Dear Dr. Chalmers

We write with regard to manuscript PONE-D-15-28816 entitled ‘Physical activity patterns and clusters in 1001 patients with COPD’.

Thank you very much for the chance to submit a revised version.

We have tried to improve our manuscript based on the valuable comments provided. Below we present a response to each of these comments. Important to mention, any changes described in our responses refer to the 'Revised Manuscript with Track Changes’ file. Moreover, we have also made small textual changes throughout the manuscript aiming better clarity. All the changes are highlighted with the ‘Track Changes’ function.

May we kindly ask you to take our manuscript into consideration for publication in PloS One?

Yours sincerely,

Rafael Mesquita

(on behalf of all authors)

RESPONSE TO COMMENTS:

1) The reviewers identified significant concerns with this manuscript which are listed below in the reviewer comments. Please carefully address these comments. The issue of the Italian data was identified as a significant concern by more than one reviewers and must be addressed- although the data is collected as part of routine care it is essential we are reassured that all data collection and use was approved by the ethical standards in that country. The manuscript cannot be published without a clear statement that this was done.

**Response: CHECK WITH THW ITALIANS – let them make an official statement about this!**

Reviewer #1: This is a cross-sectional analysis of physical activity patterns from a large multi-national dataset of over 1000 patients with COPD. Cluster analysis has been used and identified distinct cluster phenotypes related to physical activity patterns.  
  
Overall this is well written, presenting novel findings and technically sound, other than a couple of points below.  
  
Comments:  
  
1) 5 clusters were identified in the analysis. I am unsure how valid cluster 5 is compared with the others. Cluster 5 contains only 21 (2%) of patients. Looking at the figure, it may be that many of these patients would fit in cluster 4 but have been skewed by some outliers.

**Response:** The number of subjects in cluster 5 was also a point of concern for us when analysing the results. However, as the clusters were identified using rigorous statistical analyses we believe all five clusters represent real groups of patients with COPD. This is supported by the results presented in Table 4 and Figures 5B, C and D. Therefore, we do not think cluster 5 could be incorporated into other clusters. Cluster 5 is represented by a small number of patients probably due to the remarkable levels of physical activity achieved by this group. Previous publication have already shown that only a very small proportion of COPD patients can be considered active (Watz et al, Eur Respir J 2009;33:262-72), possibly explaining why this cluster is much smaller than the other ones.

2) One of the big advantages of this study are the large number of patients involved. It seems odd to also include a comparison with only 66 matched healthy volunteers. I do not think this adds to the manuscript.

**Response:** We have included a sample of healthy subjects in order to demonstrate the impairments in physical activity in COPD patients after matching for sex, age and body mass index (BMI). The fact that we had a considerably large sample of COPD patients allowed us to select a subgroup of patients that could be matched for potential confounders (sex, age and BMI) with the group of healthy subjects and, to our knowledge, this has not been done before. Nevertheless, as other reviewers have also raised concerns about the analyses involving this group, we preferred to remove them from our manuscript, as well as all related analyses. Most changes were done in the Results (pages 16-18, lines 296-313) and Discussion (page 24, lines 395-401).

3) The area under the curve results are presented as negative predictors (i.e. AUROC <0.5). This is unusual and would normally be presented as positive predictors.

**Response:** The area under each hourly pattern was calculated and presented with its 95% confidence intervals in order to quantitatively represent time-varying averages of the hourly patterns. Although these values were named as ‘Area Under the Curve (AUC)-values’, they were not derived from a Receiver Operating Characteristic (ROC) analysis, therefore not aiming to identify predictors. We apologise for the confusion. We have changed the text in the manuscript slightly in an attempt to clarify (page 10, lines 203-204).

Reviewer #2: General  
The strength of the paper is that it involves a cohort of 1,001 patients.   
  
1) However, the findings are not novel from what has already been previously reported in smaller cohorts. It is disappointing that such a large cohort was not used to provide information about differences in PA and its patterns between the 10 countries. We do not know what norms are for various geographical regions, and this cohort could potentially provide this needed information.

**Response:** We thank this reviewer for the input. Although we agree that it would be interesting to have in depth analyses of physical activity measures between countries, this was not the focus of our paper. Moreover, as already highlighted by other reviewers the current version has already a considerable amount of analyses. Nevertheless, comparisons between countries have been made, but they were provided in the online supplement as they are only secondary findings (S3 Table in File S2).

2) The clinical significance of the clusters is unclear as they are not associated with any clinical characteristics or outcomes such as acute exacerbations, COPD-related hospitalizations, or death. Are these phenotypes of the heterogenous COPD population?

**Response:** Unfortunately we cannot confirm whether the clusters identified in our study represent phenotypes of the heterogenous COPD population. According to Han et al. (Am J Respir Crit Care Med. 2010;182:598-604), a clinical phenotype should relate to clinically meaningful outcomes, such as rate of disease progression or death, and we were not able to test this due to the lack of follow-up measurements. This limitation is already present in the paper, but we had rephrased it for better clarity (page 29, lines 506-508).

3) There is an overwhelming abundance of data being presented, particularly in the Tables and Figures. Please focus the main messages to be conveyed and simplify the Tables and Figures.

**Response:** In an attempt to simply and have a more concise manuscript we have removed the analyses including healthy subjects (pages 16-18, lines 296-313; page 24, lines 395-401). Different reviewers have highlighted these analyses as a concern. Moreover, a potential problem related to the large amount of data could be an insufficient discussion. In order to overcome this problem we have tried to better explain some findings (pages 23-24, lines 382-394; pages 25-26, lines 427-439; and pages 27-28, lines 474-483).

Abstract  
4) What specifically is “insufficiently understood” about physical activity in COPD?

**Response:** It was insufficiently understood how different physical activity measures relate to clinical characteristics, for instance. With our study we have identified clusters based solely on physical activity measures and these clusters related to clinical characteristics. We were able to identify clusters due to the use of different physical activity measures, such as the time in bouts of activity and in different intensities, and this was an advance compared to many previous studies which presented physical activity levels as an average value only. Physical activity is a multi-dimensional construct meaning that it has much more relevant descriptors and components beside its total amount. These pieces of information were added to the abstract for better clarity.

5) It is unclear how the clusters would be useful for PA interventions. How should PA interventions target PA patterns?

**Response:** We believe future studies could provide interventions that were tailored to the needs of each group. Cluster 1, for instance, spent a median of 15 minutes only in moderate-to-vigorous intensity, which is half of the recommended 30 min·day-1 (Garber et al, Med Sci Sports Exerc. 2011;43:1334-59), and more than 15 hours in very light intensity (i.e., sedentary behaviour), which is more than two times of what other studies have considered as harmful (i.e., 7- hours) (Centers for Disease Control and Prevention, 2011; Chau et al, PLoS One. 2013;13;8:e80000). This cluster, therefore, could benefit from an intervention focusing not only on increasing the amount of time in moderate-to-vigorous intensity, but also on reducing the time in very light intensity. Cluster 4, on the other hand, seems to spend enough time in moderate-to-vigorous intensity (i.e., >2 hours), but would probably benefit from an intervention aiming to reduce the time in very light intensity, which is over 11 hours. Decreasing the time in very light intensity in this group without necessarily increasing the time in moderate-to-vigorous intensity would mean focusing on light intensity, which is supported by recent findings demonstrating the importance of increasing the amount of light intensity activities in COPD. These pieces of information were added to the discussion (page 25, lines 415-419; pages 27-28, lines 474-483; page 28, lines 486-492). Moreover, in order to highlight the novelty of having groups with mixed arrangements of physical activity we added a figure presenting the amount of time in very light, light, and moderate-to-vigorous intensities (Figure 5 in the revised version; page 19, line 333-334).

Introduction  
6) Why is it important to understand hourly patterns and temporal patterns? What is the hypothesis being entertained about temporal patterns—that it is better than summary values, intensity, or bouts? It seems the patterns are dictated more by human behavior (more active in morning/around lunch time, and less active in the afternoon as we settle down from a busy day and relax for the evening) than by COPD or its severity.

**Response:** We do not think hourly patterns are better than summary values. We believe hourly patterns have the potential to complement the information provided by summary values. For instance, in our study patients in GOLD grade 4 spent more time in very light intensity and less time in moderate-to-vigorous intensity compared to patients in other GOLD grades. However, the influence of GOLD grades on hourly patterns was only little or none (Fig. 2E and 2F). Our hypothesis was that different hourly patterns could be found in patients with different characteristics. Indeed, age, BMI, mMRC dyspnoea grades, and ADO index scores seemed to have a significant influence on hourly patterns as there was little or no overlap between the 95% confidence intervals of the hourly patterns. Only a few studies have used hourly patterns in COPD. Hecht et al. (COPD. 2009;6:121-9) observed the highest activity level during the late morning and early afternoon hours, whilst Tabak et al. (Int J Telemed Appl. 2012;2012:438736) observed a similar pattern of a dip of lower activity in the daily activity pattern in the early afternoon in both employed and unemployed COPD patients. These pieces of information were added to the discussion (page 24, lines 386-394). Hourly patterns have been shown to provide useful information in other populations as well. In patients with Parkinson’s disease Rochester et al. (Disabil Rehabil. 2006;28:1365-71) observed a delayed morning start and reduced peak of activity compared to controls. In the study by Evering et al. (J Psychosom Res. 2011;71:129-35) patients with chronic fatigue syndrome were significantly less physically active in the afternoon and evening compared to controls. In order to highlight the importance of investigating hourly patterns, we also added these pieces of information to the introduction (page 7, lines 127-130).

7) Reference 17 should be replaced or supplemented by the recent ERJ paper looking at the association between low-intensity PA and risk of COPD-hospitalization.

**Response:** The abovementioned paper was added to our manuscript.

Methods  
8) I raise concern whether it is ethical to include the 23 patients from Italy who did not give written informed consent. De-identification is important to protecting patient information and confidentiality, but this does not mean clinical data can be used for research purposes without consent.

**Response:**

9) What is the median split method? Please provide a reference.

**Response:** The median split method is simply a way of turning a continuous variable into a categorical one by splitting a sample into two, those above and those below the median. As we could not find an appropriate reference to add to this method we have changed the text in order to avoid confusion (page 9, lines 175-176).

10) How was clinical stability at the time of PA assessment defined? How long since the last acute exacerbation or hospitalization was the PA monitored?

**Response:** The criterion for clinical stability varied slightly between centres, but the vast majority included patients who were stable for at least 4 weeks.

11) Was the diagnosis of COPD based on the FEV1/FVC ratio alone? Is there possibility of misclassification of asthma for COPD?

**Response:** This criterion also varied slightly between centres but in all of them the diagnosis was based on post-bronchodilator spirometry according to international guidelines. Additionally, some centres excluded patients with history of asthma or respiratory disorders other than COPD. Therefore, a possibility of misclassification of asthma seems unlikely.

12) After discussing how PA should not be represented by summary data, the PA measures in this study represent “the average” of all valid weekdays. Please discuss.

**Response:** Summary data have been shown to provide useful information in previous studies, but we think these data can be complemented by other measures, such as physical activity hourly patterns. Therefore, summary values were used in our study due to the valuable information provided with them, but they were complemented with the information provided with physical activity hourly patterns. Please refer to our response to comment 6 for more details about the importance of hourly patterns.

13) What are the units to PA being used, if not steps per day?

**Response:** Different units were used depending on the parameter investigated. For instance, the time in different intensities was expressed as min∙day-1, while energy expenditure was expressed as METs-min∙day-1.

14) The aims stated in the last paragraph of the Introduction do not include anything on PA intensity, but the methods and results discuss PA intensity in great detail. Please clarify. There is also no aim on PA bouts, but the results seem to focus on bouts. Please explain what are physical activity measures. There is also no discussion of GOLD stage but these are presented extensively in the Figures.

**Response:** We understand physical activity measures as parameters derived from physical activity data, such as the time in different intensities and in bouts of activity. Therefore, these two parameters are indeed representations of physical activity measures. We did not mention about intensity of activity or bouts of activity in the aims as we used the term ‘physical activity measures’, which in our understanding encompasses these two parameters. Moreover, a brief discussion about the GOLD classification was added to the discussion (pages 23-24, lines 382-387).  
  
15) Healthy subjects were matched in age, gender, and BMI. Were they healthy because they did not have COPD or because they had no comorbidities? Matching on comorbidities, such as CAD, PVD, back pain, that affect PA would be important. Were healthy subjects retired or employed? What were their demographics? These factors all affect PA.

**Response:** Healthy subjects were free of COPD or any other debilitating disease. Unfortunately, data on comorbidities or on current occupation (retired or employed) were not available. Nevertheless, as other reviewers have also raised concerns about the analyses involving healthy subject, we preferred to remove them from our manuscript, as well as all related analyses. Most changes were done in the Results (pages 16-18, lines 296-313) and Discussion (page 24, lines 395-401).  
  
16) Why do models not adjust for site of study?

**Response:** In our study we wanted to investigate whether COPD patients could be clustered based solely on their levels of physical activity, i.e., without the influence of any other characteristic. That is why site of study was not considered in our analyses.  
  
Results  
17) The description of the clusters in Table 4 need full explanation in the text. What is very long, very light intensity? What is the difference between very long versus long; very short versus short?

**Response:** The descriptions were defined according to the amount of time in very light and moderate-to-vigorous intensities. This definition was added to the bottom of Table 4. Very light intensity was chosen as it can be used as a surrogate of sedentary behaviour, which has been focus of recent publications in COPD (Hill et al, Intern Med J. 2015;45:474-82), while moderate-to-vigorous was chosen due to its contribution to recommendations for physical activity (Garber et al., Med Sci Sports Exerc. 2011;43:1334-59).  
  
18) What does stratification of results by age, gender, and LTOT, DLCO, and ADO add to the overall results?

**Response:** These analyses were performed to answer the first aim, i.e., to describe physical activity measures and physical activity hourly patterns in patients with COPD after stratification for generic and COPD-specific characteristics. Up to now, we did not know the hourly pattern after stratification for the abovementioned clinically relevant characteristics.   
  
19) Figure 2. The authors state Figure 2 shows a “noticeable influence of age, BMI, mMRC, and ADO.” Was there any statistical methods applied to prove that there is a difference between the curves?

**Response:** We used the 95% confidence intervals of the hourly patterns to identify differences between them. If there was little or no overlap in most part of the day between two hourly patterns they were deemed to be significantly different. In order to clarify, these explanations were added to the manuscript (page 15, lines 276-277). This method of comparison has been used by other researchers as well (Bellettiere et al, PLoS One. 2015;10:e0136161).

20) Figure 3 has been previously shown many times. It is not presenting anything novel.

**Response:** That is true, but in our study we used a considerably large and heterogeneous sample, which was not done before. Nevertheless, to add novelty to this figure we have added to it the correlations of FEV1 with the time in very light and light intensities.   
  
21) Please discuss that Figure 4 shows no difference between COPD and healthy with respect to the pattern of PA over hours of the day. This was a main aim of this paper.

**Response:** In fact, healthy subjects performed their activities at higher intensities compared to patients with COPD as there was little or no overlap in most part of the day between the 95% confidence intervals of the two groups, but this difference was observed during weekdays only. Nevertheless, as other reviewers have also raised concerns about the analyses involving healthy subject, we preferred to remove them from our manuscript, as well as all related analyses. Most changes were done in the Results (pages 16-18, lines 296-313) and Discussion (page 24, lines 395-401).  
  
22) Please present the results of the 3 components in the main paper rather than in Supplement File 2 since this is a main aim of the paper.

**Response:** One of the main aims of the study was to identify clusters. Identifying principal components was just one step to identify clusters. Although we believe the most relevant features of each principal component are important, they are only secondary compared to the characteristic of the clusters. We believe including both results in the main text could make the reading tedious. Therefore, we prefer to keep the characterisation of the clusters in the main text and the most relevant features of each principal component in the supplementary file.  
  
23) Figure 6 looks strikingly very similar to Figures 1 and 2 in terms of hourly patterns. What additional information does Figure 6 provide?

**Response:** Although the figures might look similar, they provide different information. For instance, in Figure 2 we observed a significant influence of mMRC dyspnoea grades and BMI, but little or no influence of GOLD grades and GOLD groups, whilst Figure 6 showed less variability in intensity in more inactive clusters compared to more active clusters, suggesting that the former are similarly inactive throughout the day. We believe these figures are not interchangeable and therefore we prefer to keep them in. The explanations concerning Figure 6 were added to the manuscript (page 25, lines 429-431).  
  
24) The issue of synchronization of the waking up moment is first presented in the Figure 6 legend. What is its significance? It should be discussed in the methods and results.

**Response:** Subjects might have different waking up times and this could be a problem when averaging the values for each hour. For instance, a subject might be asleep at 09:00 AM but another subject might be awake and exercising at this time, and then the intensity of activity of these two subjects would be completely different. In order to overcome this problem we provided extra-analyses after synchronising the waking up moment (Figures 6C and 6D). Nevertheless, as these analyses are secondary in the revised version of the manuscript they are presented as supporting information (S2 Fig.), but a brief explanation was added to the legend of the figure for clarification.  
  
Reviewer #3: The authors present a manuscript that describes physical activity patterns in those with vary severity of COPD and some healthy controls. The manuscript presents an enormous data set with many comparisons being made.  
  
1) I am the first to appreciate all data being presented, but, in this case, the extreme amount makes the manuscript difficult to read and nearly impossible interpret. I was often lost at what comparison was being made and very few were actually justified in the introduction and/or elucidated in the discussion. Rather, the discussion was vague and did not address specific finding, presumably because there were too many. As such, I was lost as to what data was being referenced to justify which conclusion. I note PLOS ONE does not have a word limit, as such, I am lost as to why the discussion so vague, short and poorly developed?

**Response:** We understand that the amount of data might seem too much. In an attempt to simply and have a more concise manuscript we have removed the analyses including healthy subjects (pages 16-18, lines 296-313; page 24, lines 395-401). Different reviewers have highlighted these analyses as a concern. Moreover, we acknowledge that some results were insufficiently discussed, so in the revised version of the manuscript we have included more discussion in an attempt to overcome this problem (pages 23-24, lines 382-394; pages 25-26, lines 427-439; and pages 27-28, lines 474-483).   
  
2) There was a discrepancy between the sexes for this study. There are many well know sex difference in the respiratory system and many specific to COPD. This was not stated at all in the manuscript and needs at least to be mentioned.

**Response:** Compared to female subjects, male subjects we slightly older (67 (62 – 73) versus 65 (59 – 71) years; *P*<0.0001) and had higher BMI (26.5 (23.3 – 29.9) versus 24.5 (21.1 – 28.6) kg·m-2; *P*<0.0001), but there was no difference in FEV1, mMRC grades, or GOLD 2007 and 2011 classifications (*P*>0.05 for all) between these groups. These results were added to the manuscript (pages 11-12, lines 229-233).

3) Different countries were shown to have different physical activity levels, however, the control group was only from two specific countries. As such, how can the comparisons between the controls be justified to the COPD cohort.

**Response:** This was a point of concern when selecting COPD patients to be matched with healthy subjects. However, we preferred to consider patients from any of the countries in order to have more options for the matching, with the final aim of having more similar groups in terms of the selected characteristics. Nevertheless, as other reviewers have also raised concerns about the analyses including healthy subject we preferred to remove them from our manuscript (pages 16-18, lines 296-313; page 24, lines 395-401).  
  
4) There are a lot of statistical comparison but the specific analytical techniques were not as well described I they could be. For example, how were multiple comparisons handled? Clearly there were many instances.

**Response:** Multiple comparisons were not corrected and this constitutes a limitation of our study. This limitation is acknowledge in the ‘Strengths and Methodological Considerations’ subsection. Nevertheless, in an attempt to minimise the problem with multiple comparison we have changed the *P*-value threshold for the main comparisons from <0.05 to <0.01(page 10, line 206), although the threshold for the post-hoc comparisons was kept as <0.05 (page 10, lines 198-199).  
  
5) The methodological considerations were not complete and many of the problems were not fully elucidated. For example, “ Second, the clusters identified in our study were not validated”. Not further explanation was provided. What does this mean? How could it affect the results? Did it alter the results?

**Response:** By ‘not validated’ we meant we did not know whether they relate to relevant clinical outcomes such as COPD-related hospitalisations and deaths, due to the lack of follow-up assessments, or whether the same findings identified in our study could be replicated in an another sample. In order to clarify this fact we have added some text to the discussion (page 29, lines 506-509). We do not think that correcting this limitation would alter the results, but it could make them stronger.  
  
6) Many conclusions are not justified or even explained. Specifically, how could the results help promote physical activity in COPD?

**Response:** A similar concern was raised by reviewer 2. Please refer to our response to comment 5 by that reviewer.  
  
7) I note on their website PLOS ONE specifies:  
“PLOS ONE editorial decisions do not rely on perceived significance or impact, so authors should avoid overstating their conclusions”  
Throughout the manuscript there are many occasions where the authors both overstate their conclusion and do so without sufficient evidence to justify their conclusion. Often mention words being the “largest” “biggest, multicentered”, such remarks appear to be highlighting novelty and interest, which I believe is not necessary in this journal. Also, the data is all part of larger studies and while I do not believe the specific results may have been reported elsewhere, this ends up being in a “grey” zone of publishing ethics.

**Response:** Although we do think our manuscript has made important advances in comparison to previous studies, we have tried to downsize the tone throughout the manuscript.   
  
Reviewer #4: Review attached  
  
Physical activity patterns and clusters in 1001 patients with COPD   
  
1) This is a large data set of daily physical activity data in COPD collated from multiple previous trials. Describing the overall pattern of physical activity is helpful and the principle of performing cluster analysis is novel in this area (certainly in a large sample size). Although the authors should be congratulated in co-ordinating the data collection, there are some inherent bias’ and limitations in the data which need further discussion. The clusters overall appear to describe a range of different levels of physical activity rather than different patterns (Figure 5) – the daily patterns seem rather similar except the variability seems to diminish as people are more inactive. The figures presumably would be published in colour as perhaps some of the detail in the figures does not ‘shine’ through, but if so it is even more important to highlight any novelty of these clusters in the text. Is the main finding that the least active and sedentary cluster are obese, more breathless with more severe disease? The limiting factor is the limited number of variables available due to the study design.

**Response:** The main findings of our study exceed the characterisation of the least active cluster. Considerable variability in physical activity measures and hourly patterns after stratification for different characteristics, and identifying clusters with different arrangements of physical activity are other main findings. We agreed that the hourly patterns are similar between the clusters, despite being statistically different as there was little or no overlap between the 95% confidence intervals of different patterns. Nevertheless, other important findings can be observed. More inactive clusters seem to present less variability in intensity compared to more active clusters, suggesting that they are similarly inactive throughout the day. Also, it seems that the more inactive a cluster is, the more similar its hourly patterns of week and weekend are, which supports the findings of Lee et al. (Public Health Nutr. 2013;16:1436-1444) who observed a consistently low physical activity pattern on both weekdays and weekends in the least active cluster. Moreover, irrespective of the cluster patients seem to perform the activities with the highest intensity during the morning, which was also observed after stratification for different characteristics (Fig. 2 and S1 Fig.). This could be taken into account when planning interventions such as energy conservation techniques, which have as one of the main aims to reduce unnecessary energy expenditure associated with activities of daily living. These pieces of information were added to the revised version of the manuscript (pages 25-26, lines 427-439). Finally, the constraint of having a limited number of variables to be compared between the clusters is acknowledged as a limitation (page 29, lines 509-510).  
  
2) The construct of physical activity is complicated and the health outcomes are discussed in a rather simplistic and generalised way. It is likely that sedentary behaviour is associated with different health outcomes to intense bouts of exercise which shouldn’t diminish the importance of either -suggest adding something to this effect in the discussion. The authors include many world renowned experts in this area and as such their ‘voice’ is likely to be far reaching. It is therefore essential that only messages directly from the data are discussed. There is no evidence from the data to suggest the current advice around intensity of activity/exercise should be altered even if there may need to be additional advice around not being sedentary. Please rephrase the comments under clinical relevance and at the bottom of Pg 6 Intro; ‘The focus may be shifted …’ Understanding the patterns of outcome would not be enough to change the focus of activity/exercise interventions – understanding the difference in outcome might do but this is likely to vary between patients and should not be a ‘global’ message.

**Response:** We are not advocating that sedentary behaviour or light activities should be the ‘only’ focus of future interventions. Even if we wanted to, our findings do not allow us this conclusion as already highlighted by this reviewer. However, we do think that interventions aiming to make patients more active might be more successful if they change the focus from moderate-to-vigorous only to sedentary behaviour and light activities, at least as an initial approach. As highlighted recently by Sparling et al. (BMJ 2015;350:h100), a reduction in sedentary time and an increase in light activities may prove more realistic than just focusing on moderate-to-vigorous physical activity, and this may pave the way to more intense exercises. This philosophy was recently endorsed in a clinical concise review about pulmonary rehabilitation and physical activity in COPD (Spruit et al., Am J Respir Crit Care Med. 2015;192:924-33). These explanations were added to the discussion for clarification in an attempt to soften the previous message (page 28, lines 487-490). In addition, the sentence in the introduction was changed slightly (page 8, lines 137-139).

Major comments  
1) By nature this analysis is retrospective and therefore suffers the usual weaknesses. It needs to be described as such in the study design. There is no mention on how many patients have been excluded and whether they had similar characteristics to the population described (the large dataset does not obviate this) which is potential for systematic bias.

**Response:** We have used the term ‘post-hoc’ to describe the retrospective aspect of our study. However, as this was no clear we have replaced that term with ‘retrospective’. Compared to patients included in the current analyses, those who could not be included were slightly older (68 (61-74) versus 67 (61 – 72) years) and had slightly higher BMI (26.9 (23.2 – 30.9) versus 25.8 (22.5 – 29.6) kg·m-2) and lower FEV1 (45 (33 – 59) versus 49 (34 – 64) % predicted); all P-values were<0.05. There was also statistical difference for the mMRC scale, but the values in both groups were exactly the same (i.e., 2 (1 – 3) points). Despite statistical difference the values are very similar or even the same, therefore we believe the differences between groups are not clinically relevant.

2) How was it decided who would contribute to the dataset? Were all authors of studies involving physical activity and COPD in a certain time frame using sensewear monitors contacted? Bias could have been introduced if this wasn’t systematically approached. (perhaps this data is in the supplementary files but if not please add to the methods)

**Response:** A proper systematic research was not performed. The research groups that contributed to the current study were conveniently selected from recent publications (articles in peer-reviewed journals and abstracts presented at major respiratory congresses) using the SenseWear Armband in patients with COPD. This limitation was acknowledged in File S1 (page 1, lines 9-12).

3) The data is cross-sectional so provides no insight into how these clusters relate to health outcomes – this needs adding to the limitations.

**Response:** We agree. We have previously mentioned that ‘…the clusters identified in our study were not validated’, meaning that we were not able to show whether they relate to relevant clinical outcomes, such as COPD-related hospitalisations and deaths due to the lack of follow-up assessments, or whether they could be replicated in another sample. However, as this was not clear the sentence was rephrased (page 29, lines 506-509).

4) As the authors state the clusters haven’t been validated in a further population.

**Response:** Please refer to our response to the comment above.

5) Presumably resting energy expenditure was not taken into consideration?

**Response:** No. Measurement of resting energy expenditure was not standard in most centres and we preferred not to estimate this parameter using equations based on patient characteristics due to the potential limitation of these equations in patients with COPD. For example, the Harris-Benedict equation – one of the most frequently used equations – takes into account the sex, age, height, and weight of the subjects, and we believe other characteristics might influence resting energy expenditure in patients with COPD.

6) Presumably the same formulae for extrapolating energy expenditure were used? Are these the same between the Sensewear and Mini-sensewear?

**Response:** The internal algorithms are slightly different between the monitors. However, as the SenseWear Armband has been shown to be valid in different field and laboratory studies (Colbert et al., Med Sci Sports Exerc 2011;43:867-76; Mackey et al., J Gerontol A Biol Sci Med Sci 2011;66:1108-13; Hill et al., Thorax 2010;65:486-91; Cavalheri et al., Respir Med 2011;105:922-9), we believe the SenseWear Mini Armband is also valid in COPD since it is an improved version of the former. Moreover, we do not think this could have compromised our analyses since only a small proportion of subjects used the SenseWear Mini Armband (i.e., 58 patients).

7) The authors need to comment somewhere that these levels are all at an absolute level. Relative to their peak exercise capacity these people will be active at a much higher relative intensity – it is not fully understood whether this might still have health benefits which needs comment.

**Response:** All values represent absolute values. Measurement of peak exercise capacity was not available. This piece of information was added to the methods (page 10, lines 187-188).

8) Suggest removing FEV1/FVC as a comparator variable across the clusters as this is a descriptor only.

**Response:** Done.

9) The abstract is rather vague more precision is needed in the results about what the associations are between the variables. Please describe ‘cluster 1’ otherwise it does not make sense when reading the abstract alone.

**Response:** We have provided more details about the associations between patient characteristics and physical activity measures. Moreover, cluster 1 was described as the one with the most inactive patients.

10) Although there has been evidence (from some of the authors) that there are differences between weekday and weekend activity, is that supported in this large dataset? The patterns look rather similar in all the figures. Was there a statistical difference in the volume of activity or the pattern of activity between the weekend and weekday in any of the data. If not it might be better to combine this data for the cluster analysis?

**Response:** We agree that differences in physical activity summary values between week and weekend are important, but as this was already investigated in previous studies we preferred not to include it in our study. The analysis of hourly patterns of week and weekend however, was not done before and therefore was included in our study. Please refer to our response to the first comment by this reviewer for more details about hourly patterns during week and weekend.

11) Interesting the hourly pattern of activity looked rather similar for healthy and COPD participants – is this what was expected?

**Response:** We expected that COPD patients could have a different pattern due to the functional limitations related to the disease. However, the pattern between the groups was rather similar, although healthy subjects performed their activities at a higher intensity. Nevertheless, as other reviewers have also raised concerns about the analyses involving healthy subjects, we preferred to remove them from our manuscript, as well as all related analyses. Most changes were done in the Results (pages 16-18, lines 296-313) and Discussion (page 24, lines 395-401).

12) As the authors mention, a major limitation is that few patients had mild –moderate COPD.

**Response:** The majority of patients had very severe disease (38%), but the amount of patients with mild or moderate disease was reasonable (45%). Notwithstanding, some types of patients with COPD might be underrepresented, such as patients from primary care, which does represent a major limitation. On the other hand, having patients from different studies and countries allowed us to have a more diverse sample, which may enhance the external validity of our findings. These explanations are present in the manuscript (pages 28-29, lines 503-506).

13) Did the ‘italian’ patients agree to have their data stored and used for research purposes even if ethics permission wasn’t needed to collect the data

**Response:**   
  
Minor comments  
1) References 11,12 are in children – suggest adding ‘In other population, …’ at the start of the first paragraph so that the reader knows this is not referring to patients with COPD.

**Response:** As in the same sentence we also present one study in patients with COPD, we have added ‘Studies in different populations…’ to the beginning of the sentence instead.

2) Ref 14 isn’t a reference about physical activity but phenotyping – move the ref to after ‘cluster analysis [14] …’ – suggest changing the phrasing of ‘cluster analysis will be …’ to ‘could be’

**Response:** Done.

3) Intro Pg 6 following line from above change ‘will then’ to ‘could then lead …’

**Response:** Done.

4) Table 1 the numbers don’t seem to add up for the BMI classifications (only around 100 patients in total?) nor for the GOLD classifications. Probably don’t need both GOLD classifications, similarly probably don’t need height and weight to be described.

**Response:** The numbers for all categorical variables (eg, BMI, GOLD) are expressed as relative frequency (i.e., percentages), that is why the sum equals 100. Regarding GOLD classifications, we decided to include both in our study since recent studies have shown that these classifications are not interchangeable. Body weight and height were removed.

5) Results. The initial description is not necessary as the results are in Table 1. If left in please add ‘The majority’ and remove ‘ GOLD stage D …’ as this is a bit misleading as there was reasonable split across the GOLD stages in the small number described.

**Response:** We preferred to keep the description but we have made the changes as suggested.

6) Pg 12 line 224 add ‘total’ to ‘time and lowest energy expenditure’ and same Pg 13 line 247

**Response:** Done.

7) Remove comment regarding FEV1/FVC line 299

**Response:** Done.