Evolution of multiple omics approaches to define pathophysiology of pediatric acute respiratory distress syndrome

Jane E Whitney[1,2]*, In-Hee Lee[3], Ji-Won Lee[4], Sek Won Kong[2,3]

[1]Medical Critical Care, Pediatrics, Boston Children's Hospital, Boston, United States; [2]Department of Pediatrics, Harvard Medical School, Boston, United States; [3]Computational Health and Informatics Program, Boston Children's Hospital, Boston, United States; [4]Department of Pharmacology, Faculty and Graduate School of Dental Medicine, Hokkaido University, Sapporo, Japan

Abstract Pediatric acute respiratory distress syndrome (PARDS), though both common and deadly in critically ill children, lacks targeted therapies. The development of effective pharmacotherapies has been limited, in part, by lack of clarity about the pathobiology of pediatric ARDS. Epithelial lung injury, vascular endothelial activation, and systemic immune activation are putative drivers of this complex disease process. Prior studies have used either hypothesis-driven (e.g., candidate genes and proteins, in vitro investigations) or unbiased (e.g., genome-wide association, transcriptomic, metabolomic) approaches to predict clinical outcomes and to define subphenotypes. Advances in multiple omics technologies, including genomics, transcriptomics, proteomics, and metabolomics, have permitted more comprehensive investigation of PARDS pathobiology. However, omics studies have been limited in children compared to adults, and analyses across multiple tissue types are lacking. Here, we synthesized existing literature on the molecular mechanism of PARDS, summarized our interrogation of publicly available genomic databases to determine the association of candidate genes with PARDS phenotypes across multiple tissues and cell types, and integrated recent studies that used single-cell RNA sequencing (scRNA-seq). We conclude that novel profiling methods such as scRNA-seq, which permits more comprehensive, unbiased evaluation of pathophysiological mechanisms across tissue and cell types, should be employed to investigate the molecular mechanisms of PRDS toward the goal of identifying targeted therapies.

*For correspondence:
jane.whitney@childrens.harvard.edu

Introduction

Pediatric acute respiratory distress syndrome (PARDS) was first defined in children in 2015 as the acute onset of parenchymal lung disease on chest x-ray with severe hypoxemia not explained by cardiac disease (*Pediatric Acute Lung Injury Consensus Conference Group, 2015*). PARDS compared to ARDS affecting adults has lower incidence (2–13 *Pediatric Acute Lung Injury Consensus Conference Group, 2015*; *Schouten et al., 2016*; *Hough, 2017* vs. 17.9–81.0 *Rubenfeld et al., 2009* ; *Rubenfeld et al., 2009*; *Li et al., 2011* per 100,000 person-years) and mortality (18%–27% *Pediatric Acute Lung Injury Consensus Conference Group, 2015* vs. 27%–45% *Rubenfeld et al., 2009*; *Li et al., 2011*). Sepsis is the most common cause of ARDS affecting adults while pneumonia is the most common cause of PARDS (*Khemani et al., 2019*). Epidemiologic differences between pediatric and adult ARDS may reflect differences in lung function, immune response, and disease mechanism between children and adults.

The onset of PARDS follows a severe physiological insult, which can take a variety of forms including pneumonia, aspiration, inhalation, sepsis, pancreatitis, transfusion, or trauma. Prior studies have noted

differences in risk factors and clinical outcomes for children with distinct PARDS phenotypes (e.g., direct vs. indirect lung injury), which may reflect differences in pathobiology (*Yehya et al., 2018*; *Whitney et al., 2020a*). The fact that a relatively small proportion of children, who are at risk for PARDS after exposure to a severe physiologic insult, develop PARDS may reflect underlying biological differences in genetically determined host susceptibility.

PARDS mortality has decreased slightly over time (*Pediatric Acute Lung Injury Consensus Conference Group, 2015*; *Khemani et al., 2019*), possibly due to improvements in delivery of supportive care. Targeted pharmacologic therapies are lacking, which may reflect a lack of knowledge about the pathobiology of this complex disease process. Prior studies have provided some insights into the biological processes driving ARDS development in those at risk, clinical outcomes in pediatric and adults with ARDS, and phenotypic heterogeneity. However, limitations in methodology and scope exist, and we continue to lack clarity about which molecular mechanisms affecting which tissues and cell types at what time produce this complex disease process.

We sought to summarize the existing literature and data on molecular investigations of PARDS, which has included both hypothesis-driven and unbiased investigations. Adult ARDS literature is included where it supports or refutes pediatric data, or addresses questions not yet investigated in PARDS. We review candidate gene, candidate protein, and in vitro investigations testing hypotheses about PARDS risk, outcomes, and subphenotypes. We review genome-wide association studies (GWAS), transcriptomic, and metabolomic discovery investigations also addressing PARDS risk, outcomes, and subphenotypes. Then, using the list of candidate genes that have been investigated, we appraise the results of our own survey of publicly available gene databases to determine the concordance between genotype and phenotype across multiple tissues and cell types. Finally, we discuss limitations of the methods to date and describe some of the ways in which single-cell RNA sequencing (scRNA-seq) may address prior limitations. The overarching goal of this work is to compile what is known about PARDS using unbiased genomic techniques in order to elucidate disease mechanism with the goal of identifying potential treatment targets.

Summary of research to date
Hypothesis-driven investigations
Candidate gene studies

The development of PARDS in response to a severe physiologic insult is thought to reflect epithelial lung injury, systemic inflammation, and activation of the vascular endothelium. Candidate gene studies, therefore, have largely focused on genetic polymorphisms related to these processes, including surfactant protein function, pulmonary inflammation, systemic inflammation, and endothelial activation (*Figure 1*). While the requirement for a preceding physiologic insult precludes pedigree studies of PARDS, candidate gene studies (*Table 1*) have identified polymorphisms associated with the development of disease in some patients compared to others.

Surfactant is a mixture of phospholipids and surfactant proteins (SP-A, SP-B, SP-C, and SP-D), and is critical for lowering surface tension at the alveolar–epithelial interface and protection against pathogens (*Whitsett and Weaver, 2015*). Patients with inherited surfactant deficiency demonstrate severe if not lethal respiratory failure. Surfactant lipids and proteins are produced by type II alveolar cells, and the surfactant lipid transporter (ABCA3) is required to form lamellar bodies that are secreted into the airways. Thus, variants in the genes encoding surfactant proteins A through D and ABCA3 have been of interest in ARDS. The *SFTPA1*, *SFTPA2*, *SFTPB*, *SFTPC*, and *SFTPD* genes encode SP-A1, SP-A2, SP-B, SP-C, and SP-D proteins, respectively. Among nearly 400 children with pneumonia, of whom the most severely affected had PARDS and need for mechanical ventilation, an association was demonstrated between *SFTPB* Single nucleotide polymorphisms (SNPs) and need for mechanical ventilation (*Dahmer et al., 2011*). In another study comparing children with acute respiratory failure and healthy newborn controls, SNPs in *SFTPA1*, *SFTPA2*, and *SFTPC* were associated with acute respiratory failure and pulmonary dysfunction after hospital discharge (*Gandhi et al., 2020*).

The angiotensin I-converting enzyme (ACE) degrades bradykinin and catalyzes angiotensin I–II, a potent vasoconstrictor with proinflammatory and profibrotic effects. High levels of circulating ACE are associated with pulmonary inflammation, and ACE is also highly expressed by endothelial cells composing the pulmonary microvasculature (*Aird, 2007*). The insertion/deletion (I/D) variant
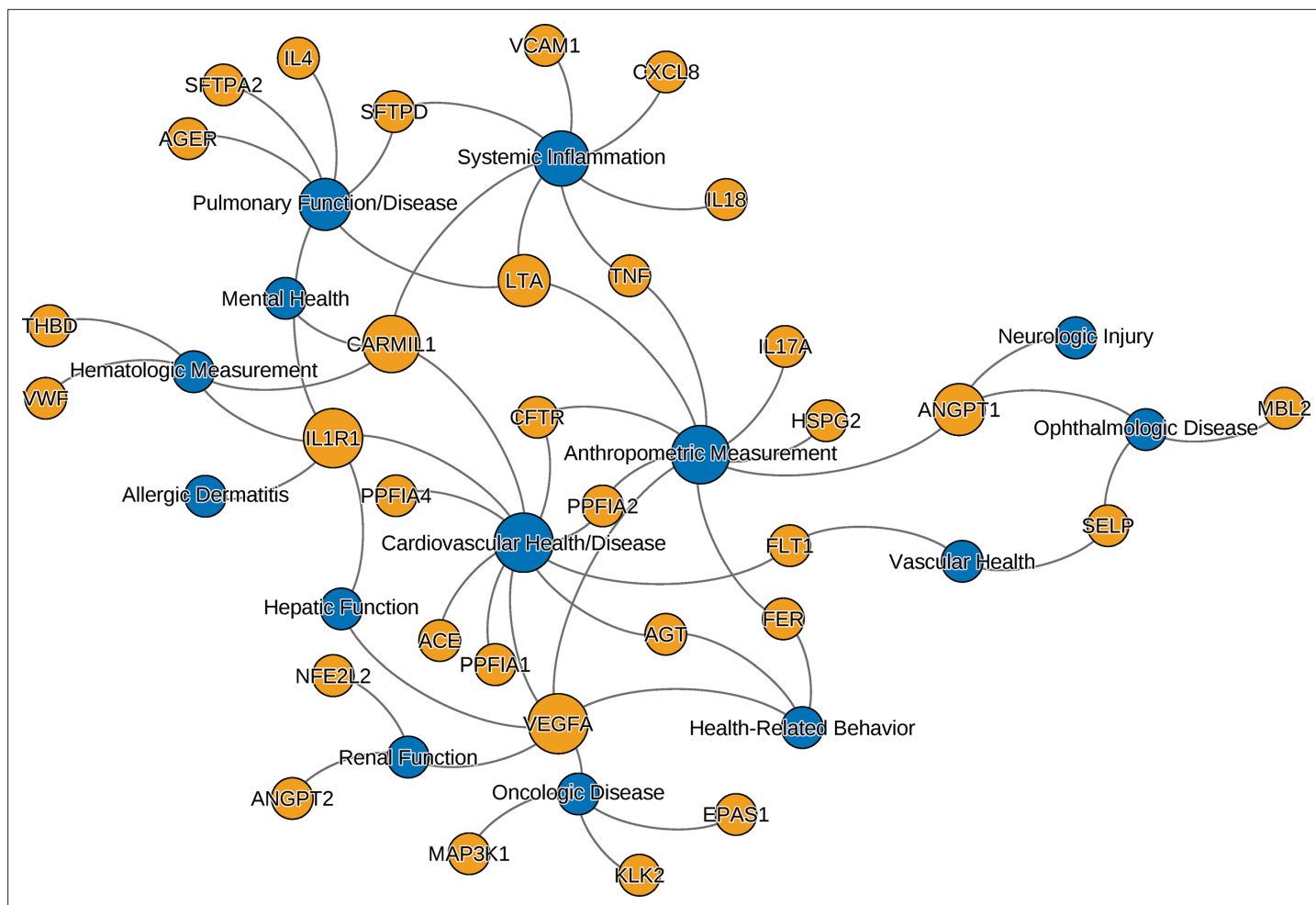
Figure 1. The relationships between ARDS candidate genes and phenotypes reported by genome-wide association studies. The nodes (circles) represent genes (light gray, gene symbols not italicized) or phenotypes (dark gray), and the associated genotype–phenotype pairs are connected by curved lines. For clear visualization, the original phenotypes reported by individual studies were simplified (**Supplementary file 2**). Node size is proportional to reported number of genotype–phenotype associations. The biological characteristics of ARDS, including pulmonary inflammation, systemic inflammation, and vascular endothelial activation, are associated with multiple candidate genes for ARDS. Abbreviations: ACE = angiotensin I-converting enzyme; AGER = advanced glycosylation end product-specific receptor; AGT = angiotensin; ANGPT = angiopoietin; ARDS = acute respiratory distress syndrome; CARMIL1 = capping protein regulator and myosin 1 linker 1; CFTR = cystic fibrosis transmembrane conductance regulator; CXCL8 = C-X-C motif chemokine ligand; EPAS1 = endothelial PAS domain protein 1; FER = feline encephalitis virus-related tyrosine kinase; FLT1 = fms-related receptor tyrosine kinase 1; HSPG = heparan sulfate proteoglycan; IL = interleukin; KLK2 = kallikrein-related peptidase 2; SFTP = surfactant protein; LTA = lymphotoxin alpha; MAP3K1 = mitogen-activated protein kinase 1; MBL = mannose binding lectin; NFE2L2 = nuclear factor, erythroid-like, BZIP transcription factor 2; PPFIA = protein tyrosine phosphatase receptor type F interacting protein alpha; SELP = selectin P; THBD = thrombomodulin; TNF = tumor necrosis factor; VCAM1 = vascular cell adhesion molecule 1; VEGFA = vascular endothelial growth factor A; vWF = von Willebrand Factor.

(rs4646994, a 287–289 basepair Alu-type repeat sequence in intron 16) in the *ACE1* gene has functional significance such that the D allele is associated with increased ACE activity and higher ACE levels in plasma and tissue. The association between D allele and ARDS risk and outcomes has been tested in several observational studies. In children, the I/D variant failed to demonstrate association with PARDS status but increased frequency of D allele was associated with PARDS severity (*Cruces et al., 2012*; *Plunkett et al., 2008*). An adult study confirmed the association between the DD genotype, increased risk for ARDS, and increased mortality from ARDS (*Adamzik et al., 2007*; *Marshall et al., 2002*) but did not test the I/D variant.

Candidate genes associated with systemic inflammation have been of long-standing interest in PARDS. Systemic inflammation often accompanies pulmonary inflammation, and systemic inflammatory processes such as sepsis are risk factors for PARDS. Tumor necrosis factor-α (TNF-α), encoded by

Table 1. Key Pediatric Studies.

Approach	Candidate Gene(s)	Sample Size	Findings (Adjusted OR and CI)	Author, Journal*, Year
HYPOTHESIS-DRIVEN				
Candidate Gene	*TNF, LTA*	490 intubated with sepsis, 610 healthy controls (total 1100)	Protective effect of *TNF-308GA* against ARDS in infants: OR 0.2, *P*=0.001	*Azevedo et al., 2012*
	ACE	60 with ARDS, 60 healthy controls (total 120, all <15 years)	I/D genotype not increased in ARDS: rate 0.4 (ARDS) vs. 0.3 (controls), *P*=NS	*Cruces et al., 2012*
	SFTPB	395 with pneumonia: 37 requiring mechanical ventilation, 26 with ALI/ARDS	Two linkage disequilibrium-tag SNPs associated with mechanical ventilation: GTGCGCG AOR = 2.62, CI 1–6.8, ATATAAG AOR = 3.1, CI 1–8.9	*Dahmer et al., 2011*
	SFTP	248 children <2 years with acute respiratory failure, 468 newborn healthy controls	34 interactions among 3 SNPs of *SFTPA1, SFTPA2, SFTPC* associated with acute respiratory failure (*P*=0.000000002–0.05) and pulmonary dysfunction after discharge (*P*=0.00002–0.03)	*Gandhi et al., 2020*
	ACE	13 Caucasian children with ARDS, 30 acute hypoxemic respiratory failure, 186 ICU controls (total 216)	I/D polymorphism not associated with acute hypoxemic respiratory failure or ARDS	*Plunkett et al., 2008*
Candidate Protein	*IL8*	480 with acute respiratory failure	Increased IL8 associated with mortality, duration mechanical ventilation, ICU LOS but not ARDS diagnosis	*Flori et al., 2019*
	many	3 cohorts: 46 with sepsis ARDS, 54 with sepsis without ARDS, and 19 ICU controls (total 119)	ANGPT2, ANGPT2/1 ratio higher in ARDS; ANGPT2, ANGPT2/1 ratio, VWF, ESM1 predicted complicated course in sepsis; in sepsis ARDS, FLT1 decreased more quickly and VWF, THBD decreased more slowly in those with complicated course	*Whitney et al., 2020a*
	ANGPT1, ANGPT2, VCAM1, vWF	2 cohorts of patients with ARDS: 52 direct, 46 indirect lung injury (total 98)	ARDS with indirect lung injury associated with increased ANGPT2/1 ratio, VCAM1, vWF (sensitivity 0.9, CI 0.8–1.0, specificity 0.8, CI 0.7–0.9)	*Whitney et al., 2020a*
	AGER, ANGPT2	82 with ARDS	Increased AGER, ANGPT2 associated with non-survival, organ failures in children with ARDS	*Yehya et al., 2016*
	CCL3, HSPA1b, IL8	153 with ARDS	Mortality associated with increased CCL3, HSPA1B, IL8, and older age in children with ARDS	*Yehya et al., 2018*
	many	235 with ARDS	Identified MMP profile associated with mortality (AOR 4.0, CI 2.1–7.6)	*Zinter et al., 2019*
	ANGPT2, VEGF, VWF	259 with ARDS, 25 status post HCST	Early ANGPT2 (OR 3.7, CI 1.1–11.5) and increasing ANGPT2 associated with mortality (AOR 3.3, CI 1.2–9.2), especially among HCST (AOR 16.3, CI 1.3–198)	*Zinter et al., 2016*
In Vitro Studies		Neutrophils and tracheal aspirates from 20 ARDS viral pneumonia with or without bacterial co-infection	In bacterial co-infection: (1) neutrophils more activated with impaired bacterial killing, respiratory burst, (2) aspirates with higher neutrophil elastase and myeloperoxidase, (3) neutrophils transmigrated into aspirate with decreased burst/killing of *H. influenzae, S. aureus*	*Grunwell et al., 2019*
		Tracheal aspirates from 42 intubated children with, 35 without ARDS (total 77)	Increased *STAT1* phosphorylation, markers of neutrophil degranulation and activation, NET release. Higher airway NETs associated with fever ventilator-free days	*Grunwell et al., 2019*
UNBIASED				
Gene Expression		28 intubated with, 26 without ARDS (total 54)	Using tracheal aspirates, a 62-gene signature to identify ARDS was developed to achieve cross-validation AUC 0.8, CI 0.6–0.9	*Grunwell et al., 2021*
		67 with sepsis and acute hypoxemic respiratory failure	Two identified endotypes differentially associated with mortality (OR 8, CI 1.6–41), complicated course (OR 4.2, 1.2–14.9)	*Yehya et al., 2019*
		96 with ARDS	Three identified sub-phentoypes associated with different clinical characteristics, outcomes	*Yehya et al., 2020*

Abbreviations: ACE = angiotensin-converting enzyme; AGER = advanced glycosylation end-product specific receptor; ALI = acute lung injury; ANGPT = angiopoietin; AOR = adjusted odds ratio; AUC = area under the receiver operator characteristic curve; ARDS = acute respiratory distress syndrome; CCL3 = C-C motif chemokine ligand 3; CI = 95% confidence interval; ESM1 = endothelial cell specific molecule 1; FLT1 = fms related receptor tyrosine kinase 1; HCST = hematopoietic stem cell transplant; HSPA1B = heat shock protein family A (Hsp70) member 1B; ICU = intensive care unit; IL = interleukin; LOS = length of stay; LTA = lymphotoxin alpha; MMP = matrix metalloproteinase; NET = neutrophil extracellular trap; OR = unadjusted odds ratio; SFTP = surfactant protein; SNP = single nucleotide polymorphism; SFTP = surfactant protein; THBD = thrombomodulin; TNF = tumor necrosis factor; TNFRSF1A = TNF receptor superfamily member 1A; VCAM1 = vascular cell adhesion molecule 1; VWF = von Willebrand Factor.

*Journal Titles are abbreviated according to U.S. National Library of Medicine convention.

the *TNFA* gene, is a pleiotropic cytokine and a primary mediator of systemic inflammation. Risk for PARDS in a cohort of intubated children with sepsis was associated with *TNFA*-863 C>A genotype (rs18000630), and the *TNFA*-308 G>A genotype (rs1800629) had a protective effect against mortality in sepsis-associated PARDS (*Azevedo et al., 2012*).

The activation of coagulation system with microthrombi causing tissue hypoxia contributes to pulmonary vascular injury in PARDS. The endothelial protein C receptor (EPCR, encoded by the *PROCR* gene) plays a cytoprotective role in sepsis by activating protein C (APC) and mediating APC effects. In the presence of thrombomodulin (TM encoded by the *THBD* gene), thrombin activates protein C to APC. We are not aware of studies investigating the association between candidate genes involved in coagulation and PARDS. Adult ARDS mortality has been associated with two *THBD* (rs1042580 and rs3716123) and one *PROCR* (rs9574) SNPs (*Sapru et al., 2016*).

A result of systemic inflammation, especially inflammation that occurs as a response to infection, is the activation of the vascular endothelium. Endothelial activation has been the focus of many candidate gene studies in adult but not yet pediatric ARDS, including those which have sought to determine its role in producing the disease in patients with risk factors. The *ANGPT2* gene encodes angiopoetin-2 (ANGPT2), which has been associated with loss of vascular integrity, increased permeability, and potentiation of vascular lung injury. In adults with sepsis, *ANGPT2* intronic SNPs (rs2442608 and rs2442630) have been associated with increased susceptibility to acute lung injury (ALI) (*Meyer et al., 2011*) or ARDS across multiple cohorts (*Su et al., 2009*), one in which plasma ANGPT2 was implicated as a causal intermediate (*Reilly et al., 2018*). In adults with trauma or sepsis as risk factors for ARDS, the development of ARDS was associated with *ANGPT2* (*Tejera et al., 2012*) and blood type A in White patients (*Reilly et al., 2021*), potentially reflecting vascular inflammatory changes. *FLT1* encodes soluble fms-like tyrosine kinase (sFLT, also known as vascular endothelial growth factor [VEGF] receptor), which competitively inhibits VEGF to promote vascular quiescence. An SNP in *FLT1* (rs9513106) was associated with reduced susceptibility to ARDS in a cohort study involving adults of Spanish and Western European descent with sepsis (*Hernandez-Pacheco et al., 2018*).

As a whole, candidate gene studies have demonstrated associations between the development of ARDS in patients with clinical risk factors and polymorphisms implicated in pulmonary inflammation, systemic inflammation, or endothelial activation. These associations have been demonstrated in multiple cohorts with varying risk factors for ARDS. However, the association of specific SNPs with ARDS has not been adequately replicated, and pediatric studies have been few. A strength of candidate gene studies is that they are highly feasible and useful when sample size is small. Notable disadvantages of candidate gene studies are false positives and poor replication. Furthermore, candidate gene studies are hypothesis-driven, investigating only genes of interest based on prior knowledge, and thus may unduly emphasize known biological themes while failing to address what has not yet been hypothesized.

Protein biomarker studies

Candidate protein studies of PARDS (*Table 1*) are numerous because they are highly feasible. Protein biomarkers are assayed from peripheral blood of patients with or at risk for PARDS, which is relatively noninvasive and low cost. Outcomes of interest for candidate protein studies have included: (1) whether PARDS develops in patients with clinical risk factors, (2) predictors of PARDS outcomes, and (3) distinguishing subphenotypes of PARDS to elucidate differences in pathobiology, clinical characteristics, and outcomes.

The development of ARDS in children and adults with clinical risk factors for the syndrome has been associated with candidate protein biomarkers reflecting endothelial activation, systemic inflammation, and epithelial lung injury. Limited number of studies and small sample sizes in pediatric ARDS reflects lower incidence of ARDS in children compared to adults as well as challenges associated with pediatric research (e.g., smaller circulating blood volume, less frequent blood draws).

Early increase in endothelial biomarkers ANGPT2 and ANGPT2/ANGPT1 ratio was associated with the development of PARDS in our prior study of children with extrapulmonary sepsis (*Whitney et al., 2020b*). Among those with PARDS, persistent organ dysfunction or death was associated with rapid decline in endothelial biomarker sFLT, and slow decline in von Willebrand Factor (vWF) and TM, which reflect the activation of the coagulation system (*Whitney et al., 2020b*). In a cohort of children, who underwent hematopoietic stem cell transplant, PARDS mortality was associated with both early

elevation in and subsequent increase of ANGPT2 (*Zinter et al., 2016*). A profile of matrix metallo-proteinases (MMPs) associated with multiple markers of inflammation and endothelial activation and predicted severe morbidity or mortality in a cohort of 235 children with PARDS (*Zinter et al., 2019*). One study of adults with ALI (*Supplementary file 1*) confirmed the association between increased ANGTP2 and mortality, modified by the presence of infection (*Calfee et al., 2012*).

Biomarkers of systemic inflammation have been studied in peripheral blood and bronchoalveolar lavage (BAL) fluid. In a cohort of children with acute respiratory failure, circulating interleukin-8 (IL-8) level was associated with worse outcomes but not PARDS status (*Flori et al., 2019*). The association of IL-8 with poor outcome from PARDS in the context of sepsis was confirmed in a larger cohort (*Yehya and Wong, 2018*).

A biomarker associated with pulmonary epithelial injury, receptor for advanced glycation end-products (RAGE), is encoded by the *AGER* gene and has activity in the lung. Yehya et al. found increased RAGE and ANGPT2 to be associated with nonsurvival and increased number of nonpulmonary organ failures in PARDS (*Yehya et al., 2016*). Three studies of septic adults (*Supplementary file 1*) supported these findings in demonstrating the association between increased RAGE and ARDS status (*Uchida et al., 2006*; *Jones et al., 2020*; *Ware et al., 2013*).

Defining endotypes with biomarkers

The concept of endophenotype was proposed as 'measurable components unseen by the unaided eye along the pathway between disease and distal genotype (*Gottesman and Gould, 2003*)' while an endotype is a subtype of disease defined by molecular pathobiology of treatment response (*Anderson, 2008*). Distinguishing endophenotypes has been of interest in ARDS, and endotypes of PARDS have been identified, which are associated with distinct clinical factors and outcomes (*Yehya et al., 2018*; *Whitney et al., 2020a*; *Yehya et al., 2018*) potentially reflecting differences in pathobiology (*Table 1*). First, Wong et al. used a large panel of biomarkers to define endotypes in pediatric sepsis, which were associated with differential mortality and response to therapies, including steroids (*Wong et al., 2015*; *Wong et al., 2012*). *Yehya et al., 2019*, recognizing sepsis as the most common cause of PARDS, used similar methods to define an endotype associated with poor outcomes from acute hypoxemic respiratory failure in septic children, identified by IL-8, C-C chemokine ligand 3 (CCL3), and heat shock protein 70 kDa 1B (HSPA1B), and age. The identifying characteristics were then tested in a PARDS cohort with good prediction of mortality (*Yehya and Wong, 2018*).

We previously reported that elevation in endothelial biomarkers ANGPT2/ANGPT1 ratio, VCAM1, and vWF distinguished children with ARDS due to indirect (e.g., extrapulmonary sepsis or trauma, shock, transfusion, pancreatitis) compared to direct (e.g., pneumonia, aspiration, drowning) lung injury (*Whitney et al., 2020a*). An adult ARDS study with complementary results (*Supplementary file 1*) showed lower ANGPT2 and higher SP-D in those with direct lung injury (*Calfee et al., 2015*). Biomarkers reflecting systemic inflammation and endothelial activation have been associated with worse clinical outcomes in adult ARDS (*Supplementary file 1*) with a differential response to therapies (*Bos et al., 2017*; *Calfee et al., 2014*; *Famous et al., 2017*; *Sinha et al., 2020*; *Sinha et al., 2018*).

In summary, protein biomarker studies in PARDS have been numerous and are highly feasible. They have described a pattern of epithelial and endothelial perturbations as well as systemic inflammation, which mirrors the findings of protein biomarker studied first conducted in adults with ARDS. Given known differences in ARDS epidemiology between children and adults and the putative influence exerted by the developing lung and immune systems, it is likely that biological differences between pediatric and adult ARDS exist. A precision medicine approach using unbiased omics methods may better define PARDS pathobiology for each patient as disease progresses or resolves.

In vitro studies

ARDS is typified by pulmonary neutrophilia, and in vitro studies (*Table 1*) have focused on clarifying the mechanisms of neutrophil activation, lack of neutrophil clearance by macrophages, and the role of neutrophil extracellular traps (NETs). NETs are webs of extracellular fibers composed of DNA with histones, myeloperoxidase (MPO) and neutrophil elastase (NE), which are important for immune response to infection but worsen inflammation when they persist (*Brinkmann et al., 2004*; *Liu et al., 2016*). Grunwell et al. examined neutrophils from children with PARDS secondary to viral lower respiratory tract infection with or without bacterial pneumonia, examining transmigration through their

model airway with cell-free tracheal aspirate fluid (*Grunwell et al., 2019*). Patients with bacterial pneumonia had neutrophils with increased markers of activation and a defective respiratory burst, with airway fluid containing higher MPO and NE activity and decreased killing of *H. influenzae* and *S. aureus* (*Grunwell et al., 2019*). In a follow-up study comparing tracheal aspirates from intubated children with and without pediatric ARDS, increased type I interferon signaling (increased phosphorylation of *STAT1*), increased NET release (*Grunwell et al., 2021*). Increased markers of neutrophil activation and degranulation were observed in children with PARDS, also associated with fewer ventilator days in patients with higher airway NE (*Grunwell et al., 2021*). Similar findings were documented in lower airway neutrophils and neutrophils exposed to airway fluid from adults with vs. without ARDS (*Supplementary file 1*), including decreased apoptosis and macrophage activity and increased NET formation in ARDS (*Grégoire et al., 2018*).

These recent investigations represent an exciting development because they elucidate some of the pathobiological changes associated with lower respiratory tract infection that make a host vulnerable to ARDS. However, experiments were limited by study samples that were small and homogeneous, the inability to see changes over time, the possibility that cell signaling happens differently in vitro and in vivo, and a priori determination of the signaling pathways of interest. While this method is not easily applied to individual patients with PARDS, in vitro studies clarify aspects of disease mechanisms where observational studies cannot. Furthermore, in vitro studies can generate hypotheses to be followed-up in future clinical studies.

Unbiased investigations

Genome-wide association studies

Several GWAS have been completed in adults but not children with ARDS (*Supplementary file 1*). In a multiphase study using GWAS to compare European-American adults with trauma-associated ALI to controls, ALI was associated with rs47191 in the *PPFIA1* gene, which encodes liprin-α-1, a protein involved in cell–matrix interactions and cell adhesion (*Christie et al., 2012*). A study of African American patients with ARDS compared to at-risk controls found ARDS to be associated with and a coding SNP (rs2228315) in the *SELPLG* gene, which regulates neutrophil adherence and diapedesis across the vascular endothelium (*Bime et al., 2018*). *SELPLG* knockout mice showed significantly reduced LPS-induced inflammatory lung injury compared with wild-type mice, and a neutralizing antibody against PSGL-1 (P-selectin glycoprotein ligand 1) reduced lung inflammation. In a comparison of European adults with ARDS and at-risk septic patients, an SNP in *FLT1* was associated with reduced susceptibility to ARDS (*Guillen-Guio et al., 2020*).

Together, these studies highlight a possible role for endothelial and pulmonary epithelial changes in genetic susceptibility to ARDS. Although GWAS permits an unbiased investigation of the genetic risks than candidate gene studies, several limitations make it difficult to implement for investigating ARDS and PARDS. First, sample sizes required for case and control groups are large relative to the prevalence of ARDS. Second, biological interpretation of risk alleles and understanding their relationship to associated genes in disease development is challenging. Third, it is important to understand the impact of genetic variants on various tissue types involved in a complex disease like ARDS which affects not only the lungs but peripheral vasculature as well.

Genome-wide gene expression studies

Analysis of microarray data from children with PARDS highlighted differential gene expression associated with biological processes implicated in disease pathogenesis and outcomes (*Table 1*). In evaluating a cohort of children with sepsis and acute hypoxemic respiratory failure, Yehya et al. identified two distinct transcriptomic profiles differentially associated with mortality and persistent organ dysfunction (*Yehya et al., 2019*). Subsequently, the same group evaluated a cohort of PARDS patients using peripheral blood transcriptome to identify three subgroups that were associated with distinct baseline clinical characteristics and ARDS outcomes (*Yehya et al., 2020*). The first subgroup was characterized by enriched adaptive immune response and persistent hypoxemia. The second subgroup was characterized by the activation of complement-related pathways, and the third exhibited suppression of adaptive immune and T-cell receptor pathways, which was associated with improved survival.

To discover transcriptomic signature in airway cells, *Grunwell et al., 2021* collected tracheal aspirate samples from 28 patients with PARDS and 24 without PARDS. The top-ranked gene in the

primary airway cells was *IL17A,* and a 62-gene prediction model for PARDS was developed, which was enriched with the genes in cytokine–cytokine receptor interaction. A neutrophil reporter assay was used to monitor gene expression changes seen in healthy donor neutrophils exposed to airway fluid patients with and without ARDS. The most significantly altered gene in the neutrophil reporter assay was *NT5E*. This preliminary study highlights a transcriptomic signature of PARDS in airway cells and neutrophils that could help identify future therapeutic targets.

Together, gene expression studies have strengthened existing hypotheses about mechanism of disease and outcomes in PARDS. Though some findings were tested in animal models, replication in humans is needed to validate the findings. Transcriptome analysis was performed on peripheral blood in most studies though Grunwell et al. examined airway epithelia. Future studies should continue evaluating differences across multiple tissue types to enhance understanding of ARDS pathogenesis in the human organism as a whole.

Metabolomics

High-resolution mass spectrometry can quantitatively measure hundreds to tens of thousands of small metabolites for diverse chemical species in tissue and liquid biopsy samples (***Kong and Hernandez-Ferrer, 2020***). To our knowledge metabolomics have not been used to study PARDS. Previous studies of adult ARDS (***Supplementary file 1***) used targeted metabolomics platforms with relatively limited chemical space coverage (***Metwaly and Winston, 2020***). Recently, Rogers et al. used an untargeted metabolomics platform to compare the profile of pulmonary edema fluid collected at the time of endotracheal intubation for 16 adults with ARDS compared to 13 with cardiogenic pulmonary edema. They found a high metabolite endotype characterized by higher concentration of lipids, amino acids, and carbohydrates to be associated with higher mortality from ARDS (***Rogers et al., 2017***).

A strength of metabolomic studies is that they are relatively unbiased and produce a large amount of data per patient. Current work is limited to a single fluid in a small cohort, which limits generalizability and replication may be difficult. Future precision medicine approaches to ARDS could include, but should not be limited to, metabolomic investigation of multiple tissue types in concert with genomic, transcriptomic, and protein biomarker analyses.

Interrogation of genomic knowledge databases

Decades of clinical research have proven the utility of ARDS candidate genes to define endophenotypes, predict clinical outcomes, and discover treatment targets. These candidate genes are essential for normal lung function and immune responses, yet many genes have pleiotropic activities in diverse cell types, tissues, and organs.

We searched the NHGRI-EBI GWAS catalog (https://www.ebi.ac.uk/gwas, last accessed: June 8, 2021) to find genotype–phenotype (GP) associations for the candidate genes. Though candidate genes were associated with ARDS in the cited works, we aimed to evaluate whether they were associated with a clinical phenotype similar to or associated with ARDS. Of the 735 GP associations initially found from catalog, we selected 140 GP associations from studies with 3000 or more individuals in both initial and replication cohort. The 140 selected GP pairs involved 32 candidate genes and 64 distinct traits (defined as having distinct identifier in Experimental Factor Ontology [https://www.ebi.ac.uk/ols/ontologies/efo]). The 64 traits were further simplified into 14 phenotypes. The list of gene–trait pair along with simplified phenotypes are shown in ***Supplementary file 2***. The network of associations between genes and simplified phenotypes confirms that the candidate genes are mostly associated with key phenotypes of ARDS: cardiovascular disease, pulmonary function, and systemic inflammatory diseases (***Figure 1***). Genetic pleiotropy may exist between individuals with higher risk for ARDS and these clinical phenotypes. Conversely, pediatric patients recovered from ARDS require close monitoring of these clinical conditions prospectively.

Summary of biomarkers by cell types at different phases

One of the most compelling reasons to characterize ARDS endotypes based on clinical factors, genetic risk, and biomarkers is to identify differences in pathobiology that may respond to different treatment strategies. For example, ANGPT2 is produced by activated endothelial cells, which destabilize the vascular junction in the setting of inflammation (***Kim et al., 2016***). Therefore, increased ANGPT2 has been proposed to have a causal role in ARDS due to endothelial activation (***Reilly et al., 2018***) and

ANGPT or ANGPT2/1 levels associate with clinical outcome (*Whitney et al., 2020a*; *Whitney et al., 2020b*; *Zinter et al., 2016*).

Specific putative biomarkers are more relevant to dysfunction of specific cell types at different phases of ARDS. Thus, we summarized ARDS biomarkers according to cell types implicated in the three major phases of ARDS progression (exudative, proliferative, and fibrotic; *Mason et al., 2016* in *Figure 2*). When considering pathobiological evolution of ARDS through these proposed phases, it is important to note that (1) patients may present to care in different phases, (2) patients may progress through phases at different rates, and (3) scientific knowledge of ARDS phases is more developed for adults than children.

Alveolar cell types 1 and 2 (AT1 and 2) maintain the structure and function of alveoli (*Figure 2*, panel A), and multiple biomarkers indicate epithelial injury. Risk alleles for ARDS are reported in *SFTPB* and *EGF* genes that are important homeostatic function of AT1 cells. In the exudative phase, injured AT1 cells release soluble RAGE, SP-B, SP-D, CC-16 (*SCGB1A*), lamininγ2 (*LAMC2*), KL-6 (*MUC1*), and KGF (*FGF7*, *Figure 2*, panel B). In the proliferative phase, TGF-β and PDGF are released from AT2 cells (*Figure 2*, panel C). Activated AT2 cells continue to secrete TGF-β, PDGF, and IGF-1 to promote interstitial and intraalveolar fibrosis, which restrict lung capacity (*Figure 2*, panel D).

Systemic infection, inflammation, and direct lung injury activate vascular endothelial cells. Genetic susceptibility to ARDS has been reported for the genes that are highly expressed in endothelial cells and have functional significance for normal cellular function, including ANGPT2 (*ANGPT2*), TIE2 (*TEK*), ACE, PBEF (*NAMPT*), VEGF (*VEGFA*), and PAI1 (*SERPINE1*) (*Acosta-Herrera et al., 2014*). Protein biomarkers released from injured endothelial cells include ANGPT1, ANGPT2, VEGF, vWF, apolipoproteins, cell-free hemoglobin, PAI1, endothelial glycocalyx, sTM (*THBD*), sICAM (*ICAM1*), and PROC (*Figure 2*, panel B).

Tissue-resident macrophages defend against pathogens and pollutants entering alveoli. Alveolar macrophages are transformed to inflammatory M1-like macrophages by the activation of NF-κB signaling pathway then cytokines and chemokines, including TNF-α, IL-1β, IL-6, IL-8, CCL2, and CCL7, are released to recruit neutrophils in the exudative phase (*Figure 2*, panel B). Activated neutrophils secrete LTB4, ROS, MMPs, histones, and elastase and form NETs. In the proliferative phase, anti-inflammatory M2-like macrophages release lipoxins, resolvins, KGF, HGF, granulocyte-macrophage colony-stimulating factor (GM-CSF, and VEGF). AT2 cells proliferate and differentiate to AT1 cells by GM-CSF in this phase (*Figure 2C*). Regulatory T cells (Treg) also limit the inflammatory response. M2-like macrophages phagocytose apoptotic neutrophils (efferocytosis).

Genetic determinants of the innate immune response may be related to ARDS susceptibility; however, no risk alleles have been reproduced in two or more independent cohorts. Gene expression studies using peripheral blood monocytes (PBMCs) showed differential expression of *IL1R2*, *FTL*, *PI3*, and *S100A2* associated with an anti-inflammatory response. Protein biomarkers from immune cells, including IL-8, IL-6, IL-1β, IL-18, GM-CSF (*CSF2*), G-CSF (*CSF3*), sTNFR1 (*TNFRSR1A*), IL-10, and TNF (*TNF*), have been studies in ARDS; however, it is difficult to distinguish which cell types are responsible for these biomarkers. G-CSF is the dominant colony-stimulating factor released from diverse lung cells in response to proinflammatory cytokines that stimulates neutrophil development and differentiation. Proinflammatory M1-like macrophages release TNF, IL-6, IL-1β, and IL-23 while anti-inflammatory M2-like macrophages release TGFB1, IL-10, and IL-13.

In patients with ARDS, biomarkers have multiple actions across cell types affected at different stages of disease. Biomarker activity also varies according to the tissue type samples. Peripheral blood, tracheal aspirates, BAL fluid, lung biopsy, and autopsy samples have been used to discover biomarkers of ARDS; however, evaluation of abnormal biomarker levels should be interpreted in the context of cell types and disease stage. To this end, recent developments in maturation of single-cell profiling techniques such as scRNA-seq will be an important for further refinement of ARDS biomarkers.

Cell type-specific expression of ARDS candidate genes in COVID-19

To demonstrate the utility of scRNA-seq and similar approaches, we use a dataset generated from adults with lethal ARDS secondary to COVID-19 (*Melms et al., 2021*). This dataset was composed of single-nucleus RNA sequencing of about 116,000 nuclei from lung autopsy samples that were highly inflamed with activated macrophages (*Figure 3A*). A total of 53 candidate genes were queried
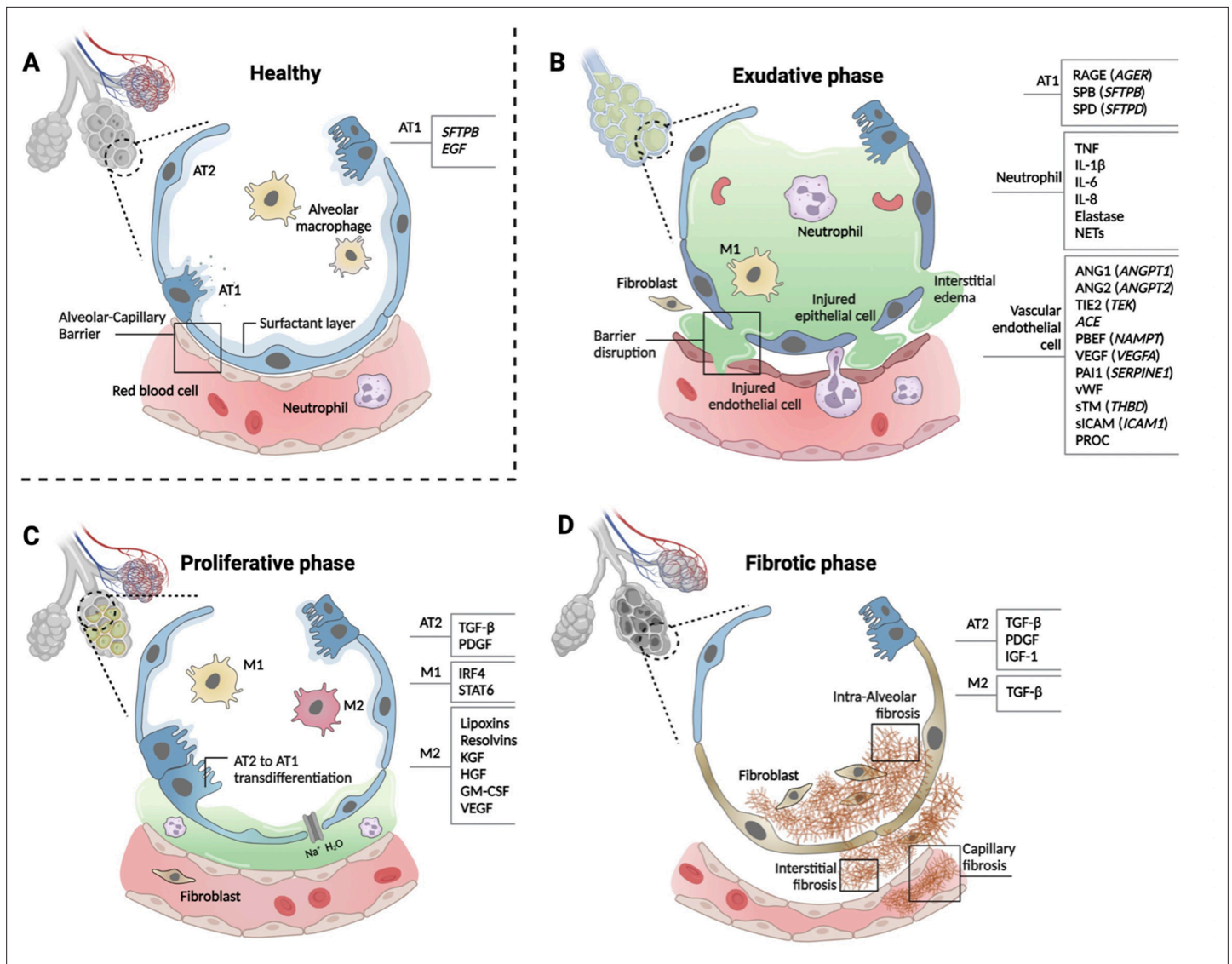
Figure 2. Phases of ARDS pathogenesis. Panel A shows alveolar–capillary interface in a state of good health. Panels B, C, and D, respectively, show the three phases of ARDS development with candidate biomarkers indicated by cell type. Candidate biomarkers indicate dysregulation following a physiologic insult with transition from (**A**) health to (**B**) exudative, (**C**) proliferative, and (**D**) fibrotic phases of illness. Patients may present to care in different phases of ARDS development, may progress through phases at different rates, and biology differ depending on cause of ARDS and patient age. Figure was created using Biorender.com. Abbreviations: ACE = angiotensin-converting enzyme; ANG1/ANGPT1 = angiopoietin 1; ANG2/ANGPT2 = angiopoietin 2; ARDS = acute respiratory distress syndrome; AT1 = type I alveolar cell; AT2 = type II alveolar cell; EGF = epidermal growth factor; GM-CSF = granulocyte-macrophage colony-stimulating factor; HGF = hepatocyte growth factor; IGF = insulin like growth factor; IL-1β = interleukin 1 beta; IL-6 = interleukin 6; IL-8 = interleukin 8; IRF4 = interferon regulatory factor 4; KGF = keratinocyte growth factor; M1 = type I macrophage; M2 = type II macrophage; NAMPT = nicotinamide phosphoribosyltransferase; NET = neutrophil extracellular trap; PAI1 = plasminogen activator inhibitor type 1; PBEF = pre-B-cell colony-enhancing factor; PDGF = platelet-derived growth factor; PROC = protein C, inactivator of coagulation factors Va and VIIIa; RAGE = receptor for advanced glycation end-products; SERPINE1 = serpin family E member 1; SFTPB = surfactant protein B; sICAM/ICAM1 = intercellular adhesion molecule 1; SPB = surfactant-associated protein B; SPD = surfactant-associated protein D; STAT6 = signal transducer and activator of transcription 6; sTM = soluble thrombomodulin; TGF-β = transforming growth factor beta; THBD = thrombomodulin; TIE2 (TEK) = tyrosine kinase with immunoglobulin- and EGF-like domains 1; TNF = tumor necrosis factor; VEGF/VEGFA = vascular endothelial growth factor A; vWF = von Willebrand Factor.

in the Single Cell portal at Broad Institute (available at https://singlecell.broadinstitute.org/single_cell/study/SCP1219/columbia-university-nyp-covid-19-lung-atlas, last accessed August 8, 2021) and scaled mean expression levels are shown for 41 detailed cell types defined by the original study. In *Figure 3B*, the same genes were highlighted in peripheral blood mononuclear cells (PBMCs) collected

Figure 3. Cell type-specific expression pattern of ARDS candidate genes. Relative expression levels of ARDS candidate genes are shown across cell types from lung autopsy sample from ARDS caused by COVID-19 (**A**) and peripheral blood mononuclear cells a generally heathy individual (**B**). Expression levels are scaled to show relative levels across all cell types from 0 (blue) to 1 (red). The size of circle is proportional to the cells expressing a gene for each cell type (from 0% to 75%). AT1 and AT2: alveolar type I and II cells; FB: fibroblasts; ECM-high: high expression of extracellular matrix components (i.e., *COL6A3, COL1A2*, and *COL3A1*); MDM: monocyte-derived macrophages. Source data for this figure are provided in a table, *Figure 3—source data 1*.

The online version of this article includes the following source data for figure 3:

Source data 1. Source data for *Figure 3*.

from healthy adults (***Ding et al., 2020***). The *CARMIL1* gene was highly expressed in diverse lung cells from severe COVID-19 patients; however, it was not expressed in PBMCs from healthy adults. The *EPAS1* gene (also known as *HIF2A*) was highly expressed in endothelial cells as well as AT1 and AT2 cells. The *FER* gene encodes a tyrosine kinase that is involved in the regulation of actin cytoskeleton. The expression of *FER* was observed in multiple cell types including AT1, AT2, alveolar macrophages, and mast cells in ARDS lung while peripheral immune cells are not actively express this gene. *IL1R1*—implicated in NET formation—was expressed in endothelial cells, epithelial cells, and fibroblasts. The

highest expression of *VEGFA* was found in AT1 cells and *vWF* expression was high across endothelial and epithelial cells. Many candidate genes for ARDS were highly expressed in lung cells while few genes such as *MAP3K1* and *NAMPT* were also expressed in PBMC suggesting PBMC as a potential source of biomarker discovery (*García-Laorden et al., 2017*).

Recent studies using autopsy samples from adults with ARDS in the setting of COVID-19 infection uncovered key pathobiological changes in lung tissue. Lung tissue samples were found to be infiltrated with aberrantly activated monocyte-derived macrophages and alveolar macrophages that produced high levels of interleukin (IL)-1β and IL-6 (*Xu et al., 2020*; *Jiang et al., 2020*). Moreover, failure to transition from AT1 to AT2 cells impaired lung regeneration and repair (*Melms et al., 2021*). These discoveries have added important information to our understanding of ARDS pathobiology and could pave the way for the development of targeted therapies.

Conclusion

Prior studies have provided insight into the pathobiological pathways that determine whether ARDS develops in patients with risk factors and what clinical outcomes are experienced by adults and children with ARDS (*Figure 4*). Ability to define ARDS endotypes based on clinical and genomic markers may indicate heterogeneity in underlying pathobiology. However, prior literature is characterized by several limitations. First, much of the existing literature is based on relatively small cohorts, heterogeneous samples, and measurements confined to one or several timepoints during disease development. Adults are better represented than children in the current ARDS literature. Second, many studies were designed to test hypotheses based on prior knowledge, and therefore investigation of disease mechanism are not unbiased and could have missed alternative conclusions. Third, most studies evaluated peripheral blood with a few more recent investigations of lower airway cells and BAL fluid. However, not all genomic changes are reflected in circulating proteins and circulating protein changes, if present, may not reflect changes at the tissue level. Evaluation of lung tissue from autopsy specimens, while offering the most specificity investigations using scRNA-seq, is not always practical in PARDS, owing to its relatively low mortality rate. Lung biopsy of children with PARDS may confer unacceptable morbidity, so tracheal aspirate or BAL specimens may be most practical. PBMCs have been subjected to scRNA-seq in several adult ARDS studies that demonstrated feasibility; however, differences between the blood and lung compartments deserve further study. A logical next step to address some of these issues is scRNA-seq evaluation of tracheal aspirate samples as well as peripheral blood from a well-characterized cohort of children with PARDS, who can be longitudinally sampled and followed. As technological advances continue, new methods should continue to be applied in order to elucidate the pathobiology of PARDS as it develops and resolves so that disease-targeted therapies can be developed.

Ideas and speculation: opportunities for new discoveries with scRNA-seq

PARDS is caused by diverse exogenous pathogens and stressors, though genetic susceptibility may play a role in which at-risk patients develop the disease. Although risk alleles in candidate genes from a limited number of studies with small sample sizes have been reported for PARDS, most studies were underpowered. Phenotypes associated with candidate genes are diverse across multiple organ systems including heart, lung, kidney, and immune systems. scRNA-seq and other unbiased characterization methods at cell level provide great potential to identify cell types and markers in genes across multiple affected tissues and circulating immune cells.
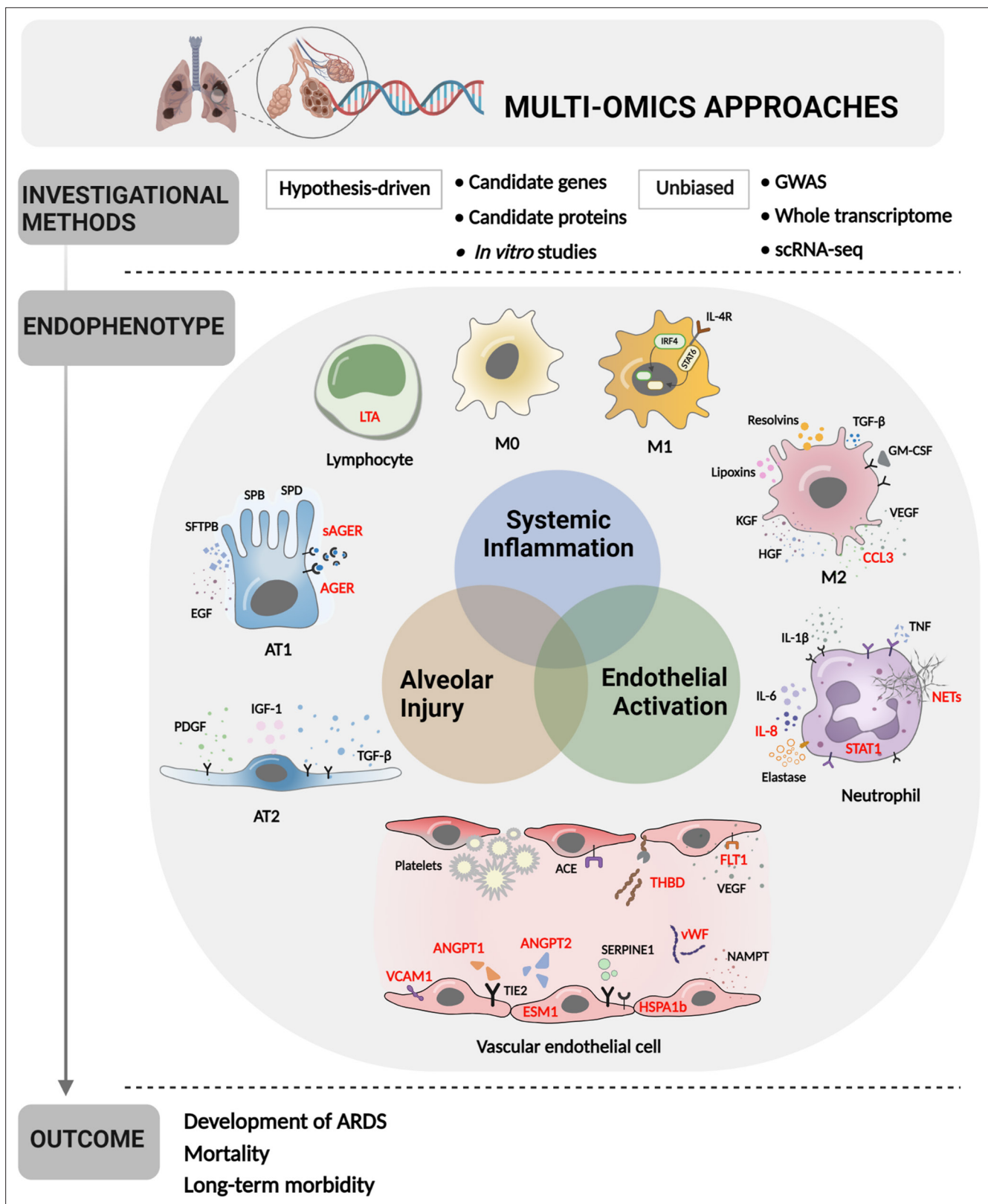
Figure 4. Multi-omics approaches to define endotypes of pediatric acute respiratory distress syndrome (PARDS). Investigational methods, including hypothesis-driven and unbiased investigations, have led to discovery of multiple contributions to ARDS pathobiology, including systemic inflammation, endothelial activation, and alveolar injury. These processes affect diverse cell and tissue types, including lymphocytes, macrophages, neutrophils, vascular endothelial cells, and two types of alveolar cells. The degree of systemic inflammation, endothelial activation, and alveolar injury across

Figure 4 continued on next page

Figure 4 continued

cell and tissue types has helped to define ARDS endotypes, which have been differentially associated with relevant clinical outcomes, including the development of ARDS in patients who have experienced clinical risk factors for the disease, mortality, and long-term morbidity from ARDS. Names of proteins implicated in PARDS as opposed to adult ARDS are depicted in red. Figure was created using Biorender.com. Abbreviations: ACE = angiotensin-converting enzyme; ANGPT1, -2 = angiopoietin 1, 2; CCL3 = C-C motif chemokine ligand 3; ESM1 = endothelial cell-specific molecule 1; FLT1 = fms-related receptor tyrosine kinase 1; GM-CSF = granulocyte-macrophage colony-stimulating factor; HGF = hepatocyte growth factor; HSPA1b = heat shock protein family A (Hsp70) member 1B; IL- = interleukin-; IRF4 = interferon regulatory factor 4; KGF = keratinocyte growth factor; LTA = lymphotoxin alpha; M0, M1, M2 = type 0, I, II macrophage; NAMPT = nicotinamide phosphoribosyltransferase; NET = neutrophil extracellular trap; SERPINE1 = serpin family E member 1; STAT1, -6 = signal transducer and activator of transcription-1,6; TGF-β = transforming growth factor beta; THBD = thrombomodulin; TIE2 = tyrosine kinase with immunoglobulin like and EGF like domains 1; VCAM1 = vascular cell adhesion molecule 1; VEGF = vascular endothelial growth factor A; vWF = von Willebrand Factor.

Additional information

Author contributions

Jane E Whitney, Conceptualization, Resources, Funding acquisition, Methodology, Writing - original draft, Writing - review and editing; In-Hee Lee, Ji-Won Lee, Visualization, Writing - review and editing; Sek Won Kong, Conceptualization, Resources, Supervision, Funding acquisition, Visualization, Methodology, Writing - review and editing

Author ORCIDs

Jane E Whitney http://orcid.org/0000-0002-5151-2626
In-Hee Lee http://orcid.org/0000-0002-8857-1355
Ji-Won Lee http://orcid.org/0000-0003-4578-2168
Sek Won Kong http://orcid.org/0000-0003-4877-7567

Additional files

Supplementary files

• Supplementary file 1. Key adult studies.

• Supplementary file 2. Phenotypes associated with ARDS candidate genes from genome-wide association studies including at least 3000 samples.

References

Acosta-Herrera M, Pino-Yanes M, Perez-Mendez L, Villar J, Flores C. 2014. Assessing the quality of studies supporting genetic susceptibility and outcomes of ARDS. *Frontiers in Genetics* 5:2008–2013. DOI: https://doi.org/10.3389/fgene.2014.00020, PMID: 24567738

Adamzik M, Frey U, Sixt S, Knemeyer L, Beiderlinden M, Peters J, Siffert W. 2007. ACE I/D but not AGT (-6)A/G polymorphism is A risk factor for mortality in ARDS. *The European Respiratory Journal* 29:482–488. DOI: https://doi.org/10.1183/09031936.00046106, PMID: 17107992

Aird WC. 2007. Phenotypic heterogeneity of the endothelium: II. Representative vascular beds. *Circulation Research* 100:174–190. DOI: https://doi.org/10.1161/01.RES.0000255690.03436.ae, PMID: 17272819

Anderson GP. 2008. Endotyping asthma: new insights into key pathogenic mechanisms in a complex, heterogeneous disease. *Lancet* **372**:1107–1119. DOI: https://doi.org/10.1016/S0140-6736(08)61452-X, PMID: 18805339

Azevedo ZM, Moore DB, Lima FC, Cardoso CC, Bougleux R, Matos GI, Luz RA, Xavier-Elsas P, Sampaio EP, Gaspar-Elsas MI, Moraes MO. 2012. Tumor necrosis factor (TNF) and lymphotoxin-alpha (LTA) single nucleotide polymorphisms: importance in ARDS in septic pediatric critically ill patients. *Human Immunology* **73**:661–667. DOI: https://doi.org/10.1016/j.humimm.2012.03.007, PMID: 22507624

Bime C, Pouladi N, Sammani S, Batai K, Casanova N, Zhou T, Kempf CL, Sun X, Camp SM, Wang T, Kittles RA, Lussier YA, Jones TK, Reilly JP, Meyer NJ, Christie JD, Karnes JH, Gonzalez-Garay M, Christiani DC, Yates CR, et al. 2018. Genome-wide association study in african americans with acute respiratory distress syndrome identifies the selectin p ligand gene as a risk factor. *American Journal of Respiratory and Critical Care Medicine* **197**:1421–1432. DOI: https://doi.org/10.1164/rccm.201705-0961OC, PMID: 29425463

Bos LD, Schouten LR, van Vught LA, Wiewel MA, Ong DSY, Cremer O, Artigas A, Martin-Loeches I, Hoogendijk AJ, van der Poll T, Horn J, Juffermans N, Calfee CS, Schultz MJ, MARS consortium. 2017. Identification and validation of distinct biological phenotypes in patients with acute respiratory distress syndrome by cluster analysis. *Thorax* **72**:876–883. DOI: https://doi.org/10.1136/thoraxjnl-2016-209719, PMID: 28450529

Brinkmann V, Reichard U, Goosmann C, Fauler B, Uhlemann Y, Weiss DS, Weinrauch Y, Zychlinsky A. 2004. Neutrophil extracellular traps kill bacteria. *Science* **303**:1532–1535. DOI: https://doi.org/10.1126/science. 1092385, PMID: 15001782

Calfee CS, Gallagher D, Abbott J, Thompson BT, Matthay MA, NHLBI ARDS Network. 2012. Plasma angiopoietin-2 in clinical acute lung injury: prognostic and pathogenetic significance. *Critical Care Medicine* **40**:1731–1737. DOI: https://doi.org/10.1097/CCM.0b013e3182451c87, PMID: 22610178

Calfee CS, Delucchi K, Parsons PE, Thompson BT, Ware LB, Matthay MA. 2014. Latent class analysis of ards subphenotypes: analysis of data from two randomized controlled trials. *The Lancet. Respiratory Medicine* **2**:611–620. DOI: https://doi.org/10.1016/S2213-2600(14)70097-9.Latent

Calfee CS, Janz DR, Bernard GR, May AK, Kangelaris KN, Matthay MA, Ware LB. 2015. Distinct molecular phenotypes of direct vs indirect ARDS in single-center and multicenter studies. *Chest* **147**:1539–1548. DOI: https://doi.org/10.1378/chest.14-2454, PMID: 26033126

Christie JD, Wurfel MM, Feng R, O'Keefe GE, Bradfield J, Ware LB, Christiani DC, Calfee CS, Cohen MJ, Matthay M, Meyer NJ, Kim C, Li M, Akey J, Barnes KC, Sevransky J, Lanken PN, May AK, Aplenc R, Maloney JP, et al. 2012. Genome wide association identifies PPFIA1 as a candidate gene for acute lung injury risk following major trauma. *PLOS ONE* **7**:e0028268. DOI: https://doi.org/10.1371/journal.pone.0028268, PMID: 22295056

Cruces P, Díaz F, Puga A, Erranz B, Donoso A, Carvajal C, Wilhelm J, Repetto GM. 2012. Angiotensin-converting enzyme insertion/deletion polymorphism is associated with severe hypoxemia in pediatric ARDS. *Intensive Care Medicine* **38**:113–119. DOI: https://doi.org/10.1007/s00134-011-2381-3, PMID: 22005825

Dahmer MK, O'cain P, Patwari PP, Simpson P, Li S-H, Halligan N, Quasney MW. 2011. The influence of genetic variation in surfactant protein B on severe lung injury in African American children. *Critical Care Medicine* **39**:1138–1144. DOI: https://doi.org/10.1097/CCM.0b013e31820a9416, PMID: 21283003

Ding J, Adiconis X, Simmons SK, Kowalczyk MS, Hession CC, Marjanovic ND, Hughes TK, Wadsworth MH, Burks T, Nguyen LT, Kwon JYH, Barak B, Ge W, Kedaigle AJ, Carroll S, Li S, Hacohen N, Rozenblatt-Rosen O, Shalek AK, Villani AC, et al. 2020. Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. *Nature Biotechnology* **38**:737–746. DOI: https://doi.org/10.1038/s41587-020-0465-8, PMID: 32341560

Famous KR, Delucchi K, Ware LB, Kangelaris KN, Liu KD, Thompson BT, Calfee CS, ARDS Network. 2017. Acute respiratory distress syndrome subphenotypes respond differently to randomized fluid management strategy. *American Journal of Respiratory and Critical Care Medicine* **195**:331–338. DOI: https://doi.org/10.1164/rccm. 201603-0645OC, PMID: 27513822

Flori H, Sapru A, Quasney MW, Gildengorin G, Curley MAQ, Matthay MA, Dahmer MK, BALI and RESTORE Study Investigators, Pediatric Acute Lung Injury and Sepsis Investigators PALISI Network. 2019. A prospective investigation of interleukin-8 levels in pediatric acute respiratory failure and acute respiratory distress syndrome. *Critical Care* **23**:128. DOI: https://doi.org/10.1186/s13054-019-2342-8, PMID: 30995942

Gandhi CK, Chen C, Wu R, Yang L, Thorenoor N, Thomas NJ, DiAngelo SL, Spear D, Keim G, Yehya N, Floros J. 2020. Association of SNP-SNP interactions of surfactant protein genes with pediatric acute respiratory failure. *Journal of Clinical Medicine* **9**:E1183. DOI: https://doi.org/10.3390/jcm9041183, PMID: 32326132

García-Laorden MI, Lorente JA, Flores C, Slutsky AS, Villar J. 2017. Biomarkers for the acute respiratory distress syndrome: how to make the diagnosis more precise. *Annals of Translational Medicine* **5**:283. DOI: https://doi. org/10.21037/atm.2017.06.49, PMID: 28828358

Gottesman II, Gould TD. 2003. The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry* **160**:636–645. DOI: https://doi.org/10.1176/appi.ajp.160.4.636, PMID: 16409270

Grégoire M, Uhel F, Lesouhaitier M, Gacouin A, Guirriec M, Mourcin F, Dumontet E, Chalin A, Samson M, Berthelot LL, Tissot A, Kerjouan M, Jouneau S, Le Tulzo Y, Tarte K, Zmijewski JW, Tadié JM. 2018. Impaired efferocytosis and neutrophil extracellular trap clearance by macrophages in ARDS. *The European Respiratory Journal* **52**:1702590. DOI: https://doi.org/10.1183/13993003.02590-2017, PMID: 29946009

Grunwell JR, Giacalone VD, Stephenson S, Margaroli C, Dobosh BS, Brown MR, Fitzpatrick AM, Tirouvanziam R. 2019. Neutrophil dysfunction in the airways of children with acute respiratory failure due to lower respiratory tract viral and bacterial coinfections. *Scientific Reports* **9**:1–13. DOI: https://doi.org/10.1038/s41598-019-39726-w, PMID: 30814584

Grunwell JR, Rad MG, Stephenson ST, Mohammad AF, Opolka C, Fitzpatrick AM, Kamaleswaran R. 2021. Machine learning-based discovery of a gene expression signature in pediatric acute respiratory distress syndrome. *Critical Care Explorations* **3**:e0431. DOI: https://doi.org/10.1097/CCE.0000000000000431, PMID: 34151274

Guillen-Guio B, Lorenzo-Salazar JM, Ma SF, Hou PC, Hernandez-Beeftink T, Corrales A, García-Laorden MI, Jou J, Espinosa E, Muriel A, Domínguez D, Lorente L, Martín MM, Rodríguez-Gallego C, Solé-Violán J, Ambrós A, Carriedo D, Blanco J, Añón JM, Reilly JP, et al. 2020. Sepsis-associated acute respiratory distress syndrome in individuals of European ancestry: a genome-wide association study. *The Lancet. Respiratory Medicine* **8**:258–266. DOI: https://doi.org/10.1016/S2213-2600(19)30368-6, PMID: 31982041

Hernandez-Pacheco N, Guillen-Guio B, Acosta-Herrera M, Pino-Yanes M, Corrales A, Ambrós A, Nogales L, Muriel A, González-Higueras E, Diaz-Dominguez FJ, Zavala E, Belda J, Ma SF, Villar J, Flores C, GEN-SEP Network. 2018. A vascular endothelial growth factor receptor gene variant is associated with susceptibility to acute respiratory distress syndrome. *Intensive Care Medicine Experimental* **6**:16. DOI: https://doi.org/10.1186/s40635-018-0181-6, PMID: 29987654

Hough RF. 2017. Recent advances in pediatric acute respiratory distress syndrome (PARDS). *Current Pediatrics Reports* **5**:228–236. DOI: https://doi.org/10.1007/s40124-017-0146-3

Jiang Y, Rosborough BR, Chen J, Das S, Kitsios GD, McVerry BJ, Mallampalli RK, Lee JS, Ray A, Chen W, Ray P. 2020. Single cell RNA sequencing identifies an early monocyte gene signature in acute respiratory distress syndrome. *JCI Insight* **5**:1–16. DOI: https://doi.org/10.1172/jci.insight.135678, PMID: 32554932

Jones TK, Feng R, Kerchberger VE, Reilly JP, Anderson BJ, Shashaty MGS, Wang F, Dunn TG, Riley TR, Abbott J, Ittner CAG, Christiani DC, Mikacenic C, Wurfel MM, Ware LB, Calfee CS, Matthay MA, Christie JD, Meyer NJ. 2020. Plasma sRAGE acts as a genetically regulated causal intermediate in sepsis-associated acute respiratory distress syndrome. *American Journal of Respiratory and Critical Care Medicine* **201**:47–56. DOI: https://doi.org/10.1164/rccm.201810-2033OC, PMID: 31487195

Khemani RG, Smith L, Lopez-Fernandez YM, Kwok J, Morzov R, Klein MJ, Yehya N, Willson D, Kneyber MCJ, Lillie J, Fernandez A, Newth CJL, Jouvet P, Thomas NJ, Pediatric Acute Respiratory Distress syndrome Incidence and Epidemiology PARDIE Investigators, Pediatric Acute Lung Injury and Sepsis Investigators PALISI Network. 2019. Paediatric acute respiratory distress syndrome incidence and epidemiology (PARDIE): an international, observational study. *The Lancet. Respiratory Medicine* **7**:115–128. DOI: https://doi.org/10.1016/S2213-2600(18)30344-8, PMID: 30361119

Kim M, Allen B, Korhonen EA, Nitschké M, Yang HW, Baluk P, Saharinen P, Alitalo K, Daly C, Thurston G, McDonald DM. 2016. Opposing actions of angiopoietin-2 on Tie2 signaling and FOXO1 activation. *The Journal of Clinical Investigation* **126**:3511–3525. DOI: https://doi.org/10.1172/JCI84871, PMID: 27548529

Kong SW, Hernandez-Ferrer C. 2020. Assessment of coverage for endogenous metabolites and exogenous chemical compounds using an untargeted metabolomics platform. Pacific Symposium on Biocomputing. 587–598. DOI: https://doi.org/10.1142/9789811215636_0052

Li G, Malinchoc M, Cartin-Ceba R, Venkata CV, Kor DJ, Peters SG, Hubmayr RD, Gajic O. 2011. Eight-year trend of acute respiratory distress syndrome: A population-based study in Olmsted County, Minnesota. *American Journal of Respiratory and Critical Care Medicine* **183**:59–66. DOI: https://doi.org/10.1164/rccm.201003-0436OC, PMID: 20693377

Liu S, Su X, Pan P, Zhang L, Hu Y, Tan H, Wu D, Liu B, Li H, Li H, Li Y, Dai M, Li Y, Hu C, Tsung A. 2016. Neutrophil extracellular traps are indirectly triggered by lipopolysaccharide and contribute to acute lung injury. *Scientific Reports* **6**:1–8. DOI: https://doi.org/10.1038/srep37252, PMID: 27849031

Marshall RP, Webb S, Bellingan GJ, Montgomery HE, Chaudhari B, McAnulty RJ, Humphries SE, Hill MR, Laurent GJ. 2002. Angiotensin converting enzyme insertion/deletion polymorphism is associated with susceptibility and outcome in acute respiratory distress syndrome. *American Journal of Respiratory and Critical Care Medicine* **166**:646–650. DOI: https://doi.org/10.1164/rccm.2108086, PMID: 12204859

Mason C, Dooley N, Griffiths M. 2016. Acute respiratory distress syndrome. *Clinical Medicine* **16**:s66–s70. DOI: https://doi.org/10.7861/clinmedicine.16-6-s66, PMID: 27956444

Melms JC, Biermann J, Huang H, Wang Y, Nair A, Tagore S, Katsyv I, Rendeiro AF, Amin AD, Schapiro D, Frangieh CJ, Luoma AM, Filliol A, Fang Y, Ravichandran H, Clausi MG, Alba GA, Rogava M, Chen SW, Ho P, et al. 2021. A molecular single-cell lung atlas of lethal COVID-19. *Nature* **595**:114–119. DOI: https://doi.org/10.1038/s41586-021-03569-1, PMID: 33915568

Metwaly SM, Winston BW. 2020. Systems biology ARDS research with a focus on metabolomics. *Metabolites* **10**:E207. DOI: https://doi.org/10.3390/metabo10050207, PMID: 32438561

Meyer NJ, Li M, Feng R, Bradfield J, Gallop R, Bellamy S, Fuchs BD, Lanken PN, Albelda SM, Rushefski M, Aplenc R, Abramova H, Atochina-Vasserman EN, Beers MF, Calfee CS, Cohen MJ, Pittet J-F, Christiani DC, O'Keefe GE, Ware LB, et al. 2011. ANGPT2 genetic variant is associated with trauma-associated acute lung injury and altered plasma angiopoietin-2 isoform ratio. *American Journal of Respiratory and Critical Care Medicine* **183**:1344–1353. DOI: https://doi.org/10.1164/rccm.201005-0701OC, PMID: 21257790

Pediatric Acute Lung Injury Consensus Conference Group. 2015. Pediatric acute respiratory distress syndrome: consensus recommendations from the Pediatric Acute Lung Injury Consensus Conference. *Pediatric Critical Care Medicine* **16**:428–439. DOI: https://doi.org/10.1097/PCC.0000000000000350, PMID: 25647235

Plunkett A, Agbeko RS, Li KW, Humphries SE, Klein NJ, Peters MJ. 2008. Angiotensin-converting enzyme D allele does not influence susceptibility to acute hypoxic respiratory failure in children. *Intensive Care Medicine* **34**:2279–2283. DOI: https://doi.org/10.1007/s00134-008-1260-z, PMID: 18787808

Reilly JP, Wang F, Jones TK, Palakshappa JA, Anderson BJ, Shashaty MGS, Dunn TG, Johansson ED, Riley TR, Lim B, Abbott J, Ittner CAG, Cantu E, Lin X, Mikacenic C, Wurfel MM, Christiani DC, Calfee CS, Matthay MA, Christie JD, et al. 2018. Plasma angiopoietin-2 as a potential causal marker in sepsis-associated ARDS development: evidence from Mendelian randomization and mediation analysis. *Intensive Care Medicine* **44**:1849–1858. DOI: https://doi.org/10.1007/s00134-018-5328-0, PMID: 30343317

Reilly JP, Meyer NJ, Shashaty MG, Anderson BJ, Ittner C, Dunn TG, Lim B, Forker C, Bonk MP, Kotloff E, Feng R, Cantu E, Mangalmurti NS, Calfee CS, Matthay MA, Mikacenic C, Walley KR, Russell J, Christiani DC, Wurfel MM, et al. 2021. The ABO histo-blood group, endothelial activation, and acute respiratory distress syndrome risk in critical illness. *The Journal of Clinical Investigation* **131**:1–12. DOI: https://doi.org/10.1172/JCI139700, PMID: 32931480

Rogers AJ, Contrepois K, Wu M, Zheng M, Peltz G, Ware LB, Matthay MA. 2017. Profiling of ARDS pulmonary edema fluid identifies a metabolically distinct subset. *American Journal of Physiology. Lung Cellular and Molecular Physiology* **312**:L703–L709. DOI: https://doi.org/10.1152/ajplung.00438.2016, PMID: 28258106

Rubenfeld GD, Caldwell E, Peabody E. 2009. Incidence and outcomes of pediatric acute lung injury. *Pediatrics* **124**:87–95. DOI: https://doi.org/10.1542/peds.2007-2462, PMID: 19564287

Sapru A, Liu KD, Wiemels J, Hansen H, Pawlikowska L, Poon A, Jorgenson E, Witte JS, Calfee CS, Ware LB, Matthay MA, NHLBI ARDS Network. 2016. Association of common genetic variation in the protein C pathway genes with clinical outcomes in acute respiratory distress syndrome. *Critical Care* **20**:151. DOI: https://doi.org/10.1186/s13054-016-1330-5, PMID: 27215212

Schouten LRA, Veltkamp F, Bos AP, van Woensel JBM, Serpa Neto A, Schultz MJ, Wösten-van Asperen RM. 2016. Incidence and mortality of acute respiratory distress syndrome in children: a systematic review and meta-analysis. *Critical Care Medicine* **44**:819–829. DOI: https://doi.org/10.1097/CCM.0000000000001388, PMID: 26509320

Sinha P, Delucchi KL, Thompson BT, McAuley DF, Matthay MA, Calfee CS, NHLBI ARDS Network. 2018. Latent class analysis of ARDS subphenotypes: a secondary analysis of the statins for acutely injured lungs from sepsis (SAILS) study. *Intensive Care Medicine* **44**:1859–1869. DOI: https://doi.org/10.1007/s00134-018-5378-3, PMID: 30291376

Sinha P, Delucchi KL, McAuley DF, O'Kane CM, Matthay MA, Calfee CS. 2020. Development and validation of parsimonious algorithms to classify acute respiratory distress syndrome phenotypes: a secondary analysis of randomised controlled trials. *The Lancet. Respiratory Medicine* **8**:247–257. DOI: https://doi.org/10.1016/S2213-2600(19)30369-8, PMID: 31948926

Su L, Zhai R, Sheu C-C, Gallagher DC, Gong MN, Tejera P, Thompson BT, Christiani DC. 2009. Genetic variants in the angiopoietin-2 gene are associated with increased risk of ARDS. *Intensive Care Medicine* **35**:1024–1030. DOI: https://doi.org/10.1007/s00134-009-1413-8, PMID: 19271210

Tejera P, Meyer NJ, Chen F, Feng R, Zhao Y, O'Mahony DS, Li L, Sheu C-C, Zhai R, Wang Z, Su L, Bajwa E, Ahasic AM, Clardy PF, Gong MN, Frank AJ, Lanken PN, Thompson BT, Christie JD, Wurfel MM, et al. 2012. Distinct and replicable genetic risk factors for acute respiratory distress syndrome of pulmonary or extrapulmonary origin. *Journal of Medical Genetics* **49**:671–680. DOI: https://doi.org/10.1136/jmedgenet-2012-100972

Uchida T, Shirasawa M, Ware LB, Kojima K, Hata Y, Makita K, Mednick G, Matthay ZA, Matthay MA. 2006. Receptor for advanced glycation end-products is a marker of type I cell injury in acute lung injury. *American Journal of Respiratory and Critical Care Medicine* **173**:1008–1015. DOI: https://doi.org/10.1164/rccm.200509-1477OC, PMID: 16456142

Ware LB, Koyama T, Zhao Z, Janz DR, Wickersham N, Bernard GR, May AK, Calfee CS, Matthay MA. 2013. Biomarkers of lung epithelial injury and inflammation distinguish severe sepsis patients with acute respiratory distress syndrome. *Critical Care* **17**:R253. DOI: https://doi.org/10.1186/cc13080, PMID: 24156650

Whitney JE, Feng R, Koterba N, Chen F, Bush J, Graham K, Lacey SF, Melenhorst JJ, Parikh SM, Weiss SL, Yehya N. 2020a. Endothelial biomarkers are associated with indirect lung injury in sepsis-associated pediatric acute respiratory distress syndrome. *Critical Care Explorations* **2**:e0295. DOI: https://doi.org/10.1097/CCE.0000000000000295, PMID: 33299985

Whitney JE, Zhang B, Koterba N, Chen F, Bush J, Graham K, Lacey SF, Melenhorst JJ, Teachey DT, Mensinger JL, Yehya N, Weiss SL. 2020b. Systemic endothelial activation is associated with early acute respiratory distress syndrome in children with extrapulmonary sepsis. *Critical Care Medicine* **48**:344–352. DOI: https://doi.org/10.1097/CCM.0000000000004091, PMID: 32058372

Whitsett JA, Weaver TE. 2015. Alveolar development and disease. *American Journal of Respiratory Cell and Molecular Biology* **53**:1–7. DOI: https://doi.org/10.1165/rcmb.2015-0128PS, PMID: 25932959

Wong HR, Salisbury S, Xiao Q, Cvijanovich NZ, Hall M, Allen GL, Thomas NJ, Freishtat RJ, Anas N, Meyer K, Checchia PA, Lin R, Shanley TP, Bigham MT, Sen A, Nowak J, Quasney M, Henricksen JW, Chopra A, Banschbach S, et al. 2012. The pediatric sepsis biomarker risk model. *Critical Care* **16**:R174. DOI: https://doi.org/10.1186/cc11652, PMID: 23025259

Wong HR, Cvijanovich NZ, Anas N, Allen GL, Thomas NJ, Bigham MT, Weiss SL, Fitzgerald J, Checchia PA, Meyer K, Shanley TP, Quasney M, Hall M, Gedeit R, Freishtat RJ, Nowak J, Shekhar RS, Gertz S, Dawson E, Howard K, et al. 2015. Developing a clinically feasible personalized medicine approach to pediatric septic

shock. *American Journal of Respiratory and Critical Care Medicine* **191**:309–315. DOI: https://doi.org/10.1164/rccm.201410-1864OC, PMID: 25489881

Xu G, Qi F, Li H, Yang Q, Wang H, Wang X, Liu X, Zhao J, Liao X, Liu Y, Liu L, Zhang S, Zhang Z. 2020. The differential immune responses to COVID-19 in peripheral and lung revealed by single-cell RNA sequencing. *Cell Discovery* **6**:73. DOI: https://doi.org/10.1038/s41421-020-00225-2, PMID: 33101705

Yehya N, Thomas NJ, Meyer NJ, Christie JD, Berg RA, Margulies SS. 2016. Circulating markers of endothelial and alveolar epithelial dysfunction are associated with mortality in pediatric acute respiratory distress syndrome. *Intensive Care Medicine* **42**:1137–1145. DOI: https://doi.org/10.1007/s00134-016-4352-1, PMID: 27101828

Yehya N, Keim G, Thomas NJ. 2018. Subtypes of pediatric acute respiratory distress syndrome have different predictors of mortality. *Intensive Care Medicine* **44**:1230–1239. DOI: https://doi.org/10.1007/s00134-018-5286-6, PMID: 29971591

Yehya N, Wong HR. 2018. Adaptation of a biomarker-based sepsis mortality risk stratification tool for pediatric acute respiratory distress syndrome. *Critical Care Medicine* **46**:e9–e16. DOI: https://doi.org/10.1097/CCM.0000000000002754, PMID: 28991832

Yehya N, Thomas NJ, Wong HR. 2019. Evidence of endotypes in pediatric acute hypoxemic respiratory failure caused by sepsis. *Pediatric Critical Care Medicine* **20**:110–112. DOI: https://doi.org/10.1097/PCC.0000000000001808, PMID: 30720645

Yehya N, Varisco BM, Thomas NJ, Wong HR, Christie JD, Feng R. 2020. Peripheral blood transcriptomic sub-phenotypes of pediatric acute respiratory distress syndrome. *Critical Care* **24**:681. DOI: https://doi.org/10.1186/s13054-020-03410-7, PMID: 33287889

Zinter MS, Spicer A, Orwoll BO, Alkhouli M, Dvorak CC, Calfee CS, Matthay MA, Sapru A. 2016. Plasma angiopoietin-2 outperforms other markers of endothelial injury in prognosticating pediatric ARDS mortality. *American Journal of Physiology. Lung Cellular and Molecular Physiology* **310**:L224–L231. DOI: https://doi.org/10.1152/ajplung.00336.2015, PMID: 26660787

Zinter MS, Delucchi KL, Kong MY, Orwoll BE, Spicer AS, Lim MJ, Alkhouli MF, Ratiu AE, McKenzie AV, McQuillen PS, Dvorak CC, Calfee CS, Matthay MA, Sapru A. 2019. Early plasma matrix metalloproteinase profiles. a novel pathway in pediatric acute respiratory distress syndrome. *American Journal of Respiratory and Critical Care Medicine* **199**:181–189. DOI: https://doi.org/10.1164/rccm.201804-0678OC, PMID: 30114376