis, this should be explained more clearly. | Thank you for this comment. We present two solutions, a six and a nine cluster specification. To make this more prominent and transparent, we added this information in the beginning of the results section. The nine-cluster solution is a purely methodological outcome, whereas the six-cluster solution is methodological *as well as* content-based. We added a paragraph in the discussion section for explanation (p.10) (see also comment 8 Reviewer 1). |
| 6 | Include more qualitative analyses for the explanation of results, residual system: high hospital admittance | P 13: regarding the residual system, the authors suggest that this system relies more upon informal care provision. They don't mention the possibility of high hospital admittance (what is often the case) - qualitative research (as mentioned above) would help to provide these insights. | We agree with the reviewer. There might be other possible ways of organizing and providing LTC not covered by our indicators. We deleted this part as we decided not to interpret the values of the indicators in the Results section. We included the reviewers point when discussing the limitations of our study, and added that LTC beds can be provided via residential care facilities but that also hospitals can play a role in the provision of LTC, which might unburden families (p. 17). |
| 7 | Why are NL and UK categorized as they are – better explanations, background, limitations | The Netherlands, the country I'm most familiar with, is indicated as a 'private need-based system'. I don't fully recognize this outcome; long term care is financed on basis of social insurance but includes hardly any co-payments, and access isn't mean tested. Yet I do recognize some other features like private service provision (as nursing homes aren't state-owned). Similarly, on page 14, it is suggested that the UK has a high performance. I'm sure experts and residents don't agree. Again, qualitative research methods would help to be more precise. Alternatively, the authors could be a bit more critical on their findings and shortcomings of the metrics they have used/categories they have identified. I moreover think that these contradictory findings teach us something about the complexity of contemporary LTC systems (which could be teased out further in the Discussion section) | Thank you for your considerations on the description of the clusters to which the Netherlands and the UK belong. We spotted a mistake in the cluster to which the Netherlands belong. Public expenditure is indeed not below average. We changed the sentence into “However, public expenditure is about average to high”. This change also made us reconsider the label of this cluster. We deleted the “private” and only refer to this as the “need based supply system”.  Concerning the UK, we agree that subjective and objective performance might differ. However, we evaluate self-perceived health as a subjective performance measure, which is close to the sample mean for the UK. |
| 8 | Discussion section: more focus on the discussion on methods | I'd suggest the authors use the discussion section to reflect on their outcomes and methods. How it is stated now, the discussion section stresses the outcomes/categories identified earlier on instead of discussing this new typology of welfare states (for instance by relating it to the work of Esping Andersen). Furthermore, they should elaborate on the complexity of identifying categories, as these can also be criticized because of the mixtures of measures and approaches that are used within the countries they classify. | Please refer to comment 2 for your suggestion to relate the discussion to Esping-Andersen’s typology. However, we want to stress, that we do not attempt to develop a new welfare state typology, but a new LTC typology. Based on your comment, we decided to move the discussion of the results in light of earlier LTC typologies from the Conclusion to the Discussion section (p.14 & 16). Furthermore, we added a critical discussion of our data and methods, stating that other/ more indicators might lead to more nuanced results. Furthermore, we discuss cluster analysis and highlight that the approach and the presentation of results we used leaves room for a less or more nuanced depiction (p.16). |
| 9 | What is qualitative on this typology | Conclusions: I don't see the qualitative comparison the authors suggest they have made; please elaborate (already earlier on in the paper) | Please refer to the our previous comment on this topic (Reviewer 1, comment 4) |
| **REVIEWER 3** | | | |
| 10 | Public-private mix of provision (beds, facilities) not included – address this as the reader expects this | The introduction and the conclusion starts with 'In the last century, marketization, commodification, and corporatization of care changed LTC systems all over the world [4], which makes a new and updated LTC typology necessary.' I was disappointed when I found out that the indicator for the public-private mix only includes the financial side. An assessment of who provides the services has been neglected: the authors do not include the share of for-profit, non-profit and public providers (in the number of beds or homes). I think I understand why - data is scarce and fragmented (e.g. there is no OECD statistics). But since the authors already collected information through their network of country experts, I believe it's a missed opportunity to not have asked for the statistics on the private-public provision mix and include this in their typology. So I would advise the authors to address this issue or mitigate the expectations of the reader. | Thank you for this comment. We agree that in a number of countries the provision of LTC services has changed, with private LTC providers gaining market shares (Ranci & Pavolini, 2008). You already point out that comparative data on this indicator is scarce. Even on the national level, data collection is difficult. The share of private providers and the shares of patients cared for by private providers might differ and are not available for every country. Furthermore, shares might be substantially different for residential and ambulatory care. We changed the introductory sentence you mention by referring more to the indicators we later use (financing, access) (p.1). We changed the sentence in the Conclusion accordingly (p.15). Furthermore, we added a note in the Materials and Methods section that comparative public-private provision data are not available (p.6). |
| 11 | Why the inclusion of performance indicators? 1. They do not measure LTC performance, 2. Not useful for categorization | I'm wondering why the authors have decided to include performance indicators in their typology. I'm not (yet?) convinced that this is a smart move. Firstly, the used performance indicators do not reflect the performance of the long-term care system, but the healthcare system as a whole and many other variables (behaviour etc). Secondly, in my (maybe old-fashioned) opinion, typologies are useful to categorise units based on structure indicators, so we (i.e. scholars) can relate this with outcome indicators. | We included performance indicators, as we perceive the international discussion in LTC to focus increasingly on how quality of LTC can be measured and achieved. The OECD publication “A Good Life in Old Age?” (OECD, 2013) underlines this evaluation.  We agree that the performance indicators (self-perceived health at age 65+ and life expectancy at age 65) do not solely measure LTC performance. As we state, indicators on pressure ulcers or unintended weight loss, would have a more narrow focus on the LTC system, but are not available as comparative indicators (and a lot of countries do not even provide these as national data) (see page 4). |
| 12 | Why are NL and UK categorized as they are  How is means-testing defined (Japan) | I was slightly surprised by the description of the clusters. E.g: (i) The Netherlands is one of the largest spenders on long-term care, but this manuscript states the following by the cluster of the Netherlands: "because public expenditure is below average." p. 14, line 18-19. (ii) Under the header 'evolving private need-based system' (p. 14), the authors mention that the performance is high. I find this troublesome since the performance of the LTC sector in the UK is far from optimal. (This brings me back to the issue that I have with including performance in the typology.) (iii) I wonder how you deal with countries such as Japan where people with higher incomes pay higher contribution rates? I highlight this issue because Japan is categorised as a country that does not have a mean-testing system. | Please refer to Reviewer 1 Comment 7 for the discussion of the Netherlands and the UK. The comments are similar in this regard and we decide to answer these comments in on instance.  Concerning Japan, we agree that higher contributions for people with higher incomes might be seen as a kind of means-test. However, the means-testing indicator we included only measures the means-testing of benefits, thus in the case of LTC provision. We made this addition to the description of the indicator (p.7). |
| 13 | Be more careful with descriptions – “tendencies” | In sum, I understand that with typologies and by clustering the countries you cannot address all the individual countries, but it does require some nuanced descriptions. The authors may want to speak more in terms of the 'tendencies' - the tendency towards x, y and z that these countries have in these clusters. | We agree, that we could phrase more hesitantly in some instances in the descriptions of system types and what this might imply. We deleted all interpretations of the values of indicators in the Results section and focus on an interpretation in the Discussion and Conclusion (see also Comment 6 Reviewer 1). |
| 14 | Make explicit what is meant by informal care | "this system shows one of the lowest shares of public expenditure and cash benefits are unbound, indicating a high level of informal care provision." I think it's important to distinguish informal care provision and non-institutional care. I define informal care as care that is decommodified, while non-institutional care means financial support for care outside the nursing home, often home-based care. I think you can easily solve this issue by making explicit what you mean by informal care. | We clarified that we understand informal care in this paper as de-commodified care by family and informal migrants (migrant in the family) (p.6). In light of your previous comment we deleted the following part of the sentence “indicating a high level of informal care provision” (p.11). |
| 15 | State clearly the observation periode | I might have missed it, but what is the observation period that is included in the analysis? 2014-2016? Can the authors make it explicit and clear? | We agree that this should be placed more prominently. We added this information to the first paragraph of the Material and Methods section (p.5). Quantitative data show the mean for the years 2014-2016 (latest years available at that time) and for institutional indicators for the year 2016, as we advised our policy experts to evaluate the institutional indicators for this year. |
| 16 | Does the choice indicator measure “marketization”? | I'm wondering if the indicator of 'choice' also measures a component of 'marketization'. What do the authors think? | We think, the indicator of choice restrictions might incorporate a component of marketization. In general, one would expect that those countries, which included (more) private provision or more private financing components into their systems, in the same instance included more freedom of choice, as now public, private, and non-profit providers compete (more) for clients. This might be the case. However, the data we use, do only show this partially, as the correlation between the overall choice index and the indicator of private expenditure is negative and insignificant (-0.19, p>0.05). In order to prove that the indicator of choice actually measures marketization as well, one need to have data on an earlier point in time (e.g. 1990s or early 2000s) in which many OECD countries only started to marketize and privatize LTC systems. |
| 17 | Figure 1 Country abbreviations | I find Figure 1 hard to read. Can the country abbreviations be a bit clearer? And would it be possible to name the clusters? | Thank you for your suggestion. The country abbreviations that we use are the two-letter country codes (ISO 3166-1 alpha-2). We added all abbreviations and the spelled-out country name in the caption of the figure. Furthermore, we shaded the background of the clusters in light grey. This should help the reader to grasp the clusters visually in one instance. We added the cluster label to the shaded areas. |
| **REVIEWER 4** | | | |
| 18 | Definition of beds – as systems are different also if beds belong to HC or LTC varies | The division between LTC and health care differs in each country. So data on what is tabulated as a "bed", and the dividing line between institutional care and community, care differ across countries. For example, nursing home care in the US includes post-acute / rehabilitation care financed by Medicare, which would included in LTC as it is delivered in nursing homes. However, in other countries post-acute / rehabilitation care would be financed by the health care system. Also, residential care in special housing tabulated as community care in that country but the care contents could correspond to institutional care in other countries. | We agree with the reviewer that division of LTC and healthcare is different in each country and it is a matter of definition where healthcare ends and LTC begins. We state this in the conclusion: “Furthermore, LTC systems have not as clear boundaries as other welfare state systems such as healthcare, unemployment, or pension systems. LTC can be provided via a separate LTC system or partially integrated in healthcare, social assistance, or pension systems, in which different access and provision rules apply [47].”  Concerning the indicator “Number of beds per 1,000 inhabitants”, we use data by the OECD. The indicator includes beds in: Long-term nursing care facilities and Other residential long-term care facilities LTC and excludes: Beds in hospitals dedicated to long-term care and Beds in residential settings such as adapted housing that can be considered as people’s home. (see definition by the OECD (OECD, 2020)). |
| 19 | Divide beds as good as possible in different functions and recalculate | The above makes it inappropriate to classify LTC based on the definitions used in each country. Rather, each item should defined as generically as possible, and then the data recalculated using this definition. | We are aware that boundaries of LTC systems are blurry and differ among countries. We are convinced that the dimensions and data we use are adequate to grasp central rules, conditions, and institutions, which define LTC systems. This confidence relies on the indicators and data, we use. The quantitative indicators are taken for the OECD database, which is an accepted and widely used source for comparative studies in LTC (e.g. Damiani et al., 2011; Colombo, 2011). Furthermore, the indicator of bed density is used in both above mentioned typologies. |
| 20 |  | Since the above was not made, the results tend to reflect the stereotype images that policy-makers tend to make based on a superficial comparison using terms that do not reflect the substance. | Indeed, we make certain generalizations as this is the case with typologies (see the discussion in the Conclusion). However, we provide a four, six, and nine cluster solution, which – from our point of view – shows that we go beyond the current state of the art (see comment 5 and 8, Reviewer 1) |