Use of a Large Language Model (LLM)-enabled Text Review and Analysis Workflow (LitExplorer) to Evaluate the Current State of the Science for Biomarkers of Acute Respiratory Distress Syndrome (ARDS)

Shreyas U. Hirway[1,2], Daniel W. Cowan[1,2], Thomas Jaworek[1,2], Teresa R. Sterner[1,3], Rebecca A. Clewell[1,4]

[1]Force Health Protection Section, 711th Human Performance Wing, Air Force Research Laboratory, Wright-Patterson Air Force Base (WPAFB), OH; [2]BlueHalo, an AV Inc. Company, Dayton, OH; [3]Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., WPAFB, OH; [4]Eagle Integrated Services, LLC, WPAFB, OH, USA

Background & Purpose

Background

- Acute Respiratory Distress Syndrome (ARDS): life-threatening respiratory condition characterized by lung edema and hypoxemia
- Joint operational causes: Blast/burn inhalation injuries, combat trauma, toxic smoke inhalation, diving-related pulmonary barotrauma, high-altitude hypoxemia, chemical/particulate exposures, infectious disease/agents

Knowledge Gap

- Few effective treatment options exist other than mechanical intervention to limit further injury (Madershahian et al., 2007; Reilly et al., 2016).
- Given the high mortality rate, it is critical to identify at-risk individuals to prevent onset and progression of ARDS (Matthay et al., 2019).
- ARDS prophylaxis is key to force health protection strategies in modern warfare environments.
- Understanding the etiology of lung injury can inform:
  - Risk mitigation strategies (quarantine versus evacuation, personal protective equipment)
  - Prophylaxis strategies (antivirals, antibiotics, etc.)
  - Treatment approaches

Objective

- Perform a comprehensive review of published peer-reviewed scientific papers to collect biomarker data for ARDS resulting from various operationally relevant physiological stressors

Approach

- Use novel large language model (LLM)-enabled LitExplorer literature review pipeline (Figure 1)
  - Accelerate the traditional manual systematic review process by automating manual abstract and full text screening and data collection

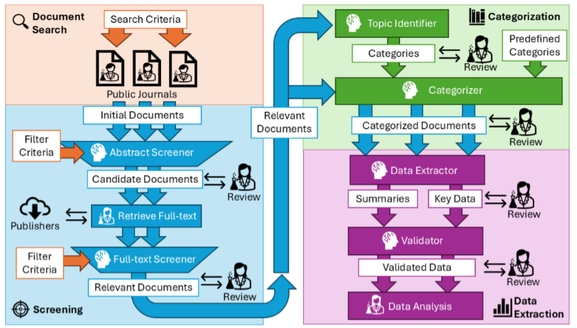

Figure 1. The LitExplorer Pipeline. The pipeline is comprised of 7 tasks: (1) initial article search; (2) abstract screening; (3) full-text retrieval and processing; (4) full-text screening; (5) biomarker data extraction; (6) validation of biomarker extraction; and (7) post-hoc data analysis.

Methods

Document Screening

- Corpus of 6,467 documents generated from a search of online scientific databases, primarily ScienceDirect and PubMed (Pulido et al., 2026)
- LLM-based Abstract Screener developed with Johns Hopkins University Applied Physics Laboratory (JHU APL) evaluated predefined inclusion and exclusion criteria
  - Designed to include human biomarker data of interest and to filter out non-relevant studies (Pulido et al., 2026)
- Screening results reviewed by subject matter experts to improve the LLM's performance
- Relevant full-text PDFs downloaded and processed using an Optical Character Recognition (OCR) engine (Google Document AI)
- Full-text screener (developed with JHU APL) used Retrieval-Augmented Generation (RAG) architecture employing more specific inclusion/exclusion criteria
  - Resulted in 901 ARDS-relevant papers (Pulido et al., 2026)

Biomarker Data Extraction

- Data extraction task identified biomarkers within each full text paper
- LLM extracted biomarker information: abbreviations, synonyms, comparison groups used in the study, statistical significance, direction of biomarker change (upregulation/ downregulation/unchanged/unclear), and quotes supporting each claim

LLM Validation Checks

- Validation checks were included to prevent instances of model hallucination or misinterpretation of data
- Each LLM-provided quote verified to ensure correct attribution in source text
- Each extracted biomarker claim validated with separate LLM prompt to verify explicit support by text in the specified paper

Identifying Top Biomarkers

- Top ARDS biomarkers selected based on frequency of mention in ARDS corpus
- Threshold of 20 mentions filtered the biomarker list
  - "Mention" = biomarker studied and showed statistically significant change in the paper
- 31 biomarkers identified

Categorizing Biomarkers by Health Effect

- Biomarkers manually grouped by standard online descriptions
- Health effect categories created for ARDS biomarkers
- LLM identified top ARDS biomarkers from corpus clustered into categories
- LLM grouped every biomarker in each paper into health effect endpoint categories:
  - Inflammation, blood-related markers, lung-related markers, infections, liver-related markers, kidney-related markers
- Biomarkers may be relevant to multiple aspects of ARDS pathophysiology
  - Categories designed to overlap
  - A given biomarker could be assigned to multiple categories

Defining Etiologies of ARDS

- LLM prompted to identify ARDS causes in each paper and group the study into major and minor categories (Figure 2)
  - Major categories: infection, inflammation, trauma, underlying comorbidities (e.g., diabetes, hypertension)
- Some papers gathered patients with ARDS from various causes
  - Each paper could be categorized into multiple etiologies

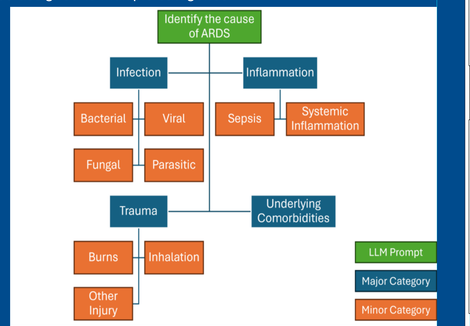

Figure 2. Flowchart Describing LLM Identification of ARDS Causes using Major and Minor Categories

Biological Matrix Sampling

- LLM prompted to determine how biomarkers in each study were measured
- Most biomarkers measured through blood/plasma
- Additional literature searches were performed to find non-invasive methods (saliva, urine)

Results

Table 1. Metrics for Top 31 ARDS Biomarkers. Number of Mentions, Changes in Biomarker Levels, and Categorization Heatmap Quantifying Biomarker Association with Target Categories. Heatmap ranges from 0 (yellow) for no biomarker association with a category to 1 (blue) for 100% association.

Name	Mentions	Down Regulated	No Significant Change	Unclear	Up Regulated	Inflammation	Blood	Lung	Infection	Liver	Kidney
C-Reactive Protein	226	25	36	22	143	1	0	0	0.053	0.004	0
Interleukin-6 (IL-6)	202	23	31	18	130	1	0	0.01	0.05	0.005	0
D-Dimer	137	12	14	15	96	0.066	1	0	0.015	0.007	0
Lactate Dehydrogenase	116	10	9	11	86	0.121	0.009	0.552	0.043	0.052	0
Procalcitonin	115	9	22	16	68	0.287	0	0.009	0.896	0	0
Lymphocyte	104	70	15	10	9	0.25	0.577	0	0.231	0	0
Interleukin-8 (IL-8)	96	11	18	4	63	1	0	0.01	0.042	0	0
Ferritin	80	14	9	11	46	0.963	0.025	0	0.088	0.013	0
Creatinine	72	6	15	4	47	0	0	0.014	0	0	0.972
Neutrophils	67	4	12	4	47	0.493	0.463	0.015	0.224	0	0
Albumin	66	44	8	4	10	0.152	0.015	0.182	0	0.182	0
White Blood Cell count	65	4	15	7	39	0.354	0.554	0	0.277	0	0
Surfactant Protein-D	64	7	16	5	36	0.094	0	0.984	0.031	0	0
Aspartate Aminotransferase	60	2	12	5	41	0	0	0	0	1	0
Platelets	58	30	16	6	6	0	0.983	0	0	0	0
Tumor Necrosis Factor-α	57	5	16	14	22	1	0	0	0.018	0	0
Interleukin-10 (IL-10)	53	3	10	5	35	1	0	0	0.038	0	0
Alanine Aminotransferase	50	0	15	5	30	0	0	0	0	1	0
Fibrinogen	47	5	9	4	28	0.106	1	0	0	0	0
Von Willebrand Factor	47	4	14	6	23	0.021	0.957	0	0.021	0.021	0
Angiopoietin-2	45	4	6	4	31	0.2	0.067	0.244	0.022	0	0
Lactate	42	4	5	4	29	0	0	0.167	0.19	0	0
Neutrophil to Lymphocyte ratio	39	3	4	1	31	0.821	0.231	0	0.128	0	0
s-RAGE	37	1	8	1	27	0.324	0	0.73	0.027	0	0
Hemoglobin	36	18	8	1	9	0	0.944	0	0	0	0
PaO2/FiO2 ratio	33	18	5	5	5	0	0	0.848	0	0	0
Plasminogen Activator Inhibitor 1	28	1	6	4	17	0.036	0.929	0.036	0.036	0	0
Krebs Von Den Lungen 6	27	0	6	3	18	0	0	1	0	0	0
Interleukin-1β (IL-1β)	25	2	7	3	13	1	0	0	0.08	0	0
RAGE	23	1	8	3	11	0.174	0	0.739	0	0	0
Total Bilirubin	22	2	5	4	11	0	0	0	0	0.955	0

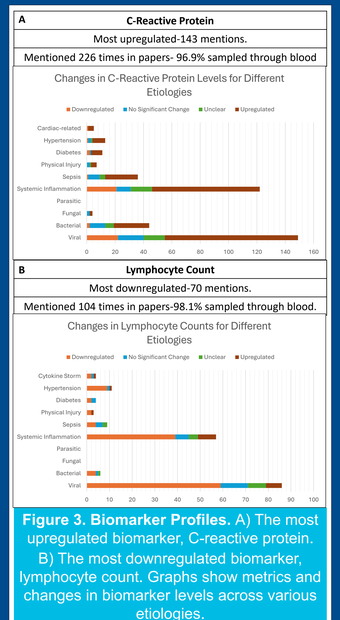

Figure 3. Biomarker Profiles. A) The most upregulated biomarker, C-reactive protein. B) The most downregulated biomarker, lymphocyte count. Graphs show metrics and changes in biomarker levels across various etiologies.

- Pipeline identified top 31 biomarkers
  - Biomarker levels and categorization based on health effects provided in Table 1
  - Top biomarker profiles shown in Figure 3
  - Changes in biomarkers by ARDS etiology shown in Figure 4

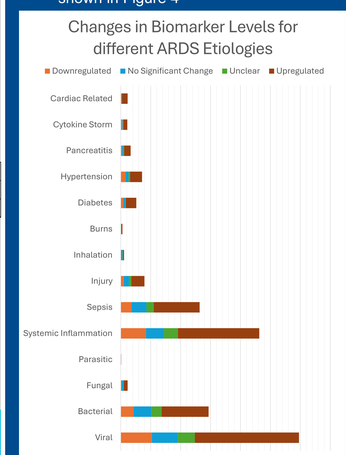

Figure 4. Changes in Biomarker Levels for Top 31 Biomarkers across Various ARDS Etiologies

Results

Biomarker Sampling

- Majority of biomarkers sampled through blood
- Alternative, non-invasive methods of sampling found for the following biomarkers:
  - Saliva (through Salimetrics kit): C-Reactive Protein, Interleukin-6 (IL-6), IL-8, Tumor Necrosis Factor-α, IL-10, IL-1β
  - Urine (through urinalysis): Lymphocyte, Creatinine, Neutrophils, Albumin, White Blood Cell count, Neutrophil to Lymphocyte ratio, Hemoglobin, Total Bilirubin

Conclusions

- **LLM pipeline accelerated and augmented the literature review process** compared to manual methods; reduced task time from months to days
- **Developed prioritized list of 31 robust ARDS biomarkers**, with data on direction of change, specific health effect endpoints, and disease etiology
- **A single biomarker will not definitively predict ARDS incidence**
  - A robust biomarker panel would include multiple blood measures along with physiological sensors.
- **Additional dose- and time-course data analysis is needed** to identify the optimal blood biomarker combination for predictive risk applications (identifying susceptible individuals or medical countermeasures for specific etiologies).
- ARDS is an on-going concern in prolonged and far-forward field care situations. **Development of a targeted multi-endpoint panel will allow improved survivability and mission success.**

Future Directions

- Biomarkers can be further narrowed down to those with greatest promise for reliable, robust use validation laboratory, field studies and/or clinical studies by analyzing dose-and time-course relationships using mathematical modeling approaches.
- Wearables hold promise for rapid identification of at-risk individuals through real-time biomarker and physiological monitoring.
- **The LLM-enabled LitExplorer pipeline provides an essential capability for rapidly transforming scientific literature into actionable insights for health service support and force health protection, while safeguarding the integrity of the data.**

References

1. Madershahian, N., Wittwer, T., Strauch, J., Franke, U. F., Wippermann, J., Kaluza, M., & Wahlers, T. (2007). Application of ECMO in multi-trauma patients with ARDS as rescue therapy. Journal of Cardiac Surgery, 22(3), 180-184.
2. Matthay, M. A., Zemans, R. L., Zimmerman, G. A., Arabi, Y. M., Beitler, J. R., Mercat, A., ... & Calfee, C. S. (2019). Acute respiratory distress syndrome. Nature Reviews Disease Primers, 5(1), 18.
3. Pulido, J.V., Marple, J.P., Gallagher, M., Ton, S., Bessling, S., Duffy, S., Floyd, T., Fragetta, K., Greene, A., Helfers, M., Hrybyk, A., Narayan, V., Ramirez, S., Klarup, K., Meehan, R., Sterner, T. R., Jaworek, T. J., Hirway, S. U., Cowan, D. W., Clewell, R.A. (2026). Automatic Search and Retrieval Tool (ASIRT): Applying Large Language Models to Enhance Systematic Literature Review. Air Force Research Laboratory, 711th Human Performance Wing, Human Effectiveness Directorate, Wright-Patterson Air Force Base OH. AFRL-RH-WP-TR-2026-0019.
4. Reilly, J. P., & Christie, J. D. (2016). Is it possible to prevent ARDS?. JAMA, 315(22), 2403-2405.

Acknowledgements / Disclaimer

We would like to thank the Johns Hopkins University Applied Physics Laboratory team (J. Vince Pulido, JP Marple, Molly Gallagher, Sarah Ton, Seneca Bessling, Summer Duffy, Taylor Floyd, Kendall Fragetta, Andrew Greene, Mika Helfers, Alex Hrybyk, Vijay Narayan, Sophia Ramirez, Kyle Klarup, Ryan Meehan) for their efforts in putting together the ARDS corpus, and helping with the abstract and full-text screening process.

The views expressed are those of the authors and do not reflect the official guidance or position of the United States Government, the Department of Defense, the United States Air Force, or the United States Space Force. The authors are contractors working with Air Force Research Laboratory/711th Human Performance Wing and the views expressed do not reflect the official guidance or position of BlueHalo, an AV Inc. Company; Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc.; Eagle Integrated Services, LLC.. Imagery in this document are property of the U.S. Air Force.