






Systematic review of diagnostic methods for acute respiratory distress syndrome

Laura A. Hagens ¹, Nanon F.L. Heijnen², Marry R. Smit¹, Marcus J. Schultz^{1,3,4}, Dennis C.J.J. Bergmans ², Ronny M. Schnabel⁴ and Lieuwe D.J. Bos ^{1,5}, on behalf of the DARTS consortium

Affiliations: ¹Dept of Intensive Care, Amsterdam UMC, location AMC, University of Amsterdam, Amsterdam, The Netherlands. ²Dept of Intensive Care, Maastricht University Medical Centre+, Maastricht, The Netherlands. ³Mahidol-Oxford Tropical Medicine Research Unit (MORU), Mahidol University, Bangkok, Thailand. ⁴Nuffield Dept of Medicine, University of Oxford, Oxford, UK. ⁵Dept of Respiratory Medicine, Amsterdam UMC, location AMC, University of Amsterdam, Amsterdam, The Netherlands.

Correspondence: Laura A. Hagens, Dept of Intensive Care, Amsterdam University Medical Center, location AMC, Meibergdreef 9, Amsterdam 1105 AZ, The Netherlands. E-mail: l.a.hagens@amsterdamumc.nl

ABSTRACT

Rationale: Acute respiratory distress syndrome (ARDS) is currently diagnosed by the Berlin definition, which does not include a direct measure of pulmonary oedema, endothelial permeability or pulmonary inflammation. We hypothesised that biomarkers of these processes have good diagnostic accuracy for ARDS.

Methods: Medline and Scopus were searched for original diagnostic studies using minimally invasive testing. Primary outcome was the diagnostic accuracy per test and was categorised by control group. The methodological quality was assessed with QUADAS-2 tool. Biomarkers that had an area under the receiver operating characteristic curve (AUROCC) of >0.75 and were studied with minimal bias against an unselected control group were considered to be promising.

Results: Forty-four articles were included. The median AUROCC for all evaluated tests was 0.80 (25th to 75th percentile: 0.72–0.88). The type of control group influenced the diagnostic accuracy ($p=0.0095$). Higher risk of bias was associated with higher diagnostic accuracy (AUROCC 0.75 for low-bias, 0.77 for intermediate-bias and 0.84 for high-bias studies; $p=0.0023$). Club cell protein 16 and soluble receptor for advanced glycation end-products in plasma and two panels with biomarkers of oxidative stress in breath showed good diagnostic accuracy in low-bias studies that compared ARDS patients to an unselected intensive care unit (ICU) population.

Conclusion: This systematic review revealed only four diagnostic tests fulfilling stringent criteria for a promising biomarker in a low-bias setting. For implementation into the clinical setting, prospective studies in a general unselected ICU population with good methodological quality are needed.



@ERSpublications

Accuracy of diagnosis of acute respiratory distress syndrome (ARDS) is associated with risk of bias. There is a lack of validated diagnostic tests in an unbiased setting, emphasising the need for quality driven diagnostic research in ARDS. <https://bit.ly/2GfPAvf>

Cite this article as: Hagens LA, Heijnen NFL, Smit MR, *et al.* Systematic review of diagnostic methods for acute respiratory distress syndrome. *ERJ Open Res* 2021; 7: 00504-2020 [<https://doi.org/10.1183/23120541.00504-2020>].



Introduction

Acute respiratory distress syndrome (ARDS) is characterised by the acute onset of non-cardiogenic pulmonary oedema and hypoxaemia and is associated with high mortality and morbidity [1, 2]. The combination of increased permeability of the endothelium and injury to the alveolar epithelium results in protein-rich alveolar fluid [3]. Procoagulatory and inflammatory proteins and metabolites of oxidative stress are abundant in the alveoli of ARDS patients [1]. Translational evidence suggests that lung injury can be initiated *via* alveolar inflammation as well as endothelial injury [3].

ARDS is currently diagnosed by means of the Berlin definition [2, 4]. The criteria utilise information that is commonly available at the bedside: it captures hypoxaemia *via* arterial oxygen tension (P_{aO_2})/inspiratory oxygen fraction (F_{IO_2}) and alveolar oedema *via* bilateral opacities on chest radiography. Interpretation of chest images gives inconsistent results, which makes diagnosing ARDS subjective and challenging [5]. But even without this limitation, chest radiography can only diagnose alveolar oedema after onset, and it will never identify the molecular mechanism resulting in alveolar oedema.

A diagnostic test that captures extravascular lung water earlier or that identifies the pathophysiological mechanisms resulting in lung injury is likely to lead to better diagnostic accuracy for the diagnosis of ARDS. Biomarkers are objective, can be derived from plasma, urine, bronchoalveolar lavage fluid or breath with minimally invasive methods, and can be reflective of key biological pathways known to be involved in ARDS development [3]. A clear diagnostic biomarker could help clinical decisions in several matters: 1) earlier recognition of pulmonary oedema can help to prevent fluid overload; and 2) identification of the biological pathways resulting in lung injury can inform targeted treatments in personalised medicine randomised controlled trials. Up to now, a clear diagnostic biomarker has not been identified [6], and because the accuracy of a biomarker may be biased by the quality of the performed study, it is important to evaluate potential bias when reviewing the currently available evidence.

The aim of this review is to give an overview of the minimally invasive diagnostic tests available for assessing the pathogenesis of ARDS in order to achieve an early and objective diagnosis of ARDS in patients on the intensive care unit (ICU). We hypothesised that tests for 1) pulmonary oedema, 2) endothelial permeability, 3) pulmonary inflammation, 4) coagulation and 5) oxidative stress have good diagnostic accuracy and that the diagnostic accuracy is lower in well-conducted studies than in biased studies.

Methods

Search

A systematic review following preferred reporting items for systematic reviews was performed [7]. We searched Medline and Scopus for potentially relevant articles up to June 9, 2020. The following terms were used: ARDS, acute lung injury (ALI), inflammation, biomarker, cytokine, breath, oedema, lung water, diagnosis, diagnostic, human and adult. The exact search can be found in the supplementary material. Two researchers (LB and LH) independently reviewed the abstracts and/or full text manuscripts and selected relevant articles. Disagreements were resolved in a consensus meeting. The review protocol is registered at PROSPERO (www.crd.york.ac.uk/prospero, CRD42020186974).

Selection criteria

Inclusion criteria were 1) original research with a diagnostic purpose that 2) reported the diagnostic accuracy of 3) a minimally invasive test for 4) pathophysiological mechanisms of ARDS 5) comparing patients having ARDS with relevant other patients. A relevant control group was defined as patients at risk for ARDS, for example receiving mechanical ventilation or respiratory support. Studies with a focus on treatment or prediction of ARDS were excluded. Other exclusion criteria were articles not available in English, animal or preclinical studies, studies in children, and unclear reference or index test. Finally, studies were excluded if the primary endpoint, diagnostic accuracy of the index test, was not available and could not be inferred from the data. Details are given below.

Reference test

The first American–European consensus criteria (AECC) date back to 1994 [8]. Studies from 1994 until 2012 were included if they used the applicable AECC definition or criteria that were closely related. Patients within the category of “acute lung injury” were included in this review as having ARDS, since in 2012 the newly introduced Berlin definition included this group as mild ARDS. For studies from 2012 onward the Berlin ARDS definition was used as the reference test [2].

Index test

The index tests were categorised into the following domains, regarding the pathophysiological mechanisms: 1) endothelial permeability, 2) pulmonary oedema, 3) inflammation, 4) coagulation or

5) oxidative stress. The index test should assess one of the pathophysiological mechanisms of ARDS, so studies looking into diagnostic tools based on cardiac function or, for example, terms in the electronic health record were excluded. Second, the tests were categorised based on the sample material: plasma, breath, alveolar fluid or other. The limit for invasiveness of the test was set at performing a bronchoalveolar lavage procedure; all tests more invasive than this method were excluded. Effectively, this excludes any type of biopsy. Finally, index tests were categorised on diagnostic accuracy: a potentially clinically relevant diagnostic accuracy was defined as an area under the receiver operating characteristic curve (AUROCC) of >0.75 .

Outcome and data extraction

The primary outcome was the AUROCC of the diagnostic test. If not available, sensitivity and specificity were used in a secondary analysis. In case both results were not reported, and the paper included a figure with individual data points, we extracted the data from the figure and recalculated the AUROCC, sensitivity and specificity. If this was unsuccessful, the study was excluded. The study population was categorised into: 1) general ICU patients, 2) cardiopulmonary surgery or cardiac ICU patients, 3) sepsis patients or 4) highly selected populations (such as only trauma patients or organ transplant patients). The control group was categorised into: 1) unselected ICU patients, 2) patients with cardiopulmonary oedema (CPE) and 3) patients with (suspected) pneumonia.

Methodological assessment and categorisation

The methodological quality of each article was assessed with the QUADAS-2 tool [9]. Risk of bias was assessed concerning patient selection, blinding and use of index test, blinding and use of reference test, and regarding patient flow. Timing of the test was considered to have a low risk of bias when the index test and reference test were performed on the same day or subsequent day. All tests that were performed later were classified as having a high risk of bias. For the assessment of the overall methodological quality of the papers, a cumulative score was calculated. The risk and concern scores were classified as follows: “High” 1 point; “Unclear” 0.5 points; “Low” 0 points, resulting in a cumulative score between 0 and 6. Based on the cumulative score, studies were categorised into tertiles: “Low”, “Intermediate” and “High” biased studies with the following cut-offs: Low: ≤ 1.5 ; Intermediate: >1.5 and ≤ 2.5 ; High: >2.5 points.

Statistical analysis

The AUROCC was summarised for each index test (so one study investigating multiple tests would provide multiple AUROCCs) and stratified for the following domains:

- pathophysiological processes: endothelial permeability, pulmonary oedema, inflammation, coagulation or oxidative stress
- population: general ICU, sepsis, cardiac care unit (CCU) or a specific group
- control group: unselected ICU, CPE or pneumonia
- sample material: plasma, breath, alveolar fluid or other
- quality of the study: low, intermediate or high risk of bias

Subsequently, the AUROCC was compared between the groups with one-way ANOVA. Significant results, defined as a p-value <0.05 , were further studied using *post hoc* analysis with pairwise t-tests. The influence of the processes resulting in changes in biomarker concentration, such as tested material, pathophysiological mechanism, population and control group on the association between bias and diagnostic accuracy, was evaluated using two-way ANOVA. For studies that reported sensitivity and specificity, meta-analysis of diagnostic accuracy was performed using the *mada* package to visually confirm any association that was found for AUROCC in the primary analysis [10]. All analyses were performed in R version 3.6.1 using the R-studio interface.

Results

The Medline and Scopus searches were last updated on June 9, 2020 and revealed 1096 articles, of which 958 remained after removing duplicates (figure 1). Title screening resulted in 143 eligible articles, of which 52 remained after reading the abstracts. After reading the full texts, 44 articles were included (figure 1 and table 1). Assessment of the included articles yielded a total of 84 index tests, including 68 different types of tests. Plasma biomarkers were most frequently studied (48 out of 84 (57%)). Categorisation based on pathophysiological mechanisms led to the following numbers: 39 tests for inflammation, 20 for endothelial permeability, 15 for pulmonary oedema, eight for oxidative stress and two for coagulation (supplementary table S1). The following populations were included in the studies: 29 studies with general ICU population (66%), nine studies with a specific population (20%), five studies with sepsis patients (11%) and one study with a CCU population (2%). The control group consisted of patients with CPE in 11 studies (25%) and

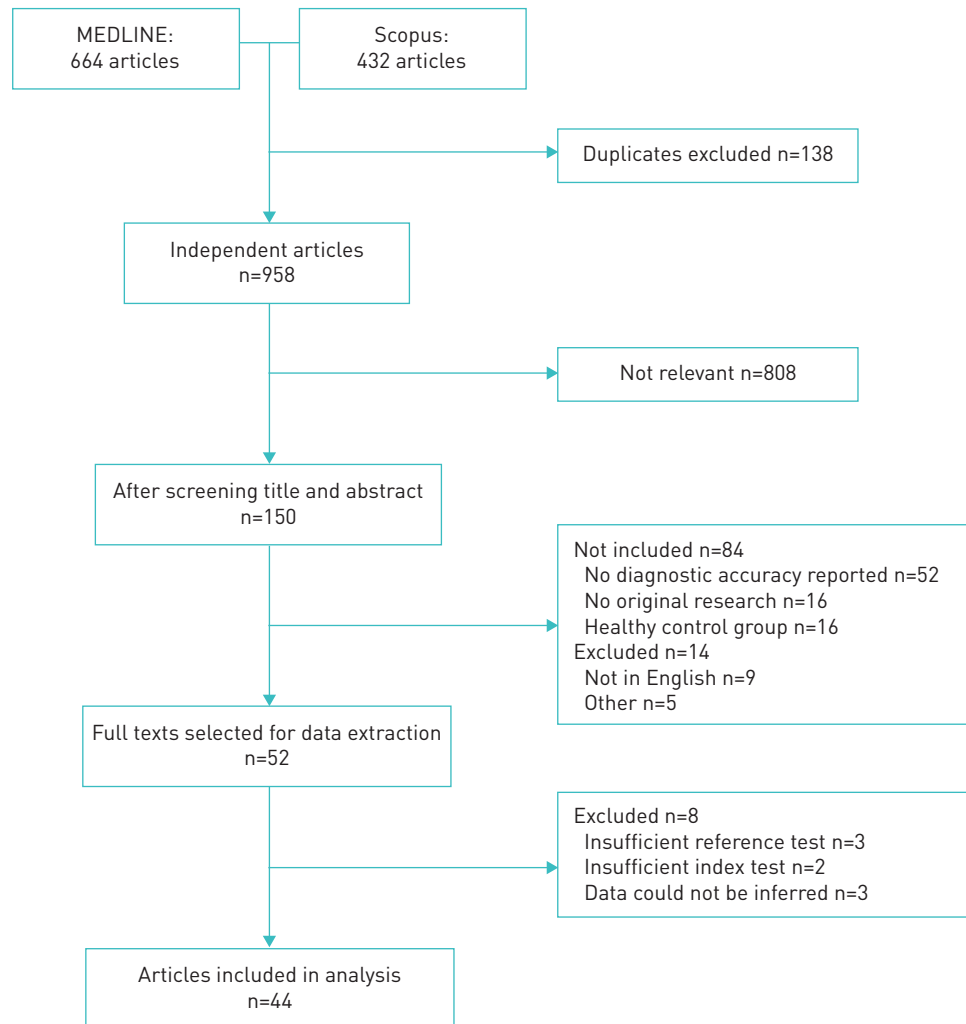


FIGURE 1 Flowchart of article selection.

pneumonia in two studies (5%). The other studies included a cohort of ICU patients that did not have CPE or pneumonia specifically (70%).

Accuracy for diagnosing ARDS

For 74 of the 84 tests (88%) the AUROCC was available. The median AUROCC was 0.80 with an interquartile range (IQR) from 0.72 to 0.88. A good diagnostic accuracy (AUROCC >0.75) was shown in 47 of the 74 tests (64%), spread over all different processes associated with ARDS development.

Diagnostic accuracy was higher in tests comparing ARDS patients with CPE patients (median AUROCC: 0.89, IQR: 0.81–0.93) than in ARDS patients compared to unselected ICU patients (median AUROCC: 0.78, IQR: 0.71–0.84, $p=0.0095$). The AUROCC was not different between studies with a control group of pneumonia patients compared with unselected ICU patients ($p=0.82$) or between pneumonia patients compared with CPE patients ($p=0.14$; figure 2). No differences in AUROCC were found for the type of studied pathophysiological mechanism ($p=0.76$), the studied biological material ($p=0.51$) and the population ($p=0.60$).

Sensitivity and specificity were available or could be calculated for 46 out of 84 (55%) studies. Similar patterns regarding the influence of the type of test, the studied biological material, the population and the control group were found when these studies were evaluated based on a single cut-off (supplementary figures S1–S4).

Assessment of bias in study methodology

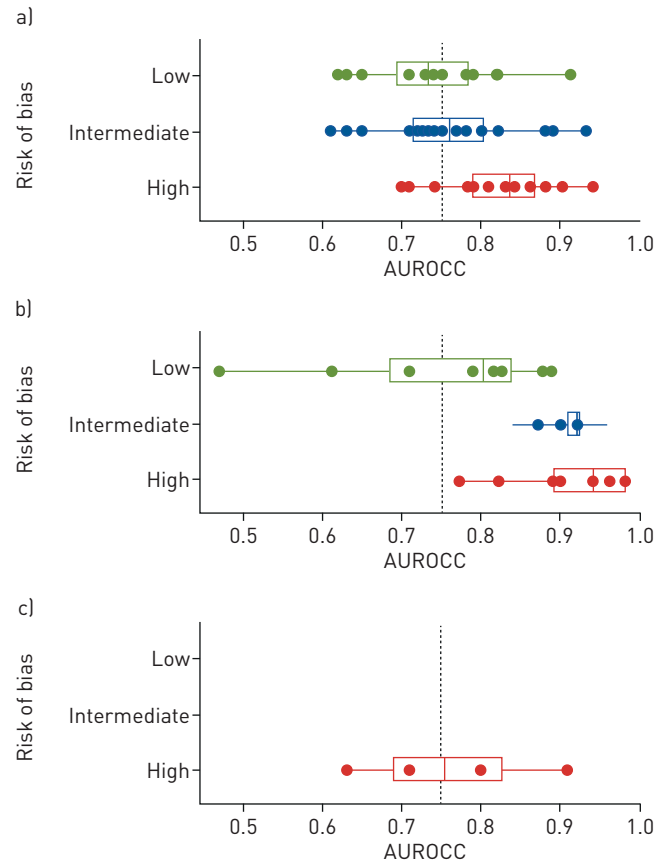
The methodological quality as assessed by the QUADAS-2 tool is shown in table 2. The final score varied among studies, with a cumulative score with a median of 2.0 (IQR: 1.5 to 3.00). Categorisation into tertiles

TABLE 1 Included studies

First author [ref.]	Year	Population	ARDS		Control group		Biomarkers
			Definition	Number	Type	Number	
ABBAS [11]	2017	ICU general	Berlin	15	Unselected ICU	19	MBG/creatinine
AMAN [12]	2011	ICU general	Other	48	Unselected ICU	14	Albumin, transferrin
ARIF [13]	2002	ICU general	Other	12	CPE	16	Transferrin, total protein, albumin, PLI
BAI [14]	2017	Specific group	Berlin	29	Unselected ICU	21	Glutamate increase
BAJWA [15]	2013	ICU general	AECC	826	CPE	209	sST2
BAUER [16]	2000	Specific group	Other	46	CPE	20	TNF- α , IL-1 β , IL-6
BERSTEN [17]	2001	ICU general	Other	34	Unselected ICU	20	SP-A, SP-B
Bos [18]	2014	ICU general	Berlin	59	Unselected ICU	42	Three metabolites identified by GCMS
Bos [19]	2014	ICU general	Berlin	92	Unselected ICU	58	Pattern recognised by e-nose 1, pattern recognised by e-nose 2
BRETT [20]	1998	ICU general	AECC	18	Unselected ICU	13	NO
BURSTEN [21]	1996	ICU general	AECC	13	Unselected ICU	17	Acyl ratio
COPETTI [22]	2008	ICU general	AECC	40	CPE	18	7 characteristics of lung ultrasound
DETERMANN [23]	2009	ICU general	AECC	15	Pneumonia	10	CC16, CC16 increase, KL-6, sRAGE, SP-D
EL SOLH [24]	2005	Specific group	AECC	34	Unselected ICU	17	PAI-1
FREMONT [25]	2010	Specific group	AECC	85	Unselected ICU	107	Model with 7 biomarkers, model with 3 biomarkers
GRISSEM [26]	2003	ICU general	AECC	6	Unselected ICU	33	PAF-AH
HERRERA [27]	1988	Cardiac care unit	Other	11	CPE	23	PPK
HOEBOER [28]	2015	Specific group	Berlin	1790	Unselected ICU	143	Albumin
HOWRYLAK [29]	2009	ICU general	Berlin	53	Unselected ICU	48	Genetic model
HUANG [30]	2019	Sepsis	AECC	21	Unselected ICU	21	IG percentage
IZQUIERDO-GARCIA [31]	2018	Specific group	AECC	31	Pneumonia	25	Metabolic biomarker panel
JABAUDON [32]	2018	ICU general	Berlin	188	Unselected ICU	34	sRAGE 0, sRAGE 1, sRAGE 1-0
JORENS [33]	1992	ICU general	Other	12	Unselected ICU	15	IL-8
KIETZMANN [34]	1993	ICU general	Other	29	Unselected ICU	7	Maximum H ₂ O ₂
KUSHIMOTO [35]	2012	ICU general	Other	59	CPE	207	PVPI, ITBV
LETOURNEAU [36]	2012	ICU general	AECC	10	Unselected ICU	19	EVLWi, P_{aO_2}/F_{iO_2} , EDI
LIN [37]	2012	ICU general	Berlin	129	Unselected ICU	83	Copeptin
LIN [38]	2013	ICU general	AECC	28	CPE	78	HBP
LIN [39]	2018	ICU general	AECC	34	CPE	87	CC16, CRP
LIU [40]	2015	Specific group	Berlin	10	Unselected ICU	18	MDA, NO, H ₂ O ₂ , 8-isoprostaglandin F ₂ α , TNF- α , IL-8, SOD, IL-10
LIU [41]	2017	Sepsis	AECC	115	Unselected ICU	19	PARK7
MONNET [42]	2007	ICU general	Other	12	CPE	36	PVPI, ELVWi/GEDVi
PARK [43]	2017	ICU general	Berlin	368	Unselected ICU	39	SP-D
SATO [44]	2004	ICU general	AECC	9	Unselected ICU	28	KL-6
SEKIGUCHI [45]	2015	ICU general	AECC	59	CPE	42	CCUS prediction model
SHAN [46]	2016	ICU general	Berlin	47	Unselected ICU	45	suPAR, hsCRP, PCT
SWEENEY [47]	2018	ICU general	Berlin	236	Unselected ICU	148	Model with 7 genes
VERHEIJ [48]	2005	ICU general	Other	13	CPE	13	PLI, PTCER, Ga/Tc slope/intercept, Ga/Tc monoexponential TER
WARE [49]	2013	Sepsis	AECC	100	Unselected ICU	100	Model with 5 biomarkers
WARE [50]	2017	Specific group	Berlin	315	Unselected ICU	78	Model with 11 biomarkers, model with 2 biomarkers
WU [51]	2019	Specific group	Berlin	60	Unselected ICU	13	CC16
XUE [52]	2015	Sepsis	Other	94	Unselected ICU	73	Tissue factor
YEH [53]	2017	Sepsis	AECC	111	Unselected ICU	18	Gas6
ZHOU [54]	2019	ICU general	Berlin	27	Unselected ICU	21	Panel with 9 metabolites

Abbreviations (in alphabetical order): AECC: American-European consensus criteria; ARDS: acute respiratory distress syndrome; CC16: Club cell protein 16; CCUS: critical care ultrasonography; CPE: cardiac pulmonary oedema; CRP: C-reactive protein; EDI: EVLW physiological dead space index; EVLWi: extravascular lung water index; F_{iO_2} : inspiratory oxygen fraction; Gas6: growth arrest-specific gene 6; Ga/Tc: gallium/technetium; GCMS: gas chromatography mass spectrometry; GEDVi: global end-diastolic volume indexed for body surface; H₂O₂: hydrogen peroxide; HBP: heparin-binding protein; hsCRP: high sensitive C-reactive protein; ICU: intensive care unit; IG: immature granulocyte; IL: interleukin; ITBV: intrathoracic blood volume; KL: Krebs von den Lungen; MBG: marinobufagenin; MDA: malondialdehyde; NO: nitric oxide; PAF-AH: platelet-activating factor acetylhydrolase; PAI-1: plasminogen activator inhibitor-1; P_{aO_2} : arterial oxygen tension; PARK7: Parkinson disease 7; PCT: procalcitonin; PLI: pulmonary leak index; PPK: prekallikrein; PTCER: pulmonary transcapillary escape rate; PVPI: extravascular lung water/pulmonary blood volume; SOD: superoxide dismutase; SP: surfactant protein; sRAGE: soluble receptor for advanced glycation end-products; sST2: soluble suppression of tumorigenicity-2; suPAR: soluble urokinase-type plasminogen activator receptor; TER: transcapillary escape rate; TNF: tumour necrosis factor.

FIGURE 2 Association between diagnostic accuracy and risk of bias, stratified for the control group. Each dot represents a diagnostic test; multiple tests could be evaluated per study. Area under the receiver operating characteristic curve (AUROCC) was higher in tests comparing acute respiratory distress syndrome (ARDS) patients with cardiopulmonary oedema (CPE) patients, then in ARDS patients compared with general intensive care unit (ICU) patients ($p=0.0095$). The AUROCC was not different between studies with control group of pneumonia patients compared to unselected ICU patients ($p=0.82$) or between pneumonia patients compared to CPE patients ($p=0.14$). a) ICU all; b) CPE; c) pneumonia.



based on the cumulative score led to 14 studies in the “Low” bias category, 17 studies in the category “Intermediate” and 13 studies in the “High” category (table 2). The risk of bias was most frequently observed for patient selection, blinding of the index test and in the timing of the index test.

Association between bias and diagnostic accuracy

The risk of bias of the study was associated with the diagnostic accuracy of the index test ($p=0.0023$). The median AUROCC was 0.75 (IQR: 0.69 to 0.82) for low-bias studies, 0.77 (IQR: 0.72 to 0.88) for intermediate-bias studies and 0.84 (IQR: 0.79 to 0.90) for high-bias studies (figure 3). Based on the pairwise comparison, the AUROCC was significantly higher for studies in the high-bias category, compared to the intermediate- and low-bias category ($p=0.020$ and $p=0.00077$, respectively). Two-way ANOVA showed that this association was consistent after correction for the type of test ($p=0.0027$), sample material ($p=0.0026$), population ($p=0.0026$) and control group ($p=0.0011$). Figure 3 shows the diagnostic accuracy per test after stratification for risk of bias. The other comparisons are visualised in supplementary figures S5–S7. The same trend was visible in the analysis of sensitivity and specificity with respect to the risk of bias (supplementary figures S1–S4).

Low-bias studies with good diagnostic accuracy

Nine tests showed a good diagnostic accuracy in the low-bias group. Of these, five compared ARDS *versus* CPE and four compared ARDS *versus* the general ICU population. The studies comparing ARDS *versus* ICU patients measured biomarkers in plasma and metabolites in exhaled breath. The plasma biomarkers assessed were Club cell protein 16 (CC16) and soluble receptor for advanced glycation end-products (sRAGE), assessing inflammation and permeability. In exhaled breath a panel with three metabolites and a panel with nine metabolites were assessed. The three metabolites describe oxidative stress and the nine metabolites most likely do too, but no clear reporting on this topic was available. The first three studies performed the test on the day of ARDS diagnosis or the day after, providing early information on diagnosis of ARDS. For the last test it was unclear at what time it was performed.

Discussion

When comparing patients with ARDS to patients who are also admitted to the ICU, only four studies yielded a good diagnostic accuracy with a limited risk of potential bias. We identified CC16 and sRAGE in

TABLE 2 Bias assessment

First author [ref.]	Year	Domain 1		Domain 2	Domain 3		Domain 4	Cumulative score	Risk of bias category
		Risk score	Concern	Risk score	Risk score	Concern	Risk score		
ABBAS [11]	2017	1	0	0.5	0.5	0	0.5	2.5	Intermediate
AMAN [12]	2011	1	0	0.5	0	0	0.5	2	Intermediate
ARIF [13]	2002	1	1	0.5	0	0	1	3.5	High
BAI [14]	2017	1	1	0.5	0.5	0	0	3	High
BAJWA [15]	2013	1	1	1	0	0	0.5	3.5	High
BAUER [16]	2000	0	0	0.5	0.5	0	0	1	Low
BERSTEN [17]	2001	0.5	1	0.5	0	0	0	2.0	Intermediate
Bos [18]	2014	0	0	0	0	0	0.5	0.5	Low
Bos [19]	2014	0	0	0	0	0	0.5	0.5	Low
BRETT [20]	1998	1	0	0.5	0	0	0.5	2	Intermediate
BURSTEN [21]	1996	0.5	0	0.5	0	0	0.5	1.5	Low
COPETTI [22]	2008	0	1	1	0	0	0	2	Intermediate
DETERMANN [23]	2009	1	1	1	0	0	0.5	3.5	High
EL SOLH [24]	2005	0	1	1	0.5	0	0	2.5	Intermediate
FREMONT [25]	2010	1	0	0.5	0.5	0	1	3	High
GRISSOM [26]	2003	0.5	0	1	0.5	0	1	3	High
HERRERA [27]	1988	0.5	0	0.5	0.5	0	0.5	2	Intermediate
HOEBOER [28]	2015	0	0	0	0	0	0	0	Low
HOWRYLAK [29]	2009	1	0	0.5	0	0	1	2.5	Intermediate
HUANG [30]	2019	0	1	0.5	0.5	0	0	2	Intermediate
IZQUIERDO-GARCIA [31]	2018	1	1	0.5	0.5	0	0	3	High
JABAUDON [32]	2018	0.5	0	0	0	0	0	0.5	Low
JORENS [33]	1992	1	0	1	0.5	0	0	2.5	Intermediate
KIETZMANN [34]	1993	0.5	0	1	0.5	0	1	3	High
KUSHIMOTO [35]	2012	1	0	0	0	0	0	1	Low
LETOURNEAU [36]	2012	0.5	0	0.5	0	0	1	2	Intermediate
LIN [37]	2012	0	0	0.5	0	0	0	0.5	Low
LIN [38]	2013	0	0	0.5	0.5	0	0	1	Low
LIN [39]	2018	0	0	0.5	0	0	0.5	1	Low
LIU [40]	2015	1	1	0.5	0.5	0	0.5	3.5	High
LIU [41]	2017	0	1	0.5	0	0.5	0	2	Intermediate
MONNET [42]	2007	1	0	1	0.5	0	0	2.5	Intermediate
PARK [43]	2017	1	0	1	0.5	0	1	3.5	High
SATO [44]	2004	1	1	1	0.5	0	1	4.5	High
SEKIGUCHI [45]	2015	1	0	0	0	0	0	1	Low
SHAN [46]	2016	1	0	1	0	0	0.5	2.5	Intermediate
SWEENEY [47]	2018	1	0	0	1	1	1	4	High
VERHEIJ [48]	2005	1	0	1	0	0	1	3	High
WARE [49]	2013	1	0	0.5	0	0	0	1.5	Low
WARE [50]	2017	1	0	0.5	0	0	1	2.5	Intermediate
Wu [51]	2019	0	1	0	0.5	0	0.5	2	Intermediate
XUE [52]	2015	1	0	1	0.5	0	0	2.5	Intermediate
YEH [53]	2017	0	0	1	0	0	0	1	Low
ZHOU [54]	2019	0.5	0	0.5	0	0	0.5	1.5	Low

plasma and two exhaled breath tests for biomarkers of oxidative stress as tests that currently have the strongest rationale for further validation. This review provides strong evidence that the diagnostic accuracy of minimally invasive tests for the diagnosis of ARDS is highly dependent on the potential bias of the study and the type of control group that is included.

Diagnostic accuracy varied widely between tests and studies included in this review. We identified that the inclusion of CPE patients as a control group consistently resulted in a higher diagnostic accuracy, suggesting that CPE can be better distinguished from ARDS than ICU patients at risk for ARDS or pneumonia patients. An attractive explanation could be that the test differentiates between protein-rich and hydrostatic oedema, but this explanation was rejected because most tests did not evaluate this phenomenon directly and could still separate these groups. For example, cardiac injury markers are also able to distinguish between CPE and ARDS, but instead of being a relevant test for ARDS, it rather signifies the homogeneity of the CPE population [55]. Importantly, ARDS patients differ from CPE

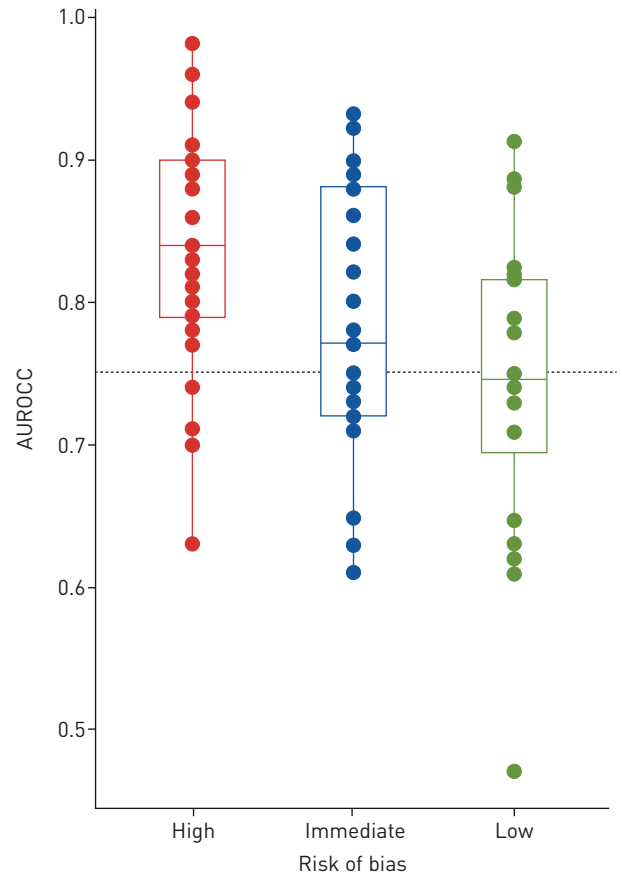


FIGURE 3 The risk of bias of a study was associated with the diagnostic accuracy of the evaluated test. Each dot represents a diagnostic test; multiple tests could be evaluated per study. The area under the receiver operating characteristic curve was significantly higher in the group with high risk of bias, compared to the intermediate-bias and low-bias groups ($p=0.020$ and $p=0.00077$, respectively). AUROCC: area under the receiver operating characteristic curve.

patients in many more aspects than the type of pulmonary oedema alone. For example, ARDS patients showed increased levels of inflammation parameters compared to CPE patients, but this is not necessarily related to ARDS but may be due to an underlying syndrome such as sepsis, pneumonia or pancreatitis. Indeed, when compared to an unselected ICU population with similar risk factors as the ARDS patients, these markers had a lower diagnostic accuracy.

The risk of bias assessed by the QUADAS-2 tool was strongly associated with the diagnostic accuracy of the study. A large part of the studies showed risk of bias due to the method of patient selection, performance and interpretation and timing of the index test. Unfortunately, this relationship and the fact that biased studies are known to overestimate diagnostic accuracy [56] makes it hard to rely on results from studies with a considerable amount of bias. It will be necessary to redo studies with tests showing good diagnostic accuracy but then in a low-bias setting before any firm conclusions can be drawn.

Focusing on studies with good diagnostic accuracy with low risk of bias, nine tests remained. Only four of them compared ARDS to an unselected ICU population. One test assessed the plasma concentration of CC16 [39]. This protein is suggested to protect the lungs against oxidative stress as well as inflammation [57]. However, CC16 is also a marker of increased permeability of the epithelial barrier and therefore seems to be involved in multiple processes of ARDS development [39, 57]. Another test assessed sRAGE in plasma, which is released by lung inflammation and leads to epithelial injury and therefore is a marker of increased permeability [32, 58]. The other two were exhaled breath tests [18, 54], with the major advantage that it can be obtained noninvasively. One test assessed a panel of three biomarkers, octane, acetaldehyde and 2/3-methylheptane, that reflect oxidative stress [18]. Of these three compounds, octane explained most of the diagnostic accuracy. Octane is generated through oxygenation of oleic acid and previous data suggest that ARDS is associated with an increased concentration of oleic acid in the circulation [59]. The other breath test assessed a larger biomarker panel, with nine exhaled breath biomarkers, not clearly reflective of one pathophysiological mechanism [54]. A drawback of exhaled breath tests is the fact that they are experimental and are therefore not directly suitable for clinical implementation. All tests seem to relate to oxidative stress, inflammation and increased permeability in the lungs, which are all known to be important in the early course of ARDS and are related to pulmonary pathophysiology directly.

This is the first review to systematically assess the diagnostic accuracy of minimally invasive techniques for ARDS while considering potential biases of each study. Our analyses show that it is pivotal to evaluate the methodological quality of the study to reveal the confounding factors while interpreting the results. This approach is one of the most important strengths of this study. Furthermore, papers not reporting diagnostic accuracy directly were not excluded when we could deduce the accuracy from figures showing individual data points. To our knowledge, no other study in critical care has utilised this approach up to now. Finally, we did not limit the definition of ARDS to those patients with a $P_{aO_2}/F_{IO_2} < 200$ mmHg by including patients who were labelled as “acute lung injury” according to the 1994 AECC definition. Since ARDS nowadays involves a heterogeneous population, of which patients with mild ARDS are a large part, it is important to recognise this group also [60]. Another strength is the exclusion of studies that used healthy volunteers as control group leaving only more relevant control groups and hopefully resulting in a more accurate comparison between similarly ill patients.

The main limitation of this review is the small number of studies that is left in each category after stratification. This sometimes led to groups with few studies, for example only one study assessed the CCU patients and only two studies compared diagnosis of ARDS with patients with pneumonia. With regard to pneumonia, it is questionable if unilateral pneumonia is the appropriate control group for ARDS, as many patients with ARDS have pneumonia and because unilateral and bilateral pneumonia in the ICU have similar outcomes [61]. Furthermore, both studies that compared ARDS to pneumonia scored high on the risk of bias. A second limitation is the fact that the AUROCC was not for all studies reported. Therefore, the analysis was performed in two parts with two different approaches, which yielded similar results. We also acknowledge that the diagnostic tests cannot be categorised into completely distinct groups, for example, there is considerable overlap between markers of oxidative stress and inflammation and our attempted separation of the two is arbitrary. Another limitation of this study is the fact that the definition of ARDS has changed over the years and therefore the “case-definition” is slightly different between studies, which might have confounded the diagnostic accuracy of specific tests. Finally, it should be noted that we assessed multiple diagnostic tests described in a single paper as independent tests, which they potentially are not. To our knowledge, there is no adequate multilevel alternative to study this phenomenon otherwise.

Results of this review show that there is no validated minimally invasive method to diagnose ARDS in an unselected ICU population. Four promising tests were identified in a low-bias setting and these warrant validation. New diagnostic studies should better attempt to minimise bias and should be reported according to Standards for Reporting Diagnostic Accuracy (STARD) guidelines [62].

A diagnostic test does not have to separate ARDS patients perfectly. This is likely impossible due to the biological heterogeneity observed in ARDS patients. Indeed, another way to evaluate these results is to appreciate the heterogeneity that is shown and advocate a personalised approach based on pathophysiological characteristics of each patient shown through the diagnostic tests that are described here [63]. A biomarker may have value when it identifies a phenotype that consistently responds to a specific type of treatment, a so-called treatable trait [64].

Conclusion

There is no minimally invasive diagnostic test for ARDS that is validated in a low-bias setting against an adequate control group. Many studies that evaluated diagnostic tests for ARDS showed risk of bias, which makes it hard to rely on the reported diagnostic accuracy. The plasma concentration of CC16, sRAGE and two panels of oxidative stress biomarkers in exhaled breath did show high diagnostic accuracy in a low-bias setting and warrant external validation. For implementation into the clinical setting, prospective studies in a general unselected ICU population with good methodological quality are needed.

Members of the DARTS consortium: Amsterdam UMC, Amsterdam, the Netherlands (Laura A. Hagens, Marry R. Smit, Marcus J. Schultz and Lieuwe D.J. Bos); Maastricht UMC, Maastricht, the Netherlands (Nanon F.L. Heijnen, Dennis C.J.J. Bergmans and Ronny M. Schnabel); and Philips Research, Eindhoven, the Netherlands (Alwin R.M. Verschueren, Tamara M.E. Nijssen and Inge Geven).

Conflict of interest: L.A. Hagens has nothing to disclose. N.F.L. Heijnen has nothing to disclose. M.R. Smit has nothing to disclose. M.J. Schultz has nothing to disclose. D.C.J.J. Bergmans has nothing to disclose. R.M. Schnabel has nothing to disclose. L.D.J. Bos reports grants from the Dutch Lung Foundation (young investigator grant, public-private partnership grant and Dirkje Postma Award) outside the submitted work.

Support statement: Lieuwe D.J. Bos is supported by Health Holland *via* the Dutch Lung Foundation (longfonds) industry-academia partnership and *via* the Dirkje Postma Award. They had no role in the design, conduct or interpretation of this review.

References

- 1 Thompson B, Chambers R, Liu K. Acute respiratory distress syndrome. *N Engl J Med* 2017; 377: 562–572.
- 2 Ranieri VM, Rubenfeld GD, Thompson BT, et al. Acute respiratory distress syndrome: the Berlin definition. *JAMA* 2012; 307: 2526–2533.
- 3 Matthay MA, Zemans RL, Zimmerman GA, et al. Acute respiratory distress syndrome. *Nat Rev Dis Primers* 2019; 5: 18.
- 4 Agrawal A, Zhuo H, Brady S, et al. Pathogenetic and predictive value of biomarkers in patients with ALI and lower severity of illness: results from two clinical trials. *Am J Physiol Lung Cell Mol Physiol* 2012; 303: L634–L639.
- 5 Sjoding MW, Hofer TP, Co I, et al. Interobserver reliability of the Berlin ARDS definition and strategies to improve the reliability of ARDS diagnosis. *Chest* 2018; 153: 361–367.
- 6 Binnie A, Tsang JLY, Dos Santos CC. Biomarkers in acute respiratory distress syndrome. *Curr Opin Crit Care* 2014; 20: 47–55.
- 7 Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; 6: e1000097.
- 8 Bernard GR, Artigas A, Brigham KL, et al. The American-European Consensus Conference on ARDS. Definitions, mechanisms, relevant outcomes, and clinical trial coordination. *Am J Respir Crit Care Med* 1994; 149: 818–824.
- 9 Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155: 529–536.
- 10 Lee J, Kim KW, Choi SH, et al. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers – Part II. Statistical methods of meta-analysis. *Korean J Radiol* 2015; 16: 1188–1196.
- 11 Abbas MMK, Patel B, Chen Q, et al. Involvement of the Bufadienolides in the detection and therapy of the acute respiratory distress syndrome. *Lung* 2017; 195: 323–332.
- 12 Aman J, van der Heijden M, van Lingen A, et al. Plasma protein levels are markers of pulmonary vascular permeability and degree of lung injury in critically ill patients with or at risk for acute lung injury/acute respiratory distress syndrome. *Crit Care Med* 2011; 39: 89–97.
- 13 Arif SK, Verheij J, Groeneveld ABJ, et al. Hypoproteinemia as a marker of acute respiratory distress syndrome in critically ill patients with pulmonary edema. *Intensive Care Med* 2002; 28: 310–317.
- 14 Bai W, Zhu WL, Ning YL, et al. Dramatic increases in blood glutamate concentrations are closely related to traumatic brain injury-induced acute lung injury. *Sci Rep* 2017; 7: 5380.
- 15 Bajwa EK, Volk JA, Christiani DC, et al. Prognostic and diagnostic value of plasma soluble suppression of tumorigenicity-2 concentrations in acute respiratory distress syndrome. *Crit Care Med* 2013; 41: 2521–2531.
- 16 Bauer TT, Montón C, Torres A, et al. Comparison of systemic cytokine levels in patients with acute respiratory distress syndrome, severe pneumonia, and controls. *Thorax* 2000; 55: 46–52.
- 17 Bersten AD, Hunt T, Nicholas TE, et al. Elevated plasma surfactant protein-B predicts development of acute respiratory distress syndrome in patients with acute respiratory failure. *Am J Respir Crit Care Med* 2001; 164: 648–652.
- 18 Bos LDJ, Weda H, Wang Y, et al. Exhaled breath metabolomics as a noninvasive diagnostic tool for acute respiratory distress syndrome. *Eur Respir J* 2014; 44: 188–197.
- 19 Bos LDJ, Schultz MJ, Sterk PJ. Exhaled breath profiling for diagnosing acute respiratory distress syndrome. *BMC Pulm Med* 2014; 14: 72.
- 20 Brett SJ, Evans TW. Measurement of endogenous nitric oxide in the lungs of patients with the acute respiratory distress syndrome. *Am J Respir Crit Care Med* 1998; 157: 993–997.
- 21 Bursten SL, Federighi DA, Parsons P, et al. An increase in serum C18 unsaturated free fatty acids as a predictor of the development of acute respiratory distress syndrome. *Crit Care Med* 1996; 24: 1129–1136.
- 22 Copetti R, Soldati G, Copetti P. Chest sonography: a useful tool to differentiate acute cardiogenic pulmonary edema from acute respiratory distress syndrome. *Cardiovasc Ultrasound* 2008; 6: 16.
- 23 Determann RM, Millo JL, Waddy S, et al. Plasma CC16 levels are associated with development of ALI/ARDS in patients with ventilator-associated pneumonia: a retrospective observational study. *BMC Pulm Med* 2009; 9: 49.
- 24 El Solh AA, Bhora M, Pineda L, et al. Alveolar plasminogen activator inhibitor-1 predicts ARDS in aspiration pneumonia. *Intensive Care Med* 2006; 32: 110–115.
- 25 Fremont RD, Koyama T, Calfee CS, et al. Acute lung injury in patients with traumatic injuries: utility of a panel of biomarkers for diagnosis and pathogenesis. *J Trauma Inj Infect Crit Care* 2010; 68: 1121–1127.
- 26 Grissom CK, Orme JFJ, Richer LD, et al. Platelet-activating factor acetylhydrolase is increased in lung lavage fluid from patients with acute respiratory distress syndrome. *Crit Care Med* 2003; 31: 770–775.
- 27 Herrera C, Velasco F, Guerrero A, et al. Contact phase of blood coagulation in cardiogenic pulmonary oedema (CPO) and adult respiratory distress syndrome (ARDS). *Intensive Care Med* 1989; 15: 99–104.
- 28 Hoeboer SH, Oudemans-van Straaten HM, Groeneveld ABJ. Albumin rather than C-reactive protein may be valuable in predicting and monitoring the severity and course of acute respiratory distress syndrome in critically ill patients with or at risk for the syndrome after new onset fever. *BMC Pulm Med* 2015; 15: 22.
- 29 Howrylak JA, Dolinay T, Lucht L, et al. Discovery of the gene signature for acute lung injury in patients with sepsis. *Physiol Genomics* 2009; 37: 133–139.
- 30 Huang Y, Xiao J, Cai T, et al. Immature granulocytes: a novel biomarker of acute respiratory distress syndrome in patients with acute pancreatitis. *J Crit Care* 2019; 50: 303–308.
- 31 Izquierdo-Garcia JL, Nin N, Jimenez-Clemente J, et al. Metabolomic profile of ARDS by nuclear magnetic resonance spectroscopy in patients with h1n1 influenza virus pneumonia. *Shock* 2018; 50: 504–510.
- 32 Jabaudon M, Berthelin P, Pranal T, et al. Receptor for advanced glycation end-products and ARDS prediction: a multicentre observational study. *Sci Rep* 2018; 8: 2603.
- 33 Jorens PG, Van Damme J, De Backer W, et al. Interleukin 8 (IL-8) in the bronchoalveolar lavage fluid from patients with the adult respiratory distress syndrome (ARDS) and patients at risk for ARDS. *Cytokine* 1992; 4: 592–597.
- 34 Kietzmann D, Kahl R, Muller M, et al. Hydrogen peroxide in expired breath condensate of patients with acute respiratory failure and with ARDS. *Intensive Care Med* 1993; 19: 78–81.

- 35 Kushimoto S, Taira Y, Kitazawa Y, *et al.* The clinical usefulness of extravascular lung water and pulmonary vascular permeability index to diagnose and characterize pulmonary edema: a prospective multicenter study on the quantitative differential diagnostic definition for acute lung injury/acute respiratory distress syndrome. *Crit Care* 2012; 16: R232.
- 36 LeTourneau JL, Pinney J, Phillips CR. Extravascular lung water predicts progression to acute lung injury in patients with increased risk*. *Crit Care Med* 2012; 40: 847–854.
- 37 Lin Q, Fu F, Chen H, *et al.* Copeptin in the assessment of acute lung injury and cardiogenic pulmonary edema. *Respir Med* 2012; 106: 1268–1277.
- 38 Lin Q, Shen J, Shen L, *et al.* Increased plasma levels of heparin-binding protein in patients with acute respiratory distress syndrome. *Crit Care* 2013; 17: R155.
- 39 Lin J, Zhang W, Wang L, *et al.* Diagnostic and prognostic values of Club cell protein 16 (CC16) in critical care patients with acute respiratory distress syndrome. *J Clin Lab Anal* 2018; 32: e22262.
- 40 Liu D, Luo G, Luo C, *et al.* Changes in the concentrations of mediators of inflammation and oxidative stress in exhaled breath condensate during liver transplantation and their relations with postoperative ARDS. *Respir Care* 2015; 60: 679–688.
- 41 Liu XW, Ma T, Cai Q, *et al.* Elevation of serum PARK7 and IL-8 levels is associated with acute lung injury in patients with severe sepsis/septic shock. *J Intensive Care Med* 2019; 34: 662–668.
- 42 Monnet X, Anguel N, Osman D, *et al.* Assessing pulmonary permeability by transpulmonary thermodilution allows differentiation of hydrostatic pulmonary edema from ALI/ARDS. *Intensive Care Med* 2007; 33: 448–453.
- 43 Park J, Pabon M, Choi AMK, *et al.* Plasma surfactant protein-D as a diagnostic biomarker for acute respiratory distress syndrome: validation in US and Korean cohorts. *BMC Pulm Med* 2017; 17: 204.
- 44 Sato H, Callister ME, Mumby S, *et al.* KL-6 levels are elevated in plasma from patients with acute respiratory distress syndrome. *Eur Respir J* 2004; 23: 142–145.
- 45 Sekiguchi H, Schenck LA, Horie R, *et al.* Critical care ultrasonography differentiates ARDS, pulmonary edema, and other causes in the early course of acute hypoxemic respiratory failure. *Chest* 2015; 148: 912–918.
- 46 Shan L, Shan F, Li J, *et al.* Association of circulating suPAR with disease severity and clinical outcomes in patients with ARDS induced by intra-abdominal infections: a prospective observational study. *Int J Clin Exp Med* 2016; 9: 12788–12795.
- 47 Sweeney TE, Thomas NJ, Howrylak JA, *et al.* Multicohort analysis of whole-blood gene expression data does not form a robust diagnostic for acute respiratory distress syndrome. *Crit Care Med* 2018; 46: 244–251.
- 48 Verheij J, Raijmakers PGHM, Lingen A, *et al.* Simple vs complex radionuclide methods of assessing capillary protein permeability for diagnosing acute respiratory distress syndrome. *J Crit Care* 2005; 20: 162–171.
- 49 Ware LB, Koyama T, Zhao Z, *et al.* Biomarkers of lung epithelial injury and inflammation distinguish severe sepsis patients with acute respiratory distress syndrome. *Crit Care* 2013; 17: R253.
- 50 Ware LB, Zhao Z, Koyama T, *et al.* Derivation and validation of a two-biomarker panel for diagnosis of ARDS in patients with severe traumatic injuries. *Trauma Surg Acute Care Open* 2017; 2: e000121.
- 51 Wu C-Y, Cheng Y-J, Hung M-H, *et al.* Association between early acute respiratory distress syndrome after living-donor liver transplantation and perioperative serum biomarkers: the role of Club cell protein 16. *Biomed Res Int* 2019; 2019: 8958069.
- 52 Xue M, Sun Z, Shao M, *et al.* Diagnostic and prognostic utility of tissue factor for severe sepsis and sepsis-induced acute lung injury. *J Transl Med* 2015; 13: 172.
- 53 Yeh L-C, Huang PW, Hsieh KH, *et al.* Elevated plasma levels of Gas6 are associated with acute lung injury in patients with severe sepsis. *Tohoku J Exp Med* 2017; 243: 187–193.
- 54 Zhou M, Sharma R, Zhu H, *et al.* Rapid breath analysis for acute respiratory distress syndrome diagnostics using a portable two-dimensional gas chromatography device. *Anal Bioanal Chem* 2019; 411: 6435–6447.
- 55 Komiya K, Akaba T, Kozaki Y, *et al.* A systematic review of diagnostic methods to differentiate acute lung injury/acute respiratory distress syndrome from cardiogenic pulmonary edema. *Crit Care* 2017; 21: 228.
- 56 Rutjes AWS, Reitsma JB, Di Nisio M, *et al.* Evidence of bias and variation in diagnostic accuracy studies. *CMAJ* 2006; 174: 469–476.
- 57 Broeckaert F, Bernard A. Clara cell secretory protein (CC16): characteristics and perspectives as lung peripheral biomarker. *Clin Exp Allergy* 2000; 30: 469–475.
- 58 Uchida T, Shirasawa M, Ware LB, *et al.* Receptor for advanced glycation end-products is a marker of type I cell injury in acute lung injury. *Am J Respir Crit Care Med* 2006; 173: 1008–1015.
- 59 Quinlan GJ, Lamb NJ, Evans TW, *et al.* Plasma fatty acid changes and increased lipid peroxidation in patients with adult respiratory distress syndrome. *Crit Care Med* 1996; 24: 241–246.
- 60 Bellani G, Laffey JG, Pham T, *et al.* Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. *JAMA* 2016; 315: 788–800.
- 61 Cilloniz C, Ferrer M, Liapikou A, *et al.* Acute respiratory distress syndrome in mechanically-ventilated patients with community-acquired pneumonia. *Eur Respir J* 2018; 51: 1702215.
- 62 Bossuyt PM, Reitsma JB, Bruns DE, *et al.* Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003; 326: 41–44.
- 63 Sinha P, Calfee CS. Phenotypes in acute respiratory distress syndrome: moving towards precision medicine. *Curr Opin Crit Care* 2019; 25: 12–20.
- 64 Agusti A, Bel E, Thomas M, *et al.* Treatable traits: toward precision medicine of chronic airway diseases. *Eur Respir J* 2016; 47: 410–419.