Estimating the attributable fraction of mortality from acute respiratory distress syndrome to inform enrichment in future randomised clinical trials


ABSTRACT

Background Efficiency of randomised clinical trials of acute respiratory distress syndrome (ARDS) depends on the fraction of deaths attributable to ARDS (AFARDS) to which interventions are targeted. Estimates of AFARDS in subpopulations of ARDS could improve design of ARDS trials.

Methods We performed a matched case-control study using the Large observational study to Understand the Global impact of Severe Acute respiratory Failure cohort. Primary outcome was intensive care unit mortality. We used nearest neighbour propensity score matching without replacement to match ARDS to non-ARDS populations. We derived two separate AFARDS estimates by matching patients with ARDS to patients with non-acute hypoxaemic respiratory failure (non-AHRF) and to patients with AHRF with unilateral infiltrates only (AHRF-UL). We also estimated AFARDS in subgroups based on severity of hypoxaemia, number of lung quadrants involved and hyperinflammatory versus hypoinflammatory phenotypes. Additionally, we derived AFARDS estimates by matching patients with AHRF to non-AHRF controls, and AFARDS-UL estimates by matching patients with AHRF-UL to non-AHRF controls.

Results Estimated AFARDS was 20.9% (95% CI 10.5% to 31.4%) when compared with AHRF-UL controls and 38.0% (95% CI 34.4% to 41.6%) compared with non-AHRF controls. Within subgroups, estimates for AFARDS compared with AHRF-UL controls were highest in patients with severe hypoxaemia (41.1% (95% CI 25.2% to 57.1%)), in those with four quadrant involvement on chest radiography (28.9% (95% CI 13.4% to 44.3%)) and in the hyperinflammatory subphenotype (26.8% (95% CI 6.9% to 46.7%)). Estimated AFARDS was 33.8% (95% CI 30.5% to 37.1%) compared with non-AHRF controls. Estimated AFARDS was 21.3% (95% CI 312.8% to 29.7%) compared with non-AHRF controls.

Conclusions Overall AFARDS mean values were between 20.9% and 38.0%, with higher AFARDS seen with severe hypoxaemia, four quadrant involvement on chest radiography and hyperinflammatory ARDS.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The excess mortality—or attributable fraction (AF)—due to acute respiratory distress syndrome (ARDS) has been estimated to range between 15% and 37%.
⇒ We do not know how this varies by severity of hypoxaemia, radiographic findings and ARDS subphenotype.

WHAT THIS STUDY ADDS

⇒ We observed a dose-response increase in AFARDS with severity of hypoxaemia, quadrants of radiographic involvement and that AFARDS was higher in the hyperinflammatory compared with the hypoinflammatory subphenotype of ARDS.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ We highlight ARDS subpopulations that can inform enrichment options in randomised clinical trials.

INTRODUCTION

Acute respiratory distress syndrome (ARDS) refers to acute hypoxaemic respiratory failure (AHRF) occurring within 1 week of a known clinical insult, with bilateral opacities on chest radiography that are not fully explained by effusions, lobar/lung collapse or nodules. Treatments with biological plausibility and strong supporting preclinical evidence, when tested within randomised clinical trials (RCTs), often report statistically indeterminate results (i.e., uncertainty highlighted by non-significant results of two-tailed tests, rather than proof of no difference between treatments (negative)). Addressing this issue remains a major clinical and methodological challenge.

There are several explanations for statistically indeterminate RCT results, aside from testing of ineffective treatments. First, as ARDS is a heterogeneous syndrome, RCTs may consist of participants who either benefit, have no effect or are harmed by the tested intervention. This explanation is supported by observations that distinct subphenotypes of ARDS respond differentially to treatments. Second, there are several design issues within ARDS RCTs. This explanation is supported by observations that sample size calculations

Correspondence to Professor Manu Shankar-Hari, The University of Edinburgh, Edinburgh, UK; manu.shankar-hari@ed.ac.uk

Received 23 March 2023 Accepted 3 July 2023 Published Online First 26 July 2023
overestimate control arm event rates, and the expected average treatment effect.\textsuperscript{12}

In this manuscript, we explore another explanation—variation in the excess fraction of mortality attributable to ARDS (AF\textsubscript{ARDS})\textsuperscript{13,14} in RCT participants. Patients with ARDS may die from ARDS (ie, AF\textsubscript{ARDS}) or with ARDS (ie, death may be due to risk factors like comorbidities and/or other organ dysfunction during critical illness). If we explicitly link the eligibility criteria of ARDS RCTs to AF\textsubscript{ARDS} estimates (ie, the excess proportion of deaths from ARDS), then efficiency of ARDS RCTs would be increased from the generic predictive and prognostic enrichment alongside increase in average treatment effect. We hypothesised that AF\textsubscript{ARDS} will vary by severity of hypoxaemia as per the Berlin ARDS definitions,\textsuperscript{1,15} by number of quadrants affected on chest radiography,\textsuperscript{16} and by subphenotype. Our hypothesis was informed by the following observations: first, two small cohort studies (online supplemental eTable 1) indicated that AF\textsubscript{ARDS} ranges between 15\% and 37\%\textsuperscript{17,18}; second, in the Berlin ARDS definitions, ARDS outcomes worsened with increasing severity of hypoxaemia.\textsuperscript{1,3} Third, the hyperinflammatory ARDS subphenotype had higher mortality and greater treatment effect within RCTs.\textsuperscript{3-11} Recently, these ARDS subphenotypes were identified within the Large observational study to UNderstand the Global impact of Severe Acute respiratory FailurE (LUNG-SAFE) cohort using machine learning models,\textsuperscript{19} are available as predefined categories within the LUNG-SAFE dataset and currently there are no AF\textsubscript{ARDS} estimates for them. Given recent proposals to include patients with AHRF (including those with unilateral infiltrates)\textsuperscript{20} within an expanded ARDS case definition,\textsuperscript{21} we also compared AF\textsubscript{AHRF} and AF\textsubscript{ARDS}.

METHODS

Data source

Our data source was the well-described LUNG-SAFE dataset. We summarise key elements of the LUNG-SAFE study design and data collection in online supplemental eMethods 1. National coordinators and site investigators of the LUNG-SAFE study are listed in the online supplemental file 1. AHRF was defined as concurrent presence of: (a) arterial oxygen tension:inspired fraction of oxygen (PaO\textsubscript{2}:FiO\textsubscript{2}) ratio \textless 300 mm Hg; (b) new pulmonary parenchymal abnormalities (either unilateral or bilateral) on chest radiography and (c) ventilatory support with continuous positive airway pressure or expiratory positive airway pressure or positive end expiratory pressure \textgreater 5 cmH\textsubscript{2}O. The diagnosis of ARDS in LUNG-SAFE studies was made by a computer algorithm in the analysis phase of the study using the ‘raw’ data that made up the various components of the Berlin ARDS definition.\textsuperscript{13}

STUDY POPULATION

Selection criteria of AHRF/ARDS cohort reported in this manuscript were described previously by Pham et al.\textsuperscript{20} We defined four populations for our matched cohort study, after excluding patients with congestive heart failure: (1) patients with ARDS-AHRF who met the Berlin ARDS criteria, (2) patients with AHRF who met criteria for AHRF (and therefore includes all patients with ARDS), (3) patients with AHRF with unilateral infiltrates (AHRF-UL)—met criteria for AHRF but not ARDS and (4) non-AHRF controls—patients receiving non-invasive or invasive ventilation who did not meet criteria for AHRF (figure 1). The LUNG-SAFE study collected only the following variables for non-AHRF controls—age, sex, intensive care unit (ICU) length of stay and ICU mortality.

ARDS subphenotypes have recently been assigned in the LUNG-SAFE cohort using a clinical classifier model with a limited selection of predictor variables\textsuperscript{19} (online supplemental eMethods 2). Of note, this model used optimised probability cutoffs and did not use latent class analysis to assign subphenotypes. Patients without ARDS do not have subphenotype allocation.

There were no patients who had unilateral infiltrates without AHRF, to act as controls for AF\textsubscript{AHRF-UL} range estimates, similar to the AF\textsubscript{ARDS} range estimates we report.

ANALYSES FRAMEWORK

The primary exposure was either ARDS, or AHRF, or AHRF-UL. The primary outcome was ICU mortality, as one of the most reported outcomes in ARDS RCTs.\textsuperscript{3} Predefined enrichment categories within the primary exposures were severity of hypoxaemia, number of quadrants affected on chest radiography and subphenotypes. Due to the low proportion of missing data (online supplemental eTable 2a), without any discernible pattern of missingness, these data were assumed to be missing at random, and complete-case analyses were used for all models.

Estimation of AF\textsubscript{ARDS} requires careful selection of controls and consideration of potential confounding variables. We estimated propensity scores for the exposures using logistic regression. We used nearest neighbour matching without replacement to match exposed patients to controls. With this approach, the mortality in the control groups within each prespecified ARDS severity category (severity of hypoxaemia, radiology and subphenotype) would also vary based on matching, enabling estimation of variation in AF\textsubscript{ARDS} within these categories, along with AF\textsubscript{ARDS}.

Propensity score models and scenarios

Model 1 scenario

AF\textsubscript{ARDS} could be reduced with treatment to mortality seen in ICU patients of similar age and sex without AHRF (one patient with ARDS was matched to two non-AHRF controls).

Model 2 scenario

AF\textsubscript{ARDS} could be reduced with treatment to mortality seen in ICU patients with AHRF after accounting for variables commonly considered as part of eligibility criteria in ARDS RCTs at the time of randomisation such as age, sex, number of comorbidities, receipt of invasive mechanical ventilation and illness severity (one patient with ARDS was matched to one AHRF-UL control).

Model 3 scenario

AF\textsubscript{AHRF} could be reduced with treatment to mortality seen in ICU patients of similar age and sex without AHRF (one patient with AHRF was matched to two non-AHRF controls).

Model 4 scenario

AF\textsubscript{AHRF-UL} could be reduced with treatment to mortality seen in ICU patients of similar age and sex without AHRF scenario (one patient with AHRF-UL was matched to two non-AHRF controls).

Additional rationale for matching methods, covariate selection and assessment of each of four different propensity score models are available in figure 1, online supplemental eMethods 3 and online supplemental eFigure 1.

We used four separate logistic regression models to estimate AF (AF\textsubscript{ARDS}, AF\textsubscript{AHRF}, AF\textsubscript{AHRF-UL}).
Critical care

We used model 1 and model 2 to estimate the variation in $AF_{\text{ARDS}}$ by severity of hypoxaemia categories (mild ($\text{PaO}_2:\text{FiO}_2 > 200$ mm Hg); moderate ($\text{PaO}_2:\text{FiO}_2$ 100–200 mm Hg) or severe ($\text{PaO}_2:\text{FiO}_2 < 100$ mm Hg)); the number of quadrants involved radiographically in the first 48 hours after ICU admission (two, three, or four), and the hyperinflammatory versus hypoinflammatory ARDS subphenotypes within LUNG-SAFE dataset.\textsuperscript{19}

Estimation of $AF_{\text{ARDS}}$ range
We used model 1 and model 2 to estimate the variation in $AF_{\text{ARDS}}$ by severity of hypoxaemia categories (mild, moderate, severe), and the number of quadrants involved radiographically in the first 48 hours after ICU admission (two, three or four).

Estimation of $AF_{\text{AHHRF}}$
We used model 3 to estimate the variation in $AF_{\text{AHHRF}}$ by severity of hypoxaemia categories (mild, moderate, severe), and the number of quadrants involved radiographically in the first 48 hours after ICU admission (two, three or four).

Estimation of $AF_{\text{AHHRF-UL}}$
We used model 4 to estimate the variation in $AF_{\text{AHHRF-UL}}$ by severity of hypoxaemia categories (mild, moderate, severe), and the number of quadrants involved radiographically in the first 48 hours after ICU admission (one, two).

Simulating $AF_{\text{ARDS}}$-based sample size estimates for different enrichment subgroups
To illustrate how $AF_{\text{ARDS}}$ can influence sample size estimation, we compared predicted sample size estimates from 28 published ARDS RCTs\textsuperscript{22–49} that used mortality as primary outcome (identified in our previous systematic review\textsuperscript{3}) to simulated sample size estimates. We simulated sample size calculations for the scenario where $AF_{\text{ARDS}} = 100\%$, and for estimates of $AF_{\text{ARDS}}$ from model 1 stratified by severity of hypoxaemia (mild, moderate or severe), maximum number of quadrants involved on chest radiography at 48 hours (two, three or four) and (c) subphenotype of ARDS (hyperinflammatory or hypoinflammatory). For all simulations, RCT control event rate was fixed at 40%, alpha at 0.05 and power at 0.80.
Table 1  Baseline characteristics of ARDS, AHRF, AHRF-UL and corresponding propensity matched control populations to derive attributable fraction

<table>
<thead>
<tr>
<th>Risk factor for ARDS n (%)</th>
<th>Model 1 AF_ARDs (n=2653)</th>
<th>Model 2 AF_AHRF (n=851)</th>
<th>Model 3 AF_AHRF-UL (n=851)</th>
<th>Model 4 AF_AHRF-UL (n=1702)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pneumonia</td>
<td>1574 (59.3)</td>
<td>510 (59.9)</td>
<td>433 (50.9)</td>
<td>2007 (57.3)</td>
</tr>
<tr>
<td>Extrapulmonary sepsis</td>
<td>421 (15.9)</td>
<td>141 (16.6)</td>
<td>115 (13.5)</td>
<td>536 (15.3)</td>
</tr>
<tr>
<td>Aspiration</td>
<td>386 (14.5)</td>
<td>114 (13.4)</td>
<td>157 (18.4)</td>
<td>543 (15.5)</td>
</tr>
<tr>
<td>Pancreatitis</td>
<td>57 (2.1)</td>
<td>16 (1.9)</td>
<td>11 (1.3)</td>
<td>68 (1.9)</td>
</tr>
<tr>
<td>Pulmonary vasculitis</td>
<td>12 (0.5)</td>
<td>2 (0.2)</td>
<td>0 (0)</td>
<td>12 (0.3)</td>
</tr>
<tr>
<td>Trauma</td>
<td>108 (4.1)</td>
<td>37 (4.3)</td>
<td>48 (5.6)</td>
<td>156 (4.5)</td>
</tr>
<tr>
<td>Inhalation</td>
<td>69 (2.6)</td>
<td>28 (3.3)</td>
<td>22 (2.6)</td>
<td>91 (2.6)</td>
</tr>
<tr>
<td>Pulmonary contusion</td>
<td>82 (3.1)</td>
<td>28 (3.3)</td>
<td>34 (4.0)</td>
<td>116 (3.3)</td>
</tr>
<tr>
<td>Burns</td>
<td>7 (0.3)</td>
<td>3 (0.4)</td>
<td>2 (0.2)</td>
<td>9 (0.3)</td>
</tr>
<tr>
<td>Non-cardiogenic shock</td>
<td>201 (7.6)</td>
<td>75 (8.8)</td>
<td>52 (6.1)</td>
<td>253 (7.2)</td>
</tr>
<tr>
<td>Drowning</td>
<td>2 (0.1)</td>
<td>1 (0.1)</td>
<td>0 (0)</td>
<td>2 (0.1)</td>
</tr>
<tr>
<td>Drug overdose</td>
<td>47 (1.8)</td>
<td>11 (1.3)</td>
<td>22 (2.6)</td>
<td>69 (2.0)</td>
</tr>
<tr>
<td>Transfusion-related</td>
<td>107 (4.0)</td>
<td>34 (4.0)</td>
<td>27 (3.2)</td>
<td>134 (3.8)</td>
</tr>
<tr>
<td>Other</td>
<td>67 (2.5)</td>
<td>17 (2.0)</td>
<td>45 (5.3)</td>
<td>112 (3.2)</td>
</tr>
<tr>
<td>No risk factor for ARDS</td>
<td>229 (8.6)</td>
<td>79 (9.3)</td>
<td>0 (0)</td>
<td>229 (6.5)</td>
</tr>
<tr>
<td>PaO2:FiO2 ratio on day 1 mean (SD)*</td>
<td>159 (67.4)</td>
<td>161 (67.6)</td>
<td>188 (63.6)</td>
<td>166 (67.8)</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF_{ARDS}</td>
<td>AF_{Non-AHRF}</td>
<td>AF_{ARDS}</td>
<td>AF_{AHRF}</td>
</tr>
<tr>
<td>ARDS (n=2653)</td>
<td>Non-AHRF (n=5306)</td>
<td>ARDS (n=851)</td>
<td>AHRF-UL (n=3504)</td>
</tr>
<tr>
<td>Subphenotype of ARDS n (%)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypoinflammatory</td>
<td>1966 (74.1)</td>
<td>616 (72.4)</td>
<td>–</td>
</tr>
<tr>
<td>Hyperinflammatory</td>
<td>687 (25.9)</td>
<td>235 (27.6)</td>
<td>–</td>
</tr>
<tr>
<td>SOFA score day 1 mean (SD)</td>
<td>9.39 (4.08)</td>
<td>9.55 (4.19)</td>
<td>8.69 (3.97)</td>
</tr>
<tr>
<td>Non-respiratory SOFA score day 1*</td>
<td>6.48 (3.96)</td>
<td>6.65 (4.06)</td>
<td>6.04 (3.87)</td>
</tr>
<tr>
<td>Ventilated on day 1 n (%)</td>
<td>2151 (81.1)</td>
<td>700 (82.3)</td>
<td>698 (82.0)</td>
</tr>
<tr>
<td>Duration of invasive mechanical ventilation mean (SD)*</td>
<td>12.0 (12.9)</td>
<td>12.2 (12.8)</td>
<td>10.6 (12.6)</td>
</tr>
<tr>
<td>ICU LOS mean (SD)</td>
<td>14.2 (14.0)</td>
<td>6.2 (8.2)</td>
<td>12.7 (13.8)</td>
</tr>
<tr>
<td>ICU mortality n (%)</td>
<td>906 (34.2)</td>
<td>730 (13.8)</td>
<td>293 (34.4)</td>
</tr>
<tr>
<td>Hospital LOS mean (SD)*</td>
<td>23.1 (20.8)</td>
<td>23.6 (20.7)</td>
<td>22.6 (20.7)</td>
</tr>
<tr>
<td>Hospital mortality n (%)*</td>
<td>1044 (39.4)</td>
<td>342 (40.2)</td>
<td>259 (30.4)</td>
</tr>
</tbody>
</table>

AF_{ARDS} in model 1 was estimated by matching 2653 patients with ARDS to non-AHRF controls that were propensity score balanced on age and sex in a 1:2 ratio. AF_{ARDS} in model 2 was estimated by matching 851 patients with ARDS to AHRF-UL controls in a 1:1 ratio and propensity score balanced on age, sex, number of comorbidities and ventilation status on day 1. AF_{AHRF} in model 3 was estimated by matching 3504 AHRF cases to non-AHRF controls that were propensity score balanced on age and sex in a 1:2 ratio. AF_{AHRF-UL} in model 4 was estimated by matching 851 AHRF-UL cases to non-AHRF controls that were propensity score balanced on age and sex in a 1:2 ratio.

Postmatching standardised mean differences for the matching covariates in each model are shown in online supplemental eTable 2A. AF, attributable fraction; AHRF, acute hypoxic respiratory failure; AHRF-UL, acute hypoxic respiratory failure with unilateral infiltrates only; ARDS, acute respiratory distress syndrome; BMI, body mass index; ICU, intensive care unit; LOS, length of stay; NMV, non-invasive mechanical ventilation; PaO2:Fio2, arterial oxygen tension:inspired fraction of oxygen; SOFA, Sequential Organ Failure Assessment.

AFARDS in model 1 was estimated by matching 2653 patients with ARDS to non-AHRF controls that were propensity score balanced on age and sex in a 1:2 ratio. AFARDS in model 2 was estimated by matching 851 patients with ARDS to AHRF-UL controls in a 1:1 ratio and propensity score balanced on age, sex, number of comorbidities and ventilation status on day 1. AFARDS in model 3 was estimated by matching 3504 AHRF cases to non-AHRF controls that were propensity score balanced on age and sex in a 1:2 ratio. AFARDS in model 4 was estimated by matching 851 AHRF-UL cases to non-AHRF controls that were propensity score balanced on age and sex in a 1:2 ratio.

Postmatching standardised mean differences for the matching covariates in each model are shown in online supplemental eTable 2A. AFARDS, attributable fraction; AHRF, acute hypoxic respiratory failure; AHRF-UL, acute hypoxic respiratory failure with unilateral infiltrates only; ARDS, acute respiratory distress syndrome; BMI, body mass index; ICU, intensive care unit; LOS, length of stay; NMV, non-invasive mechanical ventilation; PaO2:Fio2, arterial oxygen tension:inspired fraction of oxygen; SOFA, Sequential Organ Failure Assessment.

*Denotes variable with missing data. Further details of missing data are available in online supplemental eTable 2A.
We report the following sensitivity analyses: (i) unmatched analyses in all four models to assess how matching—and therefore, exclusion of controls—affect overall and subpopulation estimates and (ii) used hospital mortality as outcome measure in model 2—instead of ICU mortality—to assess how choice of mortality timepoint affected overall and subpopulation estimates for AFARDS.

The $\chi^2$ test was used to assess linear trends in ICU mortality across subpopulations, and to assess relationship between enrichment categories. Reported $p$ values are two-sided and $p$ values <0.05 were considered statistically significant. All analyses were performed using R V.3.4.2 ($\text{AFips}^{50}$, $\text{Matchit}^{51}$ and $\text{tidyverse}^{52}$ packages).

RESULTS

Among 12 906 admissions who received ventilatory support in the ICU, we identified 3504 eligible patients with AHRF; 2653 met the Berlin ARDS criteria and 851 patients had AHRF-UL (figure 1). Missing data are summarised in online supplemental eTable 2a. Baseline characteristics and outcomes for patients with non-AHRF are summarised in online supplemental eTable 2b. Patients with mild hypoxaemia most often had two quadrant infiltrates, while patients with severe hypoxaemia had four quadrant infiltrates (online supplemental eTable 3a). There was no association between severity of hypoxaemia and ARDS subphenotype (online supplemental eTable 3b).

Model 1 scenario

Model 1 scenario compared 2653 patients with ARDS matched to 5306 non-AHRF controls. Patients with ARDS had higher ICU mortality compared with controls (34.8% vs 13.8%) (table 1). Significant linear trends in mortality were seen with severity of hypoxaemia category (mild 28.6% vs 13.2%; moderate 33.2% vs 14.6%; severe 43.0% vs 13.1%; $\chi^2=32.7; p<0.001$); and with increase in number of quadrants involved (two quadrants 27.7% vs 14.4%; three quadrants 34.7% vs 14.7%; four quadrants 40.6% vs 12.7%; $\chi^2=90.9; p<0.001$). The ICU mortality was higher for hyperinflammatory (51.7% vs 12.4%) and hypoinflammatory (28.0% vs 14.4%) ARDS, compared with non-AHRF controls ($\chi^2=413.07; p<0.001$).

Model 2 scenario

Model 2 scenario compared 851 patients with ARDS with 851 AHRF-UL controls. Patients with ARDS had higher ICU mortality (34.4% vs 24.2%; $p<0.001$), with the higher control arm mortality compared with model 1 reflecting the differences

Sensitivity analysis

We report the following sensitivity analyses: (i) unmatched analyses in all four models to assess how matching—and therefore, exclusion of controls—affect overall and subpopulation estimates and (ii) used hospital mortality as outcome measure in model 2—instead of ICU mortality—to assess how choice of mortality timepoint affected overall and subpopulation estimates for AFARDS.

The $\chi^2$ test was used to assess linear trends in ICU mortality across subpopulations, and to assess relationship between enrichment categories. Reported $p$ values are two-sided and $p$ values <0.05 were considered statistically significant. All analyses were performed using R V.3.4.2 ($\text{AFips}^{50}$, $\text{Matchit}^{51}$ and $\text{tidyverse}^{52}$ packages).
in matching variables between the models (table 1). Similar to model 1, significant linear trends in mortality were seen with severity of hypoxaemia (mild 28.9% vs 19.3%; moderate 32.8% vs 28.5%; severe 44.5% vs 22.0%; \(\chi^2=58.3; p<0.001\)), and with increase in the number of quadrants in patients with ARDS compared with AHRF-UL controls, the absolute differences were lower compared with model 1 (mild 28.9% vs 19.3%; moderate 32.8% vs 28.5%; severe 44.5% vs 22.0%; \(\chi^2=58.3; p<0.001\)). The ICU mortality was higher for hyperinflammatory (50.6% vs 26.8%) and hypoinflammatory (28.2% vs 23.2%) ARDS subphenotypes, compared with matched non-AHRF controls (95% CI 147.27; p<0.001).

Range of AF\textsubscript{ARDS} and categories using model 1 and model 2

The AF\textsubscript{ARDS} ranges between 38.0% (95% CI 34.4% to 41.6%) from model 1 (figure 2A) and 20.9% (95% CI 10.5 to 31.4%) from model 2 (figure 2B).

AF\textsubscript{ARDS} varies by severity of hypoxaemia

In model 1, AF\textsubscript{ARDS} increased with worsening severity of hypoxaemia (mild=32.1% (95% CI 24.4% to 39.8%), moderate=36.7% (95% CI 29.4% to 44.1%), severe=49.7% (95% CI 43.4 to 55.9%)) (figure 2A).

Model 2 also highlighted increase in AF\textsubscript{ARDS} with worsening severity of hypoxaemia (mild=25.7% (95% CI 6.1% to 45.3%), moderate=61.9% (95% CI 45.2% to 76.1%) and severe=41.1% (95% CI 25.2% to 57.1%) (figure 2B)).

AF\textsubscript{ARDS} varies by number of quadrants involved on chest radiography

In model 1, AF\textsubscript{ARDS} increased with number of quadrants involved on chest radiography (two=29.4% (95% CI 22.7% to 36.2%), three=38.1% (95% CI 30.8% to 45.4%), four=45.9% (95% CI 40.7% to 51.2%)) (figure 2A).

Model 2 also highlighted increase in AF\textsubscript{ARDS} with number of quadrants involved on chest radiography (two=14.4% (95% CI 3.9% to 26.6%), three=17.9% (95% CI 6.7% to 29.1%), four=28.9% (95% CI 13.4% to 44.3%) (figure 2B)).

AF\textsubscript{ARDS} was higher in hyperinflammatory subphenotype

In model 1, AF\textsubscript{ARDS} was higher in hyperinflammatory subphenotype (hyperinflammatory=58.7% (95% CI 53.3% to 64.1%) vs hypoinflammatory=28.0% (95% CI 23.1% to 33.0%)) (figure 2A).

Model 2 also highlighted higher AF\textsubscript{ARDS} hyperinflammatory subphenotype (hyperinflammatory=28.6% (95% CI 6.9% to 46.7%) vs hypoinflammatory=17.4% (95% CI 4.5% to 30.2%)) (figure 2B).
Critical care

Model 3 scenario estimating AF_AHRF

Model 3 scenario compared 3504 patients with AHRF matched to 7008 non-AHRF controls. Patients with AHRF had higher ICU mortality compared with non-AHRF controls (31.7% vs 14.1%).

The estimated AF_AHRF from model 3 was 33.8% (95% CI 30.5% to 37.1%), which is lower than the AF ARDS from model 1 and higher than the AF ARDS from model 2. These differences were reflected in the severity of hypoxaemia and radiography categories (figure 3A).

Model 4 scenario estimating AF_AHRF-UL

Model 4 scenario compared 851 patients with AHRF-UL matched to 1702 non-AHRF controls. Patients with AHRF-UL had higher ICU mortality (24.2% vs 14.5%). The estimate of AF_AHRF-UL was 21.3% (95% CI 12.8% to 29.7%), which was lower than the AF ARDS from model 1 and comparable to model 2 (figure 3B). In patients with unilateral, two-quadrant involvement, who would be excluded from ARDS RCTs, the estimate of AF_AHRF-UL was 22.8% (95% CI 7.0% to 38.7%), which was comparable to AF ARDS from model 2.

Sample size requirements for ARDS RCTs change with estimated AF ARDS

As the AF ARDS increases in a RCT population, the sample size required will decrease for prespecified alpha, beta and risk reduction combinations. For example, from our current work, sample size estimates were lower for severe hypoxaemia compared with mild or moderate hypoxaemia, and lower for four quadrant radiographic involvement, compared with two or three quadrant radiographic involvement, and lower for hyperinflammatory subphenotype, compared with hypoinflammatory subphenotype (figure 4).

Sensitivity analyses

Overall unmatched estimates of AF ARDS, AF_AHRF and AF_AHRF-UL were consistent with overall estimates from matched analyses. In the unmatched analyses of model 1, model 3 and model 4—which led to an increase in number of controls—trends within enrichment categories were no longer significant. In the unmatched analysis of model 2—which led to an increase in number of exposed patients with ARDS—trends within enrichment categories were consistent with the matched analysis (table 2).

Overall estimate of AF ARDS in model 2 was lower when hospital mortality was used as the outcome measure instead of ICU mortality; subpopulation estimates were also consistently lower, but overall trends within enrichment categories remained consistent (table 2).

DISCUSSION

Using the LUNG-SAFE database, we report mean estimates of AF ARDS between 20.9% and 38.0%. We observed a dose-response increase in AF ARDS with severity of hypoxaemia and with quadrants of radiographic involvement. AF ARDS was higher in the hyperinflammatory compared with the hypoinflammatory subphenotype of ARDS. Our results are consistent with previous work on AF ARDS and we have extended these previous works by modelling distinct clinical scenarios, including the value of incorporating patients with AHRF in the extended ARDS case definitions. Of note, our AF ARDS estimates are consistent with a very different approach using marginal structural models, reported by Torres et al.17

We focused on several a priori defined ARDS subpopulations that have the potential for enrichment in clinical trials. From our previous work, higher all-cause control-arm mortality does not necessarily generate larger average treatment effects in ARDS RCTs, making us hypothesise that ARDS-specific enrichment...
subgroups may outperform generic illness severity-based prognostic enrichment.

Enrichment strategy, whether prognostic or predictive, is a trade-off between population prevalence, feasibility and expected treatment effect. In the LUNG-SAFE cohort, 23.6% of patients had severe ARDS, 36.7% had four-quadrant involvement and 36.4% were hyperinflammatory ARDS subphenotype. Severe hypoxaemia is a potentially implementable enrichment strategy for ARDS RCTs, using the approach highlighted within the Kigali modification of the ARDS definitions. Furthermore, in previous RCTs of prone positioning and extra-corporeal support, enriching on severe hypoxaemia has shown promise. Another element of the AHRF-ARDS debate is the interobserver and intraobserver reliability of chest radiology, and its feasibility in resource-limited settings. While acknowledging this debate, four-quadrant involvement appears to be an enrichment marker for high AFARDS.

Another enrichment strategy linked to precision medicine is the subphenotyping of ARDS, which would require either measuring discriminant biomarkers with near patient testing, or implementing machine learning-measuring discriminant biomarkers with near patient testing, or implementing machine learning—enriching on severe hypoxaemia which potentially broadens the implications for future RCTs and generalisability to clinical practice. Future research should explore the impact of including these populations in ARDS/AHRF RCTs.

Our analysis has strengths and limitations. We used the LUNG-SAFE dataset—a large multinational cohort recruited from 459 ICUs that was prospectively designed to enrol and follow-up patients with AHRF and which underwent systematic validation after data collection. Our assessment of enrichment categories used inclusion criteria that would be immediately applicable to inform design of ARDS RCTs. Despite the use of propensity score methods, residual confounding remains a concern, given the overall definition has implications for future RCTs and generalisability to clinical practice. Future research should explore the impact of including these populations in ARDS/AHRF RCTs.

Our findings also lead us to consider how our work informs the debate on the need for distinction between ARDS and AHRF. Specifically, the estimate of AFARDS for patients with unilateral, two-quadrant involvement, who would be excluded currently from ARDS RCTs, was comparable to AFARDS from model 2. Currently, ARDS is conceptualised as a subset of AHRF; exclusion of patients with AHRF with similar AF and overlapping biology from the overall definition has implications for future RCTs and generalisability to clinical practice. Future research should explore the impact of including these populations in ARDS/AHRF RCTs.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Sensitivity analysis—overall and subpopulation estimates of AFARDS, AF_AHRF, and AF_AHRF-UL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: AFARDS (non-AHRF controls)</td>
<td>Model 2: AFARDS (AHRF-UL controls)</td>
</tr>
<tr>
<td>Sensitivity analysis</td>
<td>Unmatched analysis</td>
</tr>
<tr>
<td>Patients with ARDS/AHRF</td>
<td>2653</td>
</tr>
<tr>
<td>Controls</td>
<td>8407</td>
</tr>
<tr>
<td>Overall AF</td>
<td>25.2% (22.7%–27.6%)</td>
</tr>
<tr>
<td>PaO2/FiO2 ratio (mm Hg)</td>
<td></td>
</tr>
<tr>
<td>&gt;200</td>
<td>8.9% (4.5%–13.4%)</td>
</tr>
<tr>
<td>100–200</td>
<td>17.2% (13.4%–21.1%)</td>
</tr>
<tr>
<td>&lt;100</td>
<td>14.5% (10.7%–18.5%)</td>
</tr>
<tr>
<td>Maximum number of quadrants involved in first 48 hours</td>
<td></td>
</tr>
<tr>
<td>One</td>
<td>–</td>
</tr>
<tr>
<td>Two</td>
<td>11.3% (7.0%–15.6%)</td>
</tr>
<tr>
<td>Three</td>
<td>10.8% (6.6%–15.1%)</td>
</tr>
<tr>
<td>Four</td>
<td>18.8% (15.1%–22.4%)</td>
</tr>
</tbody>
</table>

AF, attributable fraction; AHRF, acute hypoxaemic respiratory failure; AHRF-UL, acute hypoxaemic respiratory failure with unilateral infiltrates only; ARDS, acute respiratory distress syndrome; ICU, intensive care unit.
ARDS limitation is explicit, as the LUNG-SAFE cohort did not collect ARDS risk factors for non-ARDS controls. Our primary and sensitivity analyses focus on mortality; the impact on other outcomes used in ARDS RCTs (such as ventilator-free days) needs to be assessed. An implicit assumption in these models is that the putative treatment for ARDS/AHRF has no effect on mortality of patients with non-ARDS/non-ARDS. This assumption would not bias AF-ARDS estimates, as the control groups on commonly used ARDS RCT eligibility criteria such as severity of hypoxaemia, and number of quadrants involved in chest radiography. We show that hyperinflammatory ARDS subphenotype has higher attributable fraction.

CONCLUSIONS
ARDS is associated with excess mortality in critically ill patients. Our results highlight generic enrichment populations based on commonly used ARDS RCT eligibility criteria such as severity of hypoxaemia, and number of quadrants involved in chest radiography. We show that hyperinflammatory ARDS subphenotype has higher attributable fraction.

Author affiliations
1 Critical Care, King’s College Hospital NHS Trust, London, UK
2 School of Immunology and Microbial Sciences, King’s College London, London, UK
3 Service de médecine intensive-rénalisation, Paris-Saclay University Faculty of Medicine, Le Kremlin-Bicêtre, France
4 Equipe d’Épidémiologie respiratoire intégrative, CESP, Paris-Saclay University, Gif-sur-Yvette, France
5 Department of Anaesthesiology, Washington University in St Louis, St Louis, Missouri, USA
6 Pulmonary, Allergy and Critical Care Medicine, Stanford University, Stanford, California, USA
7 Emergency and Intensive Care, University of Milan-Bicocca, Monza, Italy
8 Interdepartmental Division of Critical Care Medicine, University of Toronto Faculty of Medicine, Toronto, Ontario, Canada
9 Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge, UK
10 School of Population Health & Environmental Sciences, King’s College London, London, UK
11 Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada
12 Department of Anesthesia and Perioperative Care, University of California, San Francisco, California, USA
13 Anaesthesia, School of Medicine, National University of Ireland Galway, Galway, Ireland
14 National Centre for Biomedical Engineering Sciences, National University of Ireland Galway, Galway, Ireland
15 YCUI, QUB, Belfast, UK
16 School of Medicine, Dentistry and Biomedical Sciences, Queen’s University Belfast, Northern Ireland
17 Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge, UK
18 School of Population Health & Environmental Sciences, King’s College London, London, UK
19 Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada
20 Department of Anesthesia and Perioperative Care, University of California, San Francisco, California, USA
21 Anaesthesia, School of Medicine, National University of Ireland Galway, Galway, Ireland
22 National Centre for Biomedical Engineering Sciences, National University of Ireland Galway, Galway, Ireland
23 Pulmonary, Allergy and Critical Care Medicine, Stanford University, Stanford, California, USA
24 School of Medicine, Dentistry and Biomedical Sciences, Queen’s University Belfast, Belfast, Northern Ireland
25 Centre for Infection Research, The Queen’s Medical Research Institute, The University of Edinburgh, Edinburgh, UK

Twitter Rohit Saha @saha_rohit, Charlotte Summers @charlot Summers and Manu Shankar-Hari @ms_hari

Collaborators


Downloaded from thorax.bmj.com on September 25, 2023 by guest. Protected by copyright.
Sackett DL. Superiority trials, non-inferiority trials, and prisoners of the 2-sided null hypothesis. Evidence-Based Medicine 2004;9:38–9.


Sinclair JC, Haynes RB. Selecting participants that raise a clinical trial’s population attributable fraction can increase the treatment effect within the trial and reduce the required sample size. J Clin Epidemiol 2011;64:893–902.


