Title

Subphenotypes in acute respiratory distress syndrome: Latent class analysis of data from two randomised controlled trials

Journal

The Lancet Respiratory Medicine, 2(8)

ISSN

2213-2600

Authors

Calfee, CS
Delucchi, K
Parsons, PE
et al.

Publication Date

2014

DOI

10.1016/S2213-2600(14)70097-9

Peer reviewed

Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials

Carolyn S Calfee, Kevin Delucchi, Polly E Parsons, B Taylor Thompson, Lorraine B Ware, Michael A Matthay, and the NHLBI ARDS Network

Summary

Background Subphenotypes have been identified within heterogeneous diseases such as asthma and breast cancer, with important therapeutic implications. We assessed whether subphenotypes exist within acute respiratory distress syndrome (ARDS), another heterogeneous disorder.

Methods We used data from two ARDS randomised controlled trials (ARMA trial and ALVEOLI trial), sponsored by the National Heart, Lung, and Blood Institute. We applied latent class modelling to identify subphenotypes using clinical and biological data. We modelled data from both studies independently. We then tested the association of subphenotypes with clinical outcomes in both cohorts and with the response to positive end-expiratory pressure (PEEP) in the ALVEOLI cohort.

Findings We analysed data for 1022 patients: 473 in the ARMA cohort and 549 in the ALVEOLI cohort. Independent latent class models indicated that a two-class (ie, two subphenotype) model was the best fit for both cohorts. In both cohorts, we identified a hyperinflammatory subphenotype (phenotype 2) that was characterised by higher plasma concentrations of inflammatory biomarkers, a higher prevalence of vasopressor use, lower serum bicarbonate concentrations, and a higher prevalence of sepsis than phenotype 1. Participants in phenotype 2 had higher mortality and fewer ventilator-free days and organ failure-free days in both cohorts than did those in phenotype 1 ($p<0·007$ for all). In the ALVEOLI cohort, the effects of ventilation strategy (high PEEP *vs* low PEEP) on mortality, ventilator-free days and organ failure-free days differed by phenotype ($p=0·049$ for mortality, $p=0·018$ for ventilator-free days, $p=0·003$ for organ-failure-free days).

Interpretation We have identified two subphenotypes within ARDS, one of which is categorised by more severe inflammation, shock, and metabolic acidosis and by worse clinical outcomes. Response to treatment in a randomised trial of PEEP strategies differed on the basis of subphenotype. Identification of ARDS subphenotypes might be useful in selecting patients for future clinical trials.

Funding National Institutes of Health.

Lancet Respir Med 2014

Published Online
May 20, 2014
http://dx.doi.org/10.1016/S2213-2600(14)70097-9

See Online/Comment
http://dx.doi.org/10.1016/S2213-2600(14)70116-X

Departments of Medicine and Anesthesia, Division of Pulmonary and Critical Care Medicine (C S Calfee MD, Prof M A Matthay MD) **and Department of Psychiatry** (Prof K Delucchi PhD), **University of California San Francisco, San Francisco, CA, USA; Department of Medicine, Division of Pulmonary and Critical Care Medicine, University of Vermont, Burlington, VT, USA** (Prof P E Parsons MD); **Department of Medicine, Pulmonary and Critical Care Medicine Unit, Massachusetts General Hospital, Boston, MA, USA** (Prof B T Thompson MD); **Biostatistics Unit, Massachusetts General Hospital, Boston, MA, USA** (Prof B T Thompson); **Department of Medicine, Division of Allergy, Pulmonary and Critical Care, Vanderbilt, University, Nashville, TN, USA** (Prof L B Ware MD); **and Cardiovascular Research Institute, San Francisco, CA, USA** (Prof M A Matthay)

Correspondence to:
Dr Carolyn S Calfee, University of California, San Francisco, Pulmonary and Critical Care Division, 505 Parnassus Avenue, Box 0111, San Francisco, CA 94143-0111, USA
carolyn.calfee@ucsf.edu

Introduction

Acute respiratory distress syndrome (ARDS) is a heterogeneous syndrome first identified in 1967 and defined by the clinical criteria of bilateral pulmonary opacities on chest radiograph, arterial hypoxaemia (partial pressure of arterial oxygen [PaO_2] to fraction of inspired oxygen [FiO_2] ratio <300), and exclusion of cardiac failure as the primary cause of the syndrome.[1–3] This definition was derived empirically on the basis of clinical experience, with the hypothesis that it would identify patients with non-cardiogenic pulmonary oedema, characterised by increased protein permeability of the alveolar–capillary membrane. Since the time of the original identification of ARDS, and increasingly during the past two decades, there has been recognition of the clinical and biological heterogeneity within ARDS;[4,5] this heterogeneity might reflect our incomplete understanding of the biology of ARDS and probably contributes to the poor track record of phase 2 and 3 trials of new treatments for patients with ARDS.[6] As a result, some investigators have proposed subdividing ARDS on the basis of clinical risk factors, or by direct versus indirect cause of lung injury. However, no consensus exists on the appropriate approach to reduce ARDS heterogeneity.

By contrast with ARDS, research in airways disease and cancer has made substantial progress towards identifying subphenotypes of disease, with important therapeutic implications. For example, subphenotypes based on the presence or absence of Th2-dependent inflammation have been identified in asthma, with important mechanistic and therapeutic implications.[7] This insight has led to new targeted treatments, such as a monoclonal antibody to interleukin-13, which is especially effective in individuals with Th2-predominant inflammation.[8] Despite widespread recognition of the heterogeneity within common critical illness syndromes such as sepsis and ARDS, and some evidence suggesting that subphenotypes might exist within severe sepsis,[6,9,10] little data are available for whether such subphenotypes exist in ARDS.

Latent class analysis is a well-validated statistical technique that uses mixture modelling to find the best-fitting model for a set of data, based on the hypothesis that the data contain several unobserved groups or

classes. The statistical approaches underlying this method were originally developed more than a century ago by investigators analysing whether a population of crabs consisted of two subspecies.[11] By contrast with traditional regression analyses, in which the goal is to understand the association between pre-specified independent variables and a known outcome, latent class analysis models ask whether there are subgroups of patients defined by a combination of the baseline variables, without mandating consideration of the outcome. Latent class-based methods have been extensively used in the social sciences and in other medical disciplines,[12,13] for instance in the identification of asthma subphenotypes,[14] but have not been used extensively in critical care. We sought to capitalise on the wealth of clinical and biological data available from two National Heart, Lung, and Blood Institute (NHLBI)-sponsored ARDS Network randomised controlled trials by using latent class analysis methods to attempt to identify and validate novel subphenotypes of ARDS and to test their association with clinical outcomes and response to treatment.

See Online for appendix

Methods

Study design

Clinical and biological data were obtained from patients enrolled in the NHLBI ARDS Network's randomised controlled trials of lower versus higher tidal volume ventilation (trial referred to here as ARMA)[15–17] and higher versus lower positive end-expiratory pressure (PEEP; trial referred to here as ALVEOLI).[18] Some patients in the ARMA trial were co-enrolled in a trial of lisofylline versus placebo;[17] after 194 patients were enrolled in that trial, the ventilator trial was discontinued because of a statistically significant reduction in mortality in the lower tidal volume group, and the 41 additional patients subsequently enrolled in the lisofyilline trial were assigned to lower tidal volume ventilation. Details of the trials are available elsewhere.[15–18] Briefly, the first ARMA was a multicentre randomised controlled trial done in the USA comparing ventilation with lower versus higher tidal volumes in patients within 36 h of ARDS onset.[15] The ALVEOLI trial was also a multicentre randomised controlled trial done in the USA, which compared ventilation with lower versus higher PEEP with similar inclusion criteria to ARMA.[18] Patients were enrolled in the ARMA study between 1996 and 1999; patients were enrolled in the ALVEOLI study between 1999 and 2002 (see appendix for further details of each trial). We excluded data for patients in the ARMA trial who were randomly allocated to higher tidal volume ventilation (n=429) because of the negative effect of higher tidal volumes on mortality, which would have precluded analysis of the association between latent class and clinical outcomes (figure 1). We first did analyses of data for patients in the ARMA trial treated with lower tidal volume (n=473, figure 1). We then independently repeated analyses with data for patients in the ALVEOLI trial (n=549; no patients excluded), to test whether the findings would be generalisable to an independent sample. All clinical data (other than outcomes) and biological data used for this analysis were collected at study baseline (pre-randomisation).

Assay Procedures

Plasma samples used for this analysis were drawn at the time of randomisation, which occurred within 36 h of meeting ARDS criteria. Plasma biomarkers were measured in duplicate (ie, two samples per patient) by use of ELISA. Details of the methods used to do the assays have been reported elsewhere.[4,19–25]

Statistical analysis

Baseline clinical data and biomarker concentrations were considered as class-defining variables in the latent class analysis model; classification was done without consideration of clinical outcomes. Details on clinical variable selection, data cleaning, and a complete list of the clinical variables included in these models are in the appendix. In addition to the baseline clinical data, we included as inputs in the latent class model the eight plasma biomarkers previously associated with poor clinical outcomes in ARDS and that were previously measured in both cohorts: surfactant protein D,[23] von Willebrand factor antigen,[19] soluble intercellular adhesion molecule-1,[24] interleukin-6 and interleukin-8,[21] soluble tumour necrosis factor receptor-1,[22] plasminogen

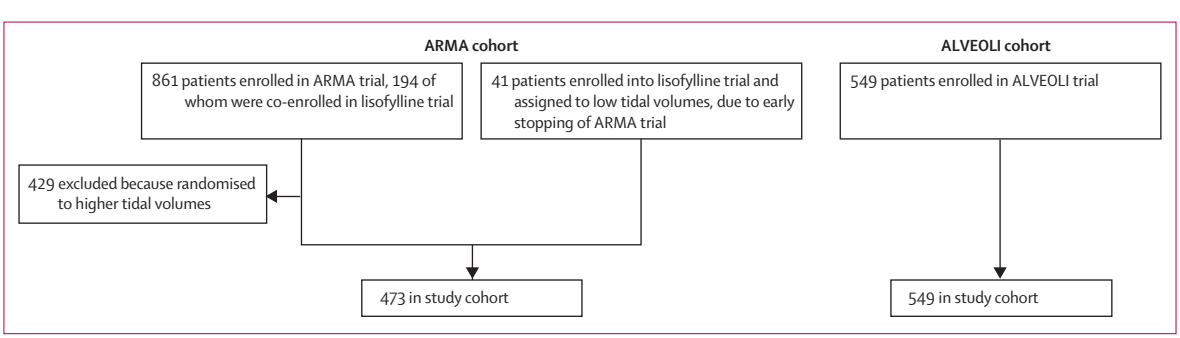

Figure 1: **Study population**

activator inhibitor-1,[20] and protein C.[20] The datasets used, by virtue of being derived from randomised controlled trials with intensive on-site auditing and data quality checks, were largely complete. However, data for some variables were missing (appendix).

We did basic two-group comparisons between the two cohorts using the *t* test, Wilcoxon rank sum, or χ^2 test, as appropriate. Next, we fitted a series of latent class models, first using the data from the ARMA cohort and then repeated independently using data from the ALVEOLI cohort. Criteria for model selection were based on the Bayesian information criteria, the Vuong-Lo-Mendell-Rubin likelihood ratio test, and the size of the smallest class. Latent class model estimation was based on full-information maximum likelihood methods. This approach allows for the use of all data for all patients, including those missing some data, in estimating the latent class models. Additional details about the latent class modelling procedures are shown in the appendix.

Once we established the number of classes, we tested the associations between class and clinical outcomes (90-day mortality, ventilator-free and organ failure-free days) using the approach developed by Lanza.[26] This method incorporates the degree of uncertainty of class membership. Finally, for the ALVEOLI cohort, in which patients were randomly allocated to lower or higher PEEP, we tested models of each outcome using class, treatment assignment, and their interaction as covariates to find out whether there was a differential treatment effect based on latent class. We did these analyses using Poisson regression for ventilator-free and organ failure-free days and logistic regression for mortality. We used Mplus (version 7.11) for the latent class analyses and SAS (version 9.3) for all other analyses.

Role of the funding source

The study funder had no role in the study design, analysis or interpretation of the data, or writing of the report for these analyses. The corresponding author had full access to all of the data and had the final responsibility to submit for publication.

Results

Baseline clinical characteristics for patients in both cohorts are shown in table 1. We noted several statistically significant differences between the two cohorts, including primary ARDS risk factor, severity of illness as measured by Acute Physiology and Chronic Health Evaluation (APACHE III) scores, and prevalence of vasopressor use at enrolment. Additionally, several ventilator variables at randomisation differed substantially between the two cohorts, probably indicative of changes in practice resulting from the publication of the results of the first trial (ARMA; of lower tidal volume ventilation) that had been adopted at the time of the start of the second trial (ALVEOLI). Baseline biomarker data

	ARMA cohort (n=473)	ALVEOLI cohort (n=549)	p value
Age in years	51 (17)	51 (17)	0·96
Sex (female)	284 (60%)	302 (55%)	0·09
Ethnic origin (white)	355 (75%)	412 (75%)	0·93
ARDS risk factor*			<0·001
Trauma	59 (13%)	45 (9%)	..
Sepsis	125 (27%)	120 (23%)	..
Aspiration	72 (16%)	84 (16%)	..
Pneumonia	145 (32%)	221 (42%)	..
Other	60 (13%)	52 (10%)	..
APACHE III score	82 (29)	94 (32)	<0·001
On vasopressors at enrolment	134/338† (40%)	144/549 (26%)	<0·001
PaO_2/FiO_2 ratio	132 (60)	128 (58)	0·25
Maximum temperature (°C)	38·5 (0·9)	38·5 (1·0)	0·09
Lowest systolic blood pressure (mm Hg)	88 (19)	88 (17)	0·90
Maximum heart rate (beats per min)	127 (22)	125 (24)	0·31
Maximum respiratory rate (breaths per min)	30 (10)	33 (10)	<0·001
Urine output in 24 h before enrolment (L)	2·02 (1·24–2·98)	1·85 (1·13–2·93)	0·06
Lowest haematocrit percentage	30 (6)	30 (6)	0·64
Peak white blood cell count (in thousands)	12·3 (8·9–17·9)	13·0 (8·7–18·2)	0·68
Platelet count (in thousands)	164 (121)	177 (124)	0·10
Lowest sodium (mEq/L)	137 (5)	137 (5)	0·86
Highest creatinine (mg/dL)	1·10 (0·80–1·70)	1·10 (0·8–1·9)	0·70
Lowest glucose (mg/L)	135 (57)	133 (64)	0·68
Lowest albumin (g/dL)	2·2 (0·6)	2·1 (0·6)	0·004
Total bilirubin (mg/dL)	1·0 (0·60–2·10)	0·80 (0·5–1·5)	<0·001
Bicarbonate (mEq/L)	21 (5)	22 (6)	0·24
Tidal volume (mL)	671 (126)	511 (119)	<0·001
Total minute ventilation (L per min)	13 (4)	12 (4)	<0·001
Positive end-expiratory pressure (cm H_20)	8·6 (3·8)	9·5 (4·3)	<0·001
Plateau pressure (cm H_20)	30 (8)	27 (7)	<0·001
Mean airway pressure (cm H_20)	16 (5)	16 (5)	0·38
$PaCO_2$ (mm Hg)	37 (8)	39 (9)	<0·001
Body-mass index	27 (7)	27 (7)	0·53

Data are mean (SD), n (%), n/N (%), or median (IQR). *Denominator was 461 for ARMA and 522 for ALVEOLI due to missing data. †Denominator was 338 due to missing data. ARDS=acute respiratory distress syndrome. APACHE=Acute Physiology and Chronic Health Evaluation.

Table 1: **Comparison of baseline characteristics of patients between the ARMA and ALVEOLI cohorts**

	ARMA cohort	ALVEOLI cohort	p value
Protein C (% control)	47 (32–66)	78 (45–122)	<0·001
Plasminogen activator inhibitor-1 (ng/mL)	70 (40–138)	61 (30–144)	0·002
Interleukin-6 (pg/mL)	264 (109–766)	238 (93–741)	0·26
Interleukin-8 (pg/mL)	43 (20–93)	40 (16–98)	0·001
Soluble tumour necrosis factor receptor-1 (pg/mL)	3255 (2128–5600)	4265 (2599–8448)	<0·001
Soluble intercellular adhesion molecule-1 (ng/mL)	627 (345–1038)	924 (605–1385)	<0·001
Surfactant protein D (ng/mL)	84 (40–162)	101 (50–218)	0·004
von Willebrand factor antigen (% control)	284 (173–436)	398 (247–624)	<0·001

Data are median (IQR).

Table 2: **Comparison of key baseline biomarker values between the ARMA and ALVEOLI cohorts**

	Bayesian Information Criterion	Entropy*	Number of individuals per class or subphenotype					p value†
			1	2	3	4	5	
ARMA cohort								
2 classes	39947·9	0·78	318	155	..	..	..	0·036
3 classes	39760·2	0·88	308	119	46	..	..	0·59
4 classes	39656·7	0·86	212	126	43	92	..	0·28
5 classes	39583·8	0·86	150	120	36	36	131	0·64
ALVEOLI cohort								
2 classes	49709·5	0·87	404	145	..	..	..	0·016
3 classes	49383·7	0·92	400	145	4	..	..	0·58
4 classes	49098·8	0·94	386	129	4	30	..	0·35
5 classes	48955·1	0·87	242	154	4	30	119	0·80

*Entropy is an index of how well the classes are separated: it ranges from zero to one and values of about 0·8 or higher are thought to be a sign of a useful model. †By Vuong-Lo-Mendell-Rubin test, testing whether the number of classes provides improved model fit compared with a model using one fewer class.

Table 3: **Fit statistics for latent class models from two to five classes**

are shown in table 2. As with the clinical data, we noted substantial differences in all baseline biomarker concentrations other than interleukin-6 between cohorts, probably due to differences in severity of illness, pre-randomisation ventilation variables, or both.

In each cohort, analysis of latent-class models suggested that a two-class model provided the best fit (table 3). In both cohorts, the p value testing the number of classes indicated that a two-class model was a significant improvement over a one-class model, but that the three-class model did not significantly increase the explanatory power. The value of the Bayesian Information Criteria continued to decrease as the number of classes increased—this decrease suggests that the addition of more classes is worth the added model complexity (table 3). This decrease was also seen in both the Akakie Information Criteria and sample-sized adjusted-Bayesian Information Criteria (data not shown). To ensure that a two-class model provided the best fit, we also explored a three-class model, which produced one class with only 46 participants in the ARMA cohort. In the ALVEOLI cohort, the third class consisted of only four participants. Although the decrease in the Bayesian Information Criteria would suggest adding additional classes to the model, on consideration of the p value (favouring a two-class model) and the small number of participants in the third class, the two-class model was retained. For simplicity, we will henceforth refer to the two classes as phenotypes 1 and 2, respectively. In the two-class model, the average latent class probabilities for the most likely class in the ARMA cohort were 0·95 for phenotype 1 and 0·92 for phenotype 2; in the ALVEOLI cohort, the analogous probabilities were 0·97 and 0·94, indicating good model fit and very strong probabilities of class assignment (appendix).

We next sought to understand the clinical and biological characteristics that distinguished each phenotype. To do this, in view of the high probabilities of class membership, we assigned study participants to their most likely phenotype and examined the mean values of the variables used in the model for each phenotype. Figure 2 shows the continuous variables for the two phenotypes in the ARMA cohort, sorted by the degree of separation between the phenotypes. Compared with phenotype 1, phenotype 2 was defined by higher plasma concentrations of interleukin-6, interleukin-8, soluble tumour necrosis factor receptor-1, and plasminogen activator inhibitor-1; higher heart rate and total minute ventilation; and lower systolic blood pressure, bicarbonate, and protein C concentration. Figure 3 shows differences in the categorical variables between the phenotypes in the ARMA cohort. Although sex and ethnic origin differed with statistical significance but not substantially between the phenotypes, vasopressor use at baseline was more than three times as common in phenotype 2 compared with phenotype 1. Furthermore, participants in phenotype 1 were more likely to have trauma-associated ARDS and less likely to have sepsis-associated ARDS than those in phenotype 2 (figure 3).

The latent class models were derived again independently in the ALVEOLI cohort, and the contribution of the key variables is shown in figures 2 and 3. The characteristics of the two subphenotypes in this cohort were similar to those in the ARMA cohort, with phenotype 2 characterised by more profound inflammation, acidosis, and shock compared with phenotype 2. Specifically, as in the ARMA cohort, phenotype 2 was characterised by higher plasma concentrations of inflammatory biomarkers, higher heart rate, and minute ventilation, and by lower systolic blood pressure, bicarbonate, and protein C concentration, compared with phenotype 2 (figure 2). As in the ARMA cohort, there were statistically significant differences in vasopressor use and in ARDS risk factors between the two phenotypes (figure 3).

To establish whether phenotype prediction would be potentially feasible with fewer variables, we used three of the measures with the greatest difference in mean absolute values between phenotypes in the ARMA cohort as predictive markers in a receiver-operator characteristic curve analysis. With three variables (interleukin-6, soluble tumour necrosis factor receptor-1, and vasopressor use [yes or no]), the area under the curve for phenotype prediction was 0·937 in the ARMA cohort and 0·929 in the ALVEOLI cohort, suggesting that phenotype can be accurately predicted with a small number of variables (appendix). The addition of one to two additional variables further increased the area under the curve slightly in both cohorts (appendix).

To establish whether the two phenotypes had different natural histories, we tested the association between probable phenotype assignment and clinical outcomes, incorporating the degree of uncertainty regarding phenotype assignment as described in the methods. In the ARMA cohort, participants in phenotype 2 had fewer organ failure-free and ventilator-free days than did those

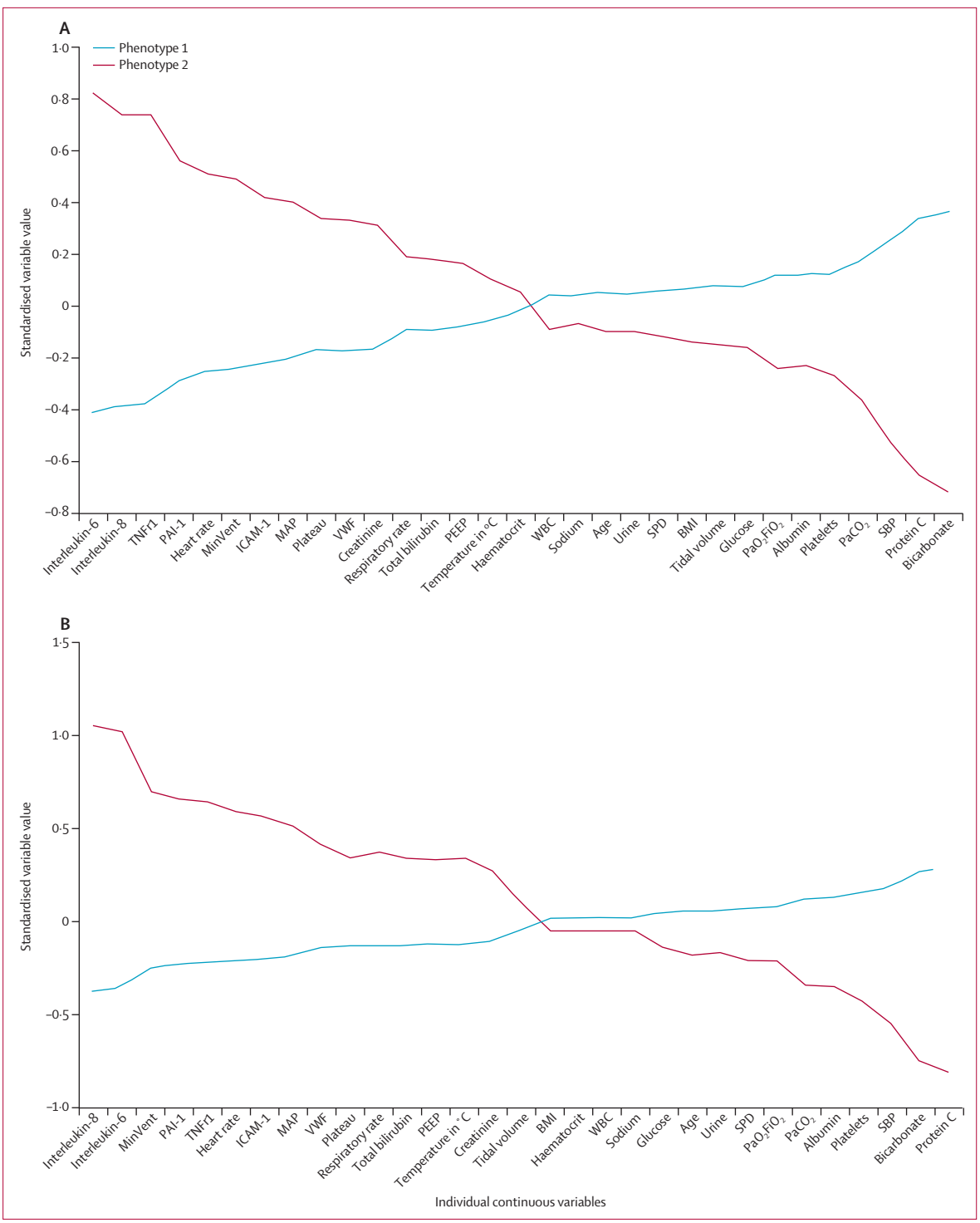

Figure 2: **Differences in standardised values of each continuous variable by phenotype in the ARMA cohort (A) and the ALVEOLI cohort (B)**
The variables are sorted on the basis of the degree of separation between the classes from maximum positive separation on the left (ie, phenotype 2 higher than phenotype 1) to maximum negative separation on the right (ie, phenotype 2 lower than phenotype 1). Variable standardisation, in which all means are scaled to zero and SDs to one, is described in the appendix. A value of +1 for the standardised variable signifies that the mean value for a given phenotype was one SD higher than the mean value in the cohort as a whole. TNFr1=tumour necrosis factor receptor-1. PAI-1=plasminogen activator inhibitor-1. MinVent=total minute ventilation. ICAM-1=intercellular adhesion molecule-1. MAP=mean airway pressure. VWF=von Willebrand factor. PEEP=positive end-expiratory pressure. Urine=urine output over prior 24 h. BMI=body-mass index. SBP=systolic blood pressure. SPD=surfactant protein D. $PaCO_2$=partial pressure of carbon dioxide in arterial blood.

in phenotype 1 (table 4). Furthermore, participants in phenotype 2 had higher mortality compared with those in phenotype 1 (44% *vs* 23%, p=0·006; table 4). Likewise, in the ALVEOLI cohort, participants in phenotype 2 had much worse clinical outcomes than those in phenotype 1, including a much higher mortality rate (51% *vs* 19%, p<0·001; table 4). Clinical outcomes analysed without adjustment for uncertainty about phenotype assignment showed a similar pattern (appendix).

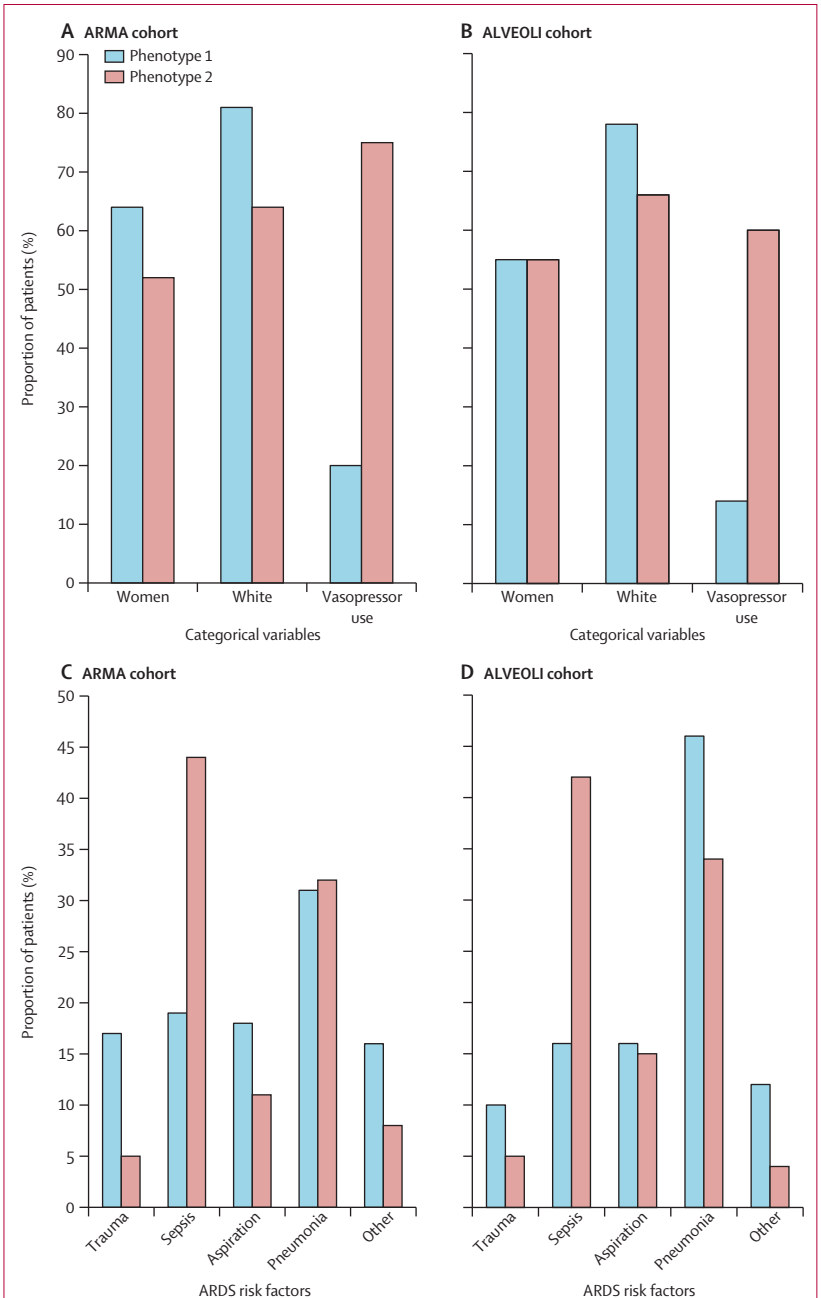

Figure 3: **Difference in categorical variables (A, B) and ARDS risk factors (C, D) by phenotype assignment**
(A) p=0·007 for comparison of sex, p<0·0001 for other comparisons. (B) p=0·96 for comparison of sex, p=0·004 for ethnic origin, p<0·0001 for vasopressor use. (C) p<0·0001 for difference across all variables. (D) p<0·0001 for difference across all variables. ARDS=acute respiratory distress syndrome.

Finally, we used data from the ALVEOLI trial to establish whether inter-phenotype differences existed in response to randomly assigned treatment (mechanical ventilation with higher *vs* lower PEEP). We saw that PEEP strategy had different effects on mortality in the two different phenotypes (p=0·049 for interaction). Specifically, within phenotype 1, 48 (24%) of 202 patients randomly allocated to the higher PEEP strategy died, compared with 33 (16%) of 202 patients randomly allocated to the low PEEP strategy. In phenotype 2, 31 (42%) of 74 patients randomly allocated to the higher PEEP strategy died, compared with 36 (51%) of 71 patients randomly allocated to the low PEEP strategy. We saw even stronger interactions between phenotype and PEEP strategy for the outcomes of ventilator-free and organ failure-free days, showing significantly differential effects of high versus low PEEP on these clinical outcomes in the two different phenotypes (table 5).

As a sensitivity analysis to determine whether general severity of illness scores could supplant phenotype identification, we tested for interactions between APACHE III score and PEEP. By contrast with the analyses using phenotype, we detected no statistically significant interactions between APACHE score and PEEP strategy for the outcomes of mortality, organ failure-free days, or ventilator-free days (p=0·58 for mortality, p=0·69 for ventilator-free days, and p=0·99 for organ failure-free days for interactions).

Discussion

Our findings suggest the existence of two different subphenotypes in patients with ARDS. These two subphenotypes have very different natural histories, clinical and biological characteristics, clinical outcomes, and response to treatment, fulfilling the criteria necessary to define a subphenotype.[7] Clinicians caring for patients with ARDS and researchers studying ARDS have long appreciated the heterogeneity within this complex syndrome, but the critical care community has not had empirical data for whether or how to refine our definitions and further subdivide ARDS. The subphenotypes identified here were evident in independent analyses of two clinical trial samples, despite substantial differences in the baseline clinical and biological profiles of these two cohorts. Furthermore, subphenotype was strongly and consistently associated with clinical outcomes in both cohorts, with pronounced differences in ventilator-free days, organ failure-free days, and mortality. Perhaps most importantly, the two subphenotypes had different responses to treatment (lower *vs* higher PEEP) in the ALVEOLI cohort, suggesting that identification of subphenotypes might be crucial for future clinical trials in ARDS.

Taken together, the variables that characterise phenotype 2 (high plasma concentrations of inflam-

matory biomarkers, severe shock, and metabolic acidosis) paint a portrait of a hyper-inflammatory ARDS subphenotype that could afflict patients across the demographic spectrum of age, sex, ethnic origin, and cause of ARDS (although there are some differences in the latter factors by subphenotype). By contrast with phenotype 2, phenotype 1 seems to be characterised by less severe inflammation and shock. Of note, neither the severity of ARDS (PaO_2/FiO_2 ratio), the severity of renal or hepatic failure, or the extent of leucocytosis distinguished the two phenotypes from each other, since the specified variables had similar values in the two phenotypes (figure 2). In concert with the results of our sensitivity analysis that incorporated APACHE III scores, these data suggest that phenotype is not just an indicator of severity of illness as measured by traditional prognostic indices.

No one clinical or biological variable was sufficient to identify subphenotype; put differently, none of the clinical characteristics typically used to subdivide ARDS—eg, ARDS risk factor, presence or absence of sepsis, direct versus indirect lung injury, or the use of vasopressors—was strictly associated with one or the other subphenotype (figure 3). When considered as a group, however, the clinical variables that characterise phenotype assignment form a coherent and plausible cluster that has face validity from a clinical and research perspective. For instance, phenotype 2 is characterised by a high prevalence of vasopressor use, more severe acidosis, and a high minute ventilation—a collection of clinical datapoints that forms a recognisable pattern of more severe ARDS and systemic injury to the practising intensivist. Hypothetically, if phenotype 2 had been characterised by a low prevalence of vasopressor use, severe acidosis, and a low minute ventilation, then it would not seem recognisable from a clinical perspective.

Of the continuous variables, the plasma protein biomarkers generally contributed more prominently to the phenotype definitions than did most of the clinical variables (including biomarkers used in clinical practice such as serum creatinine and white blood cell counts; figure 2). This finding suggests that the plasma protein biomarkers might be capturing aspects of pathophysiology that are not otherwise well-captured in our clinical data, and also that development of the capability to measure these biomarkers in an expedient point-of-care method might be necessary to incorporate subphenotype determination into clinical trials.

Two important branch points in the analytic strategy warrant further discussion. First, we deliberated carefully at the inception of the analyses about whether to include the high tidal volume patients in the analyses of the ARMA cohort, which would have enabled us to test for an interaction between tidal volume and subphenotype in that cohort. We decided not to include these patients because to do so would have precluded analysis of the association between subphenotype and mortality in that cohort. We thought this aspect of the analysis was too important to discard, and therefore decided to exclude the patients treated with high tidal volume. Second, the ultimate decision as to the optimum number of classes identified by the models requires consideration of several factors. Latent class models seek to find the best model fit, assuming that there are a given number of latent classes in the data. If there are only two classes but we fit a three-class model, the third class will be forced in by selecting a small number of cases with a more extreme or unique set of values. The p value for the Vuong-Lo-Mendell-Rubin test strongly suggests that a two-class model is preferable to a three-class model in both cohorts. That data, combined with the small size of the third class, led us to focus on a two-class model. A class that is very small, relatively, offers little information and might be more of an anomaly than a useful finding.

Although latent class modelling has not to our knowledge been previously applied in classic diverse ARDS cohorts (panel), Shah and colleagues used latent class-based models to identify subphenotypes within primary graft dysfunction, a type of acute lung injury that occurs after lung transplantation.[27] These models focused exclusively on the timing of onset and resolution of lung dysfunction, using only grades of primary graft dysfunction at various timepoints to generate latent classes, and the analyses showed that patients with severe persistent dysfunction (one of three identified classes) had the worst clinical outcomes. Similarly, Reilly and colleagues from the same research group did a latent class analysis of timing of ARDS after severe trauma.[28] In addition to their focus on a different population, these

	ARMA cohort			ALVEOLI cohort		
	Phenotype 1 (n=318)	Phenotype 2 (n=155)	p value	Phenotype 1 (n=404)	Phenotype 2 (n=145)	p value
Mortality (at 90 days)	23%	44%	0·006	19%	51%	<0·001
Ventilator-free days	17·8	7·7	<0·001	18·4	8·3	<0·001
Organ failure-free days	14·5	8·0	<0·001	16·5	8·4	<0·001

Values are estimated means that take into account the uncertainty of class membership.

Table 4: **Association between phenotype assignment and clinical outcomes, adjusted for degree of uncertainty regarding phenotype assignment**

	Phenotype 1 (n=404)		Phenotype 2 (n=145)		
	Low PEEP (n=202)	High PEEP (n=202)	Low PEEP (n=71)	High PEEP (n=74)	p value*
Mortality at 90 days	33 (16%)	48 (24%)	36 (51%)	31 (42%)	0·049
Ventilator-free days	20 (10–25)	21 (3–24)	2 (0–21)	4·5 (0–20)	0·018
Organ failure free-days	22 (11–26)	22 (9–26)	4 (0–18)	6·5 (0–21)	0·003

Data are n (%) or median (IQR). *p value for interaction between positive end-expiratory pressure (PEEP) assignment and phenotype.

Table 5: **Differences in response to PEEP strategy by phenotype (ALVEOLI cohort only)**

Panel: **Research in context**

Systematic review

We did not do a formal systematic review before the inception of these analyses; however, we knew from our engagement with the literature that there have been few attempts to use analytic mixture-model based methods to identify subphenotypes of ARDS. Since the time of the original identification of ARDS and increasingly over the past two decades, there has been recognition of the clinical and biological heterogeneity within the syndrome. As a result, some investigators have proposed subdividing ARDS on the basis of clinical risk factors, or by direct versus indirect aetiology of lung injury; however, there is no consensus in the field on the appropriate approach to reducing ARDS heterogeneity. Some previous analyses have attempted to identify ARDS subgroups using traditional regression-based methods; additionally, investigators have used latent class models to analyse the timing of lung injury after lung transplantation and after severe trauma. After completion of this study, we searched PubMed using the terms (ARDS OR "acute lung injury) AND ("latent class" OR subphenotype OR endophenotype OR endotype) and identified only one study of severe trauma and one editorial. We applied no language restrictions. We did our last search on May 9, 2014. To our knowledge, this paper is the first report of the use of latent class models to identify subphenotypes of ARDS using both clinical and biomarker data within two diverse and heterogeneous samples of patients with ARDS.

Interpretation

We identified two subphenotypes within ARDS, one of which is characterised by more severe inflammation, shock, and metabolic acidosis, substantially worse clinical outcomes, and a differential response to treatment with positive end-expiratory pressure. These findings provide proof-of-concept that the clinical syndrome of ARDS contains distinct subphenotypes and should prompt future studies aimed at further elucidating these subphenotypes with comprehensive clinical and biological data. This novel approach to identification of subphenotypes could inform the design of future randomised controlled trials of new treatments for ARDS.

approaches differ from ours in their exclusive use of timing of onset and resolution of lung injury as inputs in the latent class models; whether consideration of additional clinical data points or biomarkers as class-defining variables would lead to identification of different subphenotypes remains unknown.

The finding of differential response to PEEP by ARDS subphenotype has face validity in view of findings from other studies that have shown interactions between PEEP response and ARDS severity. Findings from a meta-analysis of 2299 patients with ARDS (including the 549 patients from the ALVEOLI cohort) showed that those with a PaO_2/FiO2 ratio of less than 200 had a 5% lower hospital mortality with higher PEEP strategies compared with lower PEEP strategies (p=0·049).[29,30] These approaches contrast with ours in two important ways. First, although PaO_2/FiO2 ratio was considered in subphenotype identification, it was not one of the variables that contributed most prominently to the classification (figure 2). Second, the interaction between PEEP and subphenotype that we identified is quantitatively larger and was statistically significant in a much smaller sample size than in the meta-analysis, suggesting that the interaction between subphenotype and PEEP response is substantially stronger than that between the PaO_2/FiO2 ratio and PEEP response. It is important to emphasise that while the interactions identified between PEEP and clinical outcomes are potentially provocative, we are reticent to make any recommendations for clinical care on the basis of a subgroup analysis. Rather, we view these results as hypothesis-generating, and think that they lend support to the need for more rapid or bedside assays of the molecular phenotype of critically ill patients to validate these findings in future trials.

Our study has several strengths. First, the latent class models were generated independently in each of the two cohorts. Specifically, findings from the ARMA cohort were not considered in the modelling strategy in the ALVEOLI cohort. In view of this approach, the similarity of the findings in the two cohorts is noteworthy. Likewise, since clinical outcomes were not considered as class-defining variables, the strengths and consistency of the associations between subphenotype and clinical outcomes are striking. Second, because we studied patients within the framework of a randomised controlled trial, we were able to draw stronger conclusions about causal associations between treatment (with PEEP) and clinical outcomes, with the usual caveats regarding subgroup analyses. Third, by virtue of using data from patients enrolled in multicentre trials, the samples studied are from demographically diverse cohorts of patients with ARDS. Fourth, the two cohorts differed substantially on many clinical and biological measures (tables 1 and 2), strengthening the generalisability of our findings and making the similarity of the subphenotypes identified in the two cohorts more significant.

This study has some limitations. First, the patients included in these analyses were drawn from randomised controlled trials of ARDS; different subphenotypes might be present in less carefully selected populations of patients with ARDS. Second, the biomarkers included in these analyses were restricted to those that had been measured already in both cohorts. Although these biomarkers have value for prognosis and pathogenesis, other informative biomarkers have emerged in ARDS research over the past several years, including angiopoietin-2,[31–33] the receptor for advanced glycation endproducts,[34] club (formerly known as Clara) cell 16,[35] brain natriuretic peptide,[36] interleukin-1 receptor antagonist,[37] and others. Consideration of these

biomarkers, or of alternative genomic or metabolomic markers, might result in more comprehensive subphenotypes being identified or may lead to the recognition of these biomarkers as important classifiers. Third, analyses of possible classifying variables was restricted to the data obtained in the original studies; clinical variables such as alcohol use,[38] cigarette smoking,[39] or other comorbidities could contribute to subphenotype identification but were not available for this analysis. Likewise, much histopathologic variability has been shown within ARDS in autopsy series; whether or how consideration of pathology findings would influence subphenotype identification remains unknown.[40,41]

We suggest that these findings provide proof-of-concept that the clinical syndrome of ARDS contains distinct subphenotypes and should prompt future studies aimed at further elucidating these subphenotypes with comprehensive clinical and biological data. In view of the differential response to treatment by subphenotype identified here, this area of research has the potential to directly inform future randomised controlled trials of novel treatments for ARDS.

Contributors

CSC designed the study, participated in data cleaning and analysis, interpreted the data, and drafted and revised the manuscript. KD did the data cleaning and analysis, contributed to data interpretation, and critically revised the manuscript. LBW, BTT, and PEP provided the biomarker data, contributed to data interpretation, and critically revised the manuscript. MAM contributed to study design, data analysis and interpretation, and critically revised the manuscript. All authors had access to the raw data. All authors provided final approval of the version to be published.

Declaration of interests

CSC reports grants from NIH/NHLBI during the conduct of the study, and grants from GlaxoSmithKline and personal fees from Cerus, outside the submitted work. BTT reports grants from the NHLBI during the conduct of the study. MAM reports grants from NHLBI during the conduct of the study, and grants from GlaxoSmithKline, data safety monitoring board fees from RocheGenetec, personal fees from Cerus, personal fees from French Society of Intensive Care, outside the submitted work. All other authors declare that they have no competing interests.

Acknowledgments

This work was supported by contracts (NO1-HR 46054-46064) with the NHLBI. CSC was supported by HL090833 and 110969. MAM was supported by HL 51856. LBW was supported by HL103836 and HL112656.

References

1 Ashbaugh DG, Bigelow DB, Petty TL, Levine BE. Acute respiratory distress in adults. *Lancet* 1967; **2:** 319–23.

2 Bernard GR, Artigas A, Brigham KL, et al. The American-European Consensus Conference on ARDS. Definitions, mechanisms, relevant outcomes, and clinical trial coordination. *Am J Respir Crit Care Med* 1994; **149:** 818–24.

3 Ranieri VM, Rubenfeld GD, Thompson BT, et al. Acute respiratory distress syndrome: the Berlin Definition. *JAMA* 2012; **307:** 2526–33.

4 Calfee CS, Eisner MD, Ware LB, et al. Trauma-associated lung injury differs clinically and biologically from acute lung injury due to other clinical disorders. *Crit Care Med* 2007; **35:** 2243–50.

5 Tejera P, Meyer NJ, Chen F, et al. Distinct and replicable genetic risk factors for acute respiratory distress syndrome of pulmonary or extrapulmonary origin. *J Med Genet* 2012; **49:** 671–80.

6 Frank AJ, Thompson BT. Pharmacological treatments for acute respiratory distress syndrome. *Curr Opin Crit Care* 2010; **16:** 62–68.

7 Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat Med* 2012; **18:** 716–25.

8 Corren J, Lemanske RF, Hanania NA, et al. Lebrikizumab treatment in adults with asthma. *N Engl J Med* 2011; **365:** 1088–98.

9 Wong HR, Cvijanovich NZ, Allen GL, et al. Validation of a gene expression-based subclassification strategy for pediatric septic shock. *Crit Care Med* 2011; **39:** 2511–17.

10 Iskander KN, Osuchowski MF, Stearns-Kurosawa DJ, et al. Sepsis: multiple abnormalities, heterogeneous responses, and evolving understanding. *Physiol Rev* 2013; **93:** 1247–88.

11 McLachlan G, Peel D. Finite mixture models. New York: John Wiley & Sons, 2000.

12 Ezzedine K, Le Thuaut A, Jouary T, Ballanger F, Taieb A, Bastuji-Garin S. Latent class analysis of a series of 717 patients with vitiligo allows the identification of two clinical subtypes. *Pigment Cell Melanoma Res* 2014; **27:** 134–39.

13 Rindskopf R, Rindskopf W. The value of latent class analysis in medical diagnosis. *Stat Med* 1986; **5:** 21–27.

14 Depner M, Fuchs O, Genuneit J, et al. Clinical and epidemiologic phenotypes of childhood asthma. *Am J Respir Crit Care Med* 2014; **189:** 129–38.

15 The Acute Respiratory Distress Syndrome Network. 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; **342:** 1301–08.

16 The Acute Respiratory Distress Syndrome Network. Ketoconazole for early treatment of acute lung injury and acute respiratory distress syndrome: a randomized controlled trial. *JAMA* 2000; **283:** 1995–2002.

17 The Acute Respiratory Distress Syndrome Network. Randomized, placebo-controlled trial of lisofylline for early treatment of acute lung injury and acute respiratory distress syndrome. *Crit Care Med* 2002; **30:** 1–6.

18 The Acute Respiratory Distress Syndrome Network. Higher versus lower positive end-expiratory pressures in patients with the acute respiratory distress syndrome. *N Engl J Med* 2004; **351:** 327–36.

19 Ware LB, Eisner MD, Thompson BT, Parsons PE, Matthay MA. Significance of von Willebrand factor in septic and nonseptic patients with acute lung injury. *Am J Respir Crit Care Med* 2004; **170:** 766–72.

20 Ware LB, Matthay MA, Parsons PE, Thompson BT, Januzzi JL, Eisner MD. Pathogenetic and prognostic significance of altered coagulation and fibrinolysis in acute lung injury/acute respiratory distress syndrome. *Crit Care Med* 2007; **35:** 1821–28.

21 Parsons PE, Eisner MD, Thompson BT, et al. Lower tidal volume ventilation and plasma cytokine markers of inflammation in patients with acute lung injury. *Crit Care Med* 2005; **33:** 1–6.

22 Parsons PE, Matthay MA, Ware LB, Eisner MD. Elevated plasma levels of soluble TNF receptors are associated with morbidity and mortality in patients with acute lung injury. *Am J Physiol Lung Cell Mol Physiol* 2005; **288:** L426–31.

23 Eisner MD, Parsons P, Matthay MA, Ware L, Greene K. Plasma surfactant protein levels and clinical outcomes in patients with acute lung injury. *Thorax* 2003; **58:** 983–88.

24 Calfee CS, Eisner MD, Parsons PE, et al. Soluble intercellular adhesion molecule-1 (sICAM-1) and clinical outcomes in patients with acute lung injury. *Intensive Care Med* 2008; **35:** 248–57.

25 Ware LB, Koyama T, Billheimer DD, et al. Prognostic and pathogenetic value of combining clinical and biochemical indices in patients with acute lung injury. *Chest* 2010; **137:** 288–96.

26 Lanza ST, Tan X, Bray BC. Latent class analysis with distal outcomes: a flexible model-based approach. *Struct Equ Modeling* 2013; **20:** 1–26.

27 Shah RJ, Diamond JM, Cantu E, et al. Latent class analysis identifies distinct phenotypes of primary graft dysfunction after lung transplantation. *Chest* 2013; **144:** 616–22.

28 Reilly JP, Bellamy S, Shashaty MG, et al. Heterogeneous phenotypes of the acute respiratory distress syndrome after major trauma. *Ann Am Thorac Soc* 2014; published online Apr 8. DOI:10.1513/AnnalsATS.201308-280OC.

29 Briel M, Meade M, Mercat A, et al. Higher vs lower positive end-expiratory pressure in patients with acute lung injury and acute respiratory distress syndrome: systematic review and meta-analysis. *JAMA* 2010; **303:** 865–73.

30 Rubenfeld GD. How much peep in acute lung injury. *JAMA* 2010; **303:** 883–84.

31 Calfee CS, Gallagher D, Abbott J, Thompson BT, Matthay MA. Plasma angiopoietin-2 in clinical acute lung injury: prognostic and pathogenetic significance. *Crit Care Med* 2012; **40:** 1731–37.

32 Agrawal A, Matthay MA, Kangelaris KN, et al. Plasma angiopoietin-2 predicts the onset of acute lung injury in critically ill patients. *Am J Respir Crit Care Med* 2013; **187:** 736–42.

33 Meyer NJ, Li M, Feng R, et al. ANGPT2 genetic variant is associated with trauma-associated acute lung injury and altered plasma angiopoietin-2 isoform ratio. *Am J Respir Crit Care Med* 2011; **183:** 1344–53.

34 Calfee CS, Ware LB, Eisner MD, et al. Plasma receptor for advanced end-products and clinical outcomes in acute lung injury. *Thorax* 2008; **63:** 1083–89.

35 Kropski JA, Fremont RD, Calfee CS, Ware LB. Clara cell protein (CC16), a marker of lung epithelial injury, is decreased in plasma and pulmonary edema fluid from patients with acute lung injury. *Chest* 2009; **135:** 1440–47.

36 Fremont RD, Koyama T, Calfee CS, et al. Acute lung injury in patients with traumatic injuries: utility of a panel of biomarkers for diagnosis and pathogenesis. *J Trauma* 2010; **68:** 1121–27.

37 Meyer NJ, Feng R, Li M, et al. IL1RN coding variant is associated with lower risk of acute respiratory distress syndrome and increased plasma IL-1 receptor antagonist. *Am J Respir Crit Care Med* 2013; **187:** 950–59.

38 Moss M, Bucher B, Moore FA, Moore EE, Parsons PE. The role of chronic alcohol abuse in the development of acute respiratory distress syndrome in adults. *JAMA* 1996; **275:** 50–54.

39 Calfee CS, Matthay MA, Eisner MD, et al. Active and passive cigarette smoking and acute lung injury after severe blunt trauma. *Am J Respir Crit Care Med* 2011; **183:** 1660–65.

40 Thille AW, Esteban A, Fernandez-Segoviano P, et al. Comparison of the Berlin definition for acute respiratory distress syndrome with autopsy. *Am J Respir Crit Care Med* 2013; **187:** 761–67.

41 Thille AW, Esteban A, Fernandez-Segoviano P, et al. Chronology of histological lesions in acute respiratory distress syndrome with diffuse alveolar damage: a prospective cohort study of clinical autopsies. *Lancet Respir Med* 2013; **1:** 395–401.