PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Assessment of PaO2/FiO2 for stratification of patients with moderate and severe acute respiratory distress syndrome
AUTHORS	Villar, Jesús; Blanco, Jesús; del Campo, Rafael; Andaluz-Ojeda, David; Díaz-Domínguez, Francisco; Muriel, Arturo; Córcoles, Virgilio; Suárez-Sipmann, Fernando; Tarancón, Concepción; González-Higueras, Elena; López, Julia; Blanch, Lluis; Pérez-Méndez, Lina; Fernández, Rosa; Kacmarek, Robert

VERSION 1 - REVIEW

REVIEWER	Robinder G. Khemani MD, MsCI Children's Hospital Los Angeles University of Southern California Keck School of Medicine
REVIEW RETURNED	18-Nov-2014

GENERAL COMMENTS	Dr. Villar and collogues prevent a very interesting analysis of existing observational data regarding standardized assessment of lung injury severity for adults with ARDS. They obtained PF ratios 24 hours after ARDS diagnosis (AECC and Berlin), on standardized ventilator settings and conclude that using these PF values while on standardized ventilator settings better stratifies risk of mortality, pulmonary mortality, ventilator free days, and non-pulmonary organ failure compared to the initial PF ratio based classification using Berlin criteria at ARDS onset. This paper highlights some of the unanswered and difficult questions that exist with ARDS diagnosis and risk stratification, in particular evolution of disease process, and variability in ventilator management as a confounder for the assessment of disease severity. It is an incredibly important area to clarify, particularly when considering risk benefit profiles for higher risk therapies and for clinical trials. In general, the paper is well written and methodology is reasonably clear. The strengths of the suggested standardized ventilator approach is of course minimizing the elements of variability in care, in particular with application of PEEP and FiO2. The limitations surround the requirement for practitioners at the bedside to change their behavior and perform such a maneuver, which is unlikely to be done as part of routine clinical care. Certainly such a maneuver could be done for the purposes of a study, or perhaps before contemplating a higher risk therapy, such as prone positioning or high frequency oscillatory ventilation. As such, when approaching defining an already under-recognized clinical syndrome such as ARDS, one has to find the balance between attempting to be inclusive and simple, to ensure that all patients with the disease are recognized, and requiring a specific maneuver to specifically risk stratify patients, which may ultimately contribute to the under-recognition of the disease if practitioners do not perform the maneuvers. This is particularly

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

important in the face of growing data to suggest that using automated early detection systems to alert providers of patients who have ARDS, or who are at risk of developing ARDS may improve outcome. The authors allude to some of these points in the discussion, but it would benefit from being explicit about the potential down-side of such an approach with regards to recognition of the disease, in particular for application of relatively low-risk therapies (i.e. lung protective VT management, fluid management etc…). If one cannot be labeled as having ARDS until 24 hours into the course of disease, there is potential for practitioners to pay less attention to these relatively low-risk but beneficial therapies- potentially worsening outcomes for these patients.

To justify such a maneuver, it would be important for the authors to demonstrate that this approach is necessary, and cannot be accomplished using existing data, without a specific maneuver. Fundamentally, there are two important potential explanations for the author's findings, which are worth trying to tease out specifically to strengthen their argument. One finding is that patients change over the first 24 hours of the disease, either in response (or lack of response) to conventional therapies, or as part of resolution of disease. Therefore, the 24 hour value is in many ways, measuring response to therapies, and may be better predictive of outcome. Others have found similar results (we found similar findings in pediatrics) where the average value over the first 24 hours after ALI diagnosis for all oxygenation parameters (Oxygenation index, Oxygen Saturation Index, SF ratio, and PF ratio) was more predictive than the initial value (Khemani ICM 2014). The second is that standardized ventilator settings may improve the predictive ability of PF ratio, because it accounts for lung recruitment and "cost" from the ventilator (others have also demonstrated this, specifically examining Oxygenation Index over PF ratio.. Flori 2005, Trachsel 2004). To specifically test which of these is most important, better justifying the requirement for a specific maneuver, the authors should

1. Examine the predictive ability of the PF ratios at 24 hours prior to changing to standard ventilator settings. It may be that the Berlin Definition applied at this time frame is as good as the author's classification schema. This would support that the standardized ventilator settings may be less important, which would facilitate consideration for clinical trials and care, because practitioners can simply monitor how PF ratio has changed over the first 24 hours. If this value is not predictive, it further support's the authors method.

2. Examine oxygenation index (OI= Mean Airway Pressure * FiO2 *100/PaO2) at time of ARDS diagnosis and 24 hours- as this may account for the variability in ventilator management both at ARDS onset and at 24 hours- which would eliminate the need for a specific maneuver. If mean airway pressure data is available, this would be an important analysis. If it is not available, the authors could compute an estimate using PEEP, Plateau Pressure, Rate, and inspiratory time. The new pediatric specific definition of ARDS has used OI instead of PF ratio, precisely because of the issues highlighted in this paper. As such, the authors have identified a crucial problem—but given the author's methods require a specific maneuver, it would be important to demonstrate this method adds information over using existing, available data at the same time frame, which will not require a specific maneuver.

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

	Some specific points to improve the clarity of the manuscript. 1. Comments in the introduction and discussion about a more homogenous disease process are a bit difficult to understand. Inherently, ARDS is a heterogeneous syndrome which is complicated by multiple potential variables like cause, co-morbidities, inflammatory response, etc. We are attempting to measure the severity of hypoxemia with PF ratio or other markers of oxygenation to identify a group more likely to benefit from specific therapies targeted at improving the hypoxemia or minimizing the VILI which comes with such hypoxemia. Homogeneity will be dependent on things other than hypoxemia—as these in fact may have a more relevant association with outcome than even hypoxemia (co-morbidities, age etc). So instead of using homogenous, it may be best to use terms like persistent hypoxemia which may benefit from different types of therapies. 2. Methods a. Please clarify how pulmonary versus non pulmonary cause of death was ascertained. It is unclear that this should be a main outcome- as the authors allude to with their discussion of the imprecision of berlin at identifying a pulmonary death. In truth, this assessment is subjective, and the hypoxemia and inflammation severity which is captured by PF ratio may contribute substantially to the final cause of death, even when it is not primarily deemed to be pulmonary. This outcome should be minimized. b. It would be worth reporting Areas Under the Curve of the ROC plot on the outcome of hospital mortality for initial Berlin PF, Berlin PF at 24 hours (before standard ventilator settings…see above), as well as PF based on the author's standard ventilator settings at 24 hours. The authors could use PF as a continuous variable, or use the classifications proposed by berlin for this analysis. This will confirm and quantify the potential improvement in risk stratification which can be seen with a 24 hour approach, and a 24 hour standardized approach. If OI is also available, it would be good to include these as well. 3. Results- a. in table 2 explicitly state extra-pulmonary organ dysfunction b. In table 3 add percentages to categorical variable such as gender and cause of ARDS c. Table 4- VFDs are generally not normally distributed- so please confirm distribution and report median if needed. d. Table 5- it is unclear if PEEP, LIS, PPlat is on standard vent settings, or prior to standard vent settings. Give PEEP is a component of LIS, and that was standardized, it would be best to include values prior to standard vent settings. Same holds for PEEP- its unclear how PEEP in the Non ARDS group can be less than 10 since requirement was for it to be changed to 10…. Unless these are values before the standardized settings were applied. Please be explicit as to what values were before settings were standardized. e. Figure 1- I think the percentages are a proportion of the total patients initially classified as moderate and severe respectively—please add this to the legend as it took me some time to figure that out. The discussion should be re-worked a bit to address some of the findings after the new analysis is performed, and specifically address the potential harm of requiring a delayed diagnosis.

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

REVIEWER	Carmen Sílvia Valente Barbas Respiatory ICU -University of São Paulo, Brazil
REVIEW RETURNED	19-Nov-2014

GENERAL COMMENTS	This is an interesting paper by Villar and colleagues that analyzed data from 478 consecutive patients included in two observational cohort studies with moderate/severe ARDS (Berlin definition). Their patients were ventilated using a low-tidal volume strategy with positive end-expiratory pressure (PEEP) ≥5 cmH2O. They examined physiologic and ventilator parameters in association with the PaO2/FiO2 at ARDS onset and 24 hours later under a standardized ventilator setting. At 24 hours, patients were reclassified as severe, moderate, mild (200<PaO2/FiO2≤300) ARDS, and non-ARDS (PaO2/FiO2>300). Outcomes of interest included group severity and hospital mortality.The authors observed that overall hospital mortality was 42.2%. At ARDS onset, 173 patients had a PaO2/FiO2≤100 but only 38.7% of them met severe ARDS criteria after 24 hours of usual care. Classification using baseline PaO2/FiO2 did not identify ARDS patients with distinct lung injury severity. When assessed under standardized ventilator settings, most patients (61.3%) were reclassified as moderate, mild, and non ARDS, while lung severity and hospital mortality changed markedly with every PaO2/FiO2 category (p=2.04E-12). They concluded that baseline PaO2/FiO2 failed to identify subgroups of patients with distinct lung injury severity. Assessment of oxygenation defect under standardized ventilator settings at 24 hours after ARDS onset represents a better method for optimizing risk stratification of ARDS patients. Specific comments: 1. Do the authors think that if they could have collected the first arterial blood gases (at ARDS onset) under standardized ventilator settings they would have the same results? Please comment ….. 2. Do the authors think that if they could have collected the 24 hours after ARDS onset arterial blood gases under non-standardized ventilator settings they would have the same results? Please comment…… 3. In the methods section the authors stated that approximately 24 hours after meeting moderate/severe ARDS criteria, oxygenation was assessed under the following standardized ventilator settings: VT=7 ml/kg PBW, PEEP=10 cmH2O and FiO2=0.5. When patients required PEEP>10 or FiO2>0.5 and could not tolerate a decrease in PEEP or FiO2 for maintaining the oxygenation target, a set of rules for setting PEEP and FiO2 were applied only during the standardized setting assessment (Table 1). The authors did not mention respiratory rate nor the PaCO2 of their patients. Dead space is an important predictor of ARDS severity and mortality…Why didn`t the authors analyze minute ventilation and PaCO2 levels…please clarify…… 4. What is the authors explanation to the fact that some ARDS patients improve after intubation and others deteriorate…… distinct spectrum of the syndrome……. right antibiotic and treatment for the infection cause of ARDS…… Genetics? Some groups showed that patients had a genetic predisposition for ARDS and for its mortality ratio…(Bajwa EK, Cremer PC, Gong MN, Zhai R, Su L, Thompson BT, Christiani DC. An NFKB1 promoter insertion/deletion polymorphism influences risk and outcome in acute respiratory distress syndrome among Caucasians.PLoS One. 2011 May 9;6(5):e19469.).please comment ……

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

VERSION 1 – AUTHOR RESPONSE

Reviewer #1:

General Comments
Dr. Villar and colleagues prevent a very interesting analysis of existing observational data regarding standardized assessment of lung injury severity for adults with ARDS. They obtained PF ratios 24 hours after ARDS diagnosis (AECC and Berlin), on standardized ventilator settings and conclude that using these PF values while on standardized ventilator settings better stratifies risk of mortality, pulmonary mortality, ventilator free days, and non-pulmonary organ failure compared to the initial PF ratio based classification using Berlin criteria at ARDS onset. This paper highlights some of the unanswered and difficult questions that exist with ARDS diagnosis and risk stratification, in particular evolution of disease process, and variability in ventilator management as a confounder for the assessment of disease severity. It is an incredibly important area to clarify, particularly when considering risk benefit profiles for higher risk therapies and for clinical trials. In general, the paper is well written and methodology is reasonably clear.
The strength of the suggested standardized ventilator approach is of course minimizing the elements of variability in care, in particular with application of PEEP and FiO2. The limitations surround the requirement for practitioners at the bedside to change their behavior and perform such a maneuver, which is unlikely to be done as part of routine clinical care. Certainly such a maneuver could be done for the purposes of a study, or perhaps before contemplating a higher risk therapy, such as prone positioning or high frequency oscillatory ventilation. As such, when approaching defining an already under-recognized clinical syndrome such as ARDS, one has to find the balance between attempting to be inclusive and simple, to ensure that all patients with the disease are recognized, and requiring a specific maneuver to specifically risk stratify patients, which may ultimately contribute to the under-recognition of the disease if practitioners do not perform the maneuvers. This is particularly important in the face of growing data to suggest that using automated early detection systems to alert providers of patients who have ARDS, or who are at risk of developing ARDS may improve outcome. The authors allude to some of these points in the discussion, but it would benefit from being explicit about the potential down-side of such an approach with regards to recognition of the disease, in particular for application of relatively low-risk therapies (i.e. lung protective VT management, fluid management etc…). If one cannot be labeled as having ARDS until 24 hours into the course of disease, there is potential for practitioners to pay less attention to these relatively low-risk but beneficial therapies-potentially worsening outcomes for these patients.

R- We thank the reviewer for the positive comments about our study. We agree with the reviewer that we need a better definition for screening ARDS patients for determining epidemiological data, for allocation of health care resources, for guiding appropriate medical therapy, and for enrolling patients into clinical trials. However, we are not recommending that one does not manage suspect patients with lung protective mechanical ventilation from the onset. Many patients without ARDS who present with severe hypoxemia and reduced lung volumes should be managed with lung protective mechanical ventilation from the time of intubation. What we are saying is that all hypoxemia is not ARDS, and that inflammation from ARDS is not reversed in 24 hrs, and if a patient's oxygenation status changes in 24 hrs they should not be classified as ARDS nor should they be enrolled and randomized into a clinical trial focusing on treatment of ARDS.
We have included some of these comments in the discussion of the revised manuscript.

To justify such a maneuver, it would be important for the authors to demonstrate that this approach is necessary, and cannot be accomplished using existing data, without a specific maneuver.
Fundamentally, there are two important potential explanations for the author's findings, which are

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

worth trying to tease out specifically to strengthen their argument. One finding is that patients change over the first 24 hours of the disease, either in response (or lack of response) to conventional therapies, or as part of resolution of disease. Therefore, the 24 hour value is in many ways, measuring response to therapies, and may be better predictive of outcome. Others have found similar results (we found similar findings in pediatrics) where the average value over the first 24 hours after ALI diagnosis for all oxygenation parameters (Oxygenation index, Oxygen Saturation Index, SF ratio, and PF ratio) was more predictive than the initial value (Khemani ICM 2014). The second is that standardized ventilator settings may improve the predictive ability of PF ratio, because it accounts for lung recruitment and "cost" from the ventilator (others have also demonstrated this, specifically examining Oxygenation Index over PF ratio (Flori 2005, Trachsel 2004).
To specifically test which of these is most important, better justifying the requirement for a specific maneuver, the authors should:

Q1) Examine the predictive ability of the PF ratios at 24 hours prior to changing to standard ventilator settings. It may be that the Berlin Definition applied at this time frame is as good as the author's classification schema. This would support that the standardized ventilator settings may be less important, which would facilitate consideration for clinical trials and care, because practitioners can simply monitor how PF ratio has changed over the first 24 hours. If this value is not predictive, it further support's the authors method.

R1- Thank you for this observation. As the reviewer has pointed out, lung pathology in ARDS can evolve within 24 hours of usual critical care. Since the criteria for assessing hypoxemia has not been standardized in any of the current definitions for ARDS, patients can be enrolled in clinical trials with the same range of PaO2/FiO2 but calculated under different levels of ventilator support. The main purpose of our study was to assess the degree of hypoxemia under standardized VT, FiO2 and PEEP settings at 24 h after ARDS diagnosis. The major contribution of our study is that the assessment of hypoxemia under a standardized ventilator setting after 24 h of usual care identified four subgroups of patients with a distinct degree of lung severity and hospital outcome.
As the reviewer is aware, changes in FiO2 and PEEP will alter the measurement of PaO2/FiO2. Depending on the clinician's selection of PEEP, a patient will be moved from one range of PaO2/FiO2 severity to another within the first 24 h of usual care. In our study, PaO2 values at baseline (ARDS diagnosis) were measured under a wide range of individualized FiO2 (0.4 to 1) and PEEP (5 to 20 cmH2O) that were chosen by physicians for each individual patient, as it happens in current adult ICUs. As reported in our manuscript, a high proportion of patients had a transient PaO2/FiO2 <100 or <200 mmHg when categorized at the time of ARDS diagnosis since after 24 h of usual care, the PaO2/FiO2 was >100 or >200 under a standardized ventilator assessment with a VT 7 ml/kg PBW, PEEP 10 and FiO2 0.5.
Although the reviewer recognizes that it is plausible that our classification/stratification at 24 h could not vary if patients were not assessed under standardized ventilator settings, it is equally important to emphasize that the PaO2 values recorded by physicians at 24 h were also measured under a wide range of individualized FiO2 (0.3 to 1) and PEEP [0 (only temporarily in 2 patients with bronchopleural fistula) to 22 cmH2O]. When analyzing patients at 24 hrs with the clinician selected PEEP and FIO2, the bulk of patients with a PaO2/FiO2≤100 had a higher mortality rate than those with a 100
We have added the results of the 24 hr assessment at clinician selected PEEP and FiO2 to the Results section of the paper: "PaO2 values recorded by physicians at 24 h were measured under a wide range of FiO2 (0.3 to 1) and PEEP [0 (only temporarily in 2 patients with bronchopleural fistula) to 22 cmH2O]. Based on those values, the bulk of patients with a PaO2/FiO2≤100 had a hospital mortality of 53.1% whereas patients with a 100

Q2) Examine oxygenation index (OI= Mean Airway Pressure * FiO2 *100/PaO2) at time of ARDS diagnosis and 24 hours- as this may account for the variability in ventilator management both at

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

ARDS onset and at 24 hours- which would eliminate the need for a specific maneuver. If mean airway pressure data is available, this would be an important analysis. If it is not available, the authors could compute an estimate using PEEP, Plateau Pressure, Rate, and inspiratory time. The new pediatric specific definition of ARDS has used OI instead of PF ratio, precisely because of the issues highlighted in this paper. As such, the authors have identified a crucial problem—but given the author's methods require a specific maneuver, it would be important to demonstrate this method adds information over using existing, available data at the same time frame, which will not require a specific maneuver.

R2- We did not measure the mean airway pressure because our goal was not to compare the correlation between OI and PaO2/FiO2 in defining/stratifying ARDS in our adult population. The AECC and the Berlin definitions do not consider OI for defining and stratifying ARDS, neither contemplated the use of any combination of values or thresholds for compliance, plateau pressure, ventilator rate or inspiratory time. OI is also not a standard measurement used in the management of adult patients as it is in pediatrics
As we stated in the manuscript, and in the new title of the manuscript, our goal was to investigate whether the value of PaO2 measured under a set of standardized ventilator settings would be better for stratifying and identifying subsets of adult ARDS patients with different degrees of lung injury that were associated with a different prognosis, and thus, for being at the bedside as a guide for early prediction of outcome and for identifying those subgroups of ARDS patients that could benefit the most from the appropriate ventilator or adjunctive techniques for improving gas exchange and outcome.
We appreciate the interest of the reviewer for further examining different ways to identify subgroups of patients with uniform severity of lung injury. However, the analysis proposed by the reviewer is beyond the scope of this study and requires further investigation in a separate study.

Specific points to improve the clarity of the manuscript.

Q3) Comments in the introduction and discussion about a more homogenous disease process are a bit difficult to understand. Inherently, ARDS is a heterogeneous syndrome which is complicated by multiple potential variables like cause, co-morbidities, inflammatory response, etc. We are attempting to measure the severity of hypoxemia with PF ratio or other markers of oxygenation to identify a group more likely to benefit from specific therapies targeted at improving the hypoxemia or minimizing the VILI which comes with such hypoxemia. Homogeneity will be dependent on things other than hypoxemia—as these in fact may have a more relevant association with outcome than even hypoxemia (co-morbidities, age etc). So instead of using homogenous, it may be best to use terms like persistent hypoxemia which may benefit from different types of therapies.

R3- Thank you. As the reviewer has stated, ARDS is a non-specific, heterogeneous syndrome that develops in the context of pulmonary and systemic disease processes. Independently of what causes acute lung injury, patients suffer from the same unique syndrome and are treated similarly. There is no data in the literature that link a particular PaO2/FiO2 ratio to predictable structural changes in the alveolar-capillary membrane, most likely because ARDS represents a common pathway of diverse events and disease entities.
Following the reviewer's recommendation, we have replaced the term "homogeneous" in several paragraphs of the revised manuscript with established hypoxemia.

Q4) Methods
a) Please clarify how pulmonary versus non pulmonary cause of death was ascertained. It is unclear that this should be a main outcome- as the authors allude to with their discussion of the imprecision of Berlin at identifying a pulmonary death. In truth, this assessment is subjective, and the hypoxemia and

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

inflammation severity which is captured by PF ratio may contribute substantially to the final cause of death, even when it is not primarily deemed to be pulmonary. This outcome should be minimized.

R4a- In our data collection form, physicians were asked to mark simply what they considered was the cause of death, pulmonary or non-pulmonary. We agree with the reviewer that this assessment could be subjective in many instances. Since we do not have data to confirm conclusively the degree of certainty or precision in assessing the cause of death, we have eliminated this information in the revised manuscript.

b) It would be worth reporting Areas Under the Curve of the ROC plot on the outcome of hospital mortality for initial Berlin PF, Berlin PF at 24 hours (before standard ventilator settings…see above), as well as PF based on the author's standard ventilator settings at 24 hours. The authors could use PF as a continuous variable, or use the classifications proposed by Berlin for this analysis. This will confirm and quantify the potential improvement in risk stratification which can be seen with a 24 hour approach, and a 24 hour standardized approach. If OI is also available, it would be good to include these as well.

R4b- We thank the reviewer for this suggestion.
Following the reviewer recommendation, we determined the predictive hospital mortality receiver operating characteristic (ROC) curve. The area under the ROC curve for stratification at T0 was 0.583 (95% CI: 0.525-0.636), at T24 without standardization was 0.605 (95%CI: 0.552-0.658), and at T24 under standardized ventilator settings was 0.693 (95%CI: 0.645-0.742) ($p<0.00001$).
We have reported these new analysis and data in the Methods and Results sections of the revised manuscript, as well as a new Figure 2 with the ROC curves.

Q5) Results
a) In table 2 explicitly state extra-pulmonary organ dysfunction

R5a- We have stated in the Methods section that physicians were asked to record at baseline, day 1, and subsequent days the total number of extrapulmonary organ failures included in the SOFA scale. For the purpose of our study, documentation of the specific type of organ failure was beyond the scope of our study design.

b) In table 3 add percentages to categorical variable such as gender and cause of ARDS

R5b- Done. Thank you.

c) Table 4- VFDs are generally not normally distributed- so please confirm distribution and report median if needed.

R5c- Thanks. We now report median and IQR in the revised manuscript.

d) Table 5- it is unclear if PEEP, LIS, PPlat is on standard vent settings, or prior to standard vent settings. Give PEEP is a component of LIS, and that was standardized, it would be best to include values prior to standard vent settings. Same holds for PEEP- its unclear how PEEP in the Non ARDS group can be less than 10 since requirement was for it to be changed to 10…. Unless these are values before the standardized settings were applied. Please be explicit as to what values were before settings were standardized.

R5d- Thank you. The reviewer is correct. The values of LIS, PEEP, VT, FiO2, and plateau pressure in

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

Table 5 represent the mean values before assessing the patient under standardized ventilator setting. Mean PaO2/FiO2 values are calculated during standardized assessment. We have now clarified this issue in Table 5 of the revised manuscript.

e) Figure 1- I think the percentages are a proportion of the total patients initially classified as moderate and severe respectively—please add this to the legend as it took me some time to figure that out.

R5e- Thank you. Done.

Q6) The discussion should be re-worked a bit to address some of the findings after the new analysis is performed, and specifically address the potential harm of requiring a delayed diagnosis.

R6- Thank you. Sections of the discussion have been rewritten addressing the findings after new analysis, as recommended by the reviewer.

Reviewer #2:

General comments:
This is an interesting paper by Villar and colleagues that analyzed data from 478 consecutive patients included in two observational cohort studies with moderate/severe ARDS (Berlin definition). Their patients were ventilated using a low-tidal volume strategy with positive end-expiratory pressure (PEEP) ≥5 cmH2O. They examined physiologic and ventilator parameters in association with the PaO2/FiO2 at ARDS onset and 24 hours later under a standardized ventilator setting. At 24 hours, patients were reclassified as severe, moderate, mild (200300). Outcomes of interest included group severity and hospital mortality. The authors observed that overall hospital mortality was 42.2%. At ARDS onset, 173 patients had a PaO2/FiO2≤100 but only 38.7% of them met severe ARDS criteria after 24 hours of usual care. Classification using baseline PaO2/FiO2 did not identify ARDS patients with distinct lung injury severity. When assessed under standardized ventilator settings, most patients (61.3%) were reclassified as moderate, mild, and non ARDS, while lung severity and hospital mortality changed markedly with every PaO2/FiO2 category (p=2.04E-12). They concluded that baseline PaO2/FiO2 failed to identify subgroups of patients with distinct lung injury severity. Assessment of oxygenation defect under standardized ventilator settings at 24 hours after ARDS onset represents a better method for optimizing risk stratification of ARDS patients.

Specific comments:
Q1) Do the authors think that if they could have collected the first arterial blood gases (at ARDS onset) under standardized ventilator settings they would have the same results? Please comment.

R1. Thank you for this observation.
In a previous publication (Villar et al, Am J Respir Crit Care Med 2007, 176:795-804), we studied 170 ARDS patients and assessed arterial blood gases at ARDS onset under the following standardized ventilator settings: VT 7 ml/kg PBW, I:E <1:1, ventilator rate to maintain PaCO2 35 to 50 mmHg, plus the following PEEP and FiO2 settings applied in the following order: (i) PEEP≥5 cmH2O and FiO2 ≥0.5, (ii) PEEP≥5 cmH2O and FiO2=1.0, (iii) PEEP≥10 cmH2O and FiO2≥0.5, and (iv) PEEP≥10 and FiO2=1.0.
Regardless of the standardized ventilatory setting, we found that some patients who originally were classified as having a PaO2/FiO2<200 did not continue to meet the AECC definition. Under some of those standardized ventilator settings, some of these patients changed to a PaO2/FiO2 between 201

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

and 300 mmHg and others to a PaO2/FiO2>300 mmHg. However, none of the settings that categorized patients into severe/moderate ARDS, mild ARDS, and non-ARDS was associated with mortality. We can speculate that two major reasons may explain our findings. First, it is well accepted that in a high proportion of patients, lung function improves dramatically within the initial 12-24 h of usual care. Second, it is likely that within the initial 24 h, patients are not stable when routine therapies (i.e. fluid resuscitation, sedation, muscle paralysis, antibiotics, insulin, catecholamine, blood transfusion, body positioning, intravascular catheterization, repetitive and aggressive suctioning or secretions, insertion of chest tubes, etc.) are implemented and during the process of selecting the most appropriate ventilatory support parameters (FiO2, PEEP, pressure, VT, respiratory rate, etc). These findings support that risk stratification of ARDS based on baseline PaO2/FiO2, with or without standardization, is not clinically useful.

Q2) Do the authors think that if they could have collected the 24 hours after ARDS onset arterial blood gases under non-standardized ventilator settings they would have the same results? Please comment.

R2- As replied to Reviewer #1 to Q1 and Q4b, the categorization of patients according to the PaO2/FiO2 value at 24 h without standardization was incapable of identifying subgroups of patients with a distinct set of characteristics of lung injury severity and overall mortality. In the revised manuscript we now provide ROC curves comparing the three approaches: categorization at baseline, categorization at 24 h, and categorization at 24 h on standardized ventilatory settings. The data confirmed the superiority of our stratification approach.

Q3) In the methods section the authors stated that approximately 24 hours after meeting moderate/severe ARDS criteria, oxygenation was assessed under the following standardized ventilator settings: VT=7 ml/kg PBW, PEEP=10 cmH2O and FiO2=0.5. When patients required PEEP>10 or FiO2>0.5 and could not tolerate a decrease in PEEP or FiO2 for maintaining the oxygenation target, a set of rules for setting PEEP and FiO2 were applied only during the standardized setting assessment (Table 1). The authors did not mention respiratory rate nor the PaCO2 of their patients. Dead space is an important predictor of ARDS severity and mortality…Why didn`t the authors analyze minute ventilation and PaCO2 levels… Please clarify.

R3- Thank you for this concern.
We agree with the reviewer that many studies have reported that dead space is an important predictor of ARDS severity and outcome. Minute ventilation is a surrogate for dead space. However, most of those studies were performed in the years prior to the implementation of lung protective ventilation. As stated in the manuscript, patients in our study were ventilated with a mean VT of 7 ml/kg at a ventilatory rate to maintain PaCO2 at 35-50 mmHg. Under those circumstances, in our series of 478 patients there were no statistically significant differences in the baseline mean values of PaCO2 between survivors and non-survivors (44±10 vs. 45±12 mmHg). At 24 h, mean PaCO2 values ranged between 44 mmHg in the non-ARDS and mild ARDS patients to 47 mmHg in the severe ARDS among groups (p=0.88). Regarding the application of ≥10 L/min of minute ventilation within the first 24 h, there were no statistical differences in the percentage of patients among subgroups, ranging from about 38% in the non-ARDS and mild ARDS patients to 46% in the moderate and severe ARDS. Since our study was not designed for identifying or validating predictors of mortality, other than the value of standardized PaO2/FiO2 ratio, we do not think that including data on minute ventilation and mean PaCO2 values provides any additional, relevant information to our study.

Q4) What is the authors explanation to the fact that some ARDS patients improve after intubation and others deteriorate…… distinct spectrum of the syndrome……. right antibiotic and treatment for the

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

infection cause of ARDS…… Genetics? Some groups showed that patients had a genetic predisposition for ARDS and for its mortality ratio… (Bajwa EK, Cremer PC, Gong MN, Zhai R, Su L, Thompson BT, Christiani DC. An NFKB1 promoter insertion/deletion polymorphism influences risk and outcome in acute respiratory distress syndrome among Caucasians. PLoS One 2011 May 9;6(5):e19469.). Please comment.

R4- Our study design precludes us from being able to explain all of these important concerns. Certainly, predisposing genetic factors can interact with the environment to determine the diversity of clinical manifestations, the response to treatment, and outcome among ARDS patients (Villar et al, Crit Care 2004, 8:180-9; Rahim et al, Genome Biol 2008, 9:215; Acosta-Herrera et al, Front Genet 2014, 5:20).
What we do believe is that all that appears to be ARDS at first sight is NOT ARDS. Many patients present with severe hypoxemia and decreased lung volume but when assessed at 24 hrs on standardized ventilator settings they do not meet ARDS criteria. We suspect that the cause of the hypoxemia was not inflammation since we would not expect generalized lung inflammation to resolve in a 24 hr period. However, these issues are beyond the scope of our study.

VERSION 2 – REVIEW

REVIEWER	Robinder G. Khemani MD, MsCI Children's Hospital Los Angeles, United States of America
REVIEW RETURNED	27-Jan-2015

GENERAL COMMENTS	In general the revision provided by Villar and colleges is a significant improvement. The additional analysis strengthens their argument about the need for standardization of ventilator management to assess disease severity, rather than simply measuring the value at 24 hours. This point, in and of itself, has important implications for the diagnosis and management of ARDS in adults and children, and I would highly recommend the authors make this the focus, and explicitly state this throughout the manuscript. While I understand this was in response to a request from the reviewer, it only strengthens your argument. They could do this by a priori specifying three groups for analysis of the outcome of mortality – Berlin critiera applied to (1) initial PF, (2) 24 hour Pf under routine clinical care and (3) 24 hour Pf under standard ventiatlor settings (you already do this but be explicit about it throughout). Adding this explicitly to the introduction, methods, presentation of the results, and the discussion would make their paper clearer, and allow for the reader to understand that the standard ventilator settings are the most crucial piece for the risk stratification. To that end, in the intro or discussion it is worth highlighting that there is a lot of variability in practice even at 24 hours- so that we must have some way to account for this variability (and your proposal is to standardize the ventilator settings). To that end, the conclusions could be consistently rewritten to state (throughout the paper when baseline is used) that baseline as well as 24 hour PF ratios under usual clinical practice were inferior to PF ratio using a standard ventilator management approach at 24 hours at stratifying mortality risk. Then it is clear to the reader who even glances at the paper that you need the ventilator standardization, and not just the 24 hour value. Specific comments Methods- make it clear that the continuous variable of PF ratio was used for the ROC- not the berlin groupings. That is clear from your figure when examining all the points- but will not be clear to the

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

	naïve reader. A figure or table which clearly illustrates the mortality rates using the 3 classification schema will be important and can very clearly articulate the point. It graphically illustrates the concept of discrimination to the clinical reader which you get in an ROC plot (but with the groups rather than the PF ratio as a continuous variable) Unclear that the Chi squared test for homogeneity adds much to the argument- really you get that if you talk about discrimination ability using the Berlin breakpoints or with the figure suggested above. However, if the authors wish to keep it, please add it to the statistical methods. Results- page 9 line 37 instead of had PF ratio should have having a PF ratio When discussing "progression" of ARDS from initial to 24 hour values, again it is important to distinguish whether you are talking about 24 hours without standard ventilator settings versus 24 hours on standard ventilator settings. It is difficult to talk about "progression" when one process does not mandate a specific therapy to risk stratify (baseline) while another (24 hours under standard) does. So, really it may be that they never had that lung disease severity to begin with- in which case its not a progression or a resolution, but a more precise estimation of their lung injury (as you claim). I would make this explicit. It is then important to consider whether this has to be done at 24 hours, or whether it can be done within the first few hours- to establish what "baseline" really is under standard ventilator settings. Why did you use the phase "bulk of patients" – this is unclear to me… Tables 3 and 4- I am struggling to understand why baseline PF ratio is no different between survivors and non survivors in table 3, but the group with severe ARDS, stratified by PF ratio has significantly higher mortality than those with moderate ARDS. The standard deviation for PF ratio is not very large to say that there was a lot of variability in the PF values. Please just double check these calculations—it implies in some ways that within the severity strata survivors and non survivors have no real difference in PF ratio- which may mean the mortality risk is coming from something other than the PF ratio. This may be true- Discussion- I think the discussion could be significantly improved if the authors are explicit about separating timing of observation and practice variation. The 24 hour non standardized approach is equally bad- so it is not so much the timing of 24 hours that is crucial. Instead it is the standard ventilator settings. This is because practice variability introduces bias in the assessment of lung disease severity which you seek to control through a standard ventilator approach. So, why does it have to be at 24 hours? Why couldn't it be at 12? Or at ARDS onset but under standard ventilator settings? I understand you didn't study that here, but your paper clearly supports that it is the standardization of ventilator settings that is crucial- not really the timing. So make that explicitly clear. This will strengthen the impact of the paper.

REVIEWER	Carmen Silvia Valente Barbas Respiratory ICU, University of São paulo Medical Scholl, São paulo, Brazil
REVIEW RETURNED	31-Jan-2015

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

GENERAL COMMENTS	The authors adequately answered the reviewers` queries and added the needed information to the revised version of the manuscript.

VERSION 2 – AUTHOR RESPONSE

Reviewer #1

General Comments

In general the revision provided by Villar and colleges is a significant improvement. The additional analysis strengthens their argument about the need for standardization of ventilator management to assess disease severity, rather than simply measuring the value at 24 hours. This point, in and of itself, has important implications for the diagnosis and management of ARDS in adults and children, and I would highly recommend the authors make this the focus, and explicitly state this throughout the manuscript.
While I understand this was in response to a request from the reviewer, it only strengthens your argument. They could do this by a priori specifying three groups for analysis of the outcome of mortality – Berlin criteria applied to (1) initial PF, (2) 24 hour PF under routine clinical care and (3) 24 hour PF under standard ventilator settings (you already do this but be explicit about it throughout). Adding this explicitly to the introduction, methods, presentation of the results, and the discussion would make their paper clearer, and allow for the reader to understand that the standard ventilator settings are the most crucial piece for the risk stratification. To that end, in the introduction or discussion it is worth highlighting that there is a lot of variability in practice even at 24 hours- so that we must have some way to account for this variability (and your proposal is to standardize the ventilator settings). To that end, the conclusions could be consistently rewritten to state (throughout the paper when baseline is used) that baseline as well as 24 hour PF ratios under usual clinical practice were inferior to PF ratio using a standard ventilator management approach at 24 hours at stratifying mortality risk. Then it is clear to the reader who even glances at the paper that you need the ventilator standardization, and not just the 24 hour value.

R- We thank the reviewer for the positive comments about our revised manuscript and for the new recommendations to strengthen our findings. We have modified the manuscript accordingly throughout the different sections of the text.

Specific comments

Q1) Methods- make it clear that the continuous variable of PF ratio was used for the ROC- not the Berlin groupings. That is clear from your figure when examining all the points- but will not be clear to the naïve reader.

R1- Thank you. Done.

Q2) A figure or table which clearly illustrates the mortality rates using the 3 classification schema will be important and can very clearly articulate the point. It graphically illustrates the concept of discrimination to the clinical reader which you get in an ROC plot (but with the groups rather than the PF ratio as a continuous variable).

R2- Thank you. In the revised manuscript we have added a new Table 6 with that information.

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

Q3) Unclear that the Chi squared test for homogeneity adds much to the argument- really you get that if you talk about discrimination ability using the Berlin breakpoints or with the figure suggested above. However, if the authors wish to keep it, please add it to the statistical methods.

R3- Thank you. We agree with the reviewer. We have deleted that sentence in the revised manuscript.

Q4) Results- page 9 line 37 instead of had PF ratio should have having a PF ratio

R4- Thanks. Done.

Q5) When discussing “progression” of ARDS from initial to 24 hour values, again it is important to distinguish whether you are talking about 24 hours without standard ventilator settings versus 24 hours on standard ventilator settings. It is difficult to talk about “progression” when one process does not mandate a specific therapy to risk stratify (baseline) while another (24 hours under standard) does. So, really it may be that they never had that lung disease severity to begin with- in which case it is not a progression or a resolution, but a more precise estimation of their lung injury (as you claim). I would make this explicit. It is then important to consider whether this has to be done at 24 hours, or whether it can be done within the first few hours- to establish what “baseline” really is under standard ventilator settings.

R5- We agree with the Reviewer. We have clarified those issues in the revised manuscript.

Q6) Why did you use the phrase “bulk of patients” – this is unclear to me…

R6- We have deleted the term “bulk” in the revised manuscript.

Q7) Tables 3 and 4- I am struggling to understand why baseline PF ratio is no different between survivors and non survivors in table 3, but the group with severe ARDS, stratified by PF ratio has significantly higher mortality than those with moderate ARDS. The standard deviation for PF ratio is not very large to say that there was a lot of variability in the PF values. Please just double check these calculations—it implies in some ways that within the severity strata survivors and non survivors have no real difference in PF ratio- which may mean the mortality risk is coming from something other than the PF ratio. This may be true-

R7- We have checked all calculations from Table 3 and all of them are correct. The mean value of PaO2/FiO2 is higher in survivors (120±39 mmHg) than in non-survivors (114±39 mmHg) although the difference is not statistically significant (p=0.0973).
It is important to emphasize that those mean values refer to the baseline PaO2/FiO2 ratios. As in our study, Caser at el (Crit Care Med 2014, 42:574-82) found no significant differences when comparing the mean PaO2/FiO2 ratio between survivors and non-survivors at baseline, and during the first 5 days of ARDS, probably because of its wide variability (see Figure 2, Caser et al paper, Ref #31 in our manuscript). As reported in our manuscript, baseline PaO2/FiO2 ratio is useless for predicting severity and outcome.

Q8) Discussion- I think the discussion could be significantly improved if the authors are explicit about separating timing of observation and practice variation. The 24 hour non-standardized approach is

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.

equally bad- so it is not so much the timing of 24 hours that is crucial. Instead it is the standard ventilator settings. This is because practice variability introduces bias in the assessment of lung disease severity which you seek to control through a standard ventilator approach. So, why does it have to be at 24 hours? Why couldn't it be at 12? Or at ARDS onset but under standard ventilator settings? I understand you didn't study that here, but your paper clearly supports that it is the standardization of ventilator settings that is crucial- not really the timing. So make that explicitly clear. This will strengthen the impact of the paper.

R8- Thank you. We agree with the Reviewer. But we have to make it clear that it is both 24 h and the use on standardized ventilator settings that are important. In our previous paper we demonstrated the use of standardized ventilator settings at baseline does not predict outcome, whereas the same settings at 24 h does predict outcome (Villar et al, Am J Respir Crit Care Med 2007, 176:795-804). We have clarified those issues in the revised manuscript.

Reviewer #2

The authors adequately answered the reviewers' queries and added the needed information to the revised version of the manuscript.
R- We thank the reviewer for the approval of the revised manuscript.

BMJ Open: first published as 10.1136/bmjopen-2014-006812 on 27 March 2015. Downloaded from http://bmjopen.bmj.com/ on December 4, 2021 by guest. Protected by copyright.