PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Multimorbidity patterns in patients with heart failure: an observational Spanish study based on electronic health records.
AUTHORS	Gimeno-Miguel, A; Gracia Gutiérrez, Anyuli; Poblador-Plou, Beatriz; Coscollar-Santaliestra, Carlos; Pérez-Calvo, J Ignacio; Divo, Miguel J; Calderón-Larrañaga, Amaia; Prados-Torres, Alexandra; Ruiz-Laiglesia, F.-J

VERSION 1 - REVIEW

REVIEWER	Miguel-Angel Muñoz IDIAPJgol. Institut Català de la Salut. Spain
REVIEW RETURNED	31-Jul-2019

GENERAL COMMENTS	This is a well designed and interesting study which contributes with relevant Information to the general knowledge of heart failure. It represents an example of collaboration between different healthcare levels and Institutions. The paper is aimed at answering three objectives: 1. To characterize the comorbidities of heart failure (HF), 2. to explore their clustering into multimorbidity patterns, 3. and to measure the impact of such patterns on the risk of hospitalization and mortality The first and second one are necessarily descriptive and the third one is a hard end point. I have some queries and recommendations to better understand the manuscript: Prevalence of multimorbidity is 98%. This high percentage, consistent with others found in other studies about multimorbidity should deserve a small discussion about the accuracy of the current definition of multimorbidity used everywhere and their impact on health. Authors should explain the definition used to consider multimorbidity in the methods section (before statistical analyses) I suppose that it must be due to the statistical analyses but from the clinical point of view I do not understand very well why the authors separate the coronary from the cardiovascular pattern and why they include Hypertension and obesity into degenerative pattern. It would be nice for the authors explain a bit more about the clusters.

	Since in the text the authors refer to cardiovascular comorbidities it would be helpful to classify them as a separate part in the table, due to the special relevance of these comorbidities in the case of HF. I find too much information in the table 1…is it really necessary for the purpose of the study? Is it related to the classification used to consider multimorbidity? Regarding graphics it would be nice to improve the quality format of Kaplan Meir curves in order to better understand them (ie: better differentiate the different patterns drawing lines in dots or discontinuous lines etc…). Discussion: to affirm the first paragraph "…The population with HF of our study suffered a higher morbidity burden than that found in the general population of the EpiChron Cohort …." authors should write down in the results section the morbidity burden of the former cohort. It must be clarify what authors consider as a cardiovascular comorbidity because in one part includes hypertension and in other not..… I like very much the discussion section.

REVIEWER	Mayra Tisminetzky University of Massachusetts
REVIEW RETURNED	02-Oct-2019

GENERAL COMMENTS	The manuscript examines an important topic that is to characterize patterns of comorbidities in HF patients and measure the impact of these patterns in death and hospitalizations. It is a very interesting manuscript with important public health implications. See below my comments/suggestions One major comment that I have about the manuscript is that the authors did not mention anything about comparing the patterns of comorbidities and their impact on outcomes in men and women until reaching the statistical analysis section. If this is one of the major points of the manuscript it should be clear from the beginning (abstract, intro, etc). In the Introduction, the authors mentioned the importance of cardiovascular and noncardiovascular comorbidity but this concept was not mentioned in the analysis, results or discussion. There was never any examination of these two different groups of comorbidities so it is unclear why is mentioned in the introduction. Methods section. There is no information about the criteria used to select the chronic conditions included in the study. Did the authors included all the conditions available, the most prevalent, the ones that were better predictors of adverse outcomes?? In the statistical analysis, it is unclear to me how they decided to label the different patterns. Is there any rationale besides the consensus of the clinical experts? Finally, it feels like the discussion section is a bit too long, investigators give a lot of details to describe their results for each pattern. Maybe selecting a few of the most important findings and

	giving some clinical implications will be appreciated by the readership.

REVIEWER	Nathalie Conrad Department of Public Health and Primary Care, KU Leuven, Belgium
REVIEW RETURNED	17-Oct-2019

GENERAL COMMENTS	Thank you for the opportunity to review this manuscript. I enjoyed reading it. The methodology is original and findings are relevant. I have some comments, which I include below. Multimorbidity patterns in patients with heart failure: an observational population study based on electronic health records. Review by Dr. Nathalie Conrad, Department of Public Health and Primary Care, KU Leuven, Belgium. Summary This study sought to characterize the comorbidities among heart failure (HF) patients, to explore their clustering into multimorbidity patterns, and to measure the impact of such patterns on the risk of hospitalization and mortality. For that purpose, authors have analysed electronic health record data from 14,670 patients with HF in the EpiChron Cohort (Aragón, Spain). Authors identified six different multimorbidity patterns, named cardiovascular, neurovascular, coronary, metabolic, degenerative, and respiratory. All patterns were associated with higher risk of hospitalization; and the cardiovascular, neurovascular and respiratory patterns significantly increased the likelihood of three-year mortality. The methodology applied by the authors is original and their findings are relevant. I have some comments, which I include below. Major issues 1. The timeframe for the identification of comorbidities is unclear. Start and end dates for the identification of diagnoses must be clearly specified. E.g. were comorbidities restricted to those recorded by a healthcare professional up to Jan 2011 (and if so, what was the look-back period (e.g. 2-3 years)? Details must be provided here. 2. Authors mention “diagnoses of chronic conditions registered in both primary and hospital care” Does this mean comorbidities had to be diagnosed by patient’s primary care physician as well as during a hospital admission to be included in the study? One would assume that one diagnosis in either setting is sufficient, but in any case, this must be clarified. 3. Authors state that “diagnoses were originally coded according to the International Classification of Primary Care (ICPC)”. This implies that diagnoses made in the ICPC system were mapped to ICD codes. Details of methods of mapping and references should be provided. 4. The full list of ICD/ICPC codes used to identify patients with specific comorbidities and EDC should also be published, in the appendix, or at least refer the reader to a published and invariant version of such a table. 5. Authors should explain what surgical aftercare refers to (does it refer to cardiac surgery, or is it unspecific?). It would be good to see sensitivity analyses investigating the impact of surgical

	aftercare on mortality/hospitalisation. E.g. do increased hospitalisation rates in women observed in cvd, resp and ci patterns remain significant when surgical aftercare is excluded? 6. Comparisons of crude mortality rates by sex are misleading (e.g. line 221). To investigate differences in mortality between men and women, one must consider differences in baseline characteristics, at least age. 7. Line 287 – "Whereas it had no effect in women, it increased the risk of hospitalization by 62% in men. This finding was unexpected, especially bearing in mind that both women and men suffering from this pattern were considerably older." Have sensitivity analyses been performed to investigate whether this difference could be attributed to differences in the pattern definition (i.e. individual diseases that would be included in men's neurovascular pattern, but no in women's) 8. The structure of the manuscript is unconventional and, in my view, needs revising. The discussion section is particularly lengthy, and mixes results and their interpretation. The results on the other hand, only account for about 400 words in this 3,200 words manuscript. I suggest, authors revise the structure to clearly differentiate their findings (to be included in the results section) and their interpretation (to be included in the discussion section). 9. Strengths and limitations of the study should be discussed and deserve to be mentioned in the discussion section. In my view, the 2-3 sentences in the article highlights are insufficient. 10. The written English of this manuscript should be reviewed. Minor issues 1. Given that the list of comorbid conditions investigated is extensive, and includes conditions that underlie the aetiology of heart failure (e.g. hypertension or ischaemic heart disease) alongside relatively benign conditions (e.g. lactose intolerance), it appears, in this case, more appropriate to define multimorbidity as the occurrence of three or more conditions. 2. Line 80-82 "However, although the existence of non-random associations among chronic HF comorbidities and their clustering into multimorbidity patterns has already been demonstrated,7" – Reference 7 does not refer to HF comorbidities and needs to be revised. 3. Line 176/177 "In men, we only identified five patterns, which were similar to those observed in women with the exception of the respiratory pattern." - Perhaps authors could elaborate on this and attempt to explain why that is. More generally, authors may want to elaborate on differences in multimorbidity patterns by sex, as this might provide relevant information towards our understanding of HF and how it affects men and women differently. 4. Line 199 "The higher prevalence of COPD in men compared to women was probably due to the higher rates of smoking among Spanish men." This type of argument needs to be backed up by appropriate references. 5. Line 200 "the estimated prevalence of COPD in the general Spanish population is 15% in men and 6% in women" – such statement must be backed up by appropriate references. 6. Line 206 "The fact that 80% of COPD cases are associated with smoking" - same here, reference needed. 7. Line 206/207 "close to 50% of HF patients have never smoked" – this sentence is inaccurate and leads readers to think that in this applies to the studied cohort. Authors also refer to one specific HF cohort to draw general conclusion, and would be advised to look at several studies and consider local specificities (i.e. perhaps

	smoking rates in the studied region / cohort are higher) before making such a broad statement. 8. Line 216 “The greater prevalence of hypothyroidism observed in HF patients compared to the general population, especially in women “– again, such statement must be backed up by appropriate references and discussed in consideration of the age of heart failure patients (hypothyroidism, like many other condition increase with age so that increased prevalence in an old cohort is perhaps unsurprising). 9. Table 1 – comorbidities appear to be randomly sorted. Sorting, for example in alphabetical or by descending prevalence, would be advisable. Other comments I am an epidemiologist, but not a clinician, and would advise that the paper is reviewed by a HF specialist.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer 1: Miguel-Angel Muñoz

1. Prevalence of multimorbidity is 98%. This high percentage, consistent with others found in other studies about multimorbidity should deserve a small discussion about the accuracy of the current definition of multimorbidity used everywhere and their impact on health.

Authors should explain the definition used to consider multimorbidity in the methods section (before statistical analyses).

We agree with the reviewer that the high prevalence of multimorbidity observed in our study, although consistent with other studies, is due in part to the definition of multimorbidity used. Although multimorbidity is consistently defined in the bibliography as the presence of two or more chronic conditions, it is true that there is a lack of consensus regarding the list (number) of diseases considered in the count. In our case, in order to study the comorbidity of heart failure as exhaustively as possible, we used the list of 114 chronic conditions proposed by Salisbury, which has been used in a number of studies worldwide. We have included a small discussion on this topic (Lines 203-210, clean version of the manuscript), and we have explained the definition used in the methods section (Lines 125-126), as suggested.

2. I suppose that it must be due to the statistical analyses but from the clinical point of view I do not understand very well why the authors separate the coronary from the cardiovascular pattern and why they include Hypertension and obesity into degenerative pattern. It would be nice for the authors explain a bit more about the clusters.

As precisely mentioned by the reviewer, it is due to the statistical analysis conducted. We must highlight that in the methodology used in this study (exploratory factor analysis) we have combined both statistical criteria (to identify the optimal number of clusters and the disease composition of each cluster) and clinical criteria (to name the clusters according to the most relevant/representative diseases included). This is the reason why there are two different cardiovascular clusters, which were named differently as “coronary” and “cardiovascular” according to the diseases conforming each one. The reason why hypertension and obesity were included in the metabolic pattern in both women and men and also in the degenerative pattern in women is because these diseases showed a loading factor of ≥0.25 in these patterns, which was the threshold established to consider a disease as part of

a cluster. We have tried to better explain the methodology used to identify the clusters of diseases (Lines 144-145, 149-152).

3. Since in the text the authors refer to cardiovascular comorbidities it would be helpful to classify them as a separate part in the table, due to the special relevance of these comorbidities in the case of HF.

We wanted to refer to both types of comorbidities (non- cardiovascular and cardiovascular) in the introduction to highlight the importance of studying not only concordant comorbidities to an index disease (i.e. heart failure), but also discordant comorbidities (i.e. non- cardiovascular), as they may still play an important role in heart failure patients. However, it was not an objective of our study to compare the impact of cardiovascular and non-cardiovascular comorbidities. We think it would be better to keep the table without separating cardiovascular comorbidities from the rest, since it might lead to a misinterpretation of the results given that both cardiovascular and non-cardiovascular diseases are grouped within the cardiovascular patterns, and some cardiovascular diseases are grouped within non-cardiovascular patterns. The patterns were differentiated according to the pathology of target organs and/or risk factors, thus establishing a pathophysiological continuum that makes it difficult to differentiate both types of comorbidities.

4. I find too much information in the table 1…is it really necessary for the purpose of the study? Is it related to the classification used to consider multimorbidity?

One of our objectives was to characterize exhaustively the comorbidity of heart failure. This is the reason why we thought appropriate to show at least the information of those chronic conditions with a prevalence greater than 5% from the list of 114 diseases (we removed from the table those chronic conditions with a prevalence lower than 5%). However, as we agree that there is a lot of information in Table 1 and in order to simplify its content, we have deleted the information regarding healthcare use rates of patients, which was not a specific objective of our study (in fact, we do not refer to these results in the text).

5. Regarding graphics it would be nice to improve the quality format of Kaplan Meir curves in order to better understand them (ie: better differentiate the different patterns drawing lines in dots or discontinuous lines etc…).

We agree with your comment. We have improved Kaplan Meier curves of Figure 1 as suggested and provided a common legend for both men and women (attached as TIFF file).

6. Discussion: to affirm the first paragraph "…The population with HF of our study suffered a higher morbidity burden than that found in the general population of the EpiChron Cohort …." authors should write down in the results section the morbidity burden of the former cohort.

We agree. We have included this information in the results section (Line 173), and slightly modified the paragraph in the discussion section (Lines 200-203).

7. It must be clarify what authors consider as a cardiovascular comorbidity because in one part includes hypertension and in other not..…

In line with our previous comment, it was not an objective of our study to differentiate cardiovascular and non-cardiovascular comorbidities. The patterns were differentiated according to the pathology of target organs and/or risk factors, thus establishing a pathophysiological continuum that makes it difficult to differentiate both types of comorbidities. Still, we have revised the manuscript to be consistent.

I like very much the discussion section.

Thank you very much for your positive comment.

Reviewer 2: Mayra Tisminetzky

1. One major comment that I have about the manuscript is that the authors did not mention anything about comparing the patterns of comorbidities and their impact on outcomes in men and women until reaching the statistical analysis section. If this is one of the major points of the manuscript it should be clear from the beginning (abstract, intro, etc).

We totally agree with your comment. One of our main objectives was to analyse comorbidity of heart failure and its impact on health outcomes according to patients' sex, since we think these may differ between men and women and that clinicians may benefit from a stratified analysis so as to differentiate the management of comorbidities according to the sex of the patient. Furthermore, the European Commission encourages the analysis of the sex/gender perspective in all the research projects funded. We have introduced this idea in the abstract (Lines 24, 30, 33, 47; clean version of the manuscript) and introduction section (Lines 87-88, 94-95, 98, 100-101).

2. In the Introduction, the authors mentioned the importance of cardiovascular and noncardiovascular comorbidity but this concept was not mentioned in the analysis, results or discussion. There was never any examination of these two different groups of comorbidities so it is unclear why is mentioned in the introduction.

A similar comment was made by Reviewer 1. We considered appropriate to refer to both types of comorbidities in the introduction to highlight the importance of studying not only concordant comorbidities of heart failure (i.e. cardiovascular), but also discordant (i.e. non-cardiovascular) comorbidities, as both may play an important role regarding prognosis, therapeutic management, etc. of patients with heart failure. That is, to highlight the importance of taking into account the whole constellation of diseases surrounding an index condition. This is the reason why we analysed 114 chronic conditions instead of limiting our study to a few cardiovascular comorbidities. In our study we found that, while some patterns share a cardiovascular aetiology, others have a non-cardiovascular nature, although comparing the specific impact of these two types of comorbidities was not an objective of our study. The patterns were differentiated according to the pathology of target organs and/or risk factors, thus establishing a pathophysiological continuum that makes it difficult to differentiate both types of comorbidities.

3. Methods section. There is no information about the criteria used to select the chronic conditions included in the study. Did the authors included all the conditions available, the most prevalent, the ones that were better predictors of adverse outcomes??

We wanted to exhaustively study the comorbidity of heart failure including in the analysis the 114 conditions considered as chronic by Salisbury et al. Since we use two different disease classification systems (ICPC and ICD-9-CM), and in order to avoid counting the same disease twice, we used the ACG System to combine all diagnoses into 260 mutually exclusive Expanded Diagnostic Clusters (EDCs), 114 of which are chronic and were therefore included in the analysis. In order to facilitate the interpretation of the patterns obtained, we only included in the factorial analysis those diseases with a prevalence greater than 5% from the list of 114. We have tried to better explain this point in the methods section (Lines 123-125, 137-139).

4. In the statistical analysis, it is unclear to me how they decided to label the different patterns. Is there any rationale besides the consensus of the clinical experts?

In this study, we have combined both statistical criteria (to identify the optimal number of clusters and the diseases composition of each cluster) and "subjective" clinical criteria (to name the clusters). The

patterns were labelled by the clinical experts according to the most relevant/representative diseases (based on their own clinical experience) and those with the highest loading factor within each pattern. We believe that the effort made by clinicians to reach a consensus to label the patterns has been key to facilitate the interpretation and discussion of the results. We have tried to explain better this aspect (Lines 147-152).

5. Finally, it feels like the discussion section is a bit too long, investigators give a lot of details to describe their results for each pattern. Maybe selecting a few of the most important findings and giving some clinical implications will be appreciated by the readership.

We agree that the discussion is a bit long because we aimed to describe all the patterns and we made an effort to try to interpret and look for possible explanations for the associations obtained. In order to respect Reviewer 1 opinion on the discussion section, we did not want to modify this section in excess. However, we have tried to shorten a little bit the discussion where possible and to provide some insight on the clinical implications (Lines 303-312), as suggested.

Reviewer 3: Nathalie Conrad

Major issues

1. The timeframe for the identification of comorbidities is unclear. Start and end dates for the identification of diagnoses must be clearly specified. E.g. were comorbidities restricted to those recorded by a healthcare professional up to Jan 2011 (and if so, what was the look-back period (e.g. 2-3 years)? Details must be provided here.

We agree that the timeframe was unclear. The diagnoses correspond to those chronic diseases that remained as active episodes in patients' electronic health record (EHRs) between January 1st, 2011 to December 31st, 2011. Consequently, comorbidities could have been recorded by a healthcare professional before January 1st, 2011 and still remain as active episodes in the EHRs. We have clarified this point in the methods section (Lines 115-118, clean version of the manuscript).

2. Authors mention "diagnoses of chronic conditions registered in both primary and hospital care". Does this mean comorbidities had to be diagnosed by patient's primary care physician as well as during a hospital admission to be included in the study? One would assume that one diagnosis in either setting is sufficient, but in any case, this must be clarified.

We agree that the explanation was ambiguous. We included in the study comorbidities diagnosed in primary care and/or in hospital admissions. Thus, a disease could have been recorded in primary care only, in hospital only, or in both settings (this is the reason why we used EDCs to combine diagnoses from the two settings/classification systems). We have clarified this point in the methods section (Lines 114-115).

3. Authors state that "diagnoses were originally coded according to the International Classification of Primary Care (ICPC)". This implies that diagnoses made in the ICPC system were mapped to ICD codes. Details of methods of mapping and references should be provided.

Diagnoses were coded as ICPC codes (in primary care) or ICD codes (in hospitals), and these were combined into EDCs using the ACG System. This software uses directly the two types of codes and combines them into EDCs, without the need of a previous mapping from ICPC to ICD. We have included a reference to the software´s website in the methods section (Line 123).

4. The full list of ICD/ICPC codes used to identify patients with specific comorbidities and EDC should also be published, in the appendix, or at least refer the reader to a published and invariant version of such a table.

Unfortunately, we cannot provide the full list of ICD/ICPC codes on which the EDCs are based, since the ACG System software uses a non-public algorithm to combine similar ICD/ICPC codes into specific EDCs. We have used all the ICD/ICPC codes recorded in the EHRs and subsequently used the software to combine them in 260 mutually exclusive EDCs, and finally we selected the 114 EDCs defined as chronic by Salisbury et al. for our study.

5. Authors should explain what surgical aftercare refers to (does it refer to cardiac surgery, or is it unspecific?). It would be good to see sensitivity analyses investigating the impact of surgical aftercare on mortality/hospitalisation. E.g. do increased hospitalisation rates in women observed in cvd, resp and ci patterns remain significant when surgical aftercare is excluded?

Unfortunately, and in line with our previous comment, we cannot know which ICPC/ICD codes are behind the EDC “surgical aftercare”. We assume that it is an unspecific code referring to any type of care after a surgery, which can be considered as a chronic condition.

In this sense, we want to highlight that although the ACG system comes with this type of limitations, it is very useful for research on multimorbidity, as it is very exhaustive, it has been internationally validated, and it generates mutually exclusive conditions avoiding potential double counting of similar conditions deriving from the use of different disease classification systems.

6. Comparisons of crude mortality rates by sex are misleading (e.g. line 221). To investigate differences in mortality between men and women, one must consider differences in baseline characteristics, at least age.

We agree that this comparison is misleading and we have therefore deleted this statement from the discussion section since this was not an objective of the study (the objective was to study the impact of multimorbidity patterns on mortality in both men and women). However, we have left the crude mortality rates in Table 1 so that the reader has an idea of the mortality rates of our patients with HF.

7. Line 287 – “Whereas it had no effect in women, it increased the risk of hospitalization by 62% in men. This finding was unexpected, especially bearing in mind that both women and men suffering from this pattern were considerably older.” Have sensitivity analyses been performed to investigate whether this difference could be attributed to differences in the pattern definition (i.e. individual diseases that would be included in men’s neurovascular pattern, but no in women’s).

We agree that this analysis would be very interesting to know whether this difference was due to the presence/absence of specific diseases within the pattern. However, we could not perform this kind of analysis as we defined a pattern as the presence of heart failure plus two any other comorbidities out of those conforming each pattern (i.e. women with different “triads” of diseases can be assigned to the same pattern, and the same occurs in men), but did not include the individual diseases in the analysis (only the presence/absence of specific patterns).

8. The structure of the manuscript is unconventional and, in my view, needs revising. The discussion section is particularly lengthy, and mixes results and their interpretation. The results on the other hand, only account for about 400 words in this 3,200 words manuscript. I suggest, authors revise the structure to clearly differentiate their findings (to be included in the results section) and their interpretation (to be included in the discussion section).

We agree with the reviewers that the discussion section is lengthy; the reason is that our first intention was to write a Results & Discussion section combining the findings obtained and their interpretation in order to avoid repetition and for the sake of fluency. Due to the journal style guidelines, we finally had

to split these two sections and we therefore decided to write the results as succinctly as possible (to avoid repetition with what is already presented in the tables). On the other hand, we tried to make an effort to discuss the composition of all the patterns and to provide possible clinical explanations. To respect both Reviewer 1 opinion (who “likes the discussion section very much”) as well as that of Reviewer 2 and yourself (who think that the discussion is too long), we have tried to keep the same format but shortening the discussion as much as possible by removing results from the discussion, among others.

9. Strengths and limitations of the study should be discussed and deserve to be mentioned in the discussion section. In my view, the 2-3 sentences in the article highlights are insufficient.

We agree. In the previous version, we tried to summarize the strengths and limitations in order to accomplish with the journal’s submission guidelines, but we have now discussed these questions more deeply in the discussion section (Lines 313-328).

10. The written English of this manuscript should be reviewed.

The manuscript has been reviewed by a native English speaker (now in acknowledgements).

Minor issues

1. Given that the list of comorbid conditions investigated is extensive, and includes conditions that underlie the aetiology of heart failure (e.g. hypertension or ischaemic heart disease) alongside relatively benign conditions (e.g. lactose intolerance), it appears, in this case, more appropriate to define multimorbidity as the occurrence of three or more conditions.

In line with your comment, Reviewer 1 also highlighted that the elevated prevalence of multimorbidity found in our study derives from the definition of multimorbidity used. We defined multimorbidity as the presence of two or more chronic conditions, since it is the most widely accepted definition in the literature, and also by the WHO. The most important question is however the list (i.e. number) of diseases considered to classify subjects as multimorbid, as the number can vary from 5 to more than 200 among studies. In our case, we used the list of 114 chronic conditions developed by Salisbury et al. because it is validated and used worldwide. We prefer to adhere to the definition of two or more conditions to be consistent with the rest of our observational studies. Nevertheless, this decision only affects the prevalence of multimorbidity; when studying the impact of patterns on health outcomes, we did use the threshold of three conditions (i.e. heart failure plus two additional diseases) as a condition to assign an individual to a specific pattern of multimorbidity.

2. Line 80-82 “However, although the existence of non-random associations among chronic HF comorbidities and their clustering into multimorbidity patterns has already been demonstrated,7” – Reference 7 does not refer to HF comorbidities and needs to be revised.

The reviewer is right that in Reference 7, the clustering of diseases into multimorbidity patterns are studied in the general population, and not only among patients with HF. We have rephrased this sentence in the introduction section (Lines 79-82).

3. Line 176/177 “In men, we only identified five patterns, which were similar to those observed in women with the exception of the respiratory pattern.” - Perhaps authors could elaborate on this and attempt to explain why that is. More generally, authors may want to elaborate on differences in multimorbidity patterns by sex, as this might provide relevant information towards our understanding of HF and how it affects men and women differently.

We agree with your comment, and we have included some ideas in the discussion section (Lines 236-238, 303-312).

4. Line 199 “The higher prevalence of COPD in men compared to women was probably due to the higher rates of smoking among Spanish men.” This type of argument needs to be backed up by appropriate references.

We agree with this comment and we have included a new reference (Line 214, Reference 18).

5. Line 200 “the estimated prevalence of COPD in the general Spanish population is 15% in men and 6% in women” – such statement must be backed up by appropriate references.

We have included a new reference to support our statement (Line 215, Reference 19).

6. Line 206 “The fact that 80% of COPD cases are associated with smoking” - same here, reference needed.

We have decided to delete this sentence from the discussion section and we have rephrased the main idea (Lines 220-222).

7. Line 206/207 “close to 50% of HF patients have never smoked” – this sentence is inaccurate and leads readers to think that in this applies to the studied cohort. Authors also refer to one specific HF cohort to draw general conclusion, and would be advised to look at several studies and consider local specificities (i.e. perhaps smoking rates in the studied region / cohort are higher) before making such a broad statement.

We agree that this sentence was ambiguous. We have deleted this statement and rephrased this part of the discussion (and added a new reference) to highlight the main idea that was that COPD could be over-diagnosed in HF patients (Lines 220-222, Reference 21).

8. Line 216 “The greater prevalence of hypothyroidism observed in HF patients compared to the general population, especially in women “– again, such statement must be backed up by appropriate references and discussed in consideration of the age of heart failure patients (hypothyroidism, like many other condition increase with age so that increased prevalence in an old cohort is perhaps unsurprising).

We have included a new reference to support the sentence (Lines 231-232, Reference 13), and rephrased the idea to conduct a fairer comparison regarding the prevalence of hypothyroidism (Lines 230-233).

9. Table 1 – comorbidities appear to be randomly sorted. Sorting, for example in alphabetical or by descending prevalence, would be advisable.

In Table 1, comorbidities are sorted by descending prevalence in the total study population, in order to facilitate the statistical comparison of their prevalence between women and men.

Other comments

I am an epidemiologist, but not a clinician, and would advise that the paper is reviewed by a HF specialist.

Several members within the study team (i.e. AGG, CCS, JIPC, MJD, APT and FJRL) are medical doctors, and AGG, JIPC and FJRL are moreover members of the IIS Aragon’s Research Group on Heart Failure.

VERSION 2 – REVIEW

REVIEWER	Miguel-Angel Muñoz Institut Català de la Salut. IDIAPJGol Universitat Autonoma de Barcelona
REVIEW RETURNED	21-Nov-2019

GENERAL COMMENTS	I do not consider necessary further revisions

REVIEWER	Mayra Tisminetzky Univ of Massachusetts USA
REVIEW RETURNED	14-Nov-2019

GENERAL COMMENTS	Authors responded to all my comments