Author's response to reviews

Title: Mortality in patients with COPD exacerbations attending emergency departments

Authors:

José M. Quintana (jmquinta@hgda.osakidetza.net)
Cristóbal Esteban (cristobal.estebangonzalez@osakidetza.net)
Anette Unzurrunzaga (anette.unzurrunzagagarate@osakidetza.net)
Susana García-Gutiérrez (susana.garciagutierrez@osakidetza.net)
Nerea González (nerea.gonzalezhernandez@osakidetza.net)
Irantzu Barrio (irantzubarrio@gmail.com)
Inmaculada Arostegui (inmaculada.arostegui@ehu.es)
Iratxe Lafuente (iratxe.lafuenteguerrero@osakidetza.net)
Marisa Baré (mbare@tauli.cat)
Nerea Fernández-de-Larrea (nerea.fernandez@salud.madrid.org)
Silvia Vidal (svidals@gmail.com)
The IRYSS-COPD Group (cristina.munozfernandez@osakidetza.net)

Version: 2
Date: 28 February 2014

Author's response to reviews: see over
Reviewer's report

Reviewer 1: Nicolas Roche

Reviewer's report:

This prospective multicenter study assessed prognostic factors of in-hospital, one week and one month post-discharge death in patients admitted for COPD exacerbations.

Using results of multivariate analyses, the authors produce a continuous and categorical DeCOPD score based on age, baseline dyspnea level, use of LTOT or LT-NIV, altered mental status and use of inspiratory accessory muscles on arrival. The AU-ROC for this score was high, 0.84, and higher than the predictive values of FEV1 % predicted and current GOLD classification.

Arterial blood gases did not improve the predictive value of the model. The authors appropriately use a derivation and a validation sample with large sample sizes. There were only rather few missing data.

Major revisions

1. Since data files appear to be quite complete, the authors could test the yield of other biological factors such as those that were used for the Pneumonia Severity Index. They could also test their score against others that include biological values such as the COPD–specific DECAF score (Steer Thorax 2012) and the generic BAP-65 score (ref 8 of their paper), both of which have shown acceptable prognostic value in previous studies.

Author’s response: Although our database includes several variables some of the variables needed to construct either the PSI (not all the PSI parameters recorded), the DECAF (though we have some of the DECAF parameters not all laboratory tests were recorded in our study, or not in similar manner) or the BAP-65 scores are, either, not available or not complete enough as to construct those scores properly. We must point out that, by the time our study started some of those new scores have not been published yet and, therefore, we were not able to properly include some of the predictor variables identified in those studies. On relation to the DECAF score, our score do not demand having an XRay or lab test to be completed. Additionally, an exclusion criterion of our study was to have a diagnosis of pneumonia at ED arrival, since we understand that those cases were not pure eCOPD. On the other hand, our score can be fulfilled quicker than others and in many other situations (not only in the hospital). Alternatively though with some missing data but, from our point of view, affordable, we have been able to construct the ADO (last version) and HADO severity scores for COPD stable patients. Nevertheless, both scores for stable COPD patients never have been tested nor validated for eCOPD patients and, therefore, the comparison must be taken with care.

Author’s action: We have complete table 4 with 2 more scores, added information
related to this table in the Results section (page 14) and comments in the Discussion (page 17).

Table 4. Comparison of different prediction scales on short-term mortality on eCOPD patients.

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Short-term mortality</th>
<th>p-value*</th>
<th>AUC</th>
<th>95% CI</th>
<th>p-value**</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECOPD Score continuous</td>
<td>Yes: 2484</td>
<td>9.03(4.27)</td>
<td>3.07(3.4)</td>
<td>≤0.0001</td>
<td>0.85</td>
</tr>
<tr>
<td>DECOPD categorical Score</td>
<td>No: 2484</td>
<td></td>
<td></td>
<td>≤0.0001</td>
<td>0.84</td>
</tr>
<tr>
<td>Mild risk</td>
<td>Yes: 1081</td>
<td>3(0.28)</td>
<td>1078(99.72)</td>
<td>≤0.0001</td>
<td>0.84</td>
</tr>
<tr>
<td>Moderate risk</td>
<td>Yes: 865</td>
<td>11(1.27)</td>
<td>854(98.73)</td>
<td>≤0.0001</td>
<td>0.82</td>
</tr>
<tr>
<td>Severe risk</td>
<td>Yes: 441</td>
<td>20(4.54)</td>
<td>421(95.46)</td>
<td>≤0.0001</td>
<td>0.82</td>
</tr>
<tr>
<td>Very severe risk</td>
<td>Yes: 97</td>
<td>24(24.74)</td>
<td>73(75.26)</td>
<td>≤0.0001</td>
<td>0.82</td>
</tr>
<tr>
<td>GOLD FEV1%</td>
<td>Yes: 2069</td>
<td></td>
<td></td>
<td>0.0133</td>
<td>0.62</td>
</tr>
<tr>
<td>FEV1% ≥80</td>
<td>Yes: 76</td>
<td>0(0.0)</td>
<td>76(100)</td>
<td>≤0.0001</td>
<td>0.82</td>
</tr>
<tr>
<td>50 ≤ FEV1% &lt;80</td>
<td>Yes: 643</td>
<td>6(0.93)</td>
<td>637(99.07)</td>
<td>≤0.0001</td>
<td>0.82</td>
</tr>
<tr>
<td>30 ≤ FEV1% &lt;50</td>
<td>Yes: 959</td>
<td>28(2.92)</td>
<td>931(97.08)</td>
<td>≤0.0001</td>
<td>0.82</td>
</tr>
<tr>
<td>FEV1% ≤30</td>
<td>Yes: 391</td>
<td>13(3.32)</td>
<td>378(96.68)</td>
<td>≤0.0001</td>
<td>0.82</td>
</tr>
<tr>
<td>GOLD COPD combined assessment</td>
<td>Yes: 2090</td>
<td></td>
<td></td>
<td>&lt;0.002</td>
<td>0.71</td>
</tr>
<tr>
<td>Low risk, low symptom burden</td>
<td>Yes: 286</td>
<td>1(0.35)</td>
<td>285(99.65)</td>
<td>≤0.0001</td>
<td>0.82</td>
</tr>
<tr>
<td>Low risk, higher symptom burden</td>
<td>Yes: 335</td>
<td>4(1.19)</td>
<td>331(98.81)</td>
<td>≤0.0001</td>
<td>0.82</td>
</tr>
<tr>
<td>High risk, low symptom burden</td>
<td>Yes: 351</td>
<td>0(0.00)</td>
<td>351(100)</td>
<td>≤0.0001</td>
<td>0.82</td>
</tr>
<tr>
<td>High risk, higher symptom burden</td>
<td>Yes: 1118</td>
<td>46(4.11)</td>
<td>1072(95.89)</td>
<td>≤0.0001</td>
<td>0.82</td>
</tr>
<tr>
<td>HADO</td>
<td>Yes: 1886</td>
<td></td>
<td></td>
<td>0.0002</td>
<td>0.68</td>
</tr>
<tr>
<td>Mild(&lt;4)</td>
<td>Yes: 188</td>
<td>0</td>
<td>188(100)</td>
<td>≤0.0001</td>
<td>0.78</td>
</tr>
<tr>
<td>Moderate(5-7)</td>
<td>Yes: 604</td>
<td>3(0.50)</td>
<td>601(99.50)</td>
<td>≤0.0001</td>
<td>0.78</td>
</tr>
<tr>
<td>Severe(&gt;8)</td>
<td>Yes: 1095</td>
<td>33(3.01)</td>
<td>1062(96.99)</td>
<td>≤0.0001</td>
<td>0.78</td>
</tr>
<tr>
<td>ADO (0-14) continuous</td>
<td>Yes: 2067</td>
<td>11.36(1.71)</td>
<td>9.29(2.19)</td>
<td>≤0.0001</td>
<td>0.78</td>
</tr>
</tbody>
</table>

*p-value of the relationship of each parameter with short-term mortality

**p-value or the comparison of areas under Receiver Operating Characteristic Curves with respect to the AUC of the DECOPD continuous scale.

Short-term mortality: mortality during hospital admission or, if discharged from the ED, in one week.

AUC: area under the ROC curve; 95% CI: 95% Confidence Intervals of the AUC.

DeCOPD: Death in eCOPD continuous and categorical scores

FEV1 classified following GOLD classes (FEV-GOLD, reference #14) or the GOLD combined assessment classes (reference #24).

ADO: ADO index that uses age, dyspnoea and FEV(1). Reference #21

HADO: Health-Activity-Dyspnoea-Obstruction (HADO) score. Reference #22.
2. Since arterial blood gases did not add anything to the proposed clinical scores, the authors could compare the predictive value of their score with other purely clinical scores (ref 20 of their paper).

Author’s response: Again, though most of the variables needed are available, unfortunately some of the data required to construct the score of ref# 20 (specifically, some of the specific signs of severity), which also was published once our study started the recruitment process, are not available or not equivalent or the minimal quality as to construct the score. To properly construct some of the existent scores, and to compare with ours, we need to be sure that we have similar data as the needed to construct the score. In other case, we prefer not to make comparisons that are not justifiable by the quality or availability of our data.

Author’s action: No changes made.

**Quality of written English:** Needs some language corrections before being published

Author’s response: The manuscript was edited by a professional native English speaking editor, as reflected in the acknowledgment section.

Author’s action: The manuscript has been reviewed again by a professional native English speaking editor by a professional native English speaking editor.
Reviewer 2: Nicholas Hart

Reviewer’s report:

General Comments
In this large prospective multi-centre observational cohort study (16 hospitals recruiting at 1.4 patients per month for 27 months), the authors have reported, based on predictive modelling, 5 factors (age, use of accessory muscle/paradoxical breathing, MRC dyspnoea score, altered GCS and use of LTOT or home NIV) that can be used to predict early mortality in patients attending the emergency department (ED) with acute exacerbations of COPD. Overall, the methodology is sound and the authors have developed a severity score to predict the mortality risk in one half of the patient cohort and then subsequently validated the score in the other second half of the cohort.

Major Comments
1. The development of a scoring system combining these five variables is novel. However, age, baseline dyspnoea, use of LTOT, NIV use and altered level of consciousness have all been reported previously as predictors of poor outcome. In this reviewer’s opinion, (a) the extent to which these data add to our current understanding of acute exacerbations of COPD, in particular, in terms of poor outcome needs to be clarified (b) the reason for excluding the ‘frequent exacerbators’ is counterintuitive and this approach needs explanation and (c) it needs to be highlighted that none of the five variables reported are modifiable, so the rationale for potential clinical strategies to be developed to reduce early mortality is limited.

Author’s response: a) We agree that in a research field, as COPD, it is difficult to add novelty. Nevertheless, as the reviewer pointed out, the importance comes from the development of a scoring system combining these five variables of all patients attended at the ED with an eCOPD. As far as we know, this is the first time a prediction rule has been developed from eCOPD patients including not only those admitted to the hospital but also those discharged to home from the ED. Therefore, the DeCOPD can be applied to all patients attending the ED (in our sample, almost 40% of patients were discharged to home from the ED). Another important practical issue is that, due to the variables finally included in our score, not also can help to the pneumologist or internist but also to the physician working at the ED, critical care units and the primary care physician as well.

b) On relation to ‘frequent exacerbators’, we must, first, clarify that those patients are not excluded from the study but just their repeated visits since we included them but just their first episode in this analysis. From a statistical point of view, including all the episodes would have needed more complex statistical models which would have added more complication to the statistical analysis section with minor changes to the final models.

c) We agree that age or previous need of LT-DOT or NIMV at home are reflecting the fragility of the patient and disease but efforts to avoid new exacerbations or readmissions in these patients can be conducted through different programs (telemedicine, continuous of care..etc), as also has been reported elsewhere (Suh ES, Mandal S, Hart N. Admission prevention in COPD: non-pharmacological management.
Severe baseline dyspnea level can be also addressed through respiratory rehabilitation programs. Finally, “use of inspiratory accessory muscle or paradoxical breathing” and “altered level of consciousness” reflect the severity of the current exacerbation (therefore are not permanent but modifiable) and should alert the ED physician, in conjoin with the previous factors, of the severity of the patient and, them, lead the treatment and follow up adequately. These two last variables make a distinction of our score to those used for stable COPD patients (GOLD classifications, ADO or HADO scores).

Author’s action: We clarify the role of the ‘frequent exacerbators’ groups in the statistical analysis part of the Methods section (page 8-9).

2. The scoring system is potentially useful but this reviewer considers it needs modification for the clinical reader. Indeed the clinician wishes to have a ‘quick and easy’ method to risk stratify the patient into risk of ICU/RICU admission and mortality risk with risks reported for these two categories. Reviewing of Figure 1, a ‘cut off’ value of #7 >7 for ICU/RICU admission and #11 >11 for mortality. The current scoring system does not facilitate risk prediction for admission (50-89% across the 4 categories) or IMV (1-5% across the 4 categories) as the prevalence is very high and very low, respectively. Hence, for the clinician a focus on ICU/RICU admission and mortality risk would be useful clinically.

Author’s response: It is almost impossible that so different outcomes (mortality, ICU/RICU admission, and, specially, for hospital admission, readmissions..etc) may share similar equally good predictors and with same weights. Having this in mind, we develop a score specifically for short term mortality. Therefore, we did not expect to have good prediction for any other outcomes. Nevertheless, we have included in Figure 1 some of the outcomes the reviewer is requesting (as a conjoint variable of ICU-with IMV- and IRCU) but we advise not to use to select people for ICU-IMV-IRCU treatment due to the low sensitivity (53.9%; Specificity: 79.05%; PPV: 34%; NPV: 89.54%) at a ‘cut off’ value of >=7.

<table>
<thead>
<tr>
<th>DECAF Score</th>
<th>ICU-IRCU or IMV YES</th>
<th>ICU-IRCU or IMV NO</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score&gt;=7</td>
<td>138 (53.9)</td>
<td>268 (20.95)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Score&lt;7</td>
<td>118 (46.1)</td>
<td>1011 (79.05)</td>
<td></td>
</tr>
</tbody>
</table>

Author’s action: New outcome parameter is included in Figure 1 and the accuracy parameters in the Results section (page 14).

3. Although information was collected during hospital admission or ED attendance, there were no data reported regarding the trajectory of physiological or symptomatic change between admission to hospital and discharge, although this information has presumably been routinely collected (e.g. oxygen saturations, respiratory rate, heart rate, breathlessness scores). In-hospital mortality or death soon after discharge may be an indicator of poor underlying health status and lung function, however, it may
also suggest failure of inpatient and post discharge treatment, which would be
demonstrated by limited change in physiological parameters, including early warning
and breathlessness scores. Finally, outcome is related to delivery of care and this
reviewer would appreciate data detailing mortality between the sites.

Author’s response: We agree with the reviewer that data recorded at several points in
time regarding the trajectory of physiological or symptomatic change from the arrival to
the hospital would improve our knowledge of the evolution of patients with adverse
outcomes. Nevertheless, the point of this score is to help ED physicians with the data
usually available at the ED. Though we have some of that information, the rate of
missing data is important.

We include here some information on those physiological or symptomatic parameters
between admission to hospital and discharge but just for those admitted to the hospital
(e.g. oxygen saturations, respiratory rate, heart rate, breathlessness scores). We must
point out that this is an observational study and, then, we collect information routinely
available.

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>Pre</th>
<th>Post</th>
<th>Change</th>
<th>n</th>
<th>Pre</th>
<th>Post</th>
<th>Change</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SatO2</td>
<td>1126</td>
<td>90.38(6.76)</td>
<td>92.48(4.87)</td>
<td>2.09(7.57)</td>
<td>29</td>
<td>91.28(5.36)</td>
<td>90.21(7.46)</td>
<td>-1.07(8.04)</td>
<td>0.01</td>
</tr>
<tr>
<td>Respiratory rate</td>
<td>457</td>
<td>22.67(5.37)</td>
<td>18.73(4.3)</td>
<td>-3.94(5.56)</td>
<td>21</td>
<td>23.95(4.94)</td>
<td>21.81(5.93)</td>
<td>-2.14(6.56)</td>
<td>0.35</td>
</tr>
<tr>
<td>Heart rate</td>
<td>778</td>
<td>90.52(15.57)</td>
<td>79.82(13)</td>
<td>-10.716</td>
<td>20</td>
<td>89.3(10.54)</td>
<td>89.6(20.53)</td>
<td>0.3(17.29)</td>
<td>0.0085</td>
</tr>
<tr>
<td>Dyspnoea level</td>
<td>1293</td>
<td>3.39(1.68)</td>
<td>2.09(1.62)</td>
<td>-1.29(1.86)</td>
<td>16</td>
<td>3.5(1.71)</td>
<td>3.19(1.68)</td>
<td>-0.31(2.47)</td>
<td>0.1683</td>
</tr>
</tbody>
</table>

Dyspnoea level measured in a scale from 1-no dyspnoea –to 7-higher level of dyspnoea –.

It can be seen that SatO2 diminished a little in those who died, while increased in
survivals, while with heart rate almost the opposite was found. Nevertheless, due to the
lack of information and the bias with died patients (very small sample size) we do not
find appropriate to include this information. As some audits have shown here, in
Europe and in the UK (1.- Pozo-Rodriguez F, López-Campos JL, Alvarez-Martinez CJ,
Fortes A, Sanchez Nieto JM, Lopez-Gabaldon E, Cosio BG, Agusti A; AUDIPOC Study
Group. Clinical audit of COPD patients requiring hospital admissions in Spain:
Lopez-Campos JL, Pozo-Rodriguez F, Hartl S; European COPD Audit team. European
hospital adherence to GOLD recommendations for chronic obstructive pulmonary
disease (COPD) exacerbation admissions. Thorax. 2013 Dec;68(12):1169-71. 3.-
Roberts CM, Ryland I, Lowe D, Kelly Y, Bucknall CE, Pearson MG; Audit Sub-
committee of the Standards of Care Committee, British Thoracic Society; Clinical
Effectiveness and Evaluation Unit, Royal College of Physicians. Audit of acute
admissions of COPD: standards of care and management in the hospital setting. Eur
Respir J. 2001 Mar;17(3):343-9) the absence of routine information on the records of
these patients is very high. That is why we found an important rate of missing data in
some important parameters. For that reason, we found not appropriate to include data with an important rate of missing information in the manuscript.

On the other hand, there were not differences in mortality between the sites (p=0.064 unadjusted; p=0.96 adjusted by the DeCOP score).

Author’s action: Information on differences in mortality between the sites was included in the Results section (page 13).

4. The relevance of re-admission data cannot be underestimated. It would be of interest to review the readmission data, or data on acute healthcare utilisation after discharge from the index event, to assess the utility of the AECOPD severity score in predicting readmissions within, for example, 28 days and 3 months. The positive and negative predictive values for early mortality and readmission are wholly important to the clinician.

Author’s response: We have included here data on readmission (at 1 and 2 months) and health care utilization (new visits to the ED at 2 months) but, again, predictors of readmissions or healthcare utilization are (and have to be) far different from the ones used in our score on mortality.

The Sensitivity (S), Specificity (Sp), Positive Predictive value (PPV) and Negative Predictive value of readmission and short-term mortality are as follows:
Short term mortality (cut-off point at >11). S: 41.38%; Sp=97%; PPV=24.74%; NPV=98.60.
Readmission at 10 days (cut-off point at >=7). S: 26.5; Sp= 78.5; PPV=4.01; NPV=96.87.
Readmission at 1 month (cut-off point at >=7). S:25.23; Sp=78.69; PPV=10.41; NPV=91.47.

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
<th>Very Severe</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Readmission 2 months</td>
<td>2484</td>
<td>220(20.35)</td>
<td>239(27.63)</td>
<td>144(32.65)</td>
<td>28(28.87)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Readmission 1 month</td>
<td>2484</td>
<td>80(7.4)</td>
<td>86(9.94)</td>
<td>46(10.43)</td>
<td>10(10.31)</td>
<td>0.1303</td>
</tr>
<tr>
<td>Readmission 10 days</td>
<td>2484</td>
<td>26(2.41)</td>
<td>35(4.05)</td>
<td>20(4.54)</td>
<td>2(2.06)</td>
<td>0.0820</td>
</tr>
<tr>
<td>New ED visit 2 months</td>
<td>2032</td>
<td>249(25.88)</td>
<td>25(35.91)</td>
<td>129(40.44)</td>
<td>20(38.46)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Author’s action: Data on readmission and mortality and accuracy measures have been included in Figure 1 (mortality), new additional file/Table 5 and text (readmission) (page 14).

5. Although we acknowledge that there has been no previous studies reporting mortality prediction models in COPD patients attending the ED, there has been previous studies that have reported predictors of re-admission and mortality in hospitalised COPD patients. Indeed, Steer et al (Thorax 2012) identified baseline dyspnoea and a composite score (the DECAF score) as a predictor of poor outcome. Comparison of the DeCOPD score and DECAF score would be interesting for the
Author’s response: Although our database includes several variables some of the variables needed to construct the DECAF score (though we have some of the DECAF parameters not all laboratory tests were recorded in our study, or not in similar manner) are not available as to construct those scores properly. But our score do not demand having an XRay or lab test to be completed. On relation to this, an exclusion criterion of our study was to have a diagnosis of pneumonia at ED arrival, since we understand that those cases were not pure eCOPD. Additionally, our score can be fulfilled quicker than others and in many other situations (not only in the hospital). We must point out that, by the time our study started some of those new scores have not been published yet and, therefore, we were not able to properly include some of the predictor variables identified in those studies. Alternatively though with some missing data but, from our point of view, affordable, we have been able to construct the ADO (last version) and HADO severity scores though both developed for COPD stable patients. Nevertheless, both scores for stable COPD patients never have been tested nor validated for eCOPD patients and, therefore, the comparison must be taken with care.

Author’s action: We have complete table 4 with 2 more scores, added information related to this table in the Results section (page 14) and comments in the Discussion (page 17).

6. The authors stratified the inclusion criteria to exclude those with (1) other major pathologies and co-morbidities (2) COPD not confirmed on spirometry (3) ‘frequent exacerbator’ phenotype attending ED on >1 occasion. With the 145 incomplete datasets, of the 3267 patients screened 24% of patients were not included. Was there a mortality difference between the multiple ED attenders (14% re-attendance rate) and the single ED attenders? Did co-morbidity confer a poor outcome as the authors hypothesised during the study design? What was the final diagnosis of the 56 patients without spirometric evidence of COPD?

Author’s response: On relation to the three questions of the reviewer: 1) there were not differences in the mortality rate between the multiple ED attenders and the single ED attenders (4.4% vs. 2.7%; p value=0.32). 2) Information on co-morbidity (as measured by the Charlson index) and mortality is included below. Though there are differences in mortality in the unadjusted analysis, it was not in the multivariate

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>No Died</th>
<th>Died</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indice Charlson&lt;2</td>
<td>1003</td>
<td>986 (98.31)</td>
<td>17 (1.69)</td>
<td>0.0181</td>
</tr>
<tr>
<td>Indice Charlson=2</td>
<td>683</td>
<td>670 (98.1)</td>
<td>13 (1.9)</td>
<td></td>
</tr>
<tr>
<td>Indice Charlson&gt;2</td>
<td>801</td>
<td>772 (96.38)</td>
<td>29 (3.62)</td>
<td></td>
</tr>
</tbody>
</table>

3) Finally, on relation to the final diagnosis of the 56 patients without spirometric evidence of COPD, the reviewers were training to check the diagnosis of those new COPD patients at the ED index visit. On those case where, at 2 months, the diagnosis
was not confirmed they were excluded from the study and we did not retrieve any standardized additional information on those patients.

Author’s action: We added information on mortality among the multiple ED attenders and the single ED attenders in the Results section (page 12).

7. Missing MRC data was treated the same as MRC Grade 5. This needs further discussion as this attracts the highest score of 5 points out of 18 points. Should this not be treated as part of the missing datasets? How was baseline MRC dyspnoea score measured? Was this from the patient on admission or from their notes?

Author’s response: Baseline dyspnoea was measured by direct interview with the patient at 24 hours after the patient attended the ED by asking patients about their basal status (previous to the current exacerbation) by using the Fletcher MRC scale (Fletcher CM, Elmes PC, Fairbairn AS, Wood CH. The significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. Br Med J 1959 Aug 29;2(5147):257-66). In the analysis, we found that patients with missing data behaved similarly as those in the most severe MRC class (mortality rates for those with missing values were similar as those in the MRC severe class). We found that finding understandable since patients more severely affected by the exacerbation are less able to answer.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>NO DEATH</th>
<th>DEATH</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRC 1</td>
<td>188</td>
<td>188(100)</td>
<td>0(0)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>MRC 2</td>
<td>600</td>
<td>599(99.83)</td>
<td>1(0.17)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>MRC 3</td>
<td>501</td>
<td>494(98.6)</td>
<td>7(1.4)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>MRC 4</td>
<td>672</td>
<td>662(98.51)</td>
<td>10(1.49)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>MRC 5</td>
<td>276</td>
<td>253(91.67)</td>
<td>23(8.33)</td>
<td>0.63</td>
</tr>
<tr>
<td>MRC Missing</td>
<td>250</td>
<td>232(92.8)</td>
<td>18(7.2)</td>
<td>REF</td>
</tr>
</tbody>
</table>

Comparison of the mortality rate of the group of patients with missing data on that question with the other MRC categories gave p values <0.0001 while with the MRC=5 group it was 0.63.

On the other hand, we tried to avoid missing data, if feasible, and if the assumptions are acceptable which we found this was the case. Additionally, the results of the analysis with and without those patients delivered similar results (AUC for the model without missing= 0.85), though missing almost a 10% more of patients.

Author’s action: We have included and explanation on this issue in the Methods (page 8-9) and Discussion sections (page 17).

8. GCS is a multiple component assessment test (E4/4, V5/5, M6/6). How altered was the GCS? Was the main changes in a specific domain e.g. Verbal – confusion. These points need clarification.
In the case of the Glasgow score just if was recorded if it was altered (<15 points) or not.

We added that information on the Methods section (page 7).

The mild, moderate, severe and very severe risk classification needs clarification in terms of risk of within hospital death, death within 1-week of discharge and death within 1-month of discharge

We have included all those outcomes (death, in-hospital death - within 1-week of ED discharge would be the difference between the 2 previous- and death within 1-month of discharge) in Figure 1.

Changes introduced in Figure 1.

Did length of use of LTOT or NIV influence the findings?

The length of previous use of LTOT or NIV was not recorded though we require a length of previous use of LTOT or NIV longer than 1 year to be included a patient in this category.

None

How was accessory muscle activity confirmed and quantified?

That information was recorded as it was provided by the ED physician in charge of the patient. Each patient was just classified as having use of inspiratory accessory muscle or not, in one hand, and, separately, as having paradoxical breathing or not at the ED evaluation (at arrival and at ED discharge).

Clarified in the Methods section how the information was gathered (page 7).

Why did ABG not add value to the scoring system? Was it because the pH values were all in the normal physiological range as a number of patients were excluded i.e. those who attended ED more than once. This needs further discussion.

As it can be seen in Table 1, most patients had a pH above 7.35, but as it can be seen below, this is not because those with repeated visits were excluded.

<table>
<thead>
<tr>
<th>Variable</th>
<th>&gt;=7.35</th>
<th>7.26-7.35</th>
<th>&lt;7.26</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple ED attenders</td>
<td>556(86.74)</td>
<td>74(11.54)</td>
<td>11(1.72)</td>
<td>0.52</td>
</tr>
<tr>
<td>Single ED attenders</td>
<td>1752(86.73)</td>
<td>219(10.84)</td>
<td>49(2.43)</td>
<td></td>
</tr>
</tbody>
</table>
13. The authors comment on ‘generalizability’ in the discussion, in particular, in the context of gender. This reviewer is concerned that 390 patients were excluded as they attended the ED on more than one occasion.

Author’s response: As I said before, those patients were not excluded from the study but just their following episodes from the analysis performed in this manuscript. We included those patients though just their first episode during the recruitment process.

Author’s action: An explanation was included in the manuscript (page 8-9) and Discussion (page 17).

Minor Comments
1. Introduction needs more detail including a number of key references in this area.

Author’s response: We have tried to improve the introduction by updating the references in this issue.

Author’s action: See changes in (page 4).

2. The English needs attention in places

Author’s response: The manuscript was edited by a professional native American-English speaking editor, as reflected in the acknowledgment section.

Author’s action: The manuscript has been reviewed again.

3. ‘Extensive bronchiectasis’ was an exclusion criteria. How was this defined?

Author’s response: We based the exclusion on the medical record of the patient. Patients who have a previous documented diagnosis of ‘Extensive bronchiectasis’ were excluded.

Author’s action: None.

4. ‘IRCU’ needs to be replaced with ‘RICU’ – Respiratory Intensive Care Unit

Author’s response: It is not clear to us if we can call RICU since they are Intermediate Respiratory Care Unit, not intensive care units. Therefore, we keep the term to avoid confusion.
Author’s action: No changes done.

5. Statistics section could be shortened or placed in the online supplement.

Author’s response: An important part of the methods of the manuscript is the statistics perform. Additionally, being an online journal we considered that space was not an issue in this case.

Author’s action: None.

6. By paradoxical movement, do the authors mean ‘Hoover’s’ sign i.e. in-drawing of the lower rib-cage during inspiration.

Author’s response: Correct.

Author’s action: None.