

BMJ Open Inter-rater reliability and prognostic value of baseline Radiographic Assessment of Lung Edema (RALE) scores in observational cohort studies of inpatients with COVID-19

Nameer Al-Yousif ^{1,2}, Saketram Komanduri,³ Hafiz Qurashi,³ Anatoliy Korzhuk,³ Halimat O Lawal,³ Nicholas Abourizk,³ Caitlin Schaefer,^{4,5} Kevin J Mitchell,⁶ Catherine M Dietz,⁶ Ellen K Hughes,⁶ Clara S Brandt,⁶ Georgia M Fitzgerald,⁶ Robin Joyce,⁶ Asmaa S Chaudhry,⁶ Daniel Kotok,⁷ Jose D Rivera,⁷ Andrew I Kim,⁷ Shruti Shettigar,⁷ Allen Lavina,⁷ Christine E Girard,⁷ Samantha R Gillenwater,⁷ Anas Hadeh,⁷ William Bain,^{4,5} Faraaz A Shah,^{4,5} Matthew Bittner,⁸ Michael Lu,⁸ Niall Prendergast,⁴ John Evankovich,^{4,5} Konstantin Golubykh,³ Navitha Ramesh,⁹ Jana J Jacobs,¹⁰ Cathy Kessinger,⁴ Barbara Methe,^{4,11} Janet S Lee,^{4,5} Alison Morris,^{4,5,11} Bryan J McVerry,^{4,5,11} Georgios D Kitsios^{4,5,11}

To cite: Al-Yousif N, Komanduri S, Qurashi H, *et al*. Inter-rater reliability and prognostic value of baseline Radiographic Assessment of Lung Edema (RALE) scores in observational cohort studies of inpatients with COVID-19. *BMJ Open* 2023;13:e066626. doi:10.1136/bmjopen-2022-066626

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2022-066626>).

Received 29 August 2022
Accepted 16 December 2022



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Dr Georgios D Kitsios;
kitsiosg@upmc.edu

ABSTRACT

Objectives To reliably quantify the radiographic severity of COVID-19 pneumonia with the Radiographic Assessment of Lung Edema (RALE) score on clinical chest X-rays among inpatients and examine the prognostic value of baseline RALE scores on COVID-19 clinical outcomes.

Setting Hospitalised patients with COVID-19 in dedicated wards and intensive care units from two different hospital systems.

Participants 425 patients with COVID-19 in a discovery data set and 415 patients in a validation data set.

Primary and secondary outcomes We measured inter-rater reliability for RALE score annotations by different reviewers and examined for associations of consensus RALE scores with the level of respiratory support, demographics, physiologic variables, applied therapies, plasma host–response biomarkers, SARS-CoV-2 RNA load and clinical outcomes.

Results Inter-rater agreement for RALE scores improved from fair to excellent following reviewer training and feedback (intraclass correlation coefficient of 0.85 vs 0.93, respectively). In the discovery cohort, the required level of respiratory support at the time of CXR acquisition (supplemental oxygen or non-invasive ventilation (n=178); invasive-mechanical ventilation (n=234), extracorporeal membrane oxygenation (n=13)) was significantly associated with RALE scores (median (IQR): 20.0 (14.1–26.7), 26.0 (20.5–34.0) and 44.5 (34.5–48.0), respectively, p<0.0001). Among invasively ventilated patients, RALE scores were significantly associated with worse respiratory mechanics (plateau and driving pressure) and gas exchange metrics (PaO₂/FiO₂ and ventilatory ratio), as well as higher plasma levels of IL-6, soluble receptor of advanced glycation end-products and soluble tumour necrosis factor receptor 1 (p<0.05). RALE scores

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ We used a larger sample size than previous studies on Radiographic Assessment of Lung Edema (RALE) score in COVID-19.
- ⇒ We developed and used a dedicated software for image analysis and RALE score annotations.
- ⇒ We used temporally and geographically independent data sets from different hospital systems, with granular clinical and research data.
- ⇒ We examined only baseline chest X-rays (CXRs) and did not evaluate trajectories of radiographic severity evolution.
- ⇒ We used portable CXR images obtained as part of routine medical care and did not standardise image acquisition protocols for this study.

were independently associated with 90-day survival in a multivariate Cox proportional hazards model (adjusted HR 1.04 (1.02–1.07), p=0.002). We replicated the significant associations of RALE scores with baseline disease severity and mortality in the independent validation data set.

Conclusions With a reproducible method to measure radiographic severity in COVID-19, we found significant associations with clinical and physiologic severity, host inflammation and clinical outcomes. The incorporation of radiographic severity assessments in clinical decision-making may provide important guidance for prognostication and treatment allocation in COVID-19.

INTRODUCTION

Infection with the SARS-CoV-2 has heterogeneous clinical presentations ranging from asymptomatic course to severe COVID-19 with



pneumonia and hypoxemia, requiring hospitalisation. Inpatients with COVID-19 may require different levels of respiratory support, ranging from low level supplementation of inspired oxygen via nasal cannula in spontaneously breathing (SB) patients on the wards, to intubation and invasive mechanical ventilation (IMV) in the intensive care unit (ICU), to extracorporeal membrane oxygenation (ECMO) support in a selected subset of the sickest patients with refractory hypoxemia.

Multiple risk stratification tools for COVID-19 have been developed, combining clinical, physiologic, laboratory or research biomarker variables. Meanwhile, diagnosis of COVID-19 pneumonia relies on presence of radiographic consolidations on chest X-ray (CXR) or computed tomography (CT). Of the two modalities, CXR is the most widely available and routinely used, and CXRs are often repeated to determine pneumonia evolution or on any new clinical indication.^{1 2} However, radiographic severity has not been systematically integrated into risk predictions for COVID-19, and severity assessments are mostly qualitative and limited to narrative descriptions in diagnostic reports. The Radiographic Assessment of Lung Edema (RALE) score was developed and validated as a semiquantitative instrument for evaluating the extent and density of radiographic opacities on CXRs in acute respiratory distress syndrome (ARDS).

RALE scores have been shown to correlate with severity of hypoxemia,^{3 4} plasma biomarker levels (such as the soluble receptor of advanced glycation end-products—sRAGE)⁵ as well as to be prognostic of clinical outcome in non-COVID ARDS.^{3 4} Nonetheless, individual studies analysed small sets of ARDS subjects and CXRs, and associations with endpoints were inconsistent.⁵ During the COVID-19 pandemic, RALE scores have been associated with COVID-19 pneumonia severity and clinical outcomes in several studies,^{6–9} but we still lack a systematic evaluation of RALE scoring reproducibility and understanding of the impact of image-related variables (such as radiographic penetration) and patient covariates on derived RALE scores. Furthermore, it remains unknown whether RALE scores capture important inter-individual variability in clinical severity when examined in the context of provided respiratory support (eg, intubated vs non-intubated patients), and whether RALE scores reflect differences in underlying biological heterogeneity of COVID-19, as represented by host–response biomarkers and subphenotypes, viral load or administered therapeutics.

We hypothesised that RALE scoring is a learnable skill among clinicians with high inter-rater reliability, and that baseline RALE scores in patients with COVID-19 have prognostic value on disease severity metrics and clinical outcomes. In this study, we investigated the reproducibility of RALE scoring by multiple independent reviewers utilising a standardised approach with a dedicated software for image analysis and RALE score annotations. We analysed CXRs in concert with detailed clinical and biological data from inpatients with COVID-19 enrolled

in four independent cohort studies. We examined associations of RALE scores with cross-sectional indices of clinical severity, physiologic variables and biomarkers and quantified the prognostic value of baseline RALE scores on COVID-19 clinical outcomes.

METHODS

Discovery data set

We analysed data obtained from hospitalised patients with COVID-19, who were enrolled from April 2020 through October 2021 in one of three independent cohort studies within the UPMC (University of Pittsburgh Medical Center) Health System (detailed description available in the online supplemental file 1):

- The Acute Lung Injury Registry (ALIR) and Biospecimen Repository**, a prospective cohort study of critically ill adult patients (18–90 years of age) with acute respiratory failure. We enrolled COVID-19 subjects following admission to the ICU and obtaining informed consent (IRB protocol STUDY19050099) and collected plasma biospecimens.
- The COVID-19 Inpatient Cohort (COVID-INC)**, a prospective cohort study of moderately ill adult inpatients with COVID-19, hospitalised mainly in dedicated inpatient wards. Following informed consent (IRB protocol STUDY20040036), we collected blood biospecimens processed similarly to the ALIR study.
- The Prognostication for COVID-19 Patients Admitted to ICUs at UPMC Pinnacle (PROCOPI) study**, a retrospective cohort study of critically ill patients with COVID-19 hospitalised in ICUs at UPMC Pinnacle hospitals. We performed retrospective chart review and data collection (IRB protocol 20E059) for patients with COVID-19 on IMV.

Clinical data collection

We extracted data on demographics, comorbid conditions and clinical test results at baseline and retrieved a portable CXR image at a baseline timepoint defined as: (1) day of hospital admission for the non-ICU patients of the COVID-INC cohort, (2) day of ICU admission for non-intubated, SB critically ill patients (ALIR and COVID-INC cohorts), (3) day of intubation for mechanically ventilated patients (ALIR, COVID-INC and PROCOPI cohorts). We scored each patient's severity of illness according to the 10-point ordinal scale of the WHO, and broadly classified baseline respiratory support in three categories: (1) SB patients, that is, not intubated subjects on various levels of oxygenation support including non-invasive ventilation, (2) IMV, intubated subjects in the ICU and (iii) ECMO, that is, intubated subjects in the ICU on ECMO support. From IMV patients, we also collected detailed physiologic data from physician-set ventilatory parameters and obtained measurements for respiratory mechanics and gas exchange (Supplement), as previously described.^{10 11} We recorded administered therapies and clinical endpoints across the COVID-19 timeline.

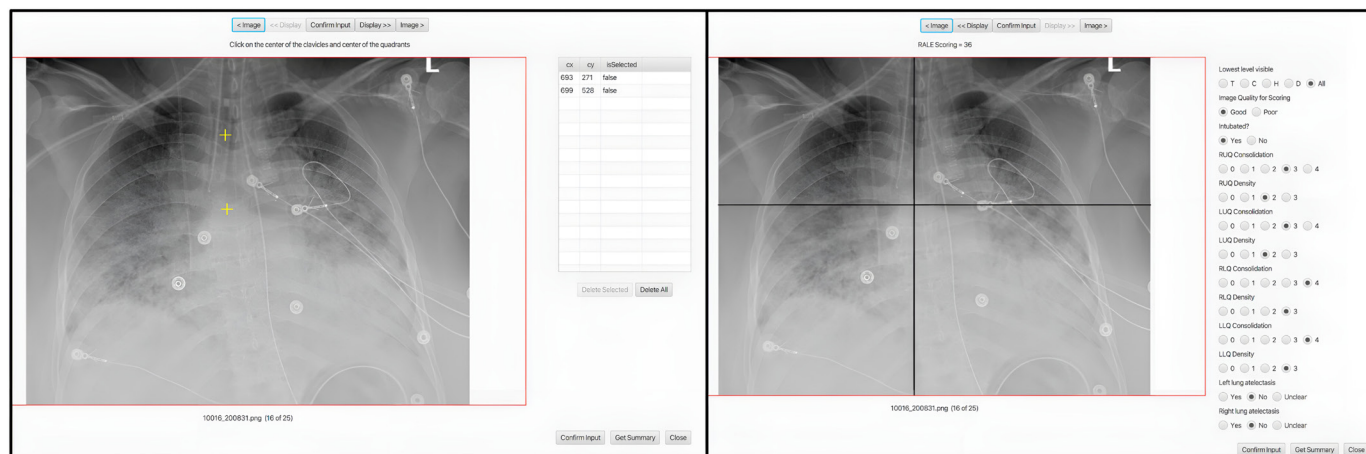


Figure 1 RALE scoring process through the Pulmo-Annotator software interface. This figure shows a screenshot of the Pulmo-Annotator software that was used to store and score the X-ray images. Left panel shows how the axis was set up with the first coordinate being assigned for the image rotation (vertical axis—mid-point of the vertebral column or mid-point of clavicles if image not rotated) and the second coordinate assigned for the horizontal axis determination at the level of the first branch of the left main bronchus. Right image shows the automated axes drawn by the software per the determined coordinates from the previous input, with options for physician annotation of image quality, penetration, presence of endotracheal tube and atelectasis in each lung, and score options for density and consolidation extent for each quadrant, to allow for automated calculation of the final RALE score. LLQ, left lower quadrant; LUQ, left upper quadrant; RALE, Radiographic Assessment of Lung Edema; RLQ, right lower quadrant; RUQ, right upper quadrant.

RALE scoring

We performed RALE score assessments by ≥ 2 independent reviewers per image with the *Pulmo-Annotator* software (Veytel, LLC) (figure 1 and details on scoring in the Supplement). In brief, we assessed radiographic penetration, image quality, presence of endotracheal tube, atelectasis and then scored the most dense radiographic opacity in each quadrant by *extent* (scores of 0 for none, 1 for $<25\%$, 2 for $25\%–50\%$, 3 for $50\%–75\%$ and 4 for $>75\%$ of quadrant involved) and *density* (scores of 1 for hazy, 2 for moderate and 3 for dense consolidation). The software allowed for easy ‘point and click’ annotations of all the anatomical mapping (eg, horizontal level of the first branch of the left main bronchus to define the horizontal axis for quadrant division), qualitative (eg, image quality), quantitative (eg, density score) and categorical features (eg, presence of endotracheal tube or atelectasis) for each image by each reviewer independently, with automated, time-stamped storage of annotations on a cloud server for subsequent data retrieval and reproducible analyses. Each quadrant’s score was automatically calculated as the product of *extent***density*, and then all four quadrant scores were summed for a final RALE score (ranging from 0 to 48).³ Following a first iteration, each reviewer was provided feedback on scores distribution and agreement with other reviewer(s), followed by a joint session with the senior reviewer (GDK) to understand sources of disagreement and then independent rescoring of CXRs with large discrepancies in total RALE scores (≥ 15 RALE score difference) or within individual quadrants (≥ 2 score difference in any quadrant extent or density). We used the RALE scores and annotated variables from the second iteration in quantitative analyses.

Plasma biomarkers

From available baseline samples from the ALIR and COVID-INC cohorts, we measured plasma biomarkers of injury and inflammation with custom-made Luminex panels as previously described.¹² We classified subjects into a hyperinflammatory versus hypoinflammatory subphenotype by using predicted probabilities for subphenotype classifications from a published parsimonious logistic regression model utilising interleukin-6 (IL-6), soluble tumour necrosis factor receptor 1 (sTNFR1) and bicarbonate.¹³ In a random subset of plasma samples ($n=63$), we quantified circulating levels of SARS-CoV-2 RNA by qPCR, as previously described.^{14 15}

Statistical analyses

We performed non-parametric comparisons for continuous (described as median and IQR) and categorical variables between clinical groups (Wilcoxon and Fisher’s exact tests, respectively). We examined for inter-reviewer agreement on RALE scores with Bland-Altman plots prefeedback and postfeedback sessions, and quantitatively by measuring inter-reviewer correlations and intra-class correlation coefficients (ICC) in two-way random effects models. For categorical variables on CXR assessments, we quantified inter-reviewer agreement with Cohen’s kappa statistics. We examined correlations of continuous variables with Pearson correlation test. We fit proportional hazards models to examine the statistical significance of baseline RALE scores on 60-day survival or time-to-liberation from IMV. We performed all analyses with the R software and a p value of <0.05 was deemed statistically significant.

Validation cohort

We obtained admission CXRs from 415 COVID-19 inpatients hospitalised within 18 different clinical sites of the Cleveland Clinic systems from March to October 2020. We collected clinical data from electronic medical records on demographics, comorbidities, physiologic and laboratory variables under an exempt review protocol (FLA 20–038) as previously described.¹⁶ We classified patients into SB and IMV groups based on the type of respiratory support by the timing of the CXR.

All findings are reported in accordance with the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) statement for observational studies.¹⁷

Patient and public involvement

Patients or the public were not involved in the design, conduct or reporting of our study.

RESULTS

Characteristics of enrolled patients in the three discovery cohorts

We analysed baseline CXRs from a total of 425 inpatients with COVID-19 (154 subjects from ALIR, 138 from COVID-INC and 133 subjects from PROCOP—(online supplemental table S1) and stratified patients by level of respiratory support at time of the CXR as SB patients (n=178), IMV (n=234) and ECMO (n=13). Our study population had a median age (IQR) of 64.0 (55.0–72.7) years, consisting mostly of men (59.7%), whites (76.0%), with high body mass index (BMI) (median 31.8 (27.0–38.2)). Overall, in-hospital mortality was 47.5%, with 58% of hospitalisation survivors discharged home, and the remainder requiring admission to inpatient rehabilitation, long term acute care or skilled nursing facilities. Detailed baseline characteristics and outcomes of the discovery dataset are presented in [table 1](#).

Inter-Rater agreement for RALE scores

In first iteration of RALE scoring, we found good inter-rater agreement between reviewers for total RALE scores (ICC 0.85, 95% confidence interval-CI [0.82 to 0.88], $p<0.0001$), with 18/425 (4%) of CXRs showing large total RALE score discrepancies (± 15 points) and 78/425 (18%) revealing large (≥ 2 points) difference in extent or density of a quadrant between two reviewers. Following feedback and independent rescoring of discrepant CXRs by the two reviewers, the inter-rater agreement on RALE scores at the second scoring iteration improved to excellent (ICC 0.93 [0.92–0.95], $p<0.0001$), with 4/425 (<1%) CXRs showing large total RALE discrepancies and 19/425 (5%) CXRs with remaining ≥ 2 point discrepancies for extent or density in a quadrant ([figure 2](#) and online supplemental table S2–S3). We then used average RALE scores from two reviewers in further quantitative analyses.

Impact of CXR image variables on RALE scores

We examined for the association between CXR image findings and RALE scores without any knowledge of

clinical data. Under-penetrated CXRs (ie, CXRs in which vertebral bodies were visible only behind the trachea) had higher median RALE scores compared with CXRs with visible vertebral bodies behind the heart ($p<0.01$, online supplemental figure S1), and right lung atelectasis (definite or possible) was associated with significantly higher scores for right lower quadrant mean density scores ($p<0.01$, online supplemental figure S1). Overall, the lower quadrants (right and left) had much higher quadrant scores compared with their corresponding upper quadrants (right and left, respectively, $p<0.0001$). Left lower quadrant scores were statistically significantly higher than right lower quadrant ones ($p<0.01$, online supplemental figure S1). Therefore, both radiographic penetration and physician-ascribed presence of atelectasis were shown to have an impact on RALE scores, with the lower quadrant scores being systematically higher than the upper quadrants.

RALE scores by baseline level of respiratory support and period of the pandemic

ECMO patients had the highest RALE scores (median (IQR): 44.5 [34.5–48.0]), followed by IMV (26.0 [20.5–34.0]) and then by SB patients (20.0 [14.1–26.7]), $p<0.0001$ ([figure 3A](#)). The association between radiographic and clinical severity was also significant for the component RALE scores in each quadrant ([figure 3B–C](#)) and by WHO ordinal scale categories ([figure 3D](#)). The COVID-INC cohort had the highest proportion of SB patients (91%) and as expected, patients in the COVID-INC cohort had lower RALE scores compared with the ALIR and PROCOP cohorts ($p<0.0001$, online supplemental figure S2A). Throughout the period of enrolment (March 2020–October 2021), we found that there was a progressive increase of baseline RALE scores over the epoch of time for IMV patients only ($R=0.16$ for RALE scores and time from March 2020 till CXR date, $p=0.017$, online supplemental figure S2B).

Baseline clinical variables and RALE scores

We then examined for associations between clinical characteristics and RALE scores at baseline, separately for SB, IMV and ECMO patients, given the significantly different RALE scores by respiratory support category. Among SB patients, men and obese patients had higher RALE scores ($p<0.05$, online supplemental figure S3) whereas among IMV patients, nursing facility residents and patients with history of chronic obstructive pulmonary disease (COPD) had significantly lower RALE scores than their counterparts ($p<0.0001$, online supplemental figure S3). Notably, for patients on IMV, age was inversely correlated with RALE scores ($p<0.0001$), whereas for both SB and IMV patients RALE scores were positively correlated with BMI ($p<0.0001$) and duration of COVID-19 symptoms ($p<0.0001$, online supplemental figure S4).

Table 1 Baseline characteristics of subjects grouped by level of respiratory support: spontaneously-breathing patients (SB), patients on invasive mechanical ventilation (IMV) and patients on extracorporeal membrane oxygenation (ECMO)

	SB (n=178)	IMV (n=234)	ECMO (n=13)	P value
Demographics				
Age, years	65.0 (55.0, 71.9)	64.3 (55.1, 73.4)	56.7 (50.7, 57.8)	0.01
Body mass index	30.8 (26.1, 36.9)	32.7 (28.5, 38.4)	35.9 (32.4, 41.8)	0.01
Sex, male	84 (47.2%)	160 (68.4%)	10 (76.9%)	<0.01
Race, White	141 (79.2%)	171 (73.1%)	11 (84.6%)	0.07
Never smokers	77 (47.2%)	90 (46.2%)	7 (53.8%)	0.86*
Resident of nursing facility	21 (12.4%)	29 (14.9%)	0 (0.0%)	0.28*
Diabetes mellitus	72 (40.4%)	104 (44.4%)	5 (38.5%)	0.69
Chronic obstructive lung disease	35 (19.7%)	37 (15.8%)	2 (15.4%)	0.58
Congestive cardiac failure	24 (13.5%)	35 (15.0%)	0 (0.0%)	0.31
Plasma biomarkers				
IL-6, pg/mL	14.5 (5.9, 47.8)	68.8 (14.0, 180.9)	220.6 (88.1, 1112.0)	<0.01
IL-8, pg/mL	16.0 (8.6, 29.7)	21.7 (14.2, 42.8)	27.8 (15.0, 46.8)	<0.01
Ang2, pg/mL	2701.7 (1426.7, 4090.9)	5634.2 (2860.3, 10913.3)	5927.3 (4170.2, 7592.2)	<0.01
Procalcitonin, pg/mL	107.5 (65.0, 283.9)	633.7 (173.3, 1956.3)	465.6 (196.6, 1261.5)	<0.01
ST2, ng/mL	87.8 (47.3, 168.0)	211.1 (96.3, 378.9)	190.1 (111.0, 253.9)	<0.01
Pentraxin-3, pg/mL	5791.2 (2621.5, 11 728.4)	9599.2 (4680.4, 21 226.8)	5525.4 (3748.7, 11 688.9)	0.01
sRAGE, pg/mL	3979.1 (2447.0, 9137.2)	6158.5 (3269.3, 12 653.0)	1754.8 (1219.6, 5102.5)	<0.01
sTNFR1, pg/mL	3438.2 (2485.5, 5359.9)	5601.7 (3375.9, 12 297.4)	7283.8 (4650.7, 8515.4)	<0.01
Hyperinflammatory phenotype	5 (3.5%)	12 (16.2%)	2 (16.7%)	<0.01*
Radiographic parameters				
RALE Score, total	20.0 (14.1, 26.7)	26.0 (20.5, 34.0)	44.5 (34.5, 48.0)	<0.01
Lower quadrants RALE Score	14.0 (10.0, 17.0)	16.0 (12.5, 20.0)	24.0 (22.0, 24.0)	<0.01
Upper quadrants RALE Score	6.0 (3.5, 9.9)	10.0 (7.0, 14.0)	22.5 (12.8, 24.0)	<0.01
Time of CXR from symptom onset	7.0 (3.0, 11.0)	9.0 (6.0, 15.0)	14.0 (11.0, 22.0)	<0.01
Clinical outcomes				
In-hospital mortality	33 (18.5%)	163 (69.7%)	6 (46.2%)	<0.01
Discharge destination				
Home	113 (63.5%)	17 (7.3%)	1 (7.7%)	<0.01
Inpatient rehabilitation	2 (1.1%)	15 (6.4%)	2 (15.4%)	<0.01
Long-term acute care	3 (1.7%)	18 (7.7%)	3 (23.1%)	<0.01
Skilled nursing facility	27 (15.2%)	21 (9.0%)	1 (7.7%)	0.154

Continuous variables are reported in median (IQR). Categorical variables are reported as n (%).

*We only included patients with available clinical data or research biomarkers in analysis. Patients with unavailable data were excluded.

Ang2, Angiotensin-2; CXR, chest X-ray; ECMO, extracorporeal membrane oxygenation; IL, interleukin; sRAGE, soluble receptor of advanced glycation end-products; ST2, suppression of tumorigenicity-2; sTNFR1, soluble Tumor Necrosis Factor Receptor 1.

Pulmonary physiology and applied therapies are associated with RALE scores

We examined physician-set ventilatory parameters, pulmonary mechanics and gas exchange metrics in IMV patients only, because such measurements are either unavailable or not reliably measured in SB patients and confounded by the extracorporeal support in ECMO patients. In terms of ventilatory parameters, RALE scores were inversely correlated with set tidal volumes (TV,

$R=-0.17$, $p=0.02$) and were higher by increasing levels of positive end-expiratory pressure (PEEP, [figure 4A,B](#)). By measured mechanics, RALE scores positively correlated both with plateau ($R=0.38$, $p<0.0001$) and driving pressures ($R=0.31$, $p<0.001$, [figure 4C,D](#)). For gas exchange, RALE scores were positively correlated with ventilatory ratios (ie, worse CO_2 clearance, $R=0.18$, $p=0.02$) and negatively correlated with $\text{PaO}_2/\text{FiO}_2$ ratios (ie, worse hypoxemia, $R=-0.3$, $p<0.0001$, [figure 4E,F](#)). Patients

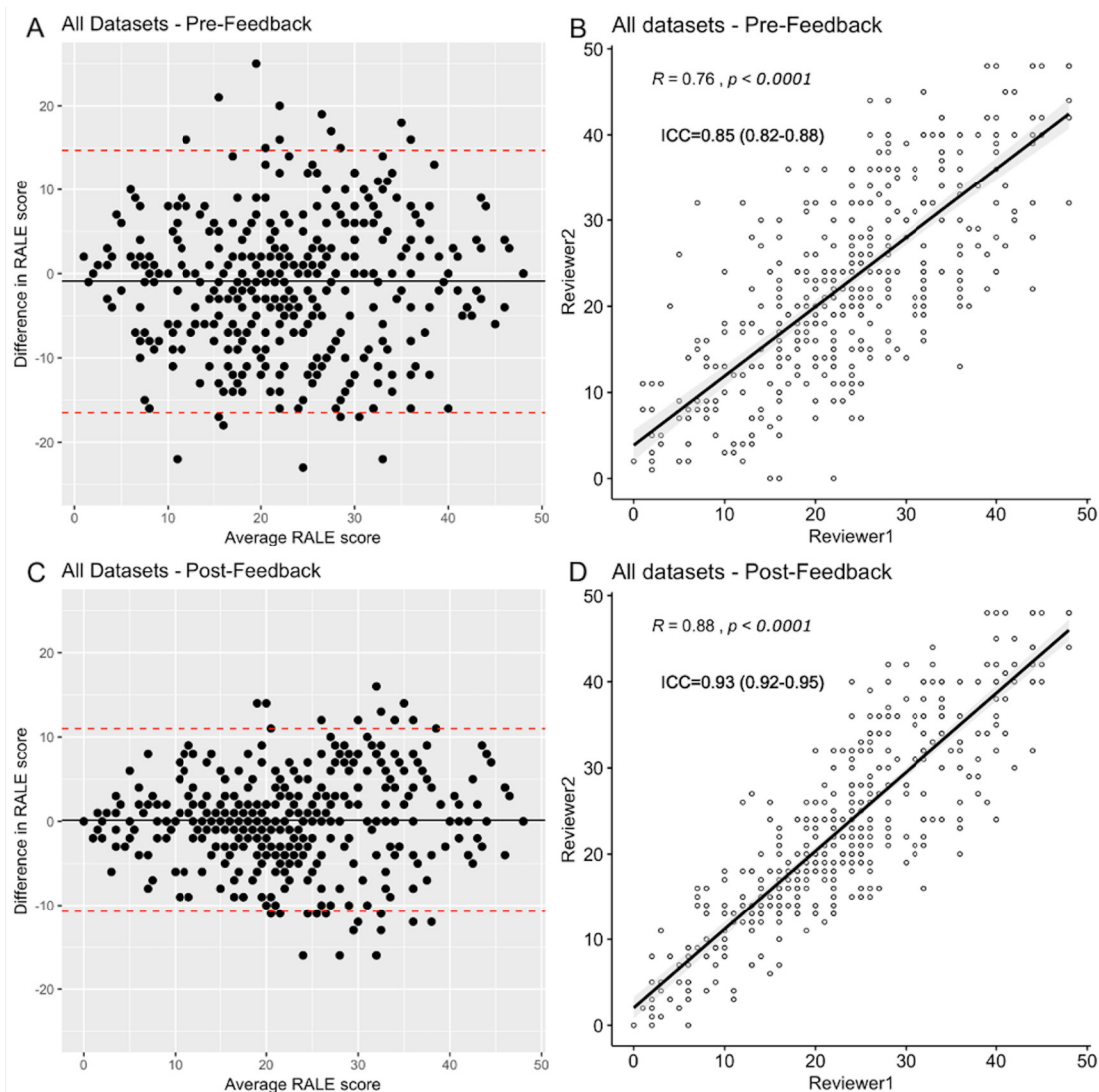


Figure 2 Inter-rater agreement for RALE scores in pre-feedback and post-feedback session assessments. Top figures show pre-feedback results of the inter-rater agreement. The left upper panel image (A) shows a Bland-Altman plot and the right image and (B) shows a scattered plot graph with a high correlation coefficient ($R=0.76$, $p<0.0001$) and good inter-rater agreement (intraclass correlation coefficient (ICC) 0.85 [0.82 – 0.88], $p<0.0001$) between the reviewers. Pre-feedback, 4% of CXRs had large RALE score discrepancies (± 15 points). The bottom panels show post-feedback results of the inter-rater agreement. The left image (C) shows a Bland-Altman plot and the right image (D) shows a scattered plot graph with a very high correlation coefficient ($R=0.88$, $p<0.0001$) and excellent inter-rater agreement (ICC 0.93 [0.92 – 0.95], $p<0.0001$) between the reviewers. Post-feedback, less than 1% of images had large RALE score discrepancies. RALE, Radiographic Assessment of Lung Edema.

on IMV who underwent prone positioning or received neuromuscular blockade had higher RALE scores than their untreated counterparts ($p<0.0001$, online supplemental figure S5).

RALE scores and plasma biomarkers

We did not examine plasma biomarker associations in ECMO patients due to small sample size. We found no significant association between RALE scores and plasma SARS-CoV-2 RNA levels ('viral RNA-emia') in either SB or IMV patients examined separately. Baseline RALE scores correlated significantly with plasma levels of IL-6 in SB patients, and with IL-6, sTNFR1 and sRAGE levels in IMV patients (figure 5A,B). When stratified into subphenotypes, hyperinflammatory patients had higher RALE

scores in both SB patients ($p=0.04$) and IMV patients ($p=0.007$, figure 5C,D).

RALE scores are prognostic of clinical outcomes

When all patients were combined (SB, IMV, ECMO), baseline RALE scores were higher among non-survivors (25.1 (19.8–33.0)) compared with survivors of hospitalisation (22.3 (15.0–31.0), $p=0.0014$, figure 6A). In a Cox proportional hazards model for 60-day survival adjusted for age, sex, BMI and COPD, RALE scores were significantly associated with worse survival (adjusted HR 1.02 (1.01–1.04) for each unit increase in RALE score, $p=0.002$). Stratified by RALE score tertiles (low<19.6, intermediate: 19.6–28.5, high>28.5), patients in the high tertile had worse 60-day survival by Kaplan-Meier curve analysis (figure 6B). When

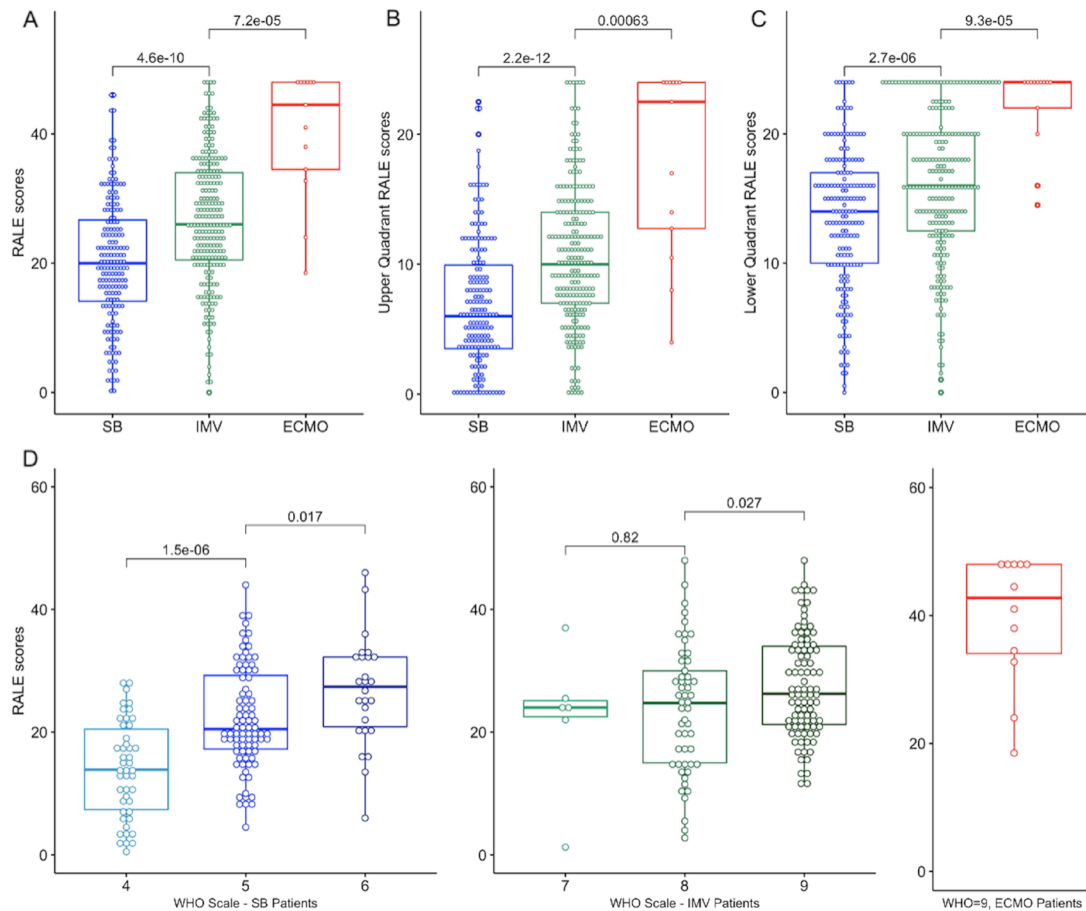


Figure 3 Patients on higher levels of respiratory support had higher RALE scores at baseline. (A): Patients on ECMO had much higher RALE scores compared with patients on IMV, who in turn had significantly higher RALE scores than SB patients. (B and C): We found similar differences in upper and lower quadrant RALE scores by level of respiratory support. (D): Total RALE scores were significantly higher by rising disease severity based on the ordinal WHO scale. ECMO, extracorporeal membrane oxygenation; IMV, invasive mechanical ventilation; RALE, Radiographic Assessment of Lung Edema; SB, spontaneously breathing.

examined separately within each group of respiratory support level, RALE scores were not significantly associated with 60-day survival in adjusted Cox proportional hazards models. Similarly, we did not find a significant association for RALE scores with time to liberation from IMV in Cox models adjusted for age, sex, BMI, COPD, TV and PEEP levels.

Among survivors of hospitalisation, higher complexity of care needs on discharge (based on disposition destination) were significantly associated with baseline RALE scores, with higher RALE scores for survivors discharged to a long-term acute care facility (33.3 (22.9–40.4)) or in-patient rehabilitation (32.0 (24.5–38.0)) compared with those discharged to a skilled-nursing facility (19.5 (13.9–27.3)) or home care (20.5 (13.5–28.0)), $p < 0.0001$ (figure 6C).

External validation of key clinical associations for RALE scores

In an independent cohort of 415 COVID-19 inpatients, online supplemental table S4, we found that baseline RALE scores were significantly different between IMV

($n=68$) and SB ($n=347$, $p < 0.0001$, online supplemental figure S6), we replicated the correlations between BMI and hypoxemia inferred by SpO_2/FiO_2 ratios and validated the association between baseline RALE scores with 90-day mortality, with non-survivors having markedly higher RALE scores than survivors ($p < 0.0001$, online supplemental figure S6).

DISCUSSION

Our study used the RALE scoring system to examine the radiographic heterogeneity of COVID-19 pneumonia among inpatients with a wide spectrum of clinical severity. With a systematic approach supported by a dedicated software, we demonstrated that RALE scoring is a learnable skill for clinicians, relatively easy to use, with excellent inter-rater agreement following appropriate training. We demonstrated that technical aspects of image quality and radiographic penetration impact RALE score assignments. Among inpatients with COVID-19, RALE scores were reflective of disease severity by level of

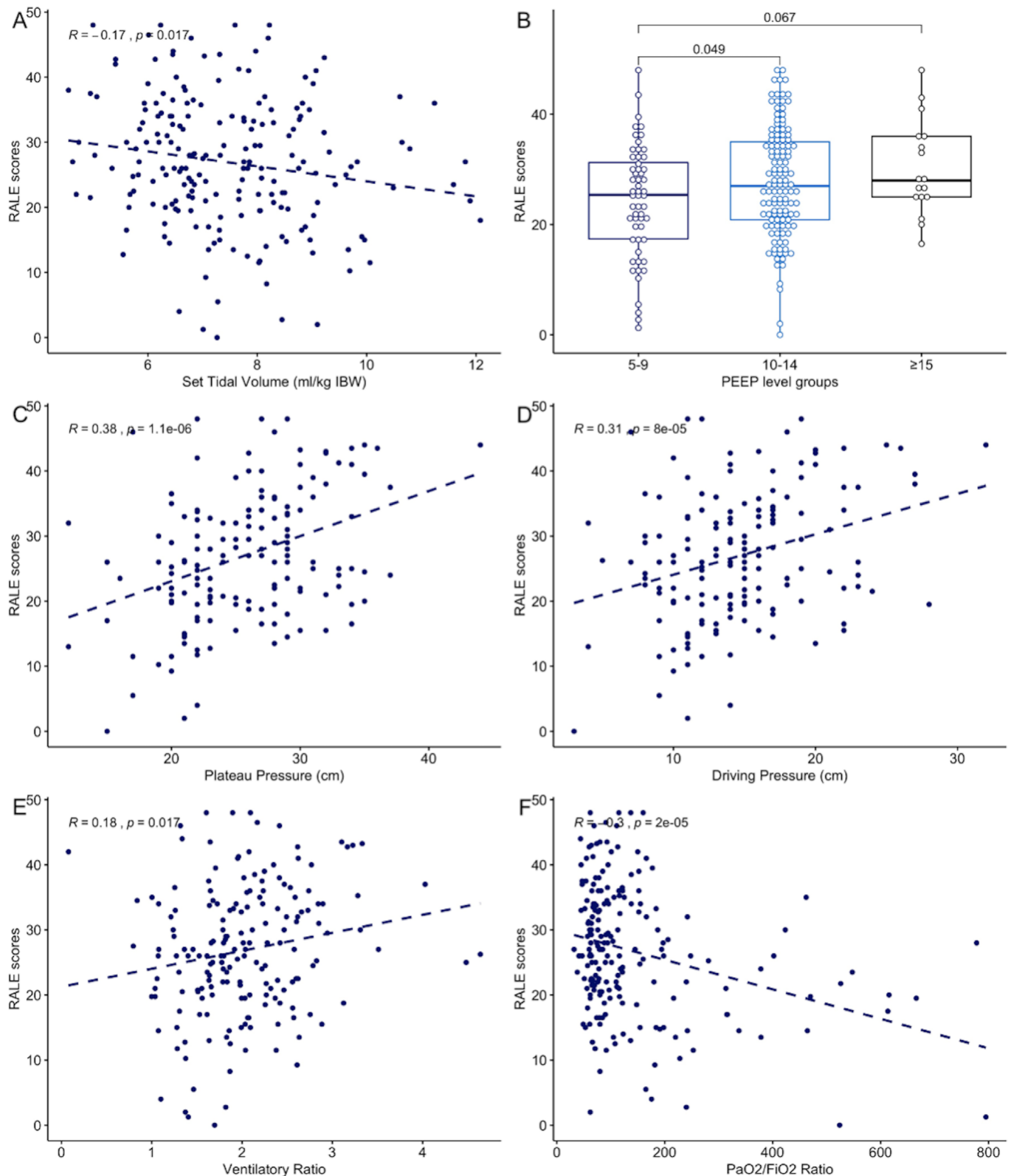


Figure 4 RALE scores were significantly associated with pulmonary dysfunction metrics in invasively mechanically ventilated patients. RALE scores were significantly associated with physician set parameters on mechanical ventilation (A: tidal volumes expressed in mL/kg of ideal body weight; B: RALE scores by positive end-expiratory pressure levels), correlated with respiratory mechanics (C: plateau pressure and D: driving pressure) and gas exchange parameters (E: positive correlation with ventilatory ratio, ie, worse CO₂ clearance and F: inverse correlation with PaO₂/FiO₂ ratio, ie, worse hypoxemia). FiO₂, fraction of inspired oxygen; PaO₂, partial pressure of oxygen; PEEP, positive end expiratory pressure; RALE, Radiographic Assessment of Lung Edema; IBW: ideal body weight.

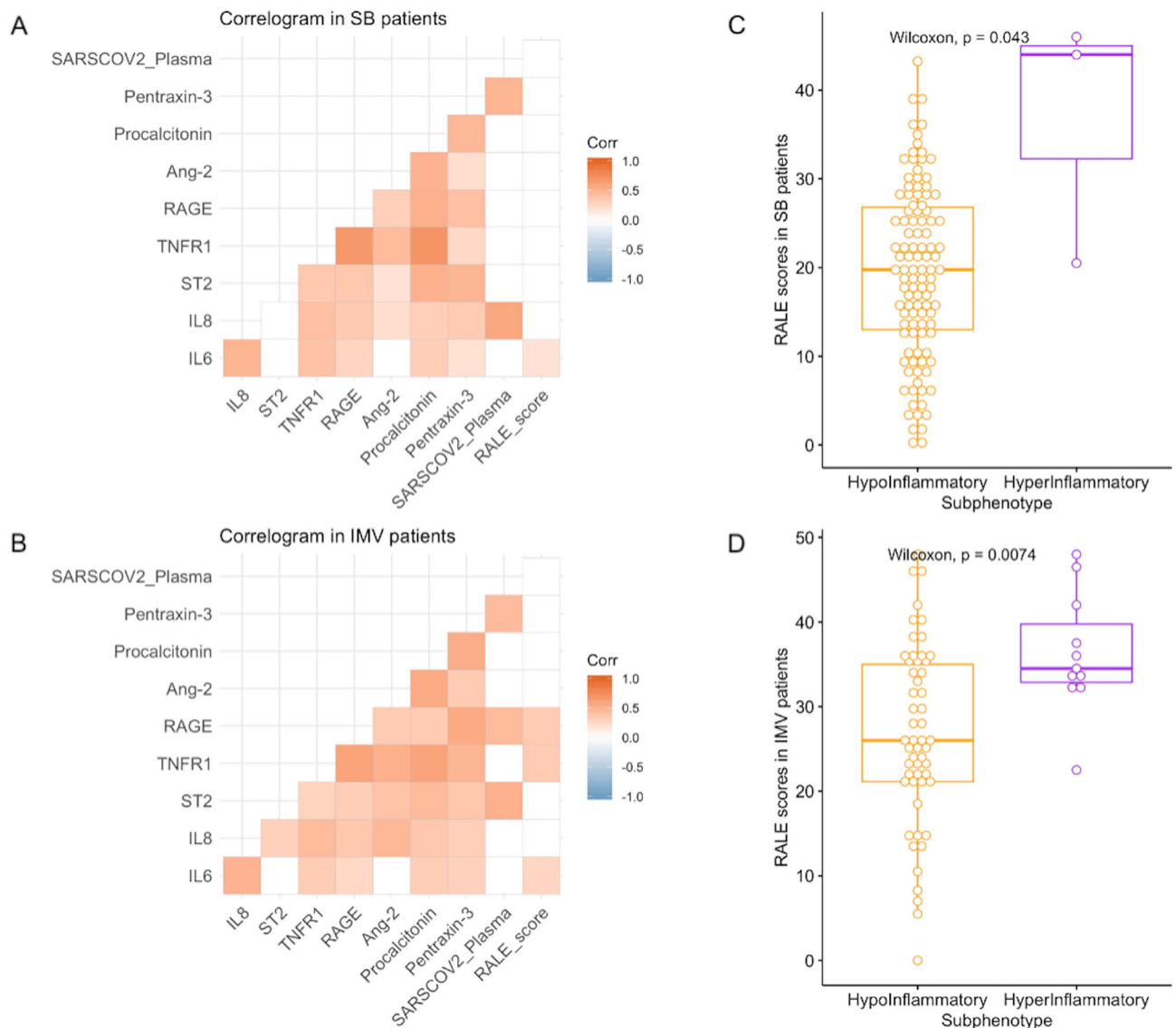


Figure 5 RALE scores were significantly correlated with biomarkers of host injury and inflammation, and significantly associated with the hyperinflammatory subphenotype. Correlograms of host response biomarkers, SARS-CoV-2 viral load and RALE scores in SB (A) and IMV (B) patients. Pearson's correlations are shown in colour code (red for positive and blue for negative) and only statistically significant correlations following adjustment for multiple testing by Benjamini-Hochberg method are shown (white boxes indicate non-significant correlations). Patients assigned to the hyperinflammatory subphenotype (based on a prediction from a parsimonious predictive model utilising IL-6, sTNFR1 and bicarbonate levels) had higher RALE scores in both SB (C) and IMV (D) patients. Ang2, angiotensin-2; IMV, invasive mechanical ventilation; IL, interleukin; RALE, Radiographic Assessment of Lung Edema; SB, spontaneously breathing; ST2, suppression of tumorigenicity-2; sTNFR1, soluble tumor necrosis factor receptor 1; sRAGE, soluble receptor of advanced glycation end-products.

respiratory support, significantly associated with patient-level pre-morbid covariates (such as age, BMI, history of COPD), correlated with respiratory dysfunction parameters (mechanics and gas exchange in IMV patients), were significantly associated with the adverse hyperinflammatory subphenotype of host responses, and shown to be prognostic of survival and discharge destination among survivors.

To study the reproducibility of RALE scoring and obtain a reliable database of radiographic assessments by expert reviewers, our team created the *Pulmo-Annotator* software, which allowed for stable storage of images/scores on a

cloud-based platform with parallel scoring from many individual reviewers. The *Pulmo-Annotator* capacities allowed us to study in depth technical aspects of image quality/penetration on resultant RALE scores as well as reviewer-related sources of variation. We were able to easily identify sporadic discordant scores or systematic patterns of deviation by reviewer, provide iterative feedback and optimise inter-rater reliability. Our exercise showed that RALE scoring is a trainable skill but requires a systematic mechanism to accomplish high inter-rater agreement. With an expansive database of expert-annotated RALE scores and image attributes, RALE scoring may

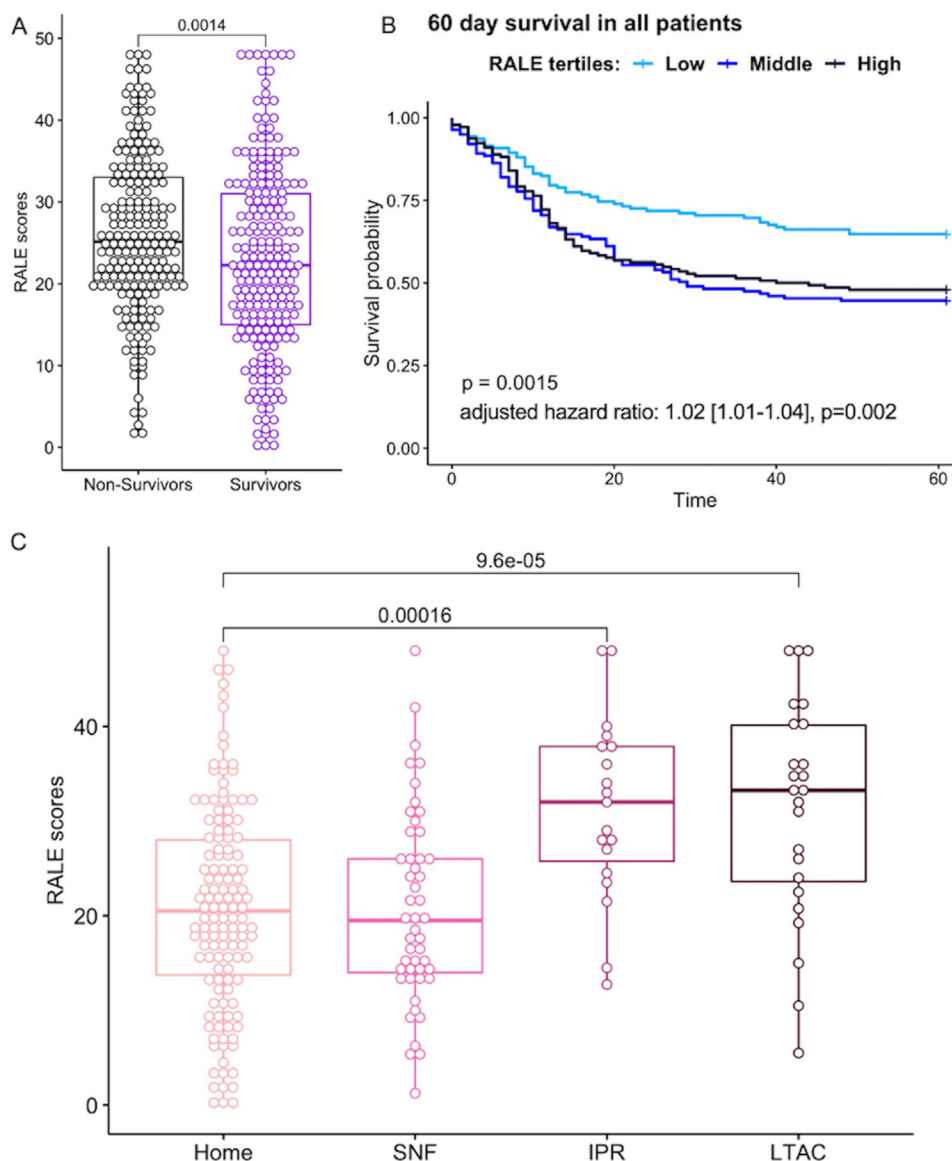


Figure 6 RALE score association with clinical outcomes. (A) Non-survivors had higher RALE scores than survivors (25.1 [19.8–33.0] vs 22.3 [15.0–31.0], $p=0.0014$). (B) By Kaplan-Meier curve analysis, patients in the low tertile of RALE scores (<19.6) had improved survival compared with patients in middle/high tertiles. (C) Higher care needs on final disposition were associated with higher RALE scores at baseline: long-term acute care facilities (LTAC) (33.3 [22.9–40.4]) or in-patient rehab (IPR) (32.0). RALE, Radiographic Assessment of Lung Edema; SNF, skilled nursing facility.

also become machine-learnable, which could transform the speed and scale of radiographic severity assessment in healthcare applications. There are multiple ongoing efforts in the field of machine learning for chest radiography,¹⁸ but any type of sophisticated model will require high-quality image annotations by clinical experts—as pursued in our study—to generate valid predictions.

We found that premorbid demographic variables were significantly associated with RALE scores at time of hospitalisation. Among IMV patients, those with possible indicators of frailty (such as older patients or nursing home residents) had significantly lower RALE scores, suggestive that such patients required a lower burden of acute respiratory illness to end up on IMV. Similarly, patients with

COPD had lower RALE scores, perhaps also indicative of their limited physiologic reserve as well as the anatomical emphysema accounting for increased radiographic lucency. On the other hand, patients with higher BMI had higher RALE scores, which may reflect both the known association of obesity with COVID-19 severity¹⁹ as well as diminished lung volumes and increased radiographic density from extrathoracic soft tissue. Therefore, such premorbid variables need to be accounted in analyses of radiographic indices with clinical endpoints.

We studied a large sample of 425 inpatients with a wide spectrum of COVID-19 severity, as illustrated by the range of WHO scale from 4 to 9 at timing of CXR and demonstrated a stepwise increase of RALE scores by

levels of respiratory support. We demonstrated significant associations of RALE scores not only with clinical severity but also with detailed metrics of pulmonary physiology (mechanics and gas exchange) as well as administered therapies used for the most severely ill patients with COVID-19 pneumonia. We found numerically higher correlations for pulmonary mechanics (eg, compliance) than gas exchange parameters (eg, ventilatory ratio), which may indicate that factors directly affecting mechanical measurements (such as pulmonary edema, atelectasis and obesity) are better reflected by radiographic densities rather than the complex and heterogeneous mechanisms of gas exchange in ARDS.²⁰ We validated our observations in an independent cohort of COVID-19 inpatients enriched for non-intubated patients. Of note, we observed a temporal correlation of RALE scores in IMV patients with the time spent from onset of the pandemic, that is, patients enrolled in 2021 having higher RALE scores than patients in the first waves of the pandemic in 2020. This temporal observation may reflect different population demographics (more frail patients hospitalised in 2020), evolving practices around initiation of IMV (more conservative criteria used as the pandemic progressed, and, thus, only sicker patients being intubated), or true, worse lung injury from emergent SARS-CoV-2 variants.

We detected novel associations of RALE scores with biomarkers of host innate immune response (IL-6 and sTNFR1) and lung epithelial injury (sRAGE) in IMV patients. The significant correlation between sRAGE levels and RALE scores validates previous findings,^{4 21 22} but the newly detected associations with innate immunity biomarkers and the hyperinflammatory subphenotype in both IMV and SB patients are suggesting that radiographic severity is not only representative of accumulated lung injury by the time of CXR but also indicative of ongoing inflammatory damage. Our findings suggest that radiographic severity assessments in severe pneumonia and ARDS may offer further insights into ongoing efforts to better characterise and understand the biological and clinical heterogeneity of such complex syndromes, and RALE scoring is an accessible tool for such purposes.

With a larger sample size than previous studies,^{6-9 23-26} and a systematic method supported by dedicated software, we validated the prognostic value of baseline RALE scores on clinical outcomes. Notably, RALE scores were predictive of 60-day survival even after adjustment of possible confounders (age, sex, history of COPD and BMI), which we chose to adjust for given their significant associations with RALE scores and known impact on COVID-19 outcomes. Nonetheless, when examined within each subgroup of levels of respiratory support (SB, IMV and ECMO), we did not find a significant prognostic effect of baseline RALE scores. Similar to our subgroup analyses, previous studies showed no prognostic value for baseline RALE score among intubated patients with COVID-19.^{27 28} Apart from small sample size considerations, such negative findings may be due to the fact the cross-sectional assessments among subjects with severe respiratory failure

to require IMV may not be sufficient to predict survival. Indeed, recent studies have shown that rising RALE scores on follow-up CXRs carry prognostic value in COVID-19,¹⁶ and we had previously shown that declining RALE scores in patients with non-COVID ARDS were associated with liberation from mechanical ventilation.⁵ Thus, although baseline RALE scores capture important cross-sectional parameters of clinical severity, reliable prognostication or assessment of treatment response may require longitudinal scoring of radiographic severity in the early period of hospitalisation.²⁹

Our study has several limitations. For logistical/feasibility reasons, we analysed only baseline CXRs from a total of 840 COVID-19 inpatients, and, thus, could not determine the trajectories of radiographic evolution that may offer important prognostic information. We analysed biospecimens only from two inpatient cohorts (ALIR and COVID-INC) and, therefore, our biomarker analyses may have had limited statistical power to detect additional significant associations. We used portable CXR images obtained as part of routine medical care and did not standardise image acquisition protocols for this study. Nonetheless, the analysed data set of images is representative of clinical practices in two major hospital systems and results are likely further generalisable.

CXRs represent the most used radiographic modality for diagnosis, monitoring severity and response to treatment among hospitalised patients with pneumonia. Although inferior in resolution and dimensionality compared with CT imaging, CXRs expose patients to substantially lower radiation dose, they are more rapid, cheaper, easily accessible and repeatable and can be used in low resource care settings. Current clinical practice involves qualitative or implicit interpretations of CXRs, for example, by narrative descriptions of densities (*focal*, *patchy* or *diffuse*) or qualifiers of progression (*improved* or *worse*). Such subjective, non-specific assessments are not reliable for objective evaluation of radiographic severity. Consequently, standard clinical practices fail to capitalise on objective imaging data provided by the most widely used modality. Our reproducible method for RALE scoring assessments offers a tool for thorough, quantitative study of radiographic severity.

With the wide availability of CXR imaging among hospitalised patients with COVID-19, incorporation of radiographic severity assessments into risk stratification may provide improved patient-level guidance on prognosis and treatment allocation.

Author affiliations

¹Internal Medicine Residency Program, UPMC Mercy, Pittsburgh, Pennsylvania, USA

²Division of Pulmonary, Critical Care, and Sleep Medicine, MetroHealth Medical Center, Cleveland, Ohio, USA

³Internal Medicine Residency Program, UPMC Pinnacle Harrisburg, Harrisburg, Pennsylvania, USA

⁴Division of Pulmonary, Allergy and Critical Care Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

⁵Acute Lung Injury Center of Excellence, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

⁶Computer Vision Group, Veytel LLC, Pittsburgh, Pennsylvania, USA

⁷Department of Pulmonary and Critical Care, Cleveland Clinic Florida, Weston, Florida, USA

⁸Internal Medicine Residency Program, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

⁹Department of Pulmonary and Critical Care, UPMC Pinnacle Harrisburg, Harrisburg, Pennsylvania, USA

¹⁰Department of Medicine, Division of Infectious Diseases, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA

¹¹Center for Medicine and the Microbiome, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

Twitter Georgios D Kitsios @KitsiosMd

Acknowledgements The authors would like to thank Olivia Glotfelty-Scheuering, a research librarian at UPMC Mercy Hospital, Manager of Library Services (MLIS), for her assistance in carrying out the literature search.

Contributors NA-Y: content guarantor, conceptualisation, methodology, validation, investigation, resources, writing—original draft, writing—review and editing, visualisation, project administration; GDK: content guarantor, conceptualisation, methodology, validation, formal analysis, investigation, resources, writing—original draft, writing—review and editing, visualisation, supervision, project administration, funding acquisition; SK: investigation, writing—review and editing; HQ: investigation, writing—review and editing; AK: investigation, writing—review and editing; HOL: investigation, writing—review and editing; NA: investigation, writing—review and editing; CS: resources, writing—review and editing; KJM: software, resources, data curation, writing—review and editing; CMD: software, resources, data curation, writing—review and editing; EKH: software, resources, data curation, writing—review and editing; CSB: software, resources, data curation, writing—review and editing; GMF: software, resources, data curation, writing—review and editing; RJ: software, resources, data curation, writing—review and editing; ASC: software, resources, data curation, writing—review and editing; DK: investigation, writing—review and editing; JDR: investigation, writing—review and editing; AIK: investigation, writing—review and editing; SS: investigation, writing—review and editing; AL: investigation, writing—review and editing; CEG: investigation, writing—review and editing; SRG: investigation, writing—review and editing; AH: investigation, writing—review and editing; WB: investigation, resources, writing—review and editing; FAS: investigation, resources, writing—review and editing; MB: investigation, resources, writing—review and editing; ML: investigation, resources, writing—review and editing; NP: investigation, resources, writing—review and editing; JE: investigation, resources, writing—review and editing; KG: investigation, writing—review and editing; NR: investigation, writing—review and editing; JJJ: investigation, writing—review and editing; CK: investigation, writing—review and editing; BM: investigation, resources, writing—review and editing; JL: investigation, resources, writing—review and editing; AM: investigation, resources, writing—review and editing; BJM: investigation, resources, writing—review and editing.

Funding Dr. Kitsios: University of Pittsburgh Clinical and Translational Science Institute, COVID-19 Pilot Award (Award/grant number: N/A); NIH (Award/grant number: K23 HL139987; R03 HL162655)

Competing interests Dr. Kitsios has received research funding from Karius, Inc. Dr. McVerry receives research funding from Bayer Pharmaceuticals, Inc. All other authors disclosed no conflict of interest

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by The Acute Lung Injury Registry (ALIR) and Biospecimen Repository: We enrolled subjects following admission to the ICU and obtained informed consent from the patients or their legally authorised representatives under the study protocol STUDY19050099 approved by the University of Pittsburgh Institutional Review Board (IRB). The COVID Inpatient Cohort (COVID-INC): We obtained consent from the patients or their legally authorised representatives under the study protocol STUDY20040036 approved by the University of Pittsburgh IRB. The Prognostication for COVID-19 Patients Admitted to Intensive Care Units at UPMC Pinnacle (PROCOPI) study: we performed a retrospective chart review and collected data from the electrical medical record (EMR) under a minimal risk study protocol (20E059) approved by the UPMC Pinnacle IRB. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. All data relevant to the study are included in the article or uploaded as supplementary information.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Nameer Al-Yousif <http://orcid.org/0000-0002-2657-2360>

REFERENCES

- Cushman D, Bennett O, Berka R. An overview of the National COVID-19 chest imaging database: data quality and cohort analysis. *Gigascience* 2021;10.
- Wu G, Li X. Mobile x-rays are highly valuable for critically ill COVID patients. *Eur Radiol* 2020;30:5217–9.
- Warren MA, Zhao Z, Koyama T, et al. Severity scoring of lung oedema on the chest radiograph is associated with clinical outcomes in ARDS. *Thorax* 2018;73:840–6.
- Jabaudon M, Audard J, Pereira B, et al. Early changes over time in the radiographic assessment of lung edema score are associated with survival in ARDS. *Chest* 2020;158:2394–403.
- Kotok D, Yang L, Evankovich JW, et al. The evolution of radiographic edema in ARDS and its association with clinical outcomes: a prospective cohort study in adult patients. *J Crit Care* 2020;56:222–8.
- Kerpel A, Apter S, Nissan N, et al. Diagnostic and prognostic value of chest radiographs for COVID-19 at presentation. *West J Emerg Med* 2020;21:1067–75.
- Sensusiati AD, Amin M, Nasronudin N, et al. Age, neutrophil lymphocyte ratio, and radiographic assessment of the quantity of lung edema (PALE) score to predict in-hospital mortality in COVID-19 patients: a retrospective study. *F1000Res* 2020;9:1286.
- Mushtaq J, Pennella R, Lavallo S, et al. Initial chest radiographs and artificial intelligence (AI) predict clinical outcomes in COVID-19 patients: analysis of 697 Italian patients. *Eur Radiol* 2021;31:1770–9.
- Ciceri F, Castagna A, Rovere-Querini P, et al. Early predictors of clinical outcomes of COVID-19 outbreak in Milan, Italy. *Clin Immunol* 2020;217:108509.
- Bain W, Yang H, Shah FA, et al. COVID-19 versus Non-COVID-19 acute respiratory distress syndrome: comparison of demographics, physiologic parameters, inflammatory biomarkers, and clinical outcomes. *Ann Am Thorac Soc* 2021;18:1202–10.
- Drohan C, Bain W, Kitsios GD. Reply: understanding COVID-19 acute respiratory distress syndrome: new pathogen, same heterogeneous syndrome. *Ann Am Thorac Soc* 2022;19:151–60.
- Drohan CM, Nouraei SM, Bain W, et al. Biomarker-Based classification of patients with acute respiratory failure into inflammatory subphenotypes: a single-center exploratory study. *Crit Care Explor* 2021;3:e0518.
- Sinha P, Delucchi KL, McAuley DF, et al. Development and validation of parsimonious algorithms to classify acute respiratory distress syndrome phenotypes: a secondary analysis of randomised controlled trials. *Lancet Respir Med* 2020;8:247–57.
- Jacobs JL, Bain W, Naqvi A, et al. Severe acute respiratory syndrome coronavirus 2 viremia is associated with coronavirus disease 2019 severity and predicts clinical outcomes. *Clin Infect Dis* 2022;74:1525–33.
- Jacobs JL, Naqvi A, Shah FA. Plasma SARS-CoV-2 RNA levels as a biomarker of lower respiratory tract SARS-CoV-2 infection in critically ill patients with COVID-19. *medRxiv* 2022.
- Kotok D, Robles JR, E Girard C, et al. Chest radiograph severity and its association with outcomes in subjects with COVID-19 presenting to the emergency department. *Respir Care* 2022;67:871–8.

- 17 Ghaferi AA, Schwartz TA, Pawlik TM. STROBE reporting guidelines for observational studies. *JAMA Surg* 2021;156:577–8.
- 18 Roberts M, Driggs D, Matthew T, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat Mach Intell* 2021.
- 19 Kompaniyets L, Goodman AB, Belay B, et al. Body Mass Index and Risk for COVID-19-Related Hospitalization, Intensive Care Unit Admission, Invasive Mechanical Ventilation, and Death - United States, March-December 2020. *MMWR Morb Mortal Wkly Rep* 2021;70:355–61.
- 20 Radermacher P, Maggiore SM, Mercat A. Fifty years of research in ARDS. gas exchange in acute respiratory distress syndrome. *Am J Respir Crit Care Med* 2017;196:964–84.
- 21 Kapandji N, Yvin E, Devriese M, et al. Importance of lung epithelial injury in COVID-19-associated acute respiratory distress syndrome: value of plasma soluble receptor for advanced glycation end-products. *Am J Respir Crit Care Med* 2021;204:359–62.
- 22 Wick KD, Leligdowicz A, Zhuo H, et al. Mesenchymal stromal cells reduce evidence of lung injury in patients with ARDS. *JCI Insight* 2021;6. doi:10.1172/jci.insight.148983. [Epub ahead of print: 22 Jun 2021].
- 23 Galloway JB, Norton S, Barker RD, et al. A clinical risk score to identify patients with COVID-19 at high risk of critical care admission or death: an observational cohort study. *J Infect* 2020;81:282–8.
- 24 Ebrahimian S, Homayounieh F, Rockenbach MABC, et al. Artificial intelligence matches subjective severity assessment of pneumonia for prediction of patient outcome and need for mechanical ventilation: a cohort study. *Sci Rep* 2021;11:858.
- 25 Cozzi D, Albanesi M, Cavigli E, et al. Chest X-ray in new coronavirus disease 2019 (COVID-19) infection: findings and correlation with clinical outcome. *Radiol Med* 2020;125:730–7.
- 26 Shen B, Hoshmand-Kochi M, Abbasi A, et al. Initial chest radiograph scores inform COVID-19 status, intensive care unit admission and need for mechanical ventilation. *Clin Radiol* 2021;76:473.e1–473.e7.
- 27 Valk CMA, Zimatore C, Mazzinari G, et al. The prognostic capacity of the radiographic assessment for lung edema score in patients with COVID-19 acute respiratory distress Syndrome-An international multicenter observational study. *Front Med* 2021;8:772056.
- 28 Herrmann J, Adam EH, Notz Q, et al. COVID-19 induced acute respiratory distress syndrome-a multicenter observational study. *Front Med* 2020;7:599533.
- 29 Sheshadri A, Shah DP, Godoy M, et al. Progression of the radiologic severity index predicts mortality in patients with parainfluenza virus-associated lower respiratory infections. *PLoS One* 2018;13:e0197418.