Divide and conquer: identifying acute respiratory distress syndrome subphenotypes

Manu Shankar-Hari, Daniel F McAuley

The acute respiratory distress syndrome (ARDS) definition identifies patients with acute onset hypoxaemia and respiratory failure, who have bilateral opacities on chest radiograph that are not fully explained by cardiac failure or fluid overload. ARDS is a common illness that accounts for approximately 10% of critical care admissions and 20% of patients requiring mechanical ventilation. The hospital mortality in patients with ARDS remains high, increasing from approximately 35% for those with mild disease to 46% for those with severe ARDS. This high mortality has remained relatively unchanged in the last 20 years. To date, despite decades of research, there is no pharmacological treatment that can modify the underlying biological mechanisms implicated in ARDS and improve patient outcomes. Within ARDS populations, there is substantial biological and outcome heterogeneity, with observed differences in dominant pathogenic mechanisms, treatment responses and outcomes. Identifying ARDS subphenotypes based on pathogenic mechanisms that determine treatment responses irrespective of ARDS severity is defined as predictive enrichment. The identification of such ARDS subphenotypes will enable improved trial design in ARDS by selecting patients based on responder characteristics to therapeutic interventions, hopefully resulting in improved outcomes.

In *Thorax*, Bos et al report a cohort study in 700 ARDS patients, testing the hypothesis that ARDS subgroups exist due to differences in biological characteristics. In this retrospective analysis of a prospectively collected cohort, 20 biomarkers were selected to represent inflammation, coagulation and endothelial activation, as hallmarks of ARDS biology. The dataset was divided into a training cohort (n=454 patients) and validation cohort (n=246 patients), based on the study recruitment period. Cluster analysis was used to identify homogeneous ARDS subphenotypes in the training cohort. The most predictive biomarkers were then confirmed in the validation cohort. These biological clusters were then linked to clinical and outcome characteristics of ARDS patients to derive clinical subphenotypes, namely reactive and uninfamed. These two clinical ARDS subphenotypes differed in terms of illness severity and critical care mortality, with the reactive group having a greater risk of death.

A key question for the reader is whether these associations are spurious or indirect or causal? Cluster analysis methods generate different results dependent on the variables chosen for identifying similarities between patients and the method of clustering. Bos et al chose biomarker characteristics as the variables on which the groups should be similar and used Ward’s method of agglomerative hierarchical clustering to identify two potentially generalisable ARDS clusters. Hierarchical clustering is a commonly used iterative method to identify homogeneous groups or clusters based on specific characteristics. The basic algorithm starts with assigning each ARDS patient a ‘value’ based on their individual biomarker profile. Then patients with similar ‘values’ are grouped together to form clusters. The underlying principle is that ARDS patients within each cluster will have similar biomarker profiles and that between clusters biomarker profiles will be different. Depending on the parameters specified, the same dataset can result in potentially different results with different clustering algorithms and there are no universally agreed optimal rule(s) for clustering. Another potential limitation is that only patients with data on all chosen biomarkers were included and missing data in clinical variables were imputed, which has the potential for selection and information bias. The

blood sampling window for biomarker measurement in this cohort was wide and drawn either on the day of ARDS diagnosis or the day before or the day after, challenging the time-based arguments for causal relationships. Despite these challenges, Bos et al provide important data with strong associations, that are consistent with our current knowledge, have biological plausibility and external validity.

Calfée and colleagues have led the field in defining ARDS subphenotypes. Using latent class analysis of clinical and biomarker data from patients enrolled in ARDS randomised controlled trials, Calfee et al have originally identified two ARDS subphenotypes. The reactive subphenotype identified by Bos et al shares many of the features of the hyperinflammatory ARDS subphenotype reported previously, although the proportion of patients in the reactive group is much higher than the hyperinflammatory subphenotype. This suggests that the hyperinflammatory and reactive groups may represent a similar subphenotype, although this is unproven. The findings from Bos et al are significant in that they have identified comparable subphenotypes in an observational cohort of patients with ARDS using a different analytic approach. While Calfee et al identified these ARDS subphenotypes using clinical and biomarker data, Bos et al identified them purely on biomarker data. The blood sampling window for biomarker measurement in these studies was defined from trial enrolment (which could be up to 2 days after meeting ARDS criteria in the ARDSnet trials), which also challenges the time-based arguments for causal relationships. Furthermore, it is possible that in the study by Bos et al, ARDS subjects were sampled earlier than in the ARDSnet trials, which may be a potential explanation for the higher proportion of ‘reactive’ subphenotype in this study. It would be important to test whether similar subphenotypes emerge after harmonising these different study datasets and performing both cluster and latent class analyses. Table 1 provides a comparative summary of these three studies.

Several important questions remain unanswered. First, assuming the hyperinflammatory/reactive subphenotype represents a common subphenotype, further work is needed to identify the key discriminant makers that reliably define this ARDS subset. Ideally, a minimal dataset of variables could be identified to efficiently achieve this. Second, although it remains unknown if ARDS subphenotypes

1Department of Critical Care Medicine, Guy’s and St Thomas’ NHS Foundation Trust, London, UK
2Division of Infection and Immunity, King’s College London, London, London, UK
3Centre for Experimental Medicine, Wellcome-Wolfson Institute for Experimental Medicine, Queen’s University of Belfast, Belfast, Northern Ireland
4Regional Intensive Care Unit, Royal Victoria Hospital, Belfast, Northern Ireland

Correspondence to Professor Daniel F McAuley, Centre for Experimental Medicine, Wellcome-Wolfson Institute for Experimental Medicine, Queen’s University of Belfast, Belfast BT7 1NN, UK; d.f.mcauley@qub.ac.uk
Table 1  Summary of studies that report ARDS subphenotypes

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Bos et al&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Calfee C et al&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Famous et al&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>700</td>
<td>1022</td>
<td>1000</td>
</tr>
<tr>
<td>Study design</td>
<td>Observational cohort</td>
<td>RCT analysed as cohort</td>
<td>RCT analysed as cohort</td>
</tr>
<tr>
<td>ARDS P/F criteria</td>
<td>≤300</td>
<td>&lt;300</td>
<td>&lt;300</td>
</tr>
<tr>
<td>Blood sampling</td>
<td>Around ARDS diagnosis</td>
<td>At baseline</td>
<td>At baseline</td>
</tr>
<tr>
<td>Biomarkers used for deriving subphenotypes</td>
<td>Lung epithelial: none</td>
<td>Lung epithelial: SP-D</td>
<td>Lung epithelial: SP-D</td>
</tr>
<tr>
<td></td>
<td>Endothelial: E-selectin; P-selectin; ANG1/2</td>
<td>Endothelial: ICAM-1; vWF</td>
<td>Endothelial: ICAM-1; vWF, ANG-2 and RAGE</td>
</tr>
<tr>
<td></td>
<td>Coagulation: antithrombin; D-Dimer; tPA; PAI-1; Inflammation: fractalkine; GM-CSF; ICAM-1; IFN-γ; IL-1β; IL-6; IL-8; IL-10; IL-13; TNF-α; MMP-8; TIMP-1;</td>
<td>Coagulation: protein C; PAI-1; Inflammation: sTNFR-1; IL-6; IL-8</td>
<td>Coagulation: protein C; PAI-1; Inflammation: sTNFR-1; IL-6; IL-8</td>
</tr>
<tr>
<td>Clinical variables used for deriving subphenotypes</td>
<td>None</td>
<td>Age, gender, ethnicity, BMI, respiratory*; cardiovascular; creatinine; urine output; bilirubin; temperature; haematocrit; WBC count; sodium; glucose; albumin; platelets; bicarbonate; aetiology of ARDS†</td>
<td>Age, gender, ethnicity, BMI, respiratory6; cardiovascular; creatinine; urine output; bilirubin; temperature; haematocrit; WBC count; sodium; glucose; albumin; platelets; bicarbonate; aetiology of ARDS†</td>
</tr>
<tr>
<td>Analytical approach to derive ARDS subsets</td>
<td>Cluster analyses based only on biomarker data</td>
<td>Latent class analyses based grouping based on clinical and biomarker data</td>
<td>Latent class analyses based grouping based on clinical and biomarker data</td>
</tr>
<tr>
<td>ARDS subset (prevalence %)</td>
<td>Reactive phenotype (58.0%) versus Uninflamed (42.0%)</td>
<td>Hyperinflammatory (29.4%) versus Phenotype 1 (70.6%)</td>
<td>Hyperinflammatory (27.3%) versus Phenotype 1 (72.7%)</td>
</tr>
<tr>
<td>Mortality (%) by ARDS subset</td>
<td>Reactive phenotype=36.8% versus Uninflamed=14.9%</td>
<td>Hyperinflammatory=47.3% versus Phenotype 1=19.4%</td>
<td>Hyperinflammatory=45.0% versus Phenotype 1=22.0%</td>
</tr>
<tr>
<td>Discriminant markers between phenotypes</td>
<td>IL-6; IFN-γ; ANG1/2; PAI-1</td>
<td>IL-6; sTNFR1; vasopressor use; IL-8; HCO3</td>
<td>IL-8; sTNFR1; vasopressor use; HCO3; minute ventilation</td>
</tr>
</tbody>
</table>

The table shows the summary of three recent studies that report ARDS subphenotypes. The Respiratory system variables* included minute ventilation, mean airway pressure, plateau pressure, respiratory rate, tidal volume, positive end-expiratory pressure; partial pressure PaO2 of carbon dioxide (PaCO2) and PaO2/FiO2 ratio.

The Cardiovascular† system variables include highest heart rate, lowest systolic blood pressure and vasopressor use.

The Lung epithelial† system variables included minute ventilation, mean airway pressure, plateau pressure, respiratory rate, tidal volume, positive end-expiratory pressure; partial pressure PaO2 of carbon dioxide (PaCO2) and PaO2/FiO2 ratio.

The table summary includes:

- **Characteristics linked to both risk of poor outcomes and response to the tested treatment.** This will enable personalised care of patients with ARDS.
  - **Contributors** MS-H wrote the first draft. MS-H and DFM critically reviewed the manuscript for important intellectual content and agreed on the final submitted version of the manuscript.
  - **Funding** MS-H is supported by the National Institutes of Health Research Clinician Scientist Award (NIHR-CS-2016-16-011).
  - **Disclaimer** The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the UK National Institute for Health Research or the Department of Health.
  - **Competing interests** None declared.
  - **Provenance and peer review** Commissioned; externally peer reviewed.

To cite Shankar-Hari M, McAuley DF. Thorax Published Online First: [please include Day Month Year]. doi:10.1136/thoraxjnl-2017-210422

> http://dx.doi.org/10.1136/thoraxjnl-2016-209719.R3


REFERENCES

5. Pham T, Rubenfeld GD. Fifty Years of Research in ARDS. The Epidemiology of Acute Respiratory Distress


