BJA

British Journal of Anaesthesia, xxx (xxx): xxx (xxxx)

doi: 10.1016/j.bja.2019.02.026
Advance Access Publication Date: xxx
Clinical Investigation

CLINICAL INVESTIGATION

Heterogeneous effects of alveolar recruitment in acute respiratory distress syndrome: a machine learning reanalysis of the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial

Fernando G. Zampieri[1,2,*], Eduardo L. Costa[3], Theodore J. Iwashyna[4], Carlos R. R. Carvalho[3], Lucas P. Damiani[1], Leandro U. Taniguchi[5,6], Marcelo B. P. Amato[3], Alexandre B. Cavalcanti[1], for the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial Investigators

[1]Research Institute, HCor—Hospital do Coração, São Paulo, Brazil, [2]Research Division, Hospital Alemão Oswaldo Cruz, São Paulo, Brazil, [3]Respiratory Intensive Care Unit, Hospital das Clínicas, University of São Paulo School of Medicine, São Paulo, Brazil, [4]University of Michigan and Veterans Affairs Center for Clinical Management Research, Ann Arbor, MI, USA, [5]Intensive Care Unit, Emergency Medicine Discipline, University of São Paulo, São Paulo, Brazil and [6]Research Institute, Hospital Sírio Libanês, São Paulo, Brazil

*Corresponding author. E-mail: fgzampieri@gmail.com

Background: Despite a robust physiological rationale, recruitment manoeuvres with PEEP titration were associated with harm in the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial (ART). We sought to investigate the potential heterogeneity in treatment effects in patients enrolled in the ART, using a machine learning approach.
Methods: The primary outcome was hospital mortality. Patients were clustered using baseline clinical and physiological data using the k-means for mixed large data method. The heterogeneity in treatment effect between clusters was investigated using Bayesian methods. We further investigated whether baseline driving pressure could modulate the association between treatment arm, cluster, and mortality.
Results: Data from all 1010 patients enrolled in the ART were analysed. Partitioning suggested that three clusters were present in the ART population. The largest cluster (Cluster 1) was characterised by patients with pneumonia and requiring vasopressor support. Recruitment manoeuvres with PEEP titration were associated with higher mortality in Cluster 1 (probability of harm of >98%), but this association was absent in Clusters 2 and 3 (probability of harm of 45% and 68%, respectively). Higher baseline driving pressure was associated with a progressive reduction in the association between alveolar recruitment with PEEP titration and mortality.
Conclusions: Recruitment manoeuvre with PEEP titration may be harmful in acute respiratory distress syndrome (ARDS) patients with pneumonia or requiring vasopressor support. Driving pressure appears to modulate the association between the ART study intervention, aetiology of ARDS, and mortality. This machine learning approach may help tailor future RCTs.

Editorial decision date: 27 February 2019; **Accepted:** 27 February 2019

Clinical trial registration: NCT01374022.

Keywords: acute respiratory distress syndrome; Bayesian statistics; heterogeneity in treatment effect; pneumonia; pulmonary complications; recruitment manoeuvres

Editor's key points

- The Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial (ART) failed to find benefit in the early use of alveolar recruitment manoeuvres in acute respiratory distress syndrome (ARDS), compared with the ARDS Network standard treatment.
- Conventional subgroup analyses failed to detect any specific subgroups where the ART intervention may be beneficial or detrimental.
- This *post hoc* analysis re-examined whether a subset of ARDS patients may benefit from the early use of alveolar recruitment manoeuvres by using an unsupervised machine learning approach.
- Machine learning revealed that recruitment manoeuvres with PEEP titration were associated with harm in patients with pneumonia and requiring vasopressor support.
- This precision medicine approach, which is widely applicable to perioperative and critical care medicine research, may help refine the design of future trials.

Acute respiratory distress syndrome (ARDS) is a severe complication of critical illness, being associated with important mortality and morbidity.[1] The cornerstone of ARDS management has been protective lung ventilation with low tidal volume; PEEP; and, in many cases, prone position and use of neuromuscular block.[1–4] Ventilation strategies that aimed at opening collapsed alveolar units were proposed to increase the residual functional capacity and also reduce the cyclic alveolar collapse and dynamic strain, thereby preventing further lung tissue injury.[5]

Despite promising preliminary data, a large multicentre RCT [Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial (ART)] failed to find a beneficial effect of the early use of alveolar recruitment manoeuvres with PEEP titration (ART treatment) in patients with moderate-to-severe ARDS when compared with the standard ARDS Network (ARDSNet) treatment based on a PEEP/FiO_2 table.[6] In this trial, recruitment manoeuvres guided by best lung compliance after PEEP titration were associated with higher mortality, despite a decrease in driving pressure and an improvement in oxygenation. Traditional subgroup analyses performed in the trial failed to detect any specific subgroups where the treatment effect could be different. However, traditional subgroup analyses are known to have several limitations, including lack of power, issues with multiple testing, and inability to account for possible within-group interactions.[7,8]

We therefore sought to explore the potential heterogeneity in treatment effect (HTE) in the ART using a Bayesian HTE framework. We assessed the presence of potential clusters by using machine learning algorithms considering pre-specified clinical and physiological variables, and subsequentially investigated HTE. As the ART treatment was intended to reduce the baseline driving pressure (an effect that could theoretically be more important at higher baseline driving-pressure values),[4,5] we also investigated whether there was an association between driving pressure measured at baseline, ventilatory strategy (ARDSNet or alveolar recruitment with PEEP titration—ART treatment) and mortality within the clusters.[9]

Methods

Patients

This is a secondary *post hoc* analysis of the ART (NCT01374022).[6] In brief, the ART randomised 1010 patients with moderate-to-severe ARDS to traditional low-tidal-volume ventilation with PEEP table (following the low-PEEP table of the ALVEOLI trial, which we refer to as the ARDSNet group[2]), or to a strategy based on recruitment manoeuvres followed by PEEP titration during the descending limb of titration procedure (ART group). PEEP was set 2 cm H_2O above the best compliance PEEP value. More details can be found in the previously published main paper.[6] All patients included in the ART were included in this secondary *post hoc* analysis.

Primary outcome

Our primary outcome was 28 days mortality for all analyses. We did not assess other secondary outcomes.

Analysis overview

We performed initial exploratory analyses assessing mortality according to treatment group (ARDSNet *vs* ART treatment) in quintiles of Simplified Acute Physiology Score (SAPS) 3, Po_2:FiO_2, and driving pressure. Our primary analysis investigated HTE by applying a Bayesian HTE approach on clusters.

Subgroup detection

We selected the following baseline variables for cluster analysis: SAPS 3 score, weight, PEEP, tidal volume (ml kg^{-1}), ventilatory frequency, driving pressure, Po_2, Pco_2, FiO_2, use of vasopressors (norepinephrine, adrenaline, or dopamine, but not inotropic agents, such as dobutamine) at randomisation, ARDS cause (pneumonia, other pulmonary sources, or non-pulmonary source), and presence of sepsis (other than respiratory source). We applied k-means for mixed large data (*kamila*[10]) approach to find clusters. The best number of clusters was defined by inspecting the prediction strength of clusters after 1000 cross-validations, as discussed by Tibshirani and Walther.[11] The advantages of *kamila* include capability of generating weights to categorical variables that may reduce their importance in clustering mixed data.[10] A visual display of the results of the clustering method using the Barnes–Hut

t-Distributed Stochastic Neighbor Embedding (tSNE) following Gower's distance clustering was also performed to confirm and visually display the results of *kamila*.[12]

Heterogeneity in treatment effect

The investigation of HTE using a Bayesian framework was performed with *beanz* R package.[13] Raw data from the ART were inputted to *beanz*. Subgroup treatment effect was calculated from raw data by the *beanz* package as the log risk ratio of mortality for the ART treatment *vs* ARDSNet groups. We applied a simple Bayesian regression model and sampled the posterior distribution using Markov Chain Monte Carlo simulations. In this approach, individual subgroup effects are shrunken towards a global mean effect whilst considering subgroup interactions in a regression procedure. We used non-informative priors for this analysis. Results are displayed through probability distribution of relative risks for global population and the clusters. The posterior distribution was approximated to a normal distribution for clarity on interpretation and display whenever appropriate.

Bayesian regression modelling

Given the long body of prior work on driving pressure and its central role in designing the ART,[9] we hypothesised that baseline driving pressure might modify the within-cluster association between the intervention and the outcome.[9] We assessed the association between ventilatory strategy, baseline driving pressure, cluster, illness severity (as measured by SAPS 3 score), and outcome in a Bayesian regression model using the *brms* R package.[14] We added a second-degree polynomial for driving pressure and mortality in the model because of the potential non-linear relationship between treatment group and driving pressure in preliminary univariate analyses. We allowed interactions between driving pressure and treatment group, between cluster and treatment group, and between cluster and driving pressure.

We used data from Amato and colleagues[9] as a prior for driving pressure and mortality in ARDS; data from the meta-analysis of Suzumura and colleagues[5] for the association between recruitment manoeuvres and outcome; and, finally, data from the ORCHESTRA study to obtain priors for association between SAPS 3 and mortality.[15] Priors were drawn from normal distributions. For driving pressure, the mean was set to 0.0484550 with a standard deviation of 0.0061232; for recruitment manoeuvres, the value was −0.2406429 and 0.01074838, respectively; and per points of SAPS 3 score, the values were 0.093883 and 0.0006377, respectively. Uninformative priors were used for interactions and polynomial components of the regression.

Bayesian regression was done using a Markov Chain Monte Carlo procedure with four chains, 2000 iterations for each chain. We inspected chain stability using trace plots. Results were shown as beta coefficients with 95% credible intervals and using marginal effect plots in the Supplementary material. For simplicity, we display odds ratio with 95% credible interval (i.e. the interval that has 95% chance of containing the value if all assumptions made for the model hold) for the association between the ART treatment and mortality at fixed values of driving pressure in each cluster.

Missing value policy

If less than 1% of the values were missing, we performed simple imputation by the median or applied the most common category. Whenever >1% of missing values occurred, we applied a multiple chain imputation technique using all other available data.[15,16] There was no missing information on outcomes.

Results

All 1010 patients were included in the analysis and were analysed by the intention-to-treat principle. There was no

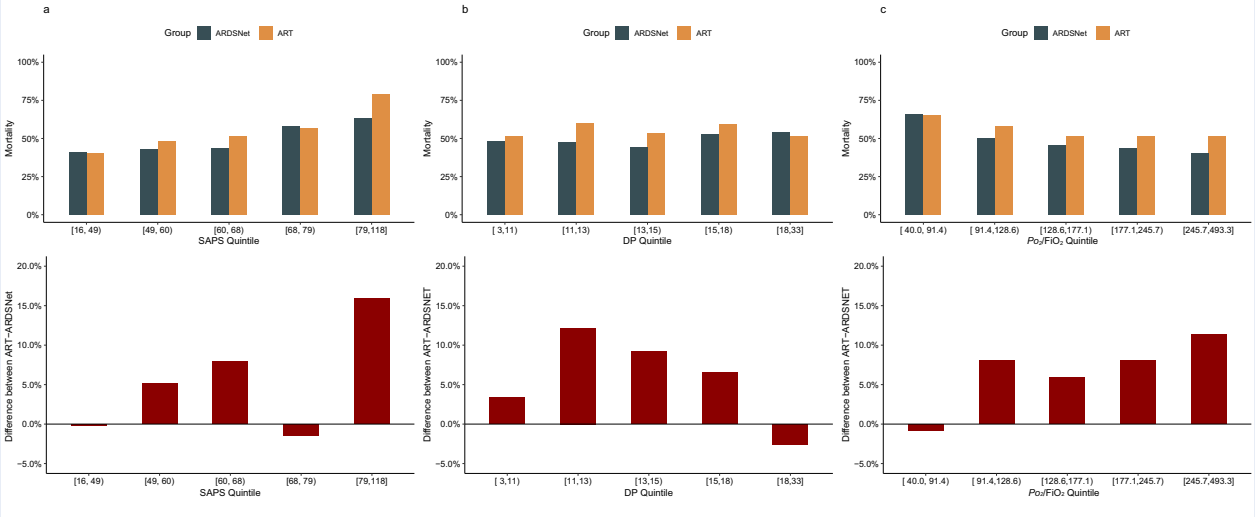

Fig 1. Top: Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial (ART) and ARDS Network (ARDSNet) mortality at quintiles of (a) Simplified Acute Physiology Score (SAPS), (b) driving pressure (DP), and (c) Po_2/FiO_2 ratio. Bottom: raw unadjusted difference in mortality between groups at every quintile. Note the non-linear relationship between baseline driving pressure and difference in mortality between groups.

clear trend of differential mortality by treatment group according to SAPS 3 quintile (Fig, 1a). Patients in the ART group had higher mortality in most quintiles of SAPS 3, Po_2/FiO_2, and driving pressure. Mortality was lower for the ART patients in the higher quintile of driving pressure (Fig. 1b) and in the worst quintile of Po_2/FiO_2 ratio (Fig. 1c). For driving pressure, the difference in mortality between the ART and ARDSNet groups decayed as driving pressure increased, but lower mortality in the ART group was only found in the highest quintile of driving pressure.

Data-based clustering analysis

Data-based clustering suggested that three clusters were appropriate in this population (prediction strength shown in Figure 1 and tSNE plot from Gower's distance shown in Supplementary Fig. S2). A description of patients on each cluster is provided in Table 1. Cluster 1 was composed of patients with ARDS attributable to pneumonia and that were using vasopressors at randomisation. Cluster 2, the smaller cluster, was composed of patients with miscellaneous ARDS causes (including pneumonia), but without the use of vasopressors. Cluster 3 was composed exclusively of patients that were using vasopressors, but that did not have pneumonia as ARDS source. Differences in respiratory and physiological variables were subtle (Table 1). Patients in Cluster 1 had higher illness severity, lower pH, and higher Pco_2. The number of patients randomised to each arm of the ART was balanced amongst clusters.

Bayesian heterogeneity in treatment effect analysis

The results of Bayesian HTE are shown in Figures 2 and 3 and Supplementary Tables S1 and S2. Figure 2 shows the posterior probability distribution of log relative risk of death according to the whole sample (grey line) and clusters; the table inside the panel represents the posterior probability of log relative risk being below 0 (i.e. a relative risk below 1, suggestive of benefit of the ART treatment) in each cluster. Figure 3 shows the relative risk with 95% credible interval for the ART

Table 1 Patient characteristics identified by unsupervised cluster analysis. Data are presented as mean (standard deviation) or *n* (%). *Shock marks vasopressor use or critical organ perfusion without a clear source, as defined by the attending physician. ARDS, acute respiratory distress syndrome; ART, Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial; SAPS, Simplified Acute Physiology Score.

	Cluster 1	Cluster 2	Cluster 3
Number of patients	475	165	370
Age (yr)	52 [18–89]	47 [19–82]	51 [18–95]
Weight (kg)	62 (11)	63 (10)	61 (10)
SAPS 3 score	64 (17)	58 (18)	64 (19)
Mean arterial pressure (mm Hg)	79 (15)	84 (15)	78 (12)
Heart rate (beats min^{-1})	102 (26)	93 (21)	104 (22)
Vasopressor use at admission, *n* (%)	475 (100.0)	0 (0.0)	370 (100.0)
Tidal volume (ml)	356 (73)	364 (66)	352 (73)
Compliance (ml cm H_2O^{-1})	29 (13)	31 (12)	30 (14)
Driving pressure (cm H_2O)	14 (4)	13 (5)	13 (4)
Plateau pressure (cm H_2O)	26 (5)	25 (5)	26 (5)
Ventilatory frequency (bpm)	26 (6)	24 (6)	25 (7)
PEEP (cm H_2O)	13 (3)	12 (3)	12 (3)
FiO_2 (mm Hg)	78 (18)	73 (17)	75 (18)
pH (units)	7.25 (0.13)	7.32 (0.10)	7.27 (0.12)
Po_2 (mm Hg)	118 (44)	119 (43)	118 (41)
Pco_2 (mm Hg)	56 (19)	49 (12)	52 (18)
Po_2/FiO_2 (ratio)	167 (87)	178 (93)	172 (85)
ART group, *n* (%)	243 (51.2)	78 (47.3)	180 (48.6)
ARDS cause, *n* (%)			
Unspecified shock*	0 (0.0)	5 (3.0)	16 (4.3)
Gastric aspiration	0 (0.0)	7 (4.2)	51 (13.8)
Sepsis, not pulmonary	0 (0.0)	28 (17.0)	168 (45.4)
Pneumonia	475 (100.0)	81 (49.1)	0 (0.0)
Orthopaedic surgery	0 (0.0)	0 (0.0)	1 (0.3)
Abdominal surgery	0 (0.0)	8 (4.8)	31 (8.4)
Cardiac surgery	0 (0.0)	1 (0.6)	2 (0.5)
Aortic surgery	0 (0.0)	0 (0.0)	3 (0.8)
Traumatic brain injury	0 (0.0)	0 (0.0)	10 (2.7)
Smoke inhalation	0 (0.0)	3 (1.8)	7 (1.9)
Near drowning	0 (0.0)	0 (0.0)	1 (0.3)
Lung contusion	0 (0.0)	5 (3.0)	6 (1.6)
Trauma	0 (0.0)	2 (1.2)	8 (2.2)
Multiple transfusions	0 (0.0)	4 (2.4)	7 (1.9)
Drug overdose	0 (0.0)	1 (0.6)	2 (0.5)
Other	0 (0.0)	20 (12.1)	57 (15.4)
Barotrauma, *n* (%)	22 (4.6)	3 (1.8)	11 (3.0)
Pneumothorax, *n* (%)	12 (2.5)	2 (1.2)	8 (2.2)
Mortality, *n* (%)	280 (58.9)	43 (26.1)	205 (55.4)

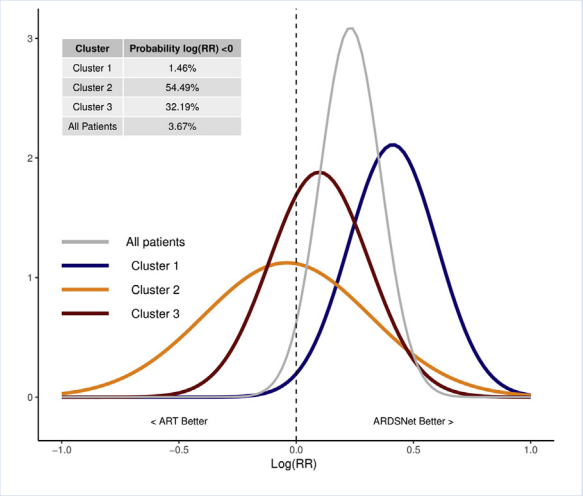

Fig 2. Results of Bayesian heterogeneity in treatment effect. Posterior probability distribution of the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial (ART) treatment effect [log(RR)] in each cluster. Table in the upper right corner contains the probability that the relative risk for mortality using the ART treatment is below 0 (i.e. a relative risk below 1, suggestive of a protective effect of the ART treatment) for each cluster. ARDSNet, ARDS Network.

treatment and mortality in both clusters (raw data in Supplementary Table S3). Detailed information on mean effect size, standard deviation, and quartiles is shown in Supplementary Table S2. Probabilities that effect sizes were different between clusters are shown in detail in Supplementary Figure S3. There was above 80% probability that effect size was lower (i.e. less harmful) in Clusters 2 and 3 when compared with Cluster 1.

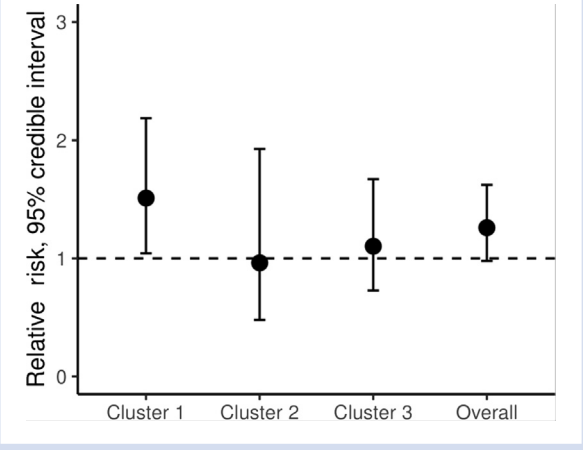

Fig 3. Results of Bayesian heterogeneity in treatment effect analysis. Relative risk (with 95% credible intervals) for the association between the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial treatment and mortality stratified according to cluster.

Interaction between baseline driving pressure and clusters

The results of Bayesian regression for the association between treatment group, driving pressure, cluster, and SAPS 3 are shown in Figure 4 and Supplementary Table S3. ART treatment, baseline driving pressure, SAPS 3 score, and cluster were all associated with mortality, with credible intervals that did not include a null effect (Supplementary Table S3). Interaction between the ART treatment and baseline driving pressure was also highly probable (both linear and quadratic associations). Mortality increased as driving pressure increased and was higher in the ART treatment group until approximately 18 cm H_2O, when a trend towards a reduction in mortality in the ART treatment appeared (Supplementary Fig. S4). However, the small number of patients with baseline driving pressures above 20 cm H_2O results in wide credible intervals with a high degree of uncertainty. The odds ratio for mortality with 95% credible interval range for the ART treatment at fixed driving pressures (10, 15, and 20 cm H_2O) stratified according to cluster is shown in Figure 4. As baseline driving pressure increases, harms associated with the ART treatment also decreased.

Discussion

In this secondary analysis of the ART, we found evidence of HTE using both cluster analysis, a form of unsupervised learning, and Bayesian regression modelling. In the three clusters identified, clinical features (specifically ARDS aetiology and use of vasopressors) were the most distinguishing characteristics. One cluster (Cluster 1) was entirely represented by patients using vasopressors and with ARDS attributable to pneumonia; the other clusters represented patients with lower illness severity (not on vasopressors at admission) and miscellaneous causes for ARDS (Cluster 2), and patients using vasopressors and with ARDS secondary to other causes

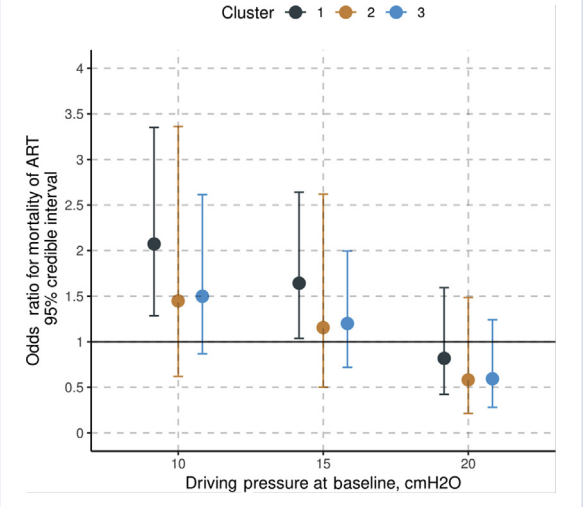

Fig 4. Odds ratio for mortality with 95% credible interval range for the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial (ART) treatment at fixed driving pressures (10, 15, and 20 cm H_2O) and stratified by cluster.

(except pneumonia: Cluster 3). Cluster 1 accounted for most of the harm associated with the intervention arm. Clinical variables that are frequently used to subgroup ARDS patients (including SAPS 3 and Po_2/FiO_2 ratio) were less discriminatory. Driving pressure modified the effect of the intervention arm, suggesting harm at low values that diminished as baseline driving pressure increased.

Investigation of HTE after a trial is published is an important way to obtain insights into subgroups of patients that may behave differently after the intervention, whilst allowing mechanistic exploration of the trial results.[17] Specifically, subgroups that were not defined *a priori* during the trial design and analysis may be found using unsupervised machine learning techniques.[18] Similar approaches have been implemented recently in *post hoc* analyses of ARDS trials. For example, a secondary analysis of the HARP-2 trial found that ARDS patients with a hyper-inflammatory profile might benefit from simvastatin treatment.[18] As no inflammatory biomarkers were collected in the ART, our initial approach was to develop a new clustering method based exclusively on clinical, ventilator, and blood gas analysis of patients using a similar unsupervised machine learning technique. Despite the apparent similarity in ventilatory and blood gas parameters, clusters had distinct clinical features and remarkably different outcomes, and different responses to the trial intervention.

Cluster 1 (pulmonary ARDS on vasopressors) accounted for most of the harm associated with the ART treatment arm. This may be because of the maleficent effects of alveolar recruitment manoeuvres in both pneumonia and in haemodynamically compromised patients. Patients with severe pneumonia commonly exhibit asymmetric and heterogeneous lung injuries, comprising areas of complete alveolar collapse regions that cannot be recruited.[19] During a recruitment manoeuvre, the distribution of gas in such a heterogeneous lung tissue may lead to harmful hyperinflation in some areas, but no recruitment at all in others.[20] Additionally, different time constants at the lung may generate pendelluft phenomenon, which may even worsen ventilator-induced lung injuries with volutrauma.[21] In animal models of pneumonia, recruitment manoeuvres can displace mucus to distal airways, which could theoretically worsen gas exchange.[22]

Recruitment manoeuvres were associated with haemodynamic impairment in the ART, despite measures designed to prevent adverse haemodynamic effects.[6] Patients with increasing doses of vasopressors (increase $\geq$0.2 $\mu g\ kg^{-1}\ min^{-1}$ within 2 h) or hypotension were excluded. Preparation for the recruitment manoeuvre followed a systematic approach, which included optimisation of fluid status and mean arterial blood pressure in all patients in the ART group. Accordingly, as reported previously,[6] there was a trend towards higher fluid balance in the ART, compared with ARDSNet (absolute difference in medians: 302 ml; P=0.06). Monitoring with invasive arterial blood pressure was mandatory, and the recruitment manoeuvre was stopped if adverse events occurred. In 15% of the patients assigned to the ART group, hypotension or other adverse events led to stopping the recruitment manoeuvre. In addition, the incidence of hypotension, or the need to start or increase vasopressors 1 h after randomisation, was higher in the ART than the ARDSNet group.

It is conceivable that recruitment manoeuvres may be particularly harmful in a scenario that combines haemodynamic instability together with lung infection, which is exactly the scenario present in Cluster 1. Our reanalysis of the ART highlights whether pulmonary *vs* extrapulmonary ARDS respond differently to therapy, including recruitment manoeuvres.[23,24] One additional important point to be highlighted is that, as the Po_2/FiO_2 ratio may provide the optimum PEEP level in ARDS,[25] one cannot exclude that both the *run-in* phase of both arms in the ART (which used the ARDSNet group PEEP table) coupled with the average PEEP level of 12–13 cm H_2O in the ARDSNet group could provide enough recruitment or, at least, the best compromise between recruitment and haemodynamic tolerance, thereby justifying the findings in the ART and in this secondary analysis.[25]

We further refined our analyses investigating whether driving pressure, a potentially prognostic factor for ARDS,[9] could modulate the effect of the ART treatment on outcome at the cluster level using a Bayesian regression model. In this analysis, an interaction between the effect of baseline driving pressure and treatment group on the outcome was observed, with higher mortality in the ART treatment group up to baseline driving pressures of 18 cm H_2O. At values of baseline driving pressures above this level, harm tended to be reduced. This finding is indicative of either a stronger (beneficial) effect size of the ART treatment or of less harm caused by this treatment in patients with higher baseline driving pressures. A beneficial effect could be explained, for example, by a higher risk of dying at high baseline driving pressures and a more pronounced reduction in driving pressures in the ART treatment arm. Conversely, the potentially deleterious haemodynamic consequences of higher airway pressures in the ART treatment arm can be mitigated when the lungs are stiff, allowing a less effective transmission of the alveolar pressure to the vascular compartment. However, the small number of patients with very high baseline driving pressure hampered more robust conclusions, with credible intervals that were from 20 cm H_2O. Taken together, our findings suggest that most of the harm from the ART treatment observed in the ART was attributable to an increased mortality in patients admitted with ARDS secondary to pneumonia and on vasopressors, and that uncertainty prevails in other patients, especially those with elevated driving pressure at baseline. This, by no means, suggests that recruitment manoeuvres were shown to be beneficial in any of the clusters and may simply aid patient selection for future RCTs in this field.

This study has several strengths. It is based on a large multicentre RCT with over 1000 patients. Thus, the treatment allocation in the subgroups is based on the intention-to-treat principle (supported by very high adherence in the original trial) and does not suffer from confounding by indication. Additionally, the analyses are supported by robust statistical methods. Baseline clusters were not arbitrarily defined—instead, they were defined by the data using an unsupervised learning technique on only pre-randomisation variables. The HTE analysis that followed clustering is knowingly not associated with information loss as a result of multiple splitting, and is less susceptible to multiple testing constraints.[26] The Bayesian regression offers good estimators of posterior probabilities, which were remarkably stable. Our analysis, does, however, suffer from some limitations, particularly the low number of patients with high baseline driving pressures. Additionally, we did not consider response to recruitment manoeuvre in the analysis, because the control arm was not exposed to recruitment manoeuvres. We have also not assessed other relevant patient-centred outcomes, such as duration of mechanical ventilation, ICU, and hospital stay. As no data on oesophageal pressure were collected in the ART, we have no information whether patients with high

driving pressure were those with abnormally high total thoracic elastance. Although it is conceivable that higher PEEP levels could counterbalance high pleural pressure in these patients, a recent trial of oesophageal-pressure-guided PEEP management failed to show clinical benefit.[27] The ART was a pragmatic trial, and diagnosis of pneumonia was made by the attending physicians. We suggested that pneumonia should be considered in patients with a productive cough or an increase in sputum volume or aspect coupled by worsening oxygenation in patients with new radiological findings. Pneumonia diagnosis could precede ARDS diagnosis and even the start of mechanical ventilation. However, as pneumonia diagnosis was not independently adjudicated, we acknowledge that a misdiagnosis might have occurred, which could impact our results. Additionally, the ART patients had ~50% mortality, which is higher than other clinical trials in ARDS patients. This could reflect specific features of the ART population or aspects of care in the participating ICUs, which could not be fully explained. It is therefore conceivable that the results of this secondary analysis could be different in other settings. Finally, as any *post hoc* analysis, our results should be interpreted with caution and validated in future ARDS cohorts.

Conclusion

One specific patient cluster was associated with higher mortality after receiving the ART intervention compared with ARDSNet low-PEEP ventilation. These patients were admitted with ARDS secondary to pneumonia and were more likely to require vasopressor support. Baseline driving pressure may be an important effect modifier of the ART intervention in ARDS patients. Our unsupervised machine learning approach is widely applicable to perioperative and critical care medicine research, and may help refine the design of future interventional trials.

Authors' contributions

Study design: FGZ.
Design and data collection of the original ART trial: ELC, MBPA, CRRC, ABC.
Statistical analysis: FGZ, LPD, ELC, LUT, TJI.
Writing paper: FGZ.
All authors reviewed the manuscript for important intellectual content.

Declaration of interest

The authors declare that they have no conflicts of interest.

Funding

Brazilian Ministry of Health through the Programa de Apoio ao Desenvolvimento Institucional do Sistema Único de Saúde, which had no role in design, analysis, or reporting of the data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bja.2019.02.026.

References

1. Acute Respiratory Distress Syndrome Network, Brower RG, Matthay MA, et al. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *N Engl J Med* 2000; **18**: 1301–8
2. Brower RG, Lanken PN, MacIntyre N, et al. Higher versus lower positive end-expiratory pressures in patients with the acute respiratory distress syndrome. *N Engl J Med* 2004; **4**: 327–36
3. Papazian L, Forel JM, Gacouin A, et al. Neuromuscular blockers in early acute respiratory distress syndrome. *N Engl J Med* 2010; **12**: 1107–16
4. Guérin C, Reignier J, Richard JC, et al. Prone positioning in severe acute respiratory distress syndrome. *N Engl J Med* 2013; **23**: 2159–68
5. Suzumura EA, Figueiró M, Normilio-Silva K, et al. Effects of alveolar recruitment manoeuvres on clinical outcomes in patients with acute respiratory distress syndrome: a systematic review and meta-analysis. *Intensive Care Med* 2014; **9**: 1227–40
6. Cavalcanti AB, Suzumura EA, Laranjeira LN, et al. Effect of lung recruitment and titrated positive end-expiratory pressure (PEEP) vs low PEEP on mortality in patients with acute respiratory distress syndrome: a randomized clinical trial. *JAMA* 2017; **14**: 1335–45
7. Henderson NC, Louis TA, Wang C, Varadhan R. Bayesian analysis of heterogeneous treatment effects for patient-centered outcomes research. *Health Serv Outcome. Res Methodol* 2016; **4**: 213–33
8. Hayward RA, Kent DM, Vijan S, Hofer TP. Multivariable risk prediction can greatly enhance the statistical power of clinical trial subgroup analysis. *BMC Med Res Methodol* 2006; **6**: 18
9. Amato MB, Meade MO, Slutsky AS, et al. Driving pressure and survival in the acute respiratory distress syndrome. *N Engl J Med* 2015; **8**: 747–55
10. Foss AH, Markatou M. kamila: clustering mixed-type data in R and Hadoop. *J Stat Softw* 2018; **13**: 1–44
11. Tibshirani R, Walther G. Cluster validation by prediction strength. *J Comput Graph Stat* 2005; **3**: 511–28
12. van der Maaten LJP, Hinton GE. Visualizing high-dimensional data using t-SNE. *J Mach Learn Res* 2008; **1**: 2579–605
13. Wang C, Louis TA, Henderson NC, Weiss CO, Varadhan R. beanz: an R package for Bayesian analysis of heterogeneous treatment effects with a graphical user interface. *J Stat Softw* 2018; **7**: 1–31
14. B1kner PC. brms: an R package for Bayesian multilevel models using Stan. *J Stat Softw* 2017; **1**: 1–28
15. Zampieri FG, Iwashyna TJ, Viglianti EM, et al. Association of frailty with short-term outcomes, organ support and resource use in critically ill patients. *Intensive Care Med* 2018; **9**: 1512–20
16. van Buuren S, Groothuis-Oudshoorn K. Multivariate imputation by chained equations in R. *J Stat Softw* 2011; **3**: 1–67
17. Iwashyna TJ, Burke JF, Sussman JB, Prescott HC, Hayward RA, Angus DC. Implications of heterogeneity of treatment effect for reporting and analysis of randomized trials in critical care. *Am J Respir Crit Care Med* 2015; **9**: 1045–51
18. Calfee CS, Delucchi KL, Sinha P, et al. Acute respiratory distress syndrome subphenotypes and differential

response to simvastatin: secondary analysis of a randomised controlled trial. *Lancet Respir Med* 2018; **9**: 691–8
19. Tomashefski Jr JF. Pulmonary pathology of the adult respiratory distress syndrome. *Clin Chest Med* 1990; **4**: 593–619
20. Constantin JM, Grasso S, Chanques G, et al. Lung morphology predicts response to recruitment maneuver in patients with acute respiratory distress syndrome. *Crit Care Med* 2010; **4**: 1108–17
21. Yoshida T, Torsani V, Gomes S, et al. Spontaneous effort causes occult pendelluft during mechanical ventilation. *Am J Respir Crit Care Med* 2013; **12**: 1420–7
22. Li Bassi G, Comaru T, Martr D, et al. Recruitment manoeuvres dislodge mucus towards the distal airways in an experimental model of severe pneumonia. *Br J Anaesth* 2018; **2**: 269–76
23. Pelosi P, D'Onofrio D, Chiumello D, et al. Pulmonary and extrapulmonary acute respiratory distress syndrome are different. *Eur Respir J* 2003; **22**(Suppl 42): 48s–56s
24. Gattinoni L, Pelosi P, Suter PM, Pedoto A, Vercedi P, Lissoni A. Acute respiratory distress syndrome caused by pulmonary and extrapulmonary disease: different syndromes? *Am J Respir Crit Care Med* 1998; **158**: 3–11
25. Chiumello D, Cressoni M, Carlesso E, et al. Bedside selection of positive end-expiratory pressure in mild, moderate, and severe acute respiratory distress syndrome. *Crit Care Med* 2014; **2**: 252–64
26. Gelman A, Tuerlinckx F. Type S error rates for classical and Bayesian single and multiple comparison procedures. *Comput Stat* 2000; **15**: 373–90
27. Beitler JR, Sarge T, Banner-Goodspeed VM, et al. Effect of titrating positive end-expiratory pressure (PEEP) with an esophageal pressure-guided strategy vs an empirical high PEEP-FiO_2 strategy on death and days free from mechanical ventilation among patients with acute respiratory distress syndrome. *JAMA* 2019; **321**: 846–57

Handling editor: G.L. Ackland