U.S. Department of Health and Human Services

National Institute of General Medical Sciences

Sepsis Human Biospecimens Investigators' Meeting

September 20, 2023 – September 21, 2023

Virtual Meeting

Meeting Report

Transcribed and drafted by:

Kimberly Rousseau and Xiaoli Zhao

Reviewed by:

Sepsis Human Biospecimens Investigators

Table of Contents

Executive Summary

The Sepsis Human Biospecimens Investigators' Meeting gathered National Institute of General Medical Sciences (NIGMS) grantees through funding opportunity announcement (FOA) PAR-21-077. The purpose of this FOA is to support the efficient collection, biobanking, and sharing of biospecimens and associated clinical data from critically ill and sepsis patients for use in future mechanistic research. It grew from the need to accelerate fundamental discoveries that provide novel insights into the heterogeneity of the pathogenesis and resolution of human sepsis, which has not been achieved by studies overly reliant on murine models of sepsis.

Throughout the meeting, investigators presented ongoing research funded by PAR-21-077; and discussed ideas, data, methods, and best practices for biorepository creation and use. Their advice to the sepsis research community is as follows:

1. There are strong benefits to using **remnant samples** including cost, feasibility, and consent. Areas to improve include reporting and normalization of the sample processing, and defined quality metrics specific for assays and scientific questions. Testing of new ways to analyze remnant samples (e.g., detecting microbial information) is a promising area to expand.
2. Novel sample storage approaches that are amenable to **emerging and future technologies** (e.g., whole-blood cryo, redox, dried blood) should be tested and incorporated into the biobanking process whenever possible. Assays for testing sample deterioration during storage would be helpful.
3. State-of-the-art technologies enabling advanced analysis using **small sample volumes** (e.g., microfluidics) will reduce the need for remnant samples and are useful for answering specific scientific questions. It is also useful to test the limits for certain advanced assays (e.g., scRNA seq, airway samples, metabolic assays) to inform the field.
4. **Clearly defined** ontology for biobanking and data repositories is the first step toward a standardized sample and data collection and processing protocol fitting specific analysis and scientific questions, which will ultimately improve resource sharing.
5. Broad biobanking with **thoughtful clinical annotation** (e.g., subgroups, timing of key events) is imperative for the effective utilization of biospecimens linked to clinical data.
6. It is important to set a reference for the **ground truth** (e.g., pre-analytic variability) before getting into the biological reality of the disease, which could be due to sample processing variability but more due to patient variability.
7. A combination of **automatic EHR screening** of patients and **retrospective syndrome adjudication** is found to be an effective workflow for prompt enrollment of critically ill patients at different disease stages without losing much fidelity.
8. It is important to identify the **proper control** groups (e.g., non-critically ill controls, non-infection controls, infection but non-septic controls) for every study and enroll those patients alongside study patients**.**
9. It is recommended to use **broad inclusion criteria** (e.g., all acutely and critically ill patients) to avoid excluding patients who do not fit neatly within the clinical definition of sepsis and who can also serve as controls. The analysis could start with the more defined patient group to serve as an anchor point and a reference.
10. **Paired samples** before and after treatment and **longitudinal specimen collection** may be good ways to overcome the vast heterogeneity that exists in the sepsis disease course. Time zero

samples would be highly valuable for finding sepsis signatures and diagnostic tools. It is of note that patient subgroups may change during longitudinal sampling and in future sepsis definitions.

11. **Pairing biospecimens with clinical metadata** should be a fundamental element when building a sepsis biorepository. The ultimate goal is to pair comprehensive longitudinal biospecimens with deep clinical phenotyping. It may be helpful to start by compiling a list of EHR elements and bedside assessments that provide high clinical value but do not add a significant burden to clinical care. Patient privacy must be adequately protected in this process.
12. There are opportunities to study **novel economic sample types** such as HME filters and urine because they are easily accessible and offer a wealth of information that may complement traditional sampling.
13. Quality assurance and data model standardization are important for the wide **sharing of EHR data** across institutions. A consensus on the types of variables to be standardized is helpful.
14. A wide range of consenting methods such as waivered consent and delayed consent are used, but local Institutional Review Board (IRB) committees have divergent interpretations of policies. It would be helpful to **publish more studies on consenting methods** to aid with the regulatory approval process. Studies to improve the rate of consent are also useful.
15. There should be a healthy balance between **open-ended discovery studies** and **hypothesis-driven studies** using a broad heterogenous biorepository or a well-defined cohort, as both are valuable to move the field forward in different ways. A top-down approach to finding the **biological signatures of clinical subgroups** or a bottom-up approach to identifying **shared underlying biological phenotypes** that explain the clinical heterogeneity both have values. A strategic way to pursue the **"biological truth**" could be layered to different depths of patient stratification.
16. There may be value in **a centralized approach** such as a structured database to deposit datasets from various sources and high-capacity computational tools, as well as biorepository centers that have the capacity and expertise in multicenter patient recruitment. It may also be beneficial to establish a searchable data and biorepository registry so investigators know about available resources for possible collaborations.

This meeting report summarizes investigators' presentations and the discussions surrounding **the current challenges in biobanking and sample utilization** facing sepsis researchers, and the **suggestions for overcoming them**.

Welcome, Introductions, and Opening Remarks

Jon Lorsch, Ph.D., Director, NIGMS
Rochelle M. Long, Ph.D., Division Director, NIGMS
Xiaoli Zhao, Ph.D., Program Director, NIGMS

Dr. Zhao, program director, National Institute of General Medical Sciences (NIGMS), opened the meeting and welcomed those on the call to the 2023 NIGMS Sepsis Human Biospecimens Investigators' Meeting. She introduced Dr. Lorsch, director, NIGMS; and Dr. Long, division director, NIGMS.

Dr. Lorsch explained the forward-looking vision set forth in the Notice of Funding Opportunities (NOFO) PAR-21-077. It originates from a 2019 report released by the National Advisory General Medical Sciences Council (NAGMSC) Sepsis Working Group, which examined the NIGMS sepsis portfolio and provided suggestions on how to accelerate research progress in the field. A key recommendation of the working group was to increase the emphasis on the use of human biospecimens and clinical data from critically ill patients over studies in animals—in particular, the cecal ligation and puncture model. In response, NIGMS released the notice NIGMS Priorities for Sepsis Research, highlighting the Institute's program priorities consistent with the NAGMSC report. To promote studies using human biospecimens and human data, PAR-21-077 was published to invest in technologies to best analyze existing human biospecimens collections and/or new methods to improve the collection, storage, and dissemination of such samples. Dr. Lorsch envisions that this effort will increase the pace of progress in the sepsis field. Furthermore, this program laid the foundation for the nascent collaboration between NIGMS and the National Heart, Lung, and Blood Institute (NHLBI) in the Acute Respiratory Distress Syndrome (ARDS), Pneumonia, and Sepsis Phenotyping (APS) Consortium, a large multicenter, longitudinal, observational study of critically ill patients. NIGMS' emphasis in this collaboration is the generation of a collection of biospecimens and clinical data from ARDS, pneumonia, and sepsis patients. Dr. Lorsch pointed out the synergy between the sepsis biospecimens program and the APS consortium and hopes that this program sets the stage for the rapid progress of the newly funded APS consortium. He is looking forward to hearing the investigators' presentations on current progress and future research directions.

Dr. Long then explained the purpose of PAR-21-077 and the charge of this sepsis biospecimens investigator meeting. The program is to support the efficient collection, biobanking, and sharing of biospecimens and associated clinical data from critically ill patients for use in future mechanistic research on sepsis. The NOFO has two main goals:

1. Determine the scientific value of existing or newly collected sepsis human biospecimen sets as testbeds for future studies on sepsis.
2. Set guidance on approaches for collecting, utilizing, and analyzing human biospecimens in order to maximize their value to the entire sepsis research community.

Dr. Long explained that this funding mechanism is broken into two phases. In the R21 phase, investigators test new methods for biospecimen acquisition, storage, and dissemination, as well as the analysis of new or existing biospecimens. The R33 phase will focus on scale-up activities.

Throughout this initiative and during both phases of the award, funded investigators are asked to participate in periodic meetings like this one and to present their scientific progress and to share their data, methods, protocols, quality standards, and scientific results. Dr. Long stated that the ultimate goal of these collective activities is to set best practices for future human biospecimen collection in order to

optimize their experimental utility and their scientific value. Over the next few days, all meeting participants will aid in setting the best practice for the future, which will inform other researchers, networks, and consortiums, and ultimately, accelerate discoveries. Through united action, we will make forward strides in understanding the development, heterogeneity, and resolution of sepsis, as well as its treatment. Dr. Long is looking forward to a productive meeting.

Dr. Zhao thanked NIGMS leadership for supporting this important program, Dr. Della White for organizing this meeting, and the Information Resources Management Branch for IT support. Dr. White is the NIGMS clinical research strategy coordinator at the Office of the Director, and she is responsible for the coordination of the sepsis biospecimens investigator meeting. Dr. Zhao explained that this is a closed meeting, and attendees include funded sepsis biospecimens investigators, selective APS investigators, and relevant National Institutes of Health (NIH) staff.

Objectives and Charge to Presenters and Participants

Meeting Co-Chairs:

Julie A. Bastarache, M.D., Vanderbilt University Medical Center (VUMC)
D. Clark Files, M.D., Wake Forest University School of Medicine

Dr. Zhao introduced the two meeting co-chairs: Dr. Bastarache from Vanderbilt University Medical Center, and Dr. Files from Wake Forest University School of Medicine.

Dr. Bastarache thanked the NIGMS leadership and program staff for recognizing and investing in an important gap in the research area. She commented that what makes this meeting truly exciting is the common purpose of funded investigators and meeting attendees in making clinically impactful headway on sepsis research. Being an investigator both in the sepsis biospecimens program and the APS consortium, she felt that important discoveries are at the cusp under the synergy of these two programs. She explained that complexity and heterogeneity are the hallmarks of sepsis and that the ongoing effort of creating biorepositories is a cornerstone to increased understanding of the patient-centered clinical and biological heterogeneity of the sepsis syndrome.

During this 2-day meeting, there will be presentations from grantees on their projects, a Q&A period following each session, and ample time for discussion led by the session and meeting chairs. Dr. Bastarache explained that the goal of this meeting is to come together and participate in vigorous discussions on the best ways to create and use sepsis biorepositories and to share the ideas, data, methods, and best practices that the investigators have developed to make important contributions to the field. She also hopes that this meeting stimulates transformative collaborations among cross-disciplinary investigators to conduct impactful sepsis research. She urged all meeting attendees to speak up and share ideas and reminded them to introduce themselves when speaking. She introduced herself as an MPI of a funded sepsis biospecimens R21 (Day 1) with Dr. Lorraine B. Ware.

Dr. Files' remarks focused on the mechanics of the meeting. He thanked the organizers of the meeting for the opportunity to learn from one another in an effort to move the field forward. He reviewed the meeting agenda and pointed out that the 2-day meeting is divided into three scientific sessions. He will start each session by introducing the session chairs, who will then moderate each session. Each presentation is allocated 20 minutes, including 5 minutes of Q&A. The session chair will introduce each speaker and moderate the Zoom chat while the meeting chair will keep track of the discussion time. At the end of each session, there is a 30-minute panel discussion led by the session chair to summarize each session and identify common themes of the session. Each day of the 2-day meeting will conclude with a closing remark from the meeting chairs.

Dr. Files highlighted a few overarching discussion points for the Facilitated Discussion session (held on Day 2, following the three scientific sessions). He asked the meeting attendees to think about how to contribute to these discussion points throughout the course of the meeting and make this session as interactive as possible. The facilitated discussion should bring all three scientific sessions together, solidify what was learned in the 2-day meeting, and focus on creating a good concept to push the sepsis field forward, both in clinical research and clinical care:

1. Discuss how the sample collection/processing/analysis process might be designed to enable future improvements in:
    a. The diagnosis and differential classification of sepsis stages
    b. The understanding of the temporal development and heterogeneity of sepsis syndrome
    c. Mechanistic studies of sepsis that subsequent investigations could pursue
2. Address challenges with biospecimen collection for research, and highlight solutions/opportunities
3. Summarize implications for active and future observational trials involving the collection of human sepsis biospecimens

Dr. Files introduced himself. His research interest is muscle wasting in critically ill patients and sepsis, and he is an MPI on a sepsis biospecimens R21 with Dr. Cristina M. Furdui (Day 2).

Session I: Novel Methods in Patient Recruitment and Sample Collection

Session Co-Chairs:
Annette Esper, M.D., M.Sc., Emory University School of Medicine
Lorraine B. Ware, M.D., Vanderbilt University Medical Center

Dr. Files introduced the two Session I co-chairs: Dr. Esper from Emory University School of Medicine, and Dr. Ware from Vanderbilt University Medical Center. Dr. Esper introduced the speakers, and Dr. Ware moderated the chat for discussion. Dr. Esper is an MPI on a funded sepsis biospecimen R21 with Dr. Rishikesan Kamaleswaran.

REMISE Study: REMnant Biospecimen Investigation in Sepsis

Christopher W. Seymour, M.D., M.Sc., University of Pittsburgh School of Medicine

Dr. Seymour introduced himself, saying that he hoped his presentation would set the stage for a great meeting. He planned to talk about why he focuses on remnant biospecimens, present preliminary data and progress during the R21 (as well as challenges and opportunities), and plans for the R33 phase. He commented that many meeting attendees have years of experience studying biospecimens, which is also a focus of his own lab. He said that he is looking forward to learning and discussing this topic.

Dr. Seymour introduced the complexity of the body's response to pattern associated molecular patterns (PAMPs) and danger associated molecular patterns (DAMPs) by displaying a classic piece of data showing the gene expression pattern after endotoxin infusion (Calvano, *Nature*, 2005). This is challenging for clinical researchers and trialists trying to tease out treatment effects in this tricky biological milieu. Many researchers focus on one of the many molecules contributing to the pathogenesis of sepsis and septic shock. Biospecimens could come in handy to break this silo because they collectively capture the complexity and heterogeneity contained in this biological milieu.

One proposed solution to the challenge of sepsis heterogeneity is to find subtypes. At this point, hundreds of classes of sepsis have been proposed, but they were developed for different purposes, using different data sources and specimens, and thus have different acceptability levels and generalizability (DeMerle K. et al., *JAMA*, 2021). A unifying theme is yet to come. Meanwhile, the conceptual model for sepsis continues to evolve, including both a host (tolerance vs. resistance) and pathogen component, manifesting as organ dysfunction in sepsis and affecting the treatment pathways of sepsis. The utilization of biospecimens offers a chance to merge these conceptual models with the biology of sepsis. Furthermore, the complexity and heterogeneity contributed to the failure of sepsis trials such as the ACCESS and the PROWESS-Shock, and the use of biorepositories is one important step toward hope in future clinical trials. However, Dr. Seymour pointed out that even the use of biospecimens brings its own heterogeneity, which set the stage for the REMISE study.

A recent paper (Brant EB et al. *NPJ Digit Med*. 2022) performed a meta-analysis on the use of artificial intelligence (AI) tools to identify sepsis from electronic health records (EHRs). Moving from a Sepsis 1 to Sepsis 3 definition, fewer variables and features were used to find sepsis patients, but the heat map indicated that these variables are used in a different way—that is, different criteria were implemented differently, and a mixed population of patients was labeled as having sepsis.

Dr. Seymour introduced remnant biospecimens, which are the remaining volumes of any human whole-blood, plasma, serum, or urine sample acquired for clinical care. About 80% of those samples were discarded after clinical tests. In a recent paper (Demerle K et al. *JAMA Network Open*, 2021), Dr. Seymour and colleagues used a virtual alert system embedded in the EHR to screen for SOFA scores and collect remnant samples from the University of Pittsburgh Medical Center (UPMC) clinical laboratory. They enrolled 1,000 patients in just 9 months.

Building on the feasibility, the main question in Dr. Seymour's R21 proposal was determining the value of remnant biospecimens in understanding sepsis heterogeneity through multi-omics. Aim 1 was to look at a multi-omic readout from remnant samples to determine the precision, accuracy, and integrity, as compared to research-grade fresh samples. The workflow goes as follows (image prepared by Biorender): 50 sepsis-enriched patients were identified within 6h of ICU followed by immediate clinical judication. Clinical samples were drawn and underwent a mandated holding period of 48 hours at 4°C in the clinical lab, and then the research team could take the samples as remnants for analysis. At the same time, the clinical coordinator approaches the same patient at the bedside and collects prospective fresh samples under informed consent. The remnant and fresh samples were compared by Dr. Seymour and collaborators for precision and accuracy in a number of analytic domains including inflammatory and endothelial biomarkers, plasma proteomics, metabolomics profiling, lipidomic, and pathogen sequencing. This protocol was recently published (Seymour et al. *Critical Care Explorations*, 2023).

Dr. Seymour presented the progress of the R21 study. The IRB was approved in 2022, with more than 8,000 patients screened at four sites using EHR alerts. Of these, 633 were manually adjudicated to be sepsis, and 47 patients consented to paired fresh and remnant samples. Most of the enrolled patients had a SOFA score of 2, but there were some patients who were more ill with a SOFA score of up to 5. The patient population is representative, given that this is a small cohort. In general, fresh samples were processed and stored at -80°C for under 2 hours, while remnant samples had a clinical testing time followed by a mandatory holding time at 4°C before being put into -80°C, which is the main difference between the life cycle of the two. More volume was collected for fresh samples in both the EDTA tube (8.8 vs. 2.6 ml) and Heparin tube (5.4 vs. 3.7 ml), and thus there were more fresh sample aliquots of >100 µl. There was very little hemolysis in both types of samples (<1%), with 37 pairs of samples being analyzed.

Preliminary results of interleukin-6 (IL-6) and angiopoietin-2 (ang-2) showed that although the average IL-6 level was higher in remnant samples than those collected at the bedside, Scatter Plot, Spaghetti Plot, and Bland-Altman Plot suggest good agreement between the two types of samples. Dr. Seymour believes that the lower inflammatory marker level in the fresh sample group may be due to delayed sample collection time waiting for informed consent when the markers had subsided. Proteomic results (chromatograms, protein mass) by collaborator Dr. Renee Robinson (Vanderbilt) showed minimal differences between sample types. Volcano plot showed that among the more than 1,200 identified proteins, only 40 were distinct between the two sample types. Notably, when mapping back to the results of a prior study, 71% of REMISE identifications were also identified in the ProCESS study.

The difference was in metabolomics. Both negative and positive polar metabolites in the remanent are quite different from those in the fresh samples with substantial variations. To date, most of the milestones of the R21 phase have been completed, except for pathogen sequencing due to the low

microbial DNA load in these samples. Dr. Seymour is collaborating with Dr. Robert Dickson (University of Michigan) to get this work off the ground.

Speaking of challenges for this study, it is difficult to coordinate multisites across three institutions, and there were concerns about the feasibility of pathogen sequencing in the remnant. Clinical judication of patients is accurate but slow so he is looking into AI tools to speed up the process for future scaling. Also, metabolomic and lipidomic studies are resource intensive but less accurate assays. Additionally, manual patient adjudication limits study scaling.

In the R33 phase, Dr. Seymour will scale up a mechanism-informed, remnant biospecimen repository featuring community hospitals, diverse samples, and integrated multi-omics analysis in collaboration with Dr. Timothy R. Billiar (University of Pittsburgh) and Dr. Nuala J. Meyer (Penn).

In conclusion, Dr. Seymour stated his belief that remnant biospecimens do have adequate integrity for analyses. Detection of protein biomarkers is accurate and precise but seems to be assaying an early time point during disease progression. There is more uncertainty with metabolomic and lipidomic analyses using remnant biospecimens, and pathogen sequencing will depend on microbial burden.

Dr. Ware asked if the difference in biomarkers (IL-6) is due to the timing of collection or different storage conditions. She suggests doing a pseudo-remanent control, in that blood is prospectively drawn but stored at the remnant versus fresh sample conditions. Dr. Seymour explained that he had not thought about simulating the remnant storage condition but could potentially look at the difference in the time of collection in Dr. Meyer's cohort, which has some natural variability from Dr. Seymour's group. A detailed analysis of the time taken to collect remnant versus fresh samples could have answered the above question.

Dr. Steven Patrie (Northwestern University) wondered if the difference between proteomics and metabolomics detection is because proteomics could detect post-translationally modified particularly truncated proteins due to active proteolysis during storage but won't be able to tell. Dr. Seymour recalled that phosphorylated protein levels are markedly different in the metabolomics versus proteomic studies, supporting this idea. He might be able to check with Dr. Robinson during the R33 phase.

In the chat, Dr. Wesley Self (Vanderbilt) noted that the APS Consortium has discussed the use of remnant biospecimens, but a lack of systematic collection and missing data have been inhibitive. Dr. Seymour responded by saying that using remnants is cost saving, and he suspects that the majority of inflammatory/endothelial biomarkers can be accurately measured in remnants. Dr. Zhao asked if remnant biospecimens are available during multiple time points of the disease course, making longitudinal studies possible. Dr. Seymour responded that remnant samples are available one to two times daily in the ICU, and trajectory work is possible. Dr. Nate Shapiro (Beth Israel Deaconess Medical Center) asked to clarity the time of the blood draw. Dr. Seymour stated that the remnant and fresh pair are drawn within 12 hours of each other, but the remnant is lab mandated to sit at -4°C for 48 hours.

The Sepsis ClinicAl Resource and Biorepository (SCARAB) Project

Julie A. Bastarache, M.D., Vanderbilt University Medical Center

Dr. Bastarache introduced the SCARAB project, a partnership between the adult (Vanderbilt) and pediatric (Meharry) intensive care units (ICUs), which addresses a few challenges in sepsis research.

The first one is the inability to noninvasively sample the distal airspace, particularly for studying sepsis-induced ARDS. The SCARAB study collects the fluid within the usually discarded heat and moisture exchanger (HME) filters collected from the ventilator circuit from intubated patients, which is reflective of the biology of the distal airspace. The second one is a biased enrollment of critically ill patients due to the time to consent and the types of patients available for consent. SCARAB uses a waiver of timely informed consent, which allows for patient enrollment without informed consent initially for minimal risk studies until time allows, enabling a broader and more equitable inclusion of patients. A third challenge is the difficulty and ambiguity of real-time patient phenotyping, as mentioned by Dr. Seymour. SCARAB enrolls all critically ill patients first and then performs post hoc phenotyping. This method increases the fidelity of the phenotyping because more information can be collected and used to make decisions; this also creates a control group. A fourth challenge is the reliance on labor-intensive data collection. SCARAB uses cutting-edge EHR phenotyping algorithms, leveraging the bioinformatics resources available at Vanderbilt. The fifth challenge is a siloed approach when studying organ dysfunction, so SCARAB collects samples and data to facilitate research on multiple organs. Finally, to address limited access to sepsis subjects for research, in the R33 expansion phase, SCARAB will develop an interactive web portal where researchers can access clinical information and/or request data specimens for use.

The goals of SCARAB are to:

- Create a two-center prospective biorepository of adult and pediatric patients with sepsis and critically ill controls
- Enroll a diverse population of patients across the age spectrum
- Collect serial biologic specimens with a broad and forward-thinking approach to facilitate future studies
- Develop automated clinical phenotyping algorithms for sepsis, ARDS, acute kidney injury (AKI), and delirium
- Assemble an external advisory board to review and approve applications to access data and use samples

SCARAB's sample collection protocol is designed to facilitate a broad array of studies. The team collects DNA for genetic studies such as candidate-gene, whole-exome sequencing, and genome-wide association study (GWAS); peripheral blood for candidate gene, miRNA, and RNAseq; plasma, HME filter fluid, and tracheal aspirates for candidate biomarkers and other discovery studies; and EHR for bioinformatics studies. SCARAB is also able to study environmental exposures and health disparities, and calculate the area of deprivation index (ADI). Urine is also collected for biomarker studies of AKI.

The inclusion criteria for SCARAB are broad: All critically ill patients aged 6 months and older admitted to the medical, surgical, or pediatric ICUs at the two sites. The exclusion criteria reflect the goal of recruiting patients early and targeted, including more-than-24-hour ICU stays prior to enrollment, imminent death, uncomplicated drug overdoses, alcohol withdrawal, gastrointestinal bleeding, diabetic ketoacidosis, and patients admitted to the ICU solely for frequency of nursing care.

Dr. Bastarache described the study timeline. The study team screens patients in the Emergency Department (ED) when they get an order for ICU admission. Upon enrollment (Day 0), the study team collects whole blood, DNA, RNA, HME filter fluid, tracheal aspirate, and urine. Blood draw is repeated on

ICU Day 2 and ICU Day 7. Daily HME filter fluid and tracheal aspirate are collected from intubated patients, and urine is collected in fully catheterized patients. Prospective clinical data are collected continually for the further development of automated EHR phenotyping. To date, SCARAB has enrolled 127 adult and 19 pediatric patients from VUMC, and 8 subjects from Meharry. The number of samples collected is pretty good, with an expected decline from Day 0 to 7. Almost all patients have EDTA and sodium citrate blood samples, serum, and PAXGene samples, half with urine samples, and 8% to 16% with HME and tracheal aspirate samples.

Dr. Bastarache described how SCARAB leverages EHR for rapid phenotyping in the ICU. In a traditional prospective format, eligible patients are identified upon ICU admission and then enrolled, samples and data are collected by trained study nurses, and then investigators adjudicate phenotypes and clinical outcomes; this strategy leads to accurate phenotyping but is labor intensive and not scalable. In the EHR phenotyping method, the EHR software finds patients and extracts both structured and unstructured data (diagnosis codes, laboratory test results, clinician notes, and imaging reports); this can be used to develop phenotyping algorithms that can be used to classify ARDS, sepsis, delirium, and AKI. Preliminary data showed that the HER-ARDS Classifier had great sensitivity and specificity for identifying investigator-adjudicated ARDS. On the other hand, the EHR-sepsis classifier had less than 0.5 sensitivity using "explicit (ICD9)" sepsis diagnosis code, and 0.67 sensitivity using "implicit sepsis codes (infection, organ failure)," while the combination of the two improved performance, so future refinement of the phenotyping algorithm includes adding additional layers of classification by NLP and AL tools.

Dr. Bastarache noted the broad impact of SCARAB by supporting other studies, including Dr. Alicia Rizzo's project (clinical fellow at Massachusetts General Hospital [MGH]) studying sex differences in critical illness using HME filter fluid, Dr. Matt Stier's project (clinical fellow at VUMC) on the metabolic profiles of immune cell subsets in critical illness, and X-zayver Smith's future project (VUMC medical student) visualizing leukocyte-endothelial interactions on skin in SCARAB patients. Alicia (Rizzo AN, et al. JCI Insight. 2022. PMID: 34874923) found that injury of lung epithelium led to glycocalyx shedding and increased glycosaminoglycan (GAG) in the HMEF, while ARDS males have much greater GAG in HMEF than females (minimal level of shedding). This is associated with severe hypoxemia and a longer duration of mechanical ventilation. Dr. Stier's unpublished single-cell metabolic data (SCENITH) showed that critically ill sepsis patients have different immune cell populations and distinct metabolic cell clusters as compared to nonseptic critically ill patients and healthy controls.

Dr. Bastarache highlighted some of the early challenges that the SCARAB study team faced. Initially, the team used a fully informed consent protocol prior to enrollment and an additional blood draw on Day 4; this led to lower-than-anticipated enrollment and introduced a bias toward patients with a family member present to give consent. As a result, the team changed the consent process to a waiver of initial timely informed consent with a provision to keep samples/data if consent was never obtained, a procedure that is adopted by the APS consortium. This, plus added weekend coverage, has improved enrollment.

Dr. Bastarache also highlighted some insights. Instituting broad enrollment criteria has led to the development of a more representative patient population and a stronger critically ill control group to study syndrome overlaps. Delayed phenotyping has allowed for a cleaner cohort and is more efficient. SCARAB could support other groundbreaking studies, even at the R21 phase. Study nurses and coordinators are essential for success in particularly building a high-quality biorepository. Dr. Bastarache

noted some innovations of SCARAB. The use of longitudinal HMEF for studying the distal airspace is a game changer for her team. Enrolling adult and pediatric patients in the same protocol will allow for interesting comparisons between age groups. The future interactive web portal and the to-be-refined EHR phenotyping algorithms are also expected to be cutting-edge and impactful. Dr. Bastarache also briefly showed additional opportunities for innovation in EHR phenotyping, including NLP mining of unstructured notes, leveraging clinic orders and laboratory tests, and incorporating the OpenAI GPT provided by Microsoft Azure to VUMC researchers (Generative AI). In addition, the plan for the R33 phase includes increased enrollment, refining phenotyping algorithms, implementing de-identification procedures, creating a web portal, and advertising and soliciting projects.

Dr. Chaz Langelier (University of California San Francisco [UCSF]) asked how radiographic data were used for ARDS adjudication. She explained that X-rays are pulled and reviewed independently by two physicians to identify bilateral infiltrates and/or pulmonary edema (RALE score). Vanderbilt is developing the capability to de-identify radiographs and EHR data, both retrospectively and prospectively, which will enable sharing with outside investigators. Dr. Ware noted that the team is using a gold-standard ARDS cohort to develop NLP text mining of radiographic reports, which was not incorporated into the EHR-ARDS algorithms but can be validated in the scale-up phase.

Dr. Andrew Lautz (pediatrician, Cincinnati Children's Hospital) asked if the SCARAB team ever ran into limitations due to sample volumes for certain assays when comparing pediatric patients to adult patients. Dr. Bastarache stated that they have not performed any analyses on pediatric patients, but they are getting white blood cells, DNA, RNA, and plasma, although less in volume. She believes that it won't be a severe limiting factor because due to advances in technologies, more can be done with smaller volumes.

Dr. Leopoldo Segal noted in the chat that it would be interesting to compare HME filter fluid to tracheal aspirate and bronchoalveolar lavage (BAL) samples because it is possible that these airway sampling methods will capture various biomarkers with different precision. Dr. Bastarache noted that the team has compared HMEF directly to aspirated pulmonary edema fluid with minimal difference but has not done a comparison to BAL samples.

Dr. Kamaleswaran asked if the NLP algorithm for radiographs has been published, and said that he found it challenging to generalize the lexicon across different radiologists due to preferences for certain terms. Dr. Bastarache responded that it is still in development and has not yet been published.

Developing a Scalable, Multicenter Pediatric Sepsis Biorepository and Clinical Database

L. Nelson Sanchez-Pinto, M.D., M.B.I., Ann & Robert H. Lurie Children's Hospital of Chicago (Lurie) and Northwestern University Feinberg School of Medicine

Dr. Sanchez-Pinto introduced his presentation, which nicely followed Dr. Lautz's question about the low volume of pediatric samples. Dr. Sanchez-Pinto is an MPI with Dr. Fran Balamuth (Children's Hospital of Philadelphia [CHOP]) on this funded R21, which has two parts: a biorepository component (led by Dr. Balamuth), and a clinical database component (led by Dr. Sanchez-Pinto). Adding to the complexity of sepsis itself, pediatric sepsis is extremely heterogeneous across the life span, from infants to adolescents. This project is aimed at studying sepsis phenotypes in children at scale. As an emergency physician and an intensivist, Dr. Sanchez-Pinto and Dr. Balamuth are interested in the concept of phenotyping patients around resuscitation in order to inform post-resuscitation therapies. Patients are flagged for suspected

sepsis in the emergency department (ED), and samples are collected pre- and post-resuscitation (standard quality improvement bundle type of resuscitation). Through endotyping, the patients are grouped based on their biological underpins and thus will receive more targeted therapy in the ICU.

Dr. Sanchez-Pinto described some of the rationale for this study design. Pediatric sepsis studies have to be scalable and multicenter because the total number of pediatric ICU beds in the U.S. is a magnitude less than adult ICU beds and thus more spread out, so more sites are needed to collect sufficient sample numbers. The study was also designed to be low cost, making it more feasible at sites with fewer resources (i.e., no 24/7 clinical coordinator coverage). The study will answer questions about the "Goldilocks" blood volumes to be collected in pediatric patients—that is, the safe amount of blood draw that can be used to explore the patients' immune and metabolic heterogeneity using cutting-edge methods. Finally, the project aims to develop a linked clinical database that captures the clinical heterogeneity of patients using EHR data in order to identify clinical signatures, which is also scalable due to the low cost.

In the first aim of the R21 phase of this project, the team will compare smaller versus larger volumes of blood draws during pre-resuscitation to test if the smaller volume is sufficient to perform the assays of interest and if the delayed processing of these samples affects data quality. Two more serial samples during the post-resuscitation period will also be collected. Aim 2 will focus on the feasibility of developing a standardized multicenter protocol for EHR extraction, transformation, and loading into a common data module to generate a rich EHR clinical database linked to samples. Aims 3 and 4 constitute the R33 phase of the project, which will expand sample collection with the right volume and processing timing based on the findings from Aim 1, and develop open-source tools for the expansion of this infrastructure to be tested at three pilot sites (CHOP, Lurie, Children's Hospital Colorado) with the idea of even larger scale expansion in the future.

Dr. Sanchez-Pinto described the workflow. Patients presented at the ED go through semi-standardized EHR triggers. A sepsis yellow trigger indicates suspected sepsis, and patients will be monitored for 3 hours before the initiation of treatment (antibiotics, etc.). A sepsis red trigger indicates septic shock, and patients will receive sepsis bundles (resuscitation, etc.) within 1 hour. Patients in these two pathways were enrolled, and EHR data were collected. These triggers also tag the patient for a blood draw for standard clinical tests before any fluids or antibiotics are given via IV. An additional volume is collected at this time for the current study (pre-resuscitation sample) under delayed consent or exception from informed consent (EFIC); if the parents do not consent to study participation, this sample is discarded.

Under Aim 1, collected samples will undergo three analyses: from basic analysis of plasma O-link proteomics to more state-of-the-art analysis such as peripheral blood mononuclear cells (PBMC) single-cell SCENITH metabolomics and flow cytometry immune profiling. Aim 1 will provide evidence to decide the optimal volume (2 mL or 5 mL) and time to processing (3 hours—immediate processing, 12 hours—overnight collection and local daytime processing, or 48 hours—shipping and central processing) to perform these assays. The time to processing may impact how scalable this is, as not all sites have the capacity for local processing or immediate processing. The same patient sample was used for the testing of different conditions to allow for comparison.

So far, the team has collected samples from seven patients. Little differences in protein expression or immune profiling were found between 2 ml and 5 ml. However, 2 ml is not sufficient for SCENITH metabolic assay for most patients, mainly due to an insufficient number of cells running the single-cell

assay, but in case enough cells were isolated from 2 ml samples, SCENITH showed a similar pattern as the 5 ml samples. As for time to processing, 3 hours versus 12 hours yields very similar immune profiling, proteomic, and metabolomics results, but after 48 hours there is significant variability in all assay types.

Moving forward with prospectively collected samples, the team is collecting samples 24 hours a day, Sunday evening through Friday midday. The team collects 5 mL samples but will accept down to 2 mL samples. Samples will be stored at 4°C and processed within 12 hours of collection. These samples are pre-resuscitation, and post-resuscitation on Days 2 to 3 and Days 5 to 7.

Dr. Sanchez-Pinto briefly introduced the clinical database quality assurance and harmonization pipeline. They are using a modified Observational Medical Outcomes Partnership (OMOP) data model for collecting EHR data and quality assurance, which allows them to leverage existing standardized tools from an organization such as Observational Health Data Sciences and Informatics (OHDSI) that uses OMOP models. They have incorporated a robust framework of quality assurance for data completeness, conformity, and plausibility and are doing data validation and verification by domain and by site, a key component for high-quality EHR data for analysis. To date, 665 patients have been enrolled into the EHR database; 13 have linked blood specimens (the first batch from two sites). Altogether, these patients have 503,371 unique values that have been mapped to 186 unique clinical measurements, 76 unique Logical Observation Identifiers Names and Codes (LOINC) codes, and 46 unique clinical concepts for organ dysfunction. The structured data quality assurance report (indicating high-quality data) will be shared using the OMOP data model via open source. Two abstracts have been submitted to the Society of Critical Care Medicine (January 2024) reporting findings from Aims 1 and 2.

For the R33 phase, the study will be expanded to include specimen collection at three sites at three time points with the 12-hour processing. It will also test the infrastructure by comparing assay results in patients stratified by data-driven organ dysfunction–based phenotype. Finally, in the R33 phase, the team will develop an open-source biorepository cohort discovery and exploration tool, providing data quality assurance and piloting cloud-based analysis.

Dr. Ware asked for clarification about storage conditions of the 48-hour time to process samples. Dr. Sanchez-Pinto confirmed that the samples were centrifuged and then stored at 4°C.

Dr. Ware asked if the OMOP database is easily transferable to other sites or if it needs to be rebuilt after transferring. Dr. Sanchez-Pinto explained that the OMOP Common Data Model has become more popular, so many sites have some version of the database already in use. However, Dr. Sanchez-Pinto noted that this system does have some limitations on incorporating ICU data such as infusion rates. They did keep some custom databases that contained more granular or source data and used the OMOP database for analysis of structured data such as organ function scores. The advantage of OMOP is built-in quality assurance tools. For example, you can compare the data characters of a new batch of data to the earlier batch, making the database more generalizable at the expense of the more granular ICU data. Dr. Bastarache noted in the chat that many big consortia, such as All of Us, have data in the OMOP format.

Dr. Segal noted that head-to-head comparisons of protocol details are rarely published, but it is important information for the research community. In addition, he noted that different analytes or different assays may be more or less sensitive to these details—for example, RNA/DNA preservation may be more sensitive to the timing. Implementation of these considerations up front will make the sampling procedure more scalable. Dr. Sanchez-Pinto agreed and explained that they chose these three types of

analyses because they believe that these are the more representative assays for the current sepsis studies and are beneficial for pediatric sepsis phenotyping. Perhaps a good direction is to study how to make these assays more tolerant in different practical storage and processing conditions.

A Systems Approach to Predicting and Classifying Neonatal Sepsis Using Biospecimens and Clinical Data

Stephanie Prescott, Ph.D., A.P.R.N., N.N.P.-B.C., University of South Florida and Inova Children's Hospital

Dr. Prescott introduced herself as a neonatal nurse practitioner, and highlighted that a challenge in studying sepsis is the lack of biospecimens collected prior to sepsis onset. In addition, research efforts in neonates are hindered by small blood volumes available for study. However, one advantage is that neonates are kept in the hospital from birth, allowing for pre-sepsis sample collection.

In the R21 phase of this grant, the study team plans to develop optimal sample collection protocols, sample preservation strategies, and archival storage practices to ensure rigor and test the quality, reliability, utility, and reproducibility of these samples. Additionally, the team will collect demographic variables and clinical data prior to, during, and after culture-positive sepsis to determine the optimal time to detect biomarkers predictive of neonatal sepsis. In the R33 phase, the team will construct multi-omic networks to develop predictive models for sepsis and sepsis recovery. Dr. Prescott introduced the research team, which includes expertise in patient recruitment, neonatology, systems biology, microbiome, proteomics, and stool metabolomics.

In this prospective observational study, the team will collect blood, stool, saliva, urine, skin swabs, feeding samples, mother's breast milk (feeding), pregnancy history, and clinical data on Days 1 and 3 of life, and weekly thereafter through 7 weeks. The inclusion criteria for enrollment are neonates born ≤ 32 weeks of gestation or with a birth weight of ≤ 1,500 grams, and with no genetic abnormalities, one-third of whom will develop sepsis. Control groups include "normal" growing preemies and critically ill preemies with no culture-positive sepsis, which will be compared to preemies diagnosed with culture-proven sepsis. An advantage of the study is the collection of samples and data before, during, and after the development of sepsis. The main goals of the study are to improve biospecimen collection, storage, and analysis to increase rigor and reproducibility in neonate sepsis research; and to integrate clinical, demographic, and biological data to predict and classify neonatal sepsis.

By the end of the 4-year study period, the team will enroll 168 mother/baby dyads and will collect more than 10,000 samples. Because these babies remain in the clinic for the duration of the study period, stool samples are readily available, but fresh blood samples are limited to 500 µl at each time point. The team is collecting remnant plasma samples from the clinical laboratory multiple times per day to overcome this challenge.

Dr. Prescott highlighted the team's milestones thus far. They have developed sample collection, preservation, and storage protocols using adult small volume samples. In addition, nurses at both clinical sites have been educated on the protocols so they handle the samples the same way, and a database of sample ontology has been developed. The team is working toward determining and documenting sample reliability, internal consistency, and validity, and will test samples at time points prior to and during sepsis events to determine the optimal time for predicting a sepsis event.

To determine the optimal collection and storage methods, the team first compared the microbial community in stool frozen immediately at -80°C (gold standard) to room temperature (all room temperature samples were held there for 7 days) stool in 95% ethanol, room temperature stool in lysis buffer, room temperature rectal swab in 95% ethanol, and room temperature rectal swab in lysis buffer. When analyzed, patient samples cluster together irrespective of storage conditions. The rectal swab group showed larger variability in the microbial community, but there were no statistical differences among different storage methods, per Bray Curtis dissimilarity analysis. When analyzing taxa, most collection and storage methods showed similar results, except that the rectal swabs had more microbes associated with the skin or mucous membranes. Because neonatal sepsis is often associated with organisms that are abundant in the mucosal surfaces of the GI tract, and necrotizing enterocolitis is a common critical illness in preterm neonates that manifests in the GI epithelial interface, the team decided to collect two longitudinal rectal swab samples allowing analysis of skin and mucous membrane–associated microbes. Because these infants have such fragile skin, the team takes extreme care to ensure patient safety when performing these procedures. The study team is conducting similar analyses for all sample types to ensure the development of stable collection and storage methods.

Dr. Prescott explained that the study is just getting started. The main IRB (USF) for prospective sample collection has been approved, and the team is waiting for the second institution (INOVA Children's) to join the single IRB protocol. Meanwhile, they are translating their consent into Spanish and setting up other logistics (freezer, sample tubes, and education for the nurses collecting the samples). Demographic and clinical data were collected into REDcap behind INOVA and USF firewalls accessed only by PI and CITI-trained, IRB-approved investigators. The team will be ready to begin collecting samples within the coming weeks.

Dr. Prescott highlighted the challenges they have encountered thus far with classifying their patients with shifting sepsis classifiers. For example, the team had a premature infant of 22-weeks of gestation born to a mother with chorioamnionitis who had received antibiotics several days prior to delivery. At birth, the baby was blood-culture positive (*Klebsiella*) but asymptomatic for sepsis (nevertheless, with possible multi-organ dysfunction such as respiratory distress syndrome and acute kidney injury as a result of prematurity). He was given 7 days of multiple broad-spectrum antibiotics, recovered from this infection, and then 4 weeks later experienced a spontaneous intestinal perforation (possibly independent of sepsis). At this time, the baby tested positive for yeast in the peritoneal fluid (treated with multiple antibiotics and fluconazole), was blood-culture negative, and was symptomatic for sepsis (requiring blood pressure support, etc.). Dr. Prescott wondered if the samples collected during these episodes (Days 1, 3, 7, and 28) should all be classified as sepsis because the patient was septic at a point, or if they should be classified based on the time points when sepsis may or may not present.

Dr. Esper asked for clarification about the scheme of sample collection. Dr. Prescott explained that the team is doing a weekly collection for 7 weeks to capture samples before, during, and after a sepsis event. However, it is not possible to predict when the sepsis event will occur, but it usually occurs within the first month of birth in approximately 35% of preemies, and they are trying to catch the sepsis episode by a weekly sampling scheme to inform the community about the best timing of sampling.

Dr. Files noted how important the classification issue is and highlighted it as a question for the current group to continue to think about. He explained that patient adjudication needs to be standardized and rigorous (present infection but not septic, septic but not in septic shock), and shifting among different

cohorts or classifications with time and overlapping among cohorts needs to be considered and noted when establishing biorepositories. Depending on the research questions, the investigators may be more interested in a sepsis-susceptible cohort than a septic cohort. Dr. Prescott agreed that some standard-care treatments to mitigate common illnesses (non-sepsis-related) for preemies tend to confound patient classification for sepsis, which makes analysis and comparison difficult. In the chat, Dr. Lautz noted that these questions about classification highlight the difference between bacteremia and sepsis, while such adjudication may be especially hard in neonates.

Dr. Matt Foster asked if the study team is collecting dried blood spots from heel sticks on these patients. Dr. Prescott answered that initially, all babies had a dried blood spot collected for state-mandated metabolic screening, but other sample types took precedence due to collaborators' interest in analyzing them, so the dried blood samples were not included in the final study protocol.

Dr. Tim McMahon asked if remnant samples from the time of sepsis being ruled out are available. Dr. Prescott confirmed that the babies have blood drawn often for clinical use; most of the blood samples collected during sepsis episodes are likely to be remnants.

In the chat, Dr. Mihir Atreya noted that this is an important cohort to study to learn about host/pathogen interactions in sepsis. He thought that it would be helpful to have pathogen metagenome sequencing data to identify those classified as culture-negative sepsis.

Scalable and Interoperable Framework for a Clinically Diverse and Generalizable Sepsis Biorepository Using Electronic Alerts for Recruitment Driven by Artificial Intelligence (SIBER-AI)

Annette Esper, M.D., M.Sc., Emory University School of Medicine

Dr. Esper introduced this MPI study (SIBER-AI) that she co-led with Dr. Kamaleswaran, who has a joint appointment at Emory and the Georgia Institute of Technology. The research team has expertise in prehospital sepsis, machine learning, clinical research in the acute setting, ethics, clinical trials, and hardware design that are needed to accomplish the goals of this project. The objectives of this study are:

1. To adapt a clinical sepsis screening algorithm to support research collection of various types of biospecimens
2. To design and test novel biospecimen collection among enriched sepsis populations in both ambulance and hospital environments to catch the early stages of the disease
3. To develop novel approaches to consent for a sepsis biorepository to make it easier to consent patients and develop a more diverse biorepository

Dr. Esper explained that a process to collect samples for sepsis biorepository already exists at Emory University, but it has a few limitations. Currently, patients are screened for sepsis in the ICU at one hospital affiliated with Emory; then the patients or surrogates are consented or a waiver of consent is in place if they cannot be reached to enable collection of data and samples within a given time frame. Additionally, blood samples are only drawn at one time point, except that in some substudies, three time points were collected. Dr. Esper identified a few areas of potential improvement, including the methods for screening, the sites and timing for enrollment, the timing of sample collection, and the consent method.

The first objective of SIBER-AI is to adapt an existing clinical sepsis screening algorithm to better support research and biospecimen collection. The Emory system currently uses a sepsis screening algorithm (developed during the 6-month design phase), but it was mainly used for clinical care rather than being incorporated into the research workflow, which is the goal of the 6-month implementation phase of this project. Using this screening algorithm, the team will identify patients with sepsis and collect biospecimens at three time points: at presentation to the ED, at ICU admission or 24 hours after ED admission, and 7 days post-ICU admission. Integrating the algorithm into Emory's research infrastructure will enable the researchers to take advantage of the real-time EHR/Redcap harmonization pipeline at Emory, allowing for more efficient patient screening, enrollment, and data harmonization.

The second objective of SIBER-AI is to design and test novel methods of biospecimen collection at appropriate time points. To pilot prehospital collection, co-investigator Dr. Carmen Polito works very closely with emergency medical services (EMS) to help them develop an EMS screening method (TeleEMS) that is in use in Atlanta (Phase I). Dr. Polito led a focus group of EMS, ED, and biotechnology personnel to design and provide feedback on the protocol. The team will start to collect environmental volatile organic compounds (VOCs) in this early septic cohort (pre-ED, sepsis suspected) hoping that volatilomics signatures could inform disease progress and outcomes, although there are a lot of unknowns about VOCs and few sepsis studies using VOCs except studies in neonates. Phase II of the study is to conduct longitudinal sample collection (skin VOCs, blood, cryopreservation of fresh blood for RNAseq) from patients identified via hospital sepsis algorithm (ED Encounter) and monitored by a real-time ICU bedside waveforms platform at Emory. VOCs can be collected from the environment, or through breath and skin, and the team is collecting both environmental and skin VOCs to test which one is the most useful for sepsis study. The team hopes that expanding the sample collection to unconventional and novel sample types will provide additional insights into the host responses and progression of sepsis.

The third objective of SIBER-AI is to develop a novel consent method reconstructing the length, timing, and content of the consent to be more appropriate for critically ill patients and their families. The team will assemble a patient advisory panel and conduct interviews with patients/family members to gauge their understanding of biorepositories and hear about their experiences when being approached for critical care research (trust, reasons for enrollment decisions, timing, etc.). The team will then incorporate the feedback and work with the patient advisory panel to develop a new consent, implement it, and conduct a survey to see whether the new consent improved enrollment and diversity in the biorepository.

Dr. Esper gave an update on SIBER-AI's progress to date. The screening algorithm has been deployed within the Emory hospital system, and the team is working to validate that it is working properly through retrospective physician adjudication. The team has chosen the device for collecting environmental VOC samples and is figuring out the best way to collect skin VOC samples with collaborators at Georgia Tech. Dr. Polito had educated the EMS team on how to collect environmental VOCs and worked out the logistics of collecting these samples before ED admission. Finally, for the consent process, the team had prior experience with a similar acute care study in cardiology and has completed the interview guide and flyers for interviews. Interview participants were approached at the time of informed consent, and their feedback regarding the consent and biorepository research was collected in real time. Patients in the post-ICU clinics were also approached for critical care and biorepository research. The team had been identified to lead a patient advisory panel to process the interview results.

Dr. Esper noted two lessons/challenges that the team has learned in this process so far. The EHR platform at Emory was changed into the EPIC system after the award, so the screening algorithm needs to be adapted to the new logistics, and there is a delay in deployment. In addition, the team is still working to create a workflow for efficient sample collection, especially how to synchronize the utility of the environmental VOCs collected at EMS with samples collected in the ED.

Dr. Esper highlighted ways in which SIBER-AI can move the field of sepsis biorepositories forward, including incorporating a validated screening algorithm to make the research workflow more automated and an expansion of the types of samples to the prehospital setting that may offer additional insights into the underlying mechanism of sepsis heterogeneity. She noted her excitement for the development of the novel consenting process, given the challenges to timely screening and enrollment, as well as collecting different samples at multiple time points.

Dr. Ware asked for the utility and analysis of environmental VOC samples. Dr. Esper explained that environmental VOC sampling uses the same sensor that their collaborators at Georgia Tech had been using for skin VOCs. The team thought that collecting air samples within the homes of patients with suspected sepsis wouldn't interrupt the more urgent EMS activities and would potentially capture a novel source for disease etiology. Dr. Esper noted that the team is still working on the analytic method and ways to integrate environmental VOCs with samples collected from the ED. In addition, they will decide later whether to use a wearable VOC sensor that touches the skin or a proximity sensor that catches skin emittance for the skin VOC sampling. In response to Dr. Segal's question, Dr. Esper clarified that the team will not collect breach VOC samples because that would interfere with EMS activities.

Dr. Sam Yang (MPI on a funded R21/R33 award) asked if the study team could perform genetic tests under the waiver of consent. Dr. Esper explained that genetic tests are not part of the proposed study and are not allowed under a waiver of consent. The team still needs to actively seek informed consent even with a waiver in place. Dr. Yang asked if IRB allows an initial over-the-board waiver of consent for all patients. Dr. Esper explained that their IRB only allows waiver of consent when they actively reach for the patient or the surrogate, but they can keep the samples collected under a waiver for those who cannot consent. She commented that their new consent process should at least address the challenge of refusal to enroll when they can find the patient/surrogate.

Dr. Files asked if any investigators are collecting blood specimens from patients who were discharged from the ED, as they represent an important control group. Dr. Seymour noted that Dr. Sachin Yende at the University of Pittsburgh has collected blood from sepsis patients after discharge. He also pointed out a publication from his team on prehospital identification of community sepsis (Brant EB. *Intensive Care Med.* 2020 Apr; 46(4): 823–824.). Dr. Foster also noted that his colleague Dr. Loretta Que from Duke University has followed patients with blood sampling for 12 months after discharge for COVID-19.

Dr. Seymour commented on the importance of EMS sampling and noted that it was done at the Piedmont Geriatric Hospital with a waiver of consent. He pointed out some important considerations for these samples, such as the sample storage en route, and the collection of similar samples in the ED. Dr. Guofei Zhou noted that a research group at the University of Michigan has tested breath VOC of critically ill COVID-19 patients (Sharma R et al. *JAMA Netw Open. 2023*). Dr. Chris Chao noted in the chat that exhaled breath has been used together with machine learning/AI for disease screening.

Session I Panel Discussion

Session Co-Chairs:
Annette Esper, M.D., M.Sc., Emory University School of Medicine
Lorraine B. Ware, M.D., Vanderbilt University Medical Center

Dr. Ware offered the following themes as a starting point for the discussion:

1. Remnants—trash or treasure?
2. Phenotyping—now or later? EHR ($) versus investigator ($$$)? Gold standards?
3. Sample timing—how early is early enough? Is automated screening the answer?
4. Consent—pros, cons, and challenges of waivers and delayed consent

Dr. Ware started the discussion by talking about remnant biospecimens. She noted that Dr. Seymour presented data that showed that protein biomarkers—but not other more complex analytes such as metabolomics and single-cell assays—are amenable to remnant samples. She invited those on the call to discuss if remnants should be collected, and if so, how the issues of variable timing in collecting and handling should be addressed. Dr. Esper noted that for most institutions, it is easy to collect remnant samples, so why aren't these samples used more? Dr. Ware noted that a pipeline for collecting remnant samples should be built first, and she doesn't have that in her institution yet.

Dr. Sanchez-Pinto explained that one benefit of using remnants is not having to obtain formal consent. However, he noted that some analyses will not be possible on remnants, and some health care–quality improvement initiatives limit the amount of excessive blood draw from pediatric patients, making remnants less available. He suggested two ways to improve current research using biorepositories: One is normalization for different assays to correct for the various lengths of time that remnant samples sit before being processed; the other one is to push for state-of-art technologies that enable advanced assays (e.g., single cell analysis) with less blood volume, reducing the need for remnant blood.

Dr. Files explained that an important factor in deciding whether to use remnant samples is setting definitions for acceptability and quality. These definitions must consider timing (when the sample is drawn from the patient and when different types of samples are processed, such as whole-blood vs. plasma); and batch effect (especially important for omic studies, and needs to be controlled). He explained that sample acceptability will vary based on the scientific questions and assays to be run. In addition, the host response states of sepsis may modify certain metabolites differently from those of healthy individuals. He hopes that the group can start to brainstorm best practices to standardize and report these factors.

Dr. Langelier (UCSF) echoed that sample and assay types are key determinants in the value of remnant samples. He explained that his team can identify microbial DNA from remnant samples from patients with sepsis and pneumonia that may be missed in culture-based assays due to antibiotic administration. They have success detecting bacterial DNA in EDTA plasma or viral causes of sepsis in remnant nasal swabs that may have been missed by targeted PCR assays.

Dr. Foster noted in the chat that his team saw a lot of pre-analytical variability in plasma, regardless of whether it was a remnant or collected in the context of a trial, likely due to processing condition and individual variability, as he spoke about later. Site-based batch effects also added to sample variability.

Dr. Bastarache raised the idea that pre-analytic variations in sample collection and processing time could be embraced as a strength rather than noise, which was coined by statisticians within the APS consortium. She explained that variability will always exist, even when following a set protocol, which adds more granularity to a single sample collection point during the sepsis continuum. In that sense, getting a total consensus on sample collection timing may not be critical.

Dr. Patrie agreed with Dr. Bastarache about embracing the benefits of timing variability, which may be better recognized with time when more data points are available. Although it is ideal to have uniform standards for sample collections across institutions, at this stage, it is important to track the timing of key events during sampling, including when the sample was drawn and the time for sample processing (e.g., the time of ultra-centrifuge). Over time, investigators could look at the impact of these variations and lock in on a certain standardized approach that fits their needs.

Dr. Lorsch proposed an idea about finding a natural internal standard in remnants that would allow researchers to gauge the quality of samples in the absence of accurate records of storage method and length. Hypothetically, there may be a molecule that presents at constant levels across patients with a well-defined decay rate, which could be used as a reference to measure the "decay age" of samples and allow comparison across samples. Dr. Seymour noted that the decay rate of different molecules will be different, and Dr. Bastarache said that it would be great to have a biochemist on board to test this. Dr. Roby Bhattacharyya (MGH and Broad Institute) commented that biomolecules may not be constant at baseline (time zero) across patients, but wondered if the ratios of well-vetted markers of organ dysfunction that might affect clearance could be useful (e.g., creatinine or cystatin-c). Dr. Lorsch noted that it might be worth it to follow molecule decays over time and under different storage conditions in controlled collection experiments, and then empirical algorithms (using a combination of the control molecule levels such as ratios, or geometric and arithmetic means) may be developed for determining the "age" of the sample by predicting the levels of other easy-to-measure molecules of interest.

Dr. Gordon Bernard commented that VUMC has a DNA remnant biorepository called BioVU, which contains DNA samples extracted from remnant clinical samples. BioVU has accumulated roughly 300,000 samples over the last 15 years, which have been successfully used for sequencing and genotyping. There was some batch variability, but this was controlled properly except for variants of some rare diseases. BioVU uses an opt-out method of consent, meaning that the samples stay in the biobank if the patient does not opt out. For more recent genetic studies, they use a formal consent process with approximately 75% of the enrollment rate. Dr. Ware asked if BioVU also collects plasma samples. Dr. Bernard explained that due to space constraints and the challenges with remnant sample quality that have been discussed, BioVU does not prospectively collect and bank plasma samples routinely, but BioVU can collect plasma samples per investigator's request if there is a defined project and patient population.

In reference to the first discussion theme, Dr. Seymour explained that remnant samples are neither trash nor treasure. Instead, they are somewhere in the middle ground. He noted that there are strong benefits to using remnant samples, including regarding cost, feasibility, and consent, but agrees that quality acceptability depends on the scientific questions being asked.

Dr. Files asked Dr. Seymour what the key issues are when setting up new partnerships with clinical labs to collect remnant samples and core standards to limit institutional variability. Dr. Seymour explained that he is preparing to engage community hospitals for the R33 phase. When working with clinical lab

directors, he found their workflows and sample processing timing were different. It is important to work with the lab directors and staff to standardize this process.

Drs. Ware and Esper noted that based on the discussion, the group recommends continued use of remnant samples.

Dr. Kamaleswaran (Emory) opened a discussion about the use of EHR data and standardization. He is particularly interested in hearing from this group about their experiences in the best practice and modeling using EHR data—that is, what has been working, what is not working, what are the ways of data interpretation that improve early identification and prediction of sepsis, and what are the opportunities for collaboration.

Dr. Sanchez-Pinto discussed the issue of data quality and assurance with EHR data, as well as the standardization of data models, so these can be eventually shared to generate data-driven phenotypes across institutions. Similar to the process of tailoring the quality of assays for a certain type of sample, his group focuses on the quality control of 50 key data elements that are needed for organ dysfunction phenotyping. They emphasize the high quality and reproducibility of these 50 variables but are open to collecting other data elements. He mentioned Dr. Seymour's collaboration with Dr. Vincent Liu (Kaiser California) on the "sepsis on FHIR" project (Brant E.B et al. *NPJ Digit Med.* 2022; 5: 44), which standardizes data across different institutions to generate national standards for EHR data collection in sepsis studies. He believes that this type of work is helpful to promote a consensus on the types of variables to be collected from EHR, methods for quality assurance, and suitability to be used for different models.

Dr. Esper shifted the gear of discussion to the automatic screening of patients that aids sample collections at different disease stages. The experience of their group is that manual screening is time-consuming and limiting for multisite recruitment, so automation of the screening is useful. She is curious about others' experiences in using automated screening and how useful they are in corresponding to different times (early, ED, ICU, etc.) of sample collection.

Dr. Bastarache explained that her team takes the opposite approach. Instead of screening patients for a predefined phenotype, they enroll patients more broadly and conduct retrospective phenotyping later. The rationale behind this approach is that it does not exclude/limit patients who may not fit the traditional definitions of sepsis or ARDS, and the group ends up with 50% of sepsis patients in the cohort. Dr. Esper followed up by saying that this approach is also appealing to those who want to study pre-sepsis. Enrolling all critically ill patients also introduces the opportunity to learn who may be prone to sepsis and which subgroup is prone to develop organ dysfunction later.

Dr. Files noted the importance of enrolling noncritically ill patients (e.g., in the ED) in addition to critically ill patients to serve as a separate study population. For instance, enrolling patients who present with an infection and then go on to resolve the infection may allow for studies on the drivers of resolution of the host responses when compared to sepsis patients. Dr. Files noted that this is a gap in the field, and colleagues from the ED department (many in the meeting) could potentially fill this gap.

Dr. Michael Filbin (MGH) described his protocol for consenting patients in the ED, which usually occurs 3 to 4 hours after arrival. His team collects a research sample at the same time as the first clinical sample using delayed consent. They then pursue consent during the following hours. He thinks these time zero

samples would be highly valuable for finding sepsis signatures and diagnostic tools. This method may lead to many blood collections from patients who are not that sick; however, Dr. Filbin went on to say that having a control cohort is a strength, and their cohort ends up with a good percentage of septic patients as well as severe sepsis. Dr. Esper asked Dr. Filbin how the ED team identifies patients to enroll, if not by automated EHR data. Dr. Filbin explained that he instructed his team to enroll any patient who appears to have organ dysfunction as the IV line is established, even if they do not outwardly appear to be septic, many of which turn out to be septic later.

Dr. Sanchez-Pinto reiterated his protocol, which is to collect a pre-resuscitation sample with delayed consent in the ED, and EHR triggers were used to identify septic patients. A research sample was drawn at the same time the clinical samples were drawn as part of the resuscitation order. Dr. Sanchez-Pinto had two comments: One programmatic consideration is that when the samples are used for the omic type of studies, the results should represent the clinical time point when that sample was drawn. For example, samples drawn at initial adjudication upon ED arrival should only be used when the research question triggers the use of samples at this time point. The other operational consideration is that sicker patients may die 12 to 24 hours before the research team has a chance to pursue informed consent. It is important to work with the IRB to include language in the waivered or delayed consent to keep the research samples when that happens.

Dr. Seymour reiterated that there is a level of healthy tension between approaches addressing several targeted gaps in biobanking (e.g., Dr. Bastarache's study) and approaches testing a big-picture concept, like his own project. For a very interdisciplinary group such as this meeting, people may not be able to come to a consensus on the precise protocol for sample collection because the ideal repository for each group differs based on the scientific questions at hand.

Dr. Ware explained that a benefit of waived or delayed consent is a less biased cohort. However, her experience in multicenter projects is that different IRBs have different restrictions regarding keeping samples collected from patients who cannot be consented. IRBs that prohibit keeping such samples negate that benefit. She would like to hear others' experiences in this.

Dr. Nate Shapiro (Beth Israel Deaconess) explained that his team has a waiver for delayed consent, but they must discard samples collected from patients who cannot consent. He asked if others have any advice on how to convince their local IRBs to allow them to keep these samples, since IRBs are subjected to the same rules, so it might be the interpretation of these rules that differs among institutions. Dr. Esper noted that their team is allowed to keep samples in this situation.

Dr. Lautz (Cincinnati Children's) asked in the chat if a central IRB helps with site differences. Dr. Files answered that in his experience, sometimes it helps, but individual sites might not accept everything described in the central IRB protocol.

Dr. Bernard had experience with IRB administration for 20 years. He explained that the criteria for impracticability used to have a very high bar. If consent is possible with enough money and time, the IRB will demand the discard of unconsented samples. Now their IRB had a shift in perspective and only asked if consent would ever happen given the practical situation. The logic that some IRBs use regarding keeping samples collected under waived or delayed consent is that by throwing away samples, you are throwing away the benefit the samples could provide despite already introducing risk to the patient.

Dr. Filbin noted that he was met with skepticism by his IRB regarding the deferred consent due to a lack of published literature (more publications on waived consent). The IRB approved his protocol but asked him to study the impact of deferred consent. He noted that genomics analysis is not "minimal risk" and needs to be consented to, but you can collect first and consent later. He hopes that as more institutions get approvals, they publish literature to arm future institutions with examples.

Dr. Esper noted that there is also a population of patients who are approached but do not give consent. She noted that this population can teach the community how to improve consenting protocols, as detailed in their study.

Dr. Yang explained that his team has the approval to draw a blood sample in ED at the time of the first clinical blood draw with waived consent (pre-resuscitation research sample), which is triggered by nurse-activated sepsis protocol. They were able to get approval for this due to the nature of minimal risk and the importance of the research. However, they cannot perform genetic analyses such as host response–related omics on samples collected under waived consent.

Dr. McMahon noted that a potential risk with blood samples drawn under delayed or waived consent is that linking with EHR data may risk loss of privacy. He wonders if the lack of consent restricts the clinical data that can be collected. Dr. Bernard responded that IRB prohibits the sharing of identifiers of patients, or generally the team is not given access to any identifiable information; instead, the specimens are assigned a unique study ID, which cannot be traced back to the patient's medical record, but deidentified EHR data were assigned to the same study ID to allow linking. This mitigates the risk of privacy breaches. Dr. Seymour noted similar approaches at their institution to protect patient privacies by firewalls, which usually satisfy the IRB.

Session II: Comprehensive Omics Analysis and Data Integration

Session Co-Chairs:

Leopoldo N. Segal, M.D., M.S., New York University School of Medicine
Mihir R. Atreya, M.D., M.P.H., Cincinnati Children's Hospital Medical Center and University of Cincinnati

Dr. Files introduced the Session II co-chairs: Dr. Segal from New York University School of Medicine, and Dr. Atreya from Cincinnati Children's Hospital Medical Center and University of Cincinnati.

Multiomic, Mass Spectrometry-Based Analysis of Dried Blood for Deep Phenotyping of Sepsis

Matt Foster, Ph.D., Duke University School of Medicine

Dr. Foster introduced the expertise of the study team, notably his background in omics and analytic chemistry and R21 MPI Dr. McMahon (Duke) in pulmonary medicine and RBC pathobiology. He explained that this study started with his involvement in a project to follow the pharmacokinetics of a single protein (EFGRvIII:CD3 therapeutic antibody) in whole blood (Schaller T.H. J Proteome Res. 2019). Traditionally, proteomics is done using plasma, but with a few changes to their bottom-up liquid chromatography-mass spectrometry (LCMS)-based proteomic protocol, the team solved this challenge of performing proteomics on whole-blood samples.

During the COVID-19 pandemic, Dr. Foster's clinical colleagues in Pulmonary Medicine at Duke began biobanking whole-blood samples from COVID-19 patients in the ICU for proteomics and metabolomics using the Neoteryx Mitra tips, which are designed for remote sampling; each tip absorbs 20 uL of blood. The convenience, ease of preservation, and small volume of blood collection are great for COVID trials, as multiple trials are competing for the same patient. The team collects samples by dipping Mitra tips into EDTA clinical samples. A total of 87 ICU patients were enrolled (50% survival rate); and longitudinal microsamples were collected on Days 1, 3, 7, 14, and 21 in the ICU. Dr. Foster's colleague (Dr. Que) also established a post-ICU clinic and collected post-ICU samples at 3 weeks, 3 months, 6 months, and 12 months.

The team performed proteomics on plasma samples and Mitra tip whole-blood samples from six COVID patients and six healthy controls. It is surprising to see that more proteins were identified in whole blood (~1,000) than in plasma (~600), but there are significant overlaps in upregulated (e.g., acute phase markers) or downregulated (e.g., lipoproteins) proteins between COVID and controls in the two types of samples (abstract by Will Thompson et al. in ASMS, Fall 2021). Dr. Foster explained that the data suggest that proteomics analysis on whole blood retains information contained in plasma and offers much more.

Dr. Foster mentioned other published studies using Mitra tips (Whelan S.A. *Anal. Chem*. 2023; Shen X et al. *Nat Biomed Eng.* 2024; Volani C. et al. *Metabolites*. 2023). Some considerations that these studies have raised are the difference between capillary and venous blood as well as the stability of certain metabolites important for drug metabolism within the Mitra tips. Snyder's group at Stanford uses the Mitra tips to inform personalized medicine (Shen X et al. *Nat Biomed Eng.* 2024).

Dr. Foster noted the major gaps that this study is aiming to address. First, most proteomic studies use plasma or serum except one preprint (Fredolini C, *MedRxiv*, 2021) he noticed, missing information contained in the cellular component. Second, plasma samples have high pre-analytical variability (noting

artifacts from platelet lysis). Finally, the landscape of post-translational modifications has not been fully explored in blood using quantitative proteomic approaches; he explained that most blood post-translational modifications are lost when cells are removed, except for glycosylation.

Aim 1 will develop methods for multi-omic phenotyping from dried blood collected via Mitra tips. The milestones are to quantify at least 1,000 proteins, 250 glycopeptides, and 400 phosphopeptides from a single tip, and more than 500 metabolites from one to two Mitra devices. In addition, the team plans to explore ways to validate these assays and establish a "ground truth" to increase generalizability, similar to the reference standards included in commercial kits to confirm that the assay is working before moving on to the real samples. Aim 2 will validate this method for 96 plate-based multi-omic phenotyping, develop reference standards for interlaboratory validation, and measure analyte stability using simulated storage and shipping conditions (e.g., different processing conditions before Mitra tip sampling, shipping of Mitra tips in room temperature, etc.).

Dr. Foster described the team's progress to date: They developed robust protocols for sample storage and processing that will allow for protein and post-translational modification (PTM) analyses, as well as metabolomics and multi-omics. The team had done some inter-lab and -platform validation and has made progress toward quality control, including developing a ground truth and assessing analyte stability. Collaborator Dr. Ian Wong (Duke) is doing additional mining of EHR from biobanked subjects, and they are looking at data integration.

Dr. Foster explained how they organize the Mitra tip samples for an easy transition to a 96-well plate format. The tips containing the blood sample are ejected into matrix tubes, which are housed within a 96-well matrix plate. A custom-fitted matrix tube decapper makes this process easier. The matrix plate and samples are then stored at -80°C. This transfer minimizes variations caused by unusual sample-collecting vessels, such as clamshells. Dr. Foster noted that this method is currently being used in a collaborative study—that is, the Sepsis Characterization in Kilimanjaro study (SICK, PI Matt Rubatch, R01AI155733), and a longitudinal (four time points) collection of Mitra tips from sepsis patients in Kenya is now available. The MPI of this R21, Dr. McMahon, is also collecting Mitra tip samples in the red blood cell ATP export and transfusion in a sepsis study funded by R01HL161071. He is capturing whole-blood samples from sepsis patients at a single time point to look at ATP manipulation in red blood cells.

The first step in the protocol for processing these samples (abstracts in US HUPO 2023, ATS 2023) is pathogen inactivation when necessary (e.g., for COVID samples), which requires heating for 15-30 min at BSL-2 plus conditions (needs biosafety approval). Then the rest of the initial processing, including trypsin digestion, can be done at the BSL-2 level and can be completed in ~4 hours. Enrichment of the peptide samples for phosphopeptide and glycopeptide takes 1 hour each. Currently, only a small portion of the sample is used for this analysis (~0.25 µL or 62.5 µg of blood for proteomics, among which ~25% is used for phosphoproteome and ~10% for N-glycoproteome).

Dr. Foster highlighted that these samples were successfully analyzed on their own instruments, such as Exploris480 or Orbitrap Astral using microflow LC, nanoLC, or Evosep at a flowrate of 100 µl/min, 400 nl/minute, or 1 µl/minute, respectively. They also sent the samples to two collaborators at 908 Devices and UC Davis, who showed that the samples can be run on different analytic tools such as a capillary electrophoresis device (ZipChip Interface) and Evosep-TimsTof2, showing the wide range of adaptability of these sample preparations.

Dr. Foster showed that for the purpose of blood proteome (n=4), samples are very stable, even left at 37°C for several days. Typically, the Mitra tips were allowed to dry for 2 hours, or overnight at 4°C, and then stored at -80°C. Dr. Foster tested protein stability when the tips were left at room temperature or 4°C for up to 15 days. Blood proteomic and single-protein abundance results showed little variability due to storage conditions, but high interperson variability. Dr. Foster noted that the glycoproteome results also showed the same pattern—that is, less than 30% coefficients of variation (CV) in the QC pool but very tight clustering for different storage conditions of each patient. Dr. Foster noted that given such a big interperson variability in glycoproteome and multi-omics in general, normalization is an important consideration across analytes. For example, Haptoglobin (Hp) levels vary significantly across subjects, and individual differences have to be normalized for when quantifying glycopeptides. The subject who has a high level of Hp also has abundant glycoforms, and normalization to total Hp helps make the glycoforms within the same experimental group more comparable.

Dr. Foster also noted that Mitra tips improved the stability of samples under different storage conditions for phosphoproteome. Wet blood samples sitting at 4°C overnight had a significant decrease in detected phosphorylation, while the Mitra tip samples left at room temperature for 1 day had very little decrease, but there were some decreases for 3 days and more for 15 days, while still better than the 1-day wet samples.

Dr. Foster noted that his team also hopes to run metabolomics and multi-omics from the same Mitra tips (Abstract, Will Thompson, 908 Device; ASMS 2023). They encountered challenges with solubilizing protein off the tips after they had been exposed to organic compounds but have seen success measuring metabolites and native peptides (e.g., Enkephalin A) in these samples.

Dr. Foster highlighted some of the lessons learned and the challenges they have faced. The team can quantify thousands of analytes from a very small amount of blood in a short amount of time. They found that some of these analytes are very stable, while others (in particular phosphopeptides and metabolites) are not stable, so the use of remote or remnant samples would be challenging to detect these. Dr. Foster reiterated the problem of high interperson variability and thought that longitudinal sampling and analysis of the same patient may be important. Nevertheless, inter-individual variability on PTMs/metabolites provides an opportunity to establish a “ground truth” dataset. Finally, Dr. Foster noted that there is a low barrier to sample collection using Mitra tips and that ongoing biobanking efforts (such as APS) could add them to their protocols easily.

Dr. Files proposed an idea of “reverse remnant,” where a small aliquot of blood is taken from a clinical lab sample before clinical testing to avoid the mandated holding time Dr. Seymour alluded to earlier. Dr. Foster confirmed that this is how Mitra tip samples used in this study were collected, and earlier presentations in this meeting suggested that the Mitra tip sampling has the potential to increase the value of biobanking for many ongoing studies. He is proposing a workflow to be developed in a clinical lab to use this method for micro amounts of whole blood before clinical lab processing.

Dr. Segal asked if the blood volume was normalized for the -omics. Dr. Foster explained that the Mitra tips should absorb 20 uL of blood, which equates to 5 mg of protein under the assumption of 250 mg/ml protein in whole blood. Another way of normalization is to use PTM. Dr. Segal added that they “spike” samples with metabolites as a way of normalization for metabolomics. Dr. Foster explained that this may have been done for Mitra tip sample analysis, especially for drug testing, but his research team did not

do it for this study since the samples were already collected at the time of the study. He thinks that it is challenging to do a prospective spiking when collecting samples in a clinical setting.

Dr. Segal noted that the collecting tube materials (plastic vs. glass) have been shown to affect the yield of DNA samples and asked if the Mitra tips are made of plastic. Dr. Foster explained that Mitra tip materials seem to be very impermeable, and they have been used for genomic DNA. However, his experience found that it is not easy to extract proteins from the Mitra tips once they have seen high levels of organic solvents.

Exploring the Sepsis-Delirium Connection Through Omics-Scale Top-Down Glycoproteomics

Steven M. Patrie, Ph.D., Northwestern University

Dr. Patrie presented his research on developing a Top-down proteomics (TDP) clinical pipeline for measuring blood-based glyco-proteoforms. He hypothesizes that this innovative proteomics data type will provide insights into the complexities associated with the onset, progression, and recovery of sepsis.

Dr. Patrie explained that proteoform investigations aim to understand a protein's microheterogeneity, which arises from various events such as alternative splicing, alternate promoter usage, alternate translation initiation, endogenous proteolysis (e.g., signal peptides), mutations, polymorphisms, or co-/post-translational modifications (Smith LM et al. *Nat Methods*. 2013). In the case of glycoproteins, proteoform microheterogeneity within a sample is due to the diverse mixture of glycans that can attach at one or more sites along the protein's backbone, resulting in tens to hundreds of co-existing glycoproteoforms with subtle variations in their sugar composition. Dr. Patrie further explained that proteoforms are best measured by TDP, which analyzes intact proteins instead of digested protein fragments as in bottom-up proteomics. Measuring intact proteoforms by mass spectrometry provides a direct snapshot of the presence of co-occurring modifications and their frequency in a single experiment, information that is lost when the protein is digested with enzymes such as trypsin.

Dr. Patrie highlighted that TDP has never been utilized in an unbiased clinical proteomics pipeline specifically designed for large-scale discovery and quantification of glycoproteoform microheterogeneity of blood glycoproteins. He elaborated that prior to the routine implementation of the workflow on sepsis-related samples from biorepositories, the R21 phase of his study will concentrate on the technical development and optimization of key components of a new analytical pipeline for this unique data type. This includes advancements in intact protein chromatography, mass spectrometry (MS), and bioinformatics, as the existing proteomics resources are not suitable for this project.

Dr. Patrie emphasized that the realization of these innovations could provide valuable insights in sepsis research, an area he aims to explore in the R33 phase of the study in a pioneering clinical proteomics study. The team seeks to assess glycoproteoform alterations in sepsis patients in the ICU, correlating changes in glycosylation patterns with the onset and resolution of septic shock. During the R33 phase, they intend to implement the workflow to study glycosylation changes on acute phase glycoproteins in blood samples from the DECODE-SEPSIS cohort from Dr. Girara at the University of Pittsburgh, and the BRAIN-ICU-1 cohort from Drs. Pandharipande and Patel at Vanderbilt. This phase will allow his team to begin exploring the feasibility of the pipeline for examining hypotheses, such as whether glycosylation memory may be correlated with long-term cognitive dysfunction following sepsis.

Dr. Patrie explained that the study is partly motivated by the observed increase in various blood acute-phase glycoproteins (such as anti-alpha chymotrypsin [AACT]) in response to infection and sepsis. He referenced a previous study that analyzed blood samples from septic shock patients following elective surgery (Caval, T. et al. *Front Immunol*. 2021). The study monitored AACT glycosylation changes before (pre-sepsis), during (onset-sepsis, ICU admittance), and upon recovery (ICU discharge). Dr. Patrie noted that while the relative amount of acute-phase glycoproteins returned to pre-sepsis levels upon ICU discharge, the spectral pattern changes detected by mass spectrometry remained at ICU discharge. This led the team to hypothesize the existence of a "glycosylation memory" in acute-phase proteins that persists after sepsis.

Dr. Patrie elaborated on the novelty and procedures of the proteoform pipeline being developed in the R21 phase. Although the team is working towards several quantitative milestones related to the pipeline, the pipeline primarily aims to obtain two sets of information not easily captured by conventional proteomics approaches. The first is the unbiased identification of blood glycoproteins in plasma or serum, along with informatics support to identify and quantify proteoform-level microheterogeneity of glycoproteoforms in an -omics environment.

Dr. Patrie explained that his team has been developing new methods for glycoprotein and glycoproteoform separations prior to mass spectrometry (MS) analysis. These methods include solution-phased isoelectric focusing (separating proteins by their isoelectric points), hydrophilic interaction liquid chromatography, capillary electrophoresis, and reversed-phase liquid chromatography. Specifically, Dr. Patrie highlighted their work using pH gradient separations (i.e., isoelectric focusing [IEF]) to probe glycoproteoforms from plasma, serum, albumin-depleted samples, or lectin pull-down samples. He demonstrated that this process helps decipher a broad range of proteoforms by LC-MS that separate based on isoelectric point (Cline E.N. et al. *Anal Chem*. 2021). Dr. Patrie also noted that these types of isolations require less than 100 µL of plasma or serum, and that the measurements may be routinely possible with only 10 µL of blood when used in combination with highly sensitive techniques such as capillary electrophoresis.

Dr. Patrie next discussed the development of a bioinformatic pipeline designed to identify and visualize glycoproteoform networks and quantify their changes across samples, known as the Proteoform Network Analysis (PNA). He explained that this work aims to address the critical limitations of common data-dependent mass spectrometry approaches, which struggle to effectively fragment proteins exhibiting high proteoform microheterogeneity. Dr. Patrie emphasized the importance of assigning glycoproteoforms and visualizing their relationships through glycoproteoform networks to begin to understand the underlying mechanisms that may alter protein glycosylation in response to disease processes, such as sepsis.

He outlined the key goals of the informatics pipeline as follows:

(1) Confidently identify glycoproteins in complex mixtures in an unbiased manner.
(2) Assign distinct sugar compositions for all observed glycoproteoforms for each protein.
(3) Quantify the relationships between observed glycoproteoforms through network analysis resources.
(4) Make site-independent predictions of what N-glycans are likely present for each protein.

Dr. Patrie detailed the team's innovative approach to identifying glycoproteins in complex samples, demonstrating their ability to readily identify glycoproteins in plasma and compare them to lectin pull-down targets. He explained that each glycoproteoform assignment represents the sum of the sugar composition of a protein, rather than glycosylation at specific sites (Hossler P et al. *Biotechnol Bioeng*. 2006). He outlined how the application of machine learning and network analysis tools on the multidimensional TDP datasets aids in making unique sugar composition assignments for each glycoprotein.

Dr. Patrie presented examples to illustrate this concept, focusing on di-N-glycosylated lipocalin-type prostaglandin D synthase (L-PGDS), which exhibits hundreds of measurable glycoproteoforms across a range of isoelectric points. The network analysis tools display the relationships between the sugar compositions (i.e., fucose, sialic acid, hexose, and GlcNAc) and intensities of the L-PGDS glycoproteoforms. He explained that these new networking tools also enable his team to make in silico predictions about the types of N-glycans expected for each glycoprotein. These predictions can then be validated by conventional glycopeptide experiments or through a novel glycoproteoform tandem mass spectrometry technique, which allows for the assignment of specific glycans at individual residues.

Finally, Dr. Patrie highlighted the potential utility of the PNA approach through a proof-of-concept workup of the AACT data collected by Dr. Albert Heck's group (Caval, T. et al. *Front Immunol*. 2021 Jan 14:11:608466), as mentioned earlier. Using this method, they were able to demonstrate unique AACT glycoproteoform network patterns for different patient cohorts (i.e., pre-sepsis, ICU admittance, ICU discharge group). The networks could readily quantify persistent changes in AACT fucosylation or LacNAcation, despite the fact that the level of the assayed protein (AACT) had returned to pre-sepsis levels after hospital discharge.

Currently, the team is working on predicting N-glycans on AACT in septic shock patients at four different time points and aims to identify glycosylation patterns to cluster the patients based on study variables. Dr. Patrie demonstrated that using these glycoprotein profiles, the team could cluster patients based on their time to sepsis onset (i.e., ICU admittance after 3 days vs. 3 weeks). When the glycoprotein and predicted glycan profiles associated with the onset of sepsis are used in predictive modeling, the team was able to predict days to sepsis recovery (out to several weeks) based on the subsiding of their glycosylation patterns. These promising results, combined with achieving many of the technical milestones associated with the TDP pipeline development, motivate the team to bring this new tool to a clinical proteomic environment in the R33 phase for further testing.

Dr. Files asked if there are any special considerations when collecting and storing samples for glycosylation analysis. Dr. Patrie referenced Dr. Foster’s presentation and noted that glycosylation is robust under different storage conditions. However, Dr. Patrie explained that the top-down approach is different from the bottom-up proteomic assays. The last year has been dedicated to determining the pipeline and analysis methods needed, but the coming year will start to address questions of sample or interpatient variability.

Dr. Zhao asked what the link between changes in glycoproteoform and sepsis outcome is, and if any mechanistic explanations for this link have been identified. Dr. Patrie explained that alterations in the glycoproteoform of many inflammatory acute-phase proteins observed in sepsis and various sepsis subtypes (e.g., meningococcal, pediatric, nosocomial, fecal peritonitis [FP], and septic shock) may be results of gene-level regulation that are not known yet. In addition, glycosyltransferase expression is

sensitive to cytokine levels. Therefore, unique cytokine signatures associated with different sepsis endotypes may modulate glycan profiles in unique ways.

Biorepository Optimization and Use for Endotyping Critically Ill SARS-CoV-2-Infected Patients

Leopoldo N. Segal, M.D., M.S., New York University School of Medicine

Dr. Segal introduced his study and noted that his lab is focused on host/microbe interactions (Singh S et al. Mucosal Immunol. 2022). Specifically, the team is focused on microbes that enter the lower airway and affect the local lung mucosa. They achieve this through a wide array of -omics techniques including microbiome methods related to DNA, RNA, metabolomics, and proteomics.

The lab has studied multiple diseases other than critical illnesses but when the pandemic hit, Dr. Segal shifted his priorities and resources to collecting and analyzing blood and lower airway samples from critically ill COVID-19 patients. He believes that this group of patients is less heterogenous than other groups with critical illness such as sepsis and ARDS, as COVID-19 is a single disease caused by a single pathogen. The team changed their existing IRB protocol to allow family consent for patients unable to consent themselves. Direct consent is obtained after the patient recovers; in the event of death, the team can keep de-identified samples and data collected from these patients. Given that New York City was an epicenter of early COVID-19 outbreaks, Dr. Segal's team was able to enroll 100 patients during April 2020 alone and built a prospectively collected biobank with associated metadata.

The main question in Dr. Segal's study was understanding what was driving poor clinical outcomes in severely ill COVID-19 patients. They wanted to understand if factors such as a secondary infection, viral cytotoxicity, or an exuberant inflammatory response were associated with poor clinical outcomes.

Under the support of a different award, the team enrolled every critically ill intubated patient in the hospital and collected samples at six time points from Day 0 to 21 of ICU or intubation. Dr. Segal explained that the study team has access to EHR extracts of daily health updates, X-rays, and MRIs. With the collected samples, the team conducted DNA and RNA sequencing to determine the viral load, host transcriptome, metatranscriptome, and antibody responses.

For the R21/R33, the team will go beyond these methods. In the first aim of the R21 phase, the team wants to make sense of the clinical metadata. They extracted prospectively collected clinical data and EPIC data into research databases and will use latent class analytical approaches to identify distinct phenotypes that are associated with poor clinical outcome. In the second aim, the team uses metabolomics on blood and lower airway samples. Finally, in the third aim, the team focused on developing single-cell RNAseq protocols. Dr. Segal explained that the quality control part will be done in the R21 phase, and the R33 phase will expand the analysis to a large cohort of patients and compare the biological assays among different clinical phenotypes (rapid/slow progressors, and poor/good responders).

Dr. Segal noted a major challenge that his team has encountered: This is a very complex cohort, and although patients are all labeled as having severe disease and were similarly intubated in the ICU, these patients have different trajectories and different outcomes. Additionally, there are many confounding factors that differentiate these patients. Dr. Segal displayed a figure representing when the samples were collected from each patient during the disease trajectory divided by good (extubated and survival) or

poor (intubation > 60 days or death) outcome. Major events were also marked on the chart, such as symptom onset, hospitalization, ICU stay, intubation, or death. Their trajectories are aligned at the time of intubation; interventions, such as antibiotics, antivirals, and steroids, can also be overlaid onto this figure. This figure graphically demonstrates the complexity of the cohort. The team didn't use 28-day ventilator-free as an indicator of good outcome as more than 50% of this cohort requires ventilation for more than 30 days.

Dr. Segal noted that timing sample collection across the cohort with respect to disease trajectory is difficult. For example, Day 1 of hospitalization may be Day 7 or Day 14 since symptom onset, depending on the patient. Because of this, collecting longitudinal samples and precisely tracking key events (symptom onset, hospitalization, intubation, etc.) is very important. With information such as this, patients can be further stratified into slow or rapid progressors in both the poor and good outcome groups, which is particularly important to follow for viral dynamics and antibody dynamics.

Dr. Segal discussed building a cross-sectional latent class analysis. The team had determined a set of metadata (baseline demographics, medical history, COVID information, pre-intubation clinical data, and interventions) to be used for the analysis, and used an unbiased approach to cluster patients based on these characteristics. By doing so, the research team can identify what variables are the best to distinguish each cluster.

The team has also adopted a similar latent class analysis method on longitudinal data (Bos LD, *Lancet Respir Med*. 2021 Dec;9(12):1377-1386) but Dr. Segal's analysis uses a lot more data variables. The team first needed to determine how much longitudinal data was needed to build an accurate model. To do so, they developed supervised analytical approaches and found that using 3 weeks of longitudinal data was best: any less, and one's confidence in predicting outcomes falls; any more, and one's confidence in predicting outcomes does not improve. Thus, Dr. Segal used 3-week longitudinal data for patient clustering, which was tested for the prediction of patient outcomes using a supervised method. He also explained that this approach could either use all variables or selected critical care domains (e.g., coagulation, hemodynamics, infection, inflammation, renal function, and ventilatory function). Clustering variables by domains and then combining these domains yields greater precision than using all variables individually. The longitudinal analysis divided patients into three classes (1–3), each class containing three subclasses (A–C) composed of two groups (high mortality, low mortality). When zooming into defined parameters like CO2/min, P/F ratio/min, or vent ratio on the longitudinal axis, one can identify the parameters that better define the different clinical trajectories of classes 1–3. Dr. Segal explained that this bioinformatics approach identified the key factors for worse prognosis in a given patient (massive cardiogenic dysfunction or fatal infections), which is important in terms of patient endotyping.

Dr. Segal noted early results of the metabolomics assays (Aim 2). Using both polar and hydrophobic LCMS, the team has identified some common metabolites in the lower airway and BAL samples and is applying quality control methods to normalize the data. Differentially expressed metabolites can be more clearly visualized.

Dr. Segal noted the challenges associated with obtaining high-quality, single-cell RNAseq data (ECCITE Seq and cell hashing) in this study. The team had used fresh samples for CITE-seq and scRNAseq under BSL-3 condition on a 10x device, but they wanted to optimize the protocol to use cryopreserved samples. Because certain cell types, such as neutrophils, are highly sensitive to freeze/thaw, the team tested many conditions to determine how to protect sensitive cell types. The team learned that samples should be

thawed gently (not in a hot water bath), and that sorting viable cells on flow cytometry prior to sequencing delayed processing and thereby reduced yield, which is not recommended.

Dr. Segal explained that the team has also been working to integrate antibody- and lipid-based cell hashing into the protocol. They can perform cell hashing successfully on PBMC (yield >1,000 cells in various samples) but have faced challenges doing so with BAL samples. Cell numbers are few; many were label-negative or labeled as doublets due to the cell debris in the sample. This creates problems when performing RNAseq, preventing efficient deconvolution into single-cell transcriptomes. Dr. Segal settled on a single prep per sample, paying attention to thawing and processing time carefully in order to get quality scRNAseq data on lower airway samples.

Dr. Segal summarized his observations thus far: It is complex to organize metadata and pair it with samples, so it is recommended to harness the integration of prospectively collected metadata and EHR data. Next, using defined critical care domains to sort clinical variables first that will be used in prediction trajectories (semisupervised) outperforms agnostic data pooling, because these variables contain clinical concepts or biases that need to be taken into account. Additionally, metabolomics with a polar LCMS approach yields quality data on a large number of metabolites in lower airway and plasma samples. Finally, single-cell RNAseq is challenging on cryopreserved lower airway samples, and the critical steps to overcome this challenge include thawing and length of processing time.

Dr. Foster asked if Dr. Segal's pipelines are available on the platform GitHub. Dr. Segal explained that all of the published methods (i.e., microbiome, metabolomics) are available on GitHub but not much on the realm of scRNAseq. He remarked that he wishes that quality control efforts could be published more frequently so they could serve as references for others.

Dr. Files noted in the chat that many interesting methods are being tested in various ways by all of the researchers on the call. He wondered if it would be possible to publish a group report on which methods worked or did not work, given how helpful it would be to future investigators.

Leveraging Multi-Omics to Maximize the Scientific Value of Pediatric Sepsis Biorepository and Advance Patient Endotyping

Mihir R. Atreya, M.D., M.P.H., Cincinnati Children's Hospital Medical Center and University of Cincinnati

Dr. Atreya introduced his study and noted that he and his team do not have much preliminary data yet, as the grant was only recently funded, so he will focus on a conceptual overview. Dr. Atreya explained that critical illnesses—including sepsis and ARDS—are inherently complex. Additionally, the genetic variants associated with these illnesses are common but have low penetrance. Thus, the cumulative burden of multiple genes associated with sepsis increases sepsis susceptibility and risk of poor outcome (Engoren M. et al. *PLoS One*. 2022).

Dr. Atreya also explained the redundancy and dynamicity of the host response in sepsis, which makes the adjudication of sepsis onset hard. Thousands of genes are differentially expressed within hours of infectious or inflammatory insult. Pediatric sepsis endotyping delineates patients who respond poorly to corticosteroids (Maslove DM and Wong HR. *Trends Mol Med*. 2014), but it doesn't provide a framework to identify the future targeted therapies based on gene expression–based endotyping. Finally, nearly 50% of patients switch sepsis endotypes during their disease course.

Dr. Atreya explained the increasing molecular complexity when trying to integrate different types of molecular data (20,000 genes code for 100,000–400,000 proteomes, per Dr. Patrie), while different molecular layers provide orthogonal evidence. Data about the genome, transcriptome, proteome, and metabolome provide information on what may happen, what is likely to happen, what is currently happening, and what happened in the past, respectively (plus the microbiome). He explained that his team is interested in epigenome-wide modifications and likens them to master puppeteers of gene expression during critical illness. These modifications include DNA methylation, histone modification, and long noncoding RNA (increasingly popular but not well established yet). Dr. Atreya explained that studying two or more of these -omics layers may provide a better understanding of the underlying biological truth about diseases (Hasin Y et al. Genome Biol. 2017). The team is also interested in how these omic layers converse with each other. For example, which epigenetic factors are more important in determining the gene expressions of the transcriptome, and how is this signal relayed to the proteome?

A critical challenge is figuring out how to generate and integrate clinical, phenotyping, and multi-omic data collected from critically ill patients across the age spectrum (Atreya MR et al. *Crit Care*. 2022). Dr. Atreya explained that this study is built on a long-running pediatric septic shock biorepository containing serum and plasma samples from 2,000 patients, and DNA and RNA samples from 800 patients at two time points. The first question this funding mechanism will answer is how to integrate DNA methylation profiles into existing RNAseq-based endotypes, which will deepen biological understanding and may inform targeted drug development based on the existing RNAseq-based endotypes (works on Latent Class Analysis to be published soon). Dr. Atreya explained that in the future, the team will do whole-genome genotyping and integrate plasma aptamer–based proteomics as well as cell-based assays (scATACseq, CITEseq) to delineate the immune cell subpopulations that drive the above-mentioned endotypes.

In the first aim of the R21 phase, the team will conduct quality control of previously banked specimens for pilot sequencing. The second aim includes a pilot-integrated analysis of host methylome and transcriptome, which will help estimate the sample size needed for the R33 phase for derivation and validation of multi-omic endotypes (Aim 3). Finally, the team wants to see if clinical variables can be used to predict these multi-omic endotypes that demonstrate reproducibility among a larger dataset. For Aim 1, the patient group containing multiorgan dysfunction, and clinical trajectory was selected for analysis. DNA samples undergo bisulfite conversion and standard DNA methylation arrays to identify patterns of CpG sites across the epigenome. Bulk RNAseq will also be performed by Inflammatix, Inc. Pilot data integration will be done through multi-omics factor analysis (Argelaguet R et al. *Mol Syst Biol*. 2018), a method of data integration that was previously used in cancer epigenetic studies.

The R33 phase will scale up using existing biospecimens to derive multi-omic endotypes and determine their biological relevance. Dr. Atreya noted that about 450 to 500 patients will have DNA methylation profiling data on Day 1 (collected within 24 hours of septic shock), and transcriptomics will be done using samples on Day 1 and Day 3. Finally, the team will validate the clinical relevance of these multi-omic endotypes. Dr. Atreya noted that given the high stability of DNA, they are interested in whether the DNA methylation profile measured on Day 1 affects the temporal shift in gene expression throughout the course of disease (change in gene expression over time), and how that influences organ dysfunction trajectories (persistent vs. resolving multiple organ dysfunction syndrome, [MODS]).

The final aim of the R33 phase of this grant is to use the clinical data in the biorepository to fit a set of previously published clinical classifiers against the multi-omic endotypes (supervised machine learning) (Sanchez-Pinto et al. *Pediatr Crit Care Med*. 2023 Jun 2; *JAMA Netw Open*. 2020; 3). These clinical classifiers could predict persistent MODS in 15,000 out of 40,000 patients in an external dataset. The idea is that clinical data alone may not be the most accurate in predicting clinical outcomes and adding simple biological data (e.g., DNA methylation pattern) would improve the model performance.

Dr. Atreya noted the team's progress to date. Bulk mRNAseq has been performed on 174 patients at Day 1 and Day 3. DNA samples are currently undergoing quality control testing prior to use in the DNA methylation arrays.

Dr. Atreya thanked the late Dr. Hector Wong for his key contribution and foresight in establishing a highly comprehensive biorepository at this institution that allowed the search for the "biological truths" in sepsis.

Dr. Segal asked if it is possible to define biological truths. Dr. Atreya explained that there are different layers of biological truths across critical illness syndromes, and he hopes to go beyond the surface that was currently scratched. Dr. Segal noted that every aspect of the multi-omic work provides a unique contribution to the biological truths, and there should not be any limitations on the methods researchers choose to use.

Dr. Foster asked what the stability of RNA is over time, and if the sample preservation methods changed during the lifetime of the biorepository. Dr. Atreya explained that the same research assistant has been working on this study since the beginning, and that they used the same methods and materials throughout. He expects RNA quality to be high and consistent. Additionally, Dr. Atreya noted that these samples were collected in PAXgene Blood RNA tubes, so the team is not able to do single-cell RNAseq on the existing specimens. They hope to move to this technique in the future and noted the challenge of collecting single cells across multiple sites.

Session II Panel Discussion

Session Co-Chairs:
Leopoldo N. Segal, M.D., M.S., New York University School of Medicine
Mihir R. Atreya, M.D., M.P.H., Cincinnati Children's Hospital Medical Center and University of Cincinnati

Dr. Atreya started the panel discussion by noting a recurring theme of this session: the need for a structured database to deposit datasets from various sources (properly linked to clinical metadata) into high-capacity computational tools able to analyze large multi-omic data sets to get a more comprehensive view. Dr. Patrie described his experience with the RUSH Alzheimer's Disease Center (RADC), which has a large biorepository with well-crafted cohorts from studies they completed over the last 20 years. They share samples with a variety of groups with the expectation that the data generated will be fed back to the RADC, which will be made available to future investigators. He noted that he has benefited from their clinically oriented infrastructure, which keeps linked patient data easily accessible. It is a nice model on which to build an infrastructure, with diverse types of data available, and it continues to expand.

Dr. Bastarache noted that other studies that successfully made sense of high-dimensional multi-omic data seem to have focused patient populations and questions to answer, while the study of sepsis is much more complex. She referenced the timeline of patient clinical trajectories that Dr. Segal presented and wondered if there are even any computational tools to integrate the multiple layers of biological heterogeneities on top of the clinical heterogeneity, and if the results are understandable even if tools are available. This is much more complex than doing multi-omics to answer focused questions. Dr. Segal added that the timeline data is only a start to grasp the high degree of heterogeneity within this cohort. Unless there are well-defined comparisons, it is unlikely to yield useful observations.

Dr. Patrie described an approach that the Michael J. Fox Foundation is using, which is a high-dimensional study of a very specific target. They are trying to establish complementary proteomics resources that could tackle the characterization of the post-translational modifications of alpha-synuclein on brain tissues. They emphasize complementary methods and will start with one brain region relevant to Parkinson's disease. The brain samples will be divided into different groups that will use these complementary proteomics methods. Dr. Patrie thinks that this approach has a high potential to generate relatable data between groups that can be integrated to reflect the whole brain. This is only possible due to the highly specific and well-thought-out study design. He noted that data generalization only comes after a large amount of data is added to a repository generated from various sources and different time points. However, he thinks it is still useful to aggregate data, as this will allow the investigators to identify trends in data and inform new hypothesis development, even if the trends are not statistically significant at the time. A new focused study could be designed based on this information.

Dr. Bastarache asked if such a complex task is possible to tackle given the heterogeneity seen in these critical illnesses, and wondered if researchers should not get into multi-omics studies until more is understood about the clinical heterogeneities. Dr. Patrie responded that his analytic expertise lies in developing tools, and he relies on people who understand the heterogeneity and mechanisms of sepsis to create a set of samples allowing for the test of a specific hypothesis. In his opinion, the definitions of sepsis endotypes and signatures of each endotype have not been clearly established. One way is to tackle a single endotype at a time using a core method in a strategic way. In the chat, Dr. Zhao asked if it

would be possible to proceed with these studies in a crude way, and with time, gain more precision as analytical techniques evolve.

Dr. Lautz noted that this discussion highlights that the challenges of building a biorepository are different from the challenges of answering a highly specific hypothesis-driven question. He noted that taking a broad enrollment approach to build a biorepository is a good strategy to catch all aspects, but the question is how to make this biorepository useful for an investigator who needs to make a specific comparison drilling down through the heterogeneity. In those cases, thoughtful clinical annotation is imperative—for example, clinicians may be able to build separate subgroups based on the clinical information associated with such a biorepository. Dr. Bastarache noted that her team takes the broad enrollment approach but carefully phenotypes patients post hoc (sepsis, organ failure, etc.). She felt this was a well-balanced approach.

Dr. Seymour discussed a provocative idea of "no biological truth," and the only endotype that matters is what links to treatment. To make the search more efficient, people should start with the patient groups that are responsive to treatment, which is the ultimate supervised approach. Searching for a biological mechanism that is not amenable to any treatment may not help clinically. Dr. Zhao noted that the Human Immunology Project Consortium does mechanistic clinical trials using interventions such as antibiotics or vaccines and has made discoveries about underlying biology using this method. Dr. Lorsch asked how a new therapy could be used to study underlying mechanisms for treatment responses without prior knowledge of its mechanisms of action. Dr. Zhao explained that approved therapies can be used in mechanistic clinical trials to study biological responses to the therapy.

Dr. Atreya added that the top-down and bottom-up approaches both have their values. Pioneering works in the field on gene expression–based endotypes were linked to differential responses to treatments (e.g., corticosteroids, fluid therapy, activated protein C, etc.) at least in retrospective analysis. Each of these studies has its own truth, but he is wondering how the investigators should build a consensus framework to better understand biology.

Dr. Bastarache noted in the chat that targeted therapies stemmed from studying biological heterogeneity in the cancer field. Studying response to existing treatment is an efficient approach, but it will not lead to the discovery of new therapies. Dr. Bhattacharyya (an infectious disease physician at MGH) agrees with Dr. Bastarache that an increased understanding of biology could lead to more rational targeted therapy; once the biological differences between endotypes are more defined, researchers would not be taking a shot in the dark when developing therapies. Dr. Bhattacharyya also echoed Dr. Atreya in that it would be more productive if investigators embed the multi-omic approaches into clinical trials. Specifically, instead of looking for latent variables predicting treatment responses post hoc, one can harness the power of randomization in randomized trials to stratify the biological responses by class in order to understand responses to therapy. Dr. Lautz added in the chat that this is the approach the SHIPSS trial is taking when looking at endotyping and response to corticosteroids in pediatric septic shock.

Dr. Foster thinks open-ended discovery studies have value. He noted that this funding mechanism was meant to encourage that. The samples used in the projects should represent sepsis, but the goal is to generate as much data as possible instead of answering a specific question. Dr. Segal agreed that a broad and agnostic approach still has its place in sepsis research.

Dr. Bernard noted that currently approved interventions (e.g., corticosteroids, mechanical ventilation, immunomodulators, etc.) have the potential to alter many biological measurements even if these are not the main targets of the intervention. He thinks that studying the naive, unaltered sepsis state combined with interventions that can change the whole biological milieu are both important. Dr. Bhattacharyya added in the chat that the biological/"omics" response to therapy could be followed to predict clinical usefulness if the measurement is simple enough.

Dr. Sanchez-Pinto explained his thoughts on a strategic way to pursue the "biological truth." There are different layers and depths to patient stratification and phenotyping, and researchers can start with the low-hanging fruits by focusing on the implementation of clinical phenotyping and applying therapies to the group with the greatest confidence. Then the deeper multi-omic endotyping work will inform additional layers of targeted therapies in a more precise way. A phenotype is useful as long as it is reproducible and has values to guide clinical practices. For example, Dr. Calfee's work on IL-6 and other biomarkers to predict the likely patient group to respond to different PEEPs, fluid, or statins would be in this category (Sinha P. *Curr Opin Crit Care*. 2019 Feb;25(1):12-20). If the conclusion needs to be validated before implementation, we should set validation and implementation as short-term goals. On the other hand, this R21/R33 mechanism provides an opportunity for researchers to build tools to pursue the next layer of more targeted therapies (e.g., anti-cytokine, etc.), moving from broad-stroke therapies (PEEP, fluid) once a big, broad phenotype is solidified (e.g., hyper-/hypo-inflammatory).

Dr. Atreya thinks that a more complete framework would be helpful to understand the biological basis of endotypes and phenotypes. He noted that Dr. Calfee's work stratifies innate immune and endothelial function well, and Dr. Bos's gene expression studies of the same two endotypes suggest changes in innate immune responses (Lieuwe DJ Bos, *Am J Respir Crit Care Med*. 2019 Jul 1;200(1):42-50.) In contrast, established pediatric septic shock endotypes (Wong et al, *Am J Respir Crit Care Med*, 2015) largely reflect the adaptive immune response. Thus, an integrated schema for endo-phenotype identification is necessary.

Dr. Patrie commented that the goal of this funding mechanism is to establish rigor, reproductivity, and utility of certain techniques. In order to achieve that, it is best to have a well-designed study with an expected difference among groups. A broad, all-inclusive study design may dilute small signals. He asked if there are any clearly definable clinical endotypes, but Dr. Bastarache didn't think so. Dr. Patrie still believes that eventually, the sepsis field may come to a generally agreed classification method for patients, so the technology experts can prove the utility of their tools in hypothesis-driven research.

Dr. Bastarache emphasized the heterogeneity within this population—almost every patient is distinct. She believes that instead of drilling down clinical phenotypes, they may be able to identify shared underlying biological phenotypes that explain the clinical heterogeneity.

Dr. Seymour noted that mathematicians in the field feel that current subtyping works are inefficient—individuals within a subtype are still diverse and on a continuum, and there are no molecular signatures to each subgroup. So, in the end, we may move to individual treatment effect (ITE) instead of targeted therapy for each subtype. Publications on ITE have started to emerge in the field. He sees testing an individual patient for a range of possible therapies and applying it to such patients as an attractive future direction for clinical trials. He is not antidiscovery, but some people do believe that all research will eventually lead to ITE.

Dr. Atreya opened a new discussion about single-cell-level data, its advantages and disadvantages, and how to integrate it into an existing framework built on bulk RNAseq data. Dr. Segal noted great excitement for single-cell technology but cautioned about the challenges associated with it. Under the support of this R21, he is trying to optimize the methods, but he worries about what he misses even with quality data. For example, sequencing results will be dominated by abundant cells (e.g., neutrophils in the airway samples of COVID patients) and highly expressed genes on the surface of each cell (sequencing depth and cell specificity). As a result, it is important to incorporate orthogonal approaches, paired clinical phenotypes, defined clinical groups, and longitudinal sampling to generate testable hypotheses.

Dr. Bhattacharyya noted that while you cannot infer single-cell transcriptional profiles from bulk, you can do the reverse. When you identify new cell states of interest, you can find marker genes unique to that substate. This method can be used to reinterpret past datasets and prospectively design studies in the context of this knowledge. Single-cell study can be a scaffold for adding on analysis, and there are many open-source tools (e.g., CIBERSORTx, xCell) to do that through bulk deconvolution. He agrees that scRNAseq does not yield high depth per cell, but he still believes that you can learn about an assembly of transcriptionally similar cells if the cell sampling is broad enough. However, this can be cost prohibitive. The cellular context contained in these data could be useful to guide therapy. For example, there may be a cell substate that is overdriven by IL-6.

Day 1 Closing Remarks

Meeting Co-Chairs:

Julie A. Bastarache, M.D., Vanderbilt University Medical Center
D. Clark Files, M.D., Wake Forest University School of Medicine

Dr. Bastarache noted that given the main focus of the R21/R33 mechanisms on biorepositories and since there are many leaders in sepsis, it is important to discuss the best practice learned here on building biorepositories. She summarized Day 1 of the meeting in the following aspects:

- Themes
  - Remnant (including dried blood spots) versus fresh blood samples
  - Immediate versus delayed phenotyping; broad versus targeted patient cohorts
  - Manual versus automated EHR data collection
- Challenges
  - Blood volumes (especially for pediatric and neonatal patients)
  - Timing, trajectory, and treatment effects
  - Data integration (integrate EHR and CRF data plus analysis), high-dimensional data (OMOP)
  - Harmonizing sample collection with advanced assays (complicated processing)
  - Clinical heterogeneity (no solution but need to be mindful of this)
- Commonalities
  - Plasma, cells, EHR phenotyping/identification, study patients across the life span
- Novelties
  - Novel consent, HME, sampling the environment, patient engagement, EAB, microsampling, multi-omics, dried blood analysis

Dr. Bastarache also presented a set of questions to guide Day 2's facilitated panel discussion:

- Should we share our **study protocols** to facilitate collaborations and identify synergy?
- Should we create a **sample quality resource** for us and the community?
- How can we ensure that the samples/data we are collecting are **forward thinking**?
- Is there a role for **centralization** (e.g., RUSH ADC)? Some centers focus on enrollment while others focus on highly technical assays.
- Do we prioritize unbiased or targeted discovery?

Dr. Files thanked NIGMS for getting the funded investigators together, as this environment has been highly collaborative. He reiterated a discussion point that permeated throughout the meeting, which is whether the proper study design needs a standard protocol and prescribed patient cohort or a broad group with specialized protocols. He believes that both are necessary. He speculates that the key issues the APS consortium needs to decide now, for example, are which patients to enroll and which samples to collect. In that case, a common protocol would be hugely helpful, and people can learn a lot from it. However, a rigid protocol may stifle many innovations discussed in that meeting. He thinks that both approaches are important, and the community will benefit from shared experiences, protocols, as well as successes and failures. Dr. Zhao thanked the speakers and session chairs for a wonderful first day.

Session III: Expanding the Horizon on Analytic Techniques

Session Co-Chairs:
Eric P. Schmidt, M.D., Massachusetts General Hospital and Harvard Medical School
Paul Bollyky, M.D., Ph.D., Stanford University Medical Center

Dr. Files introduced the Session III co-chairs: Dr. Schmidt from Massachusetts General Hospital and Harvard Medical School, and Dr. Bollyky from Stanford University Medical Center.

Circulating Bacteriophages for the Diagnosis of Sepsis

Paul Bollyky, M.D., Ph.D., Stanford University Medical Center

Dr. Bollyky introduced the title of this study and the MPI, Dr. Sam Yang, both at Stanford. He explained that clinical diagnostics for sepsis are inadequate, especially since blood culture is time-consuming and subject to both false-positive and -negative results. Other diagnostics under development, such as 16S PCR and transcriptomics, are not perfect either (false positive, no AMR and information, etc.). As a result, the field needs rapid serum–based diagnostics to identify bacterial pathogens.

Dr. Bollyky introduced circulating cell-free DNA (cfDNA), which is 50-to-200-long nucleotides present in the plasma. Most of these DNA strands are of human origin, but some reflect the microbes living in the body. CfDNA is an exciting area of medicine that changed the clinical diagnosis of perinatal testing, cancer staging, and transplant rejection (nicknamed liquid biopsy). Companies such as Karius have investigated using cfDNA in sepsis diagnostics and other infectious diseases. The cfDNA test developed by Karius showed acceptable specificity. However, the Karius platform faces challenges in differentiating the colonizing flora from pathogens causing sepsis (the background problem) and in differentiating subtypes of microbes (the resolution problem), such as coagulase-negative staphylococcus from *Staphylococcus aureus* (*S. aureus*), because of the limited amount of information contained in the short DNA fragments.

Dr. Bollyky's team wondered if it was possible to use bacteriophages to identify pathogens in sepsis. Bacteriophages are viruses that seize bacteria, which are highly specific to their host species and are abundant in circulation. Most bacteriophages are DNA viruses. Dr. Bollyky's group found that the abundance of bacteriophages reflects bacterial population dynamics (i.e., if bacterial populations increase, bacteriophages associated with that bacterial group also increase). They think that this observation may be used to identify the bacterial pathogens.

To test the theory, Dr. Bollyky's team used biospecimens collected from 62 patients with sepsis and 10 asymptomatic controls in the Stanford Emergency Room. DNA extracted from these samples was sequenced using the Illumina platform and analyzed through an in-house pipeline that the group developed for this project. Methods are detailed in a recent publication (Haddock NL. Nat Microbiol. 2023 Aug;8(8):1495-1507). The group found that the vast majority of cfDNAs are of human origin in both the control and sepsis groups (more than 95%). Among the remaining nonhuman reads, a wide diversity of origins was detected, including mammalian viruses and bacterial sequences and species that were found in both groups; however, when zooming in on the bacterial sequences, the two groups differed in the percentages of present microbes (26.66% Propionibacterium in control vs. 29.83% Enterobacterium). However, the types of bacteriophages and the phageome diversity in both groups were largely the same.

This suggests that bacterial cfDNAs and phage cfDNAs detect different types of information. They conclude that the circulating phageome is unchanged in sepsis.

When Dr. Bollyky's team looked at individual patients, they found a lot of variations in the phageome of individuals, similar to other microbiome studies. The team is excited about the data that patients with sepsis driven by a particular microbe had an overabundance of bacteriophages specific to that microbe. This was true of sepsis caused by *Escherichia coli* (*E. coli*), Staphylococcus, Streptococcus, and Klebsiella. In other words, sepsis is associated with pathogen-specific increases in phage.

Dr. Bollyky explained that the team studied a second independent larger set of data from a Karius-funded investigation (the SepSeq Study). The samples from 267 septic patients and 167 controls collected from the Stanford ED had been previously sequenced, and the team downloaded the sequencing data. The septic patients were categorized as blood-culture positive, blood-culture negative/site-culture positive, or systemic inflammatory response syndrome (SIRS, no identified infection, negative culture results). Like the first cohort, septic patients with *E. coli* infection have an overabundance of *E. coli* phages; this was also true of Streptococcus, Staphylococcus, And Klebsiella. Looking at patients subdivided by culture category, in many cases, it is possible to identify an overabundance of specific bacteriophages in the blood culture–negative/site culture–positive patients.

The team then wondered if this platform could be used to differentiate between closely related strains of bacteria, thereby solving the resolution problem of the bacterial cfDNA sequencing technique. Their results showed that phage but not bacterial cfDNA can differentiate Staphylococcus aureus and Coag-negative Staphylococcus because these two have completely different phages. They want to test whether phage could help solve the background problem by identifying commensal flora versus pathogens. Their data showed that *E. Coli* sepsis had characteristic phages as compared to asymptomatic controls, SIRS, and sepsis caused by other bacteria. In addition, phages can differentiate different *E. coli* pathology variance (e.g., EHEC, VTEC).

Dr. Bollyky explained that this tool was also used to identify infections beyond sepsis (Haddock et al. mSystems. 2023 Aug 31; 8(4):e0049723). Using a third cohort of 91 patients with site infections from UCSF (wound fluid, joint infections, urine, BAL, etc.), the team identified similar associations in pathogen-specific bacteriophages and site infections from *E coli*, Streptococcus, *S. aureus*, and Klebsiella.

Dr. Bollyky highlighted his conclusions thus far. First, people have a circulating phageome that can be studied using conventional metagenomic sequencing. The team used Illumina in this study but hopes to adapt the method to Nanopore and other more rapid sequencing in the future. Next, circulating bacteriophages can be used to identify bacterial pathogens in sepsis and other infections. He believes that phageome is a frontier in human disease.

Dr. Bollyky also highlighted his next steps. The team is working to develop a phage lexicon for antimicrobial-resistant strains. Fungi also have phages, so Dr. Bollyky is working to add them and viral pathogens to the pipeline. He hopes to validate these studies with large, external datasets and introduce improved sequencing protocols.

Dr. Ellen Burnham (University of Colorado) asked where Dr. Bollyky sees this platform going in the future, and if it could be used point-of-care to aid the start of antimicrobials. Dr. Bollyky explained that he is optimistic that this platform could be used as point-of-care one day, especially given that sequencing

methods and computational power will continue to improve and quicken. He thinks the time-limiting factor for cfDNAs is not the sequencing and analytics since 98% of the human DNAs have been depleted, but rather sample acquisition and processing. Either phage or bacterial cfDNA seq will replace bacterial culture one day in the lab.

Dr. Zhao asked for more information about the blood culture–negative/site culture–positive group. Dr. Yang explained that the prior Karius study can detect offsite infections using bacterial cfDNA.

Dr. Langelier asked how much the ability to identify microbes increases when the phageome is analyzed instead of only bacterial genomes. Dr. Bollyky explained that the Karius test is good at detecting microbes where there wasn't much commensal flora background; in these cases, adding in phageome studies does not result in new information. In other cases, when trying to discern *E. coli* or Staphylococcus subsets, phageome analysis can better discern subtypes. Dr. Bollyky explained that he envisions phageome analysis as complementary to bacterial cell-free DNA analysis. He highlighted the importance of discerning microbial subtypes (e.g., methicillin-resistant *S. aureus* [MRSA] and methicillin-sensitive *S. aureus* [MSSA] have different phages) because information on a pathogen's antimicrobial resistance patterns helps clinicians make treatment decisions.

Dr. Bhattacharyya wonders how "complete" the reference sequences are for bacteriophages. He noted the great abundance and diversity of bacteriophages in existence and asked how often Dr. Bollyky's team encounters phage DNA that cannot be assigned to a known phage, and if there is room to improve in the phage catalog. Dr. Bollyky explained that his team made the decision to limit their study to the top 24 microbial pathogens in sepsis and then assembled phage lexicons for those 24 pathogens. In that way, this was a study with a diagnostic approach, not a study of the microbiome.

Dr. Files noted that determining the difference between colonization and infection has become more complex with the clinical use of RT-qPCR as well as respiratory and blood-culture techniques.

Dr. Sanchez-Pinto asked if the bacteriophage system can distinguish an overgrowth of gut flora, such as enterococci after antibiotics, from a pathogen that did not cause the original infection. Dr. Bollyky explained that yes, phage and bacterial cell-free DNA can identify site infections in some cases. However, cell-free DNA from the gut often does not enter the circulation in high enough densities to be sequenced except in patients with heavy immune suppression. Dr. Sanchez-Pinto also noted that pediatric patients often experience viral or fungal sepsis in addition to bacterial. Dr. Bollyky expressed his interest in studying fungi and molds but remarked that few species have enough known phage sequences to make meaningful conclusions other than Aspergillus.

Dr. McMahon asked if there are phages specific to bacteria that produce extended-spectrum beta-lactamase (ESBL). Dr. Bollyky explained that ESBL organisms each have their own distinct phages, but not all these organisms have complete reference datasets. He is hopeful that as sequencing repositories expand, these gaps will be filled.

Dr. Bhattacharyya asked if there are any regional differences in bacteriophage sequence or content for a given pathogen. Dr. Bollyky noted that his team has not studied this question but plans to in the future. He noted that only studying samples collected at one site is a current limitation of the study.

Establishment of a Multi-Center Biobank of Patient-Specific Induced Pluripotent Stem Cells for Pediatric Sepsis Research

Mihir R. Atreya, M.D., M.P.H., Cincinnati Children's Hospital Medical Center and University of Cincinnati

Dr. Atreya explained that this is a high-risk, high-reward project with Dr. Lautz as the MPI. The motivation for this study was to develop a platform for *in vitro* mechanistic studies to understand the biology of sepsis endotypes and phenotypes using a human-relevant model. Specifically, he and his team wanted to leverage advances in induced pluripotent stem cells (iPSCs), which are derived from PBMCs or other somatic cells (e.g., fibroblasts) using Yamanaka factors (Oct4, Sox2, Klf3, C-myc) and can be then differentiated into various cell types from the ectoderm, mesoderm, and endoderm using specific protocols.

Dr. Atreya envisions a broad potential use of iPSCs in sepsis research. It can be a renewable cell source for both common and rare cell types and can be added on to create more complex organ-specific organoid models of sepsis using multiple cell types. Dr. Atreya explained that his team is using iPSCs as a starting point for disease modeling, but in the future, they could also be used for drug screening and for studying genome editing therapies or cell-based therapies (precision medicine). There was precedence, in that iPSC models are successfully used in other monogenic and polygenic conditions, particularly during the chronic disease state to model the postpathogen infection period.

As background, Dr. Atreya explained that a prior study from his team leverages a biorepository of serum and plasma from pediatric patients of septic shock and used a biomarker-based risk stratification approach to understand the underlying biological drivers for severe and persistent organ dysfunction versus recovery from it (Wong HR. *Sci Transl Med*. 2019 Nov 13;11(518):eaax9000). These biomarkers were precisely validated in controlled experiments.

For the current R21 project, Dr. Atreya developed microvascular organoids from healthy donor iPSCs to model endothelial function (Atreya et al. ATS 2023 abstract). 3D whole-mount confocal images showed that mesoderm-differentiated cells are positive for CD31 and PDGFRβ. Upon stimulation with LPS, the microvascular organoids have decreased in total tube length, branch count, average loop area, and occludin expression. Occludin is a key endothelial barrier protein whose loss contributes to capillary leak.

Then, the team treated the healthy donor microvascular organoids with existing risk-stratified serum collected from septic patients or healthy donors. Using the Nanostring platform, the team identified differentially expressed genes between the organoids treated with serum from healthy donors or septic individuals. The key pathways these genes belong to are involved in endothelial dysfunction, such as the angiotensin system or glycan signaling, and inflammation.

Dr. Atreya explained that the MPI, Dr. Lautz, is focused on iPSC-derived cardiomyocyte models for cardiomyocyte dysfunction in sepsis (Lautz et al. *ATS* 2023 abstract). iPSC-derived cardiomyocytes can contract in cultures that mimic the heartbeats. Dr. Lautz found that treating iPSC-derived cardiomyocytes with serum collected from high-risk sepsis patients, but not low-risk, reduced the shortening fraction and beat frequency of the cardiomyocytes. They also found differentially expressed genes between cardiomyocytes treated with high-risk versus low-risk serum. The differentially expressed genes belonged to signaling pathways such as adrenergic receptors and cellular calcium ion signaling. Dr. Atreya explained that these pathways will be mechanistic targets to pursue in future studies.

Dr. Atreya noted that the true goal of this study is to generate organoids from patient-derived iPSCs. The benefits of this are that the iPSCs would retain the genetic characteristics of the patient they were derived from and that they could be used to model endotype or phenotype-specific disease mechanisms *in vitro*. Additionally, organoids from patient-derived iPSCs may be used to perform high-throughput drug screening for endotypes or phenotypes.

Dr. Atreya presented two questions that this study hopes to answer:

1. Can patient-specific iPSCs from critically ill children with sepsis provide a more robust disease model than healthy donor iPSCs treated with pooled plasma or sera from septic patients?
2. Do patient-specific iPSCs recapitulate the primary signatures in sepsis, and can it be demonstrated in two cell types (immune and endothelial cells)?

Dr. Atreya explained how this human iPSC study works into the framework of the study he presented on Day 1 using human primary circulating cells. Both studies are founded on the basis of transcriptome-based endotype assignments. They are trying to show that the characteristics of patient-derived iPSCs are comparable to the multi-omic signatures of the circulating immune and endothelial cells with the same endotype assignment.

The aims of the R21 portion of this grant were to establish a strategy for the collection of immune and endothelial cell-enriched PBMCs for hiPSC generation, scRNAseq, as well as phenotyping and functional validation. In the R33 phase, the goal is to perform multi-omic characterization and comparison of the primary cells and iPSCs.

Dr. Atreya explained his team's enrichment strategy. Whole blood is harvested in ACD tubes (Citrate) and then immunomagnetically enriched using a CD45 marker. CD45-positive cells are reprogrammed into iPSCs, which can then be differentiated into iPSC-derived monocytes or endothelial cells. The team also collects plasma from the same group of patients for experimental treatment of the iPSCs. The CD45-negative cells undergo another positive selection step to isolate CD45-negative, CD34-positive cells, which represent the circulating endothelial cell fraction.

Dr. Atreya explained that his method can successfully catch the neutrophil subsets (IL1R1+, MGAM+, MME+ o FCGR3B/CD16+) from fresh specimens. It can also enrich for monocytes/macrophages (CD14+, CD68+, CXCL8+, LYZ+), T-cells (IL7R+, ANK3+, CAMK4+), and NK cells (GZMH+, GZMA+, GNLY+, CCL5+). Dr. Atreya noted that mature and progenitor endothelial cells can be found in circulation. These cell types are rare to find in healthy patients but increase in number to cope with endothelial injuries in critically ill patients. Subcluster analyses on the CD45-negative, CD34-positive cells collected from critically ill patients show enrichment for endothelial cells originating from the heart and lung. Dr. Atreya reasoned that these circulating endothelial cells retain their organ-specific signatures, which may help inform the identification of targets for organ-specific treatments.

Looking forward to the R33 phase, the team hopes to scale this study to do a cross-comparison of multi-omic data among circulating primary cells and iPSC-derived cells, after showing the feasibility of the iPSC protocol. Dr. Atreya believes that iPSCs treated with serum from high-risk patients would recapitulate the key biological signatures of these critically ill patients. The team hopes to not only show that iPSC is a relevant model, but that it can also retain endotype- or phenotype-specific mechanisms. An iPSC biobank that captures patient genetic diversity will be needed for work at this scale.

Dr. Atreya noted that this grant was funded in April 2023. Since then, the team has enrolled two critically ill pediatric septic shock patients. The team uses a dual consent process: one standard consent for genomic testing for sepsis, and another for iPSC generation. They have performed single-cell RNAseq on circulating leukocytes and endothelial cells from both patients. The timeline for iPSC generation is 3 to 4 months before a cell type–specific differentiation process.

Dr. Zhao asked if the timing of sample collection affects iPSC generation. Dr. Atreya explained that this is one of the milestones of the R21 phase, and his team will study if 24 hours versus 72 hours before cryopreservation makes a difference for iPSC generation. Dr. Zhao asked if sample collection respective to disease timeline affects iPSC generation. Dr. Atreya explained that the iPSCs would retain the patient's genetic characteristics regardless of collection time, but it is not clear if endotype-specific biology is merely a reflection of the patient's immune cell snapshot at a given time. Their biorepository has samples with endotype assignments at two time points, and they can test the specific effects of risk-stratified serum collected at different time points, which may help answer this question.

Dr. Sanchez-Pinto suggested that studying iPSC derived from survivors of severe sepsis may help determine the underlying biology that allows those patients to overcome their disease. Dr. Atreya agreed, and explained that understanding which molecular features drive organ dysfunction or prevent or promote organ recovery is equally important to study. Dr. Lautz added that genetic susceptibility to other disease processes has been identified using iPSC-derived cell lines.

Monitoring Pro-Resolving Leukocyte Responses in Peripheral Blood Predicts Clinical Severity During Sepsis

Joel Voldman, Ph.D., Massachusetts Institute of Technology

Dr. Voldman introduced his study on behalf of the MPI, Dr. Bruce D. Levy, and explained that his lab is interested in studying the resolution of sepsis. The team has expertise in engineering, microfluidics, proteomics/MS, physician-scientist, and biorepository. He described the idea that resolution is an active process controlled by specialized pro-resolving mediators (SPMs), not simply the absence of activation. The main hypothesis in this study is that sepsis hyperinflammatory responses result from defective endogenous resolution mechanisms. These pro-resolving pathways are disrupted during sepsis and repaired during resolution.

Dr. Voldman described his research goals. The first goal is to use microaliquots (~10 µl) of capillary blood serially collected from critically ill patients to study SPM pathway activation and identify patients with dysregulated SPM responses. The second goal is to collect larger blood samples from patients with dysregulated SPM responses to perform deep phenotyping (functional assay, phosphoproteomics, etc.). Finally, the team hopes to uncover relationships between SPM pathway function and sepsis resolution focusing on phosphoproteomics.

In the R21 phase, Dr. Voldman has focused on technical de-risking, including demonstrating the feasibility of making functional measurements on 10 µL fingerstick blood samples, showing that such measurements are as equally correlated to clinical severity measures as measurements taken from 50 µL blood volumes, and demonstrating that high-quality phosphoproteomic data can be obtained from a small number of polymorphonuclear neutrophils (PMNs).

The R33 phase will include the collection of longitudinal fingerstick samples to monitor leukocyte activation at frequent intervals, select patients with dysregulated SPM responses for additional phenotyping, and understand the function of the SPM pathways during sepsis resolution. Disease-related controls and healthy subjects will be included as a reference for functional SPM responses.

Dr. Voldman described the technology development needed to complete these studies. Given the low blood volumes and cell numbers, the team needed an efficient way to fractionate and collect neutrophils. They chose to use multidimensional double spiral (MDDS) devices, which are microfluidic devices that use inertial forces to separate the large PMNs from small red blood cells. Whole blood is diluted in PBS and then passed through the spiral chamber. Cells of different sizes are pushed along the wall of the spirals at different speeds. The team has successfully used MDDS devices on 50 uL of whole blood but wanted to be able to use the devices on 10 uL (Jeon et al., Lab on a Chip, 2020). Dr. Voldman explained that the main challenge with this is the dead volume and ensuring that the final sample is concentrated enough for subsequent assays such as flow cytometry. The sample passes through the spiral chamber eight times; with each cycle, the sample decreases in volume and increases in purity (clearer color as compared to RBC-containing waste liquid).

Dr. Voldman described the performance (recovery, purity, activation) of the MDDS devices on 10 uL whole blood samples: 80% to 90% of the cells within the final samples are neutrophils, and more than 80% of the neutrophils in the initial samples are recovered. As assessed by flow cytometry using CD62L and CD11b as markers, the rate of activation of the final sample is lower than whole-blood samples. MDDS performs as well or better than other isolation methods, such as MACS isolation or density centrifugation.

Dr. Voldman explained that his team performs functional assays that correlate to the activation of the SPM pathway, including a phagocytosis assay. Previous studies have shown that the phagocytosis of pHrodo *E. coli* bioparticles is strongly correlated with clinical measures of sepsis severity (SOFA score) and SPM pathway activation (Jundi et al., Nature Biomedical Eng, 2019; Jundi et al., JCI Insight, 2021). They showed that MDDS-isolated neutrophils from 10 µL healthy donor samples had a decent amount of pHrodo updates, although to a lesser extent than that of the whole-blood samples. IRB protocol has been approved to test this function using septic samples, and the team also needs to demonstrate reproducibility in more samples.

Another functional assay the team is using is an isodielectric separation (IDS) analysis. This is a microfluidic assay that distinguishes cells based on their electrical properties, which is a proxy measurement for resting versus activated states. Dr. Voldman showed that using cells isolated from small volumes of healthy and septic patient samples, septic neutrophils had a more activated state than the controls, as the electrical properties of these cells are more aligned with activated positive controls. The team is collecting more samples to validate this result.

The main question of the mass spectrometry studies was whether high-quality phosphoproteomic data could be generated from small numbers of neutrophils, since little prior work has achieved that from samples <5 ml. In this protocol, neutrophils are isolated from a whole-blood sample by MACS isolation, left untreated or activated with phorbol myristate acetate (PMA), and lysed. The resulting proteins are trypsinized, labeled with Tandem Mass Tag labeling, and profiled by LCMS.

Dr. Voldman showed early results of the phosphoserine/phosphothreonine (pS/pT) mass spectrometry. The team can measure more than 1,000 to 2,000 pS/pT sites on 10 to 20 µL of blood and track differential expression of proteins after PMA activation compared to untreated controls. STMN1-pS25 was used as a positive control of the PMA-activated whole blood to demonstrate the good quality of the MS detection. The team is now ready to tackle unbiased phosphoproteomics in MDDS-isolated neutrophils from fingerstick samples. Dr. Voldman explained that there is an additional immunoprecipitation step in the sample preparation to study phosphotyrosine (pTyr) due to its low abundance. However, the team is able to reproducibly measure pTyr from 500 µl PMN samples, which is the first in the field. Multiple pTyr on ERK1 and ERK2 were detected after PMA treatment, as expected.

As the team awaits the R33 phase, they have begun to investigate if resolvin D1 (RvD1), a lipid mediator of SPM pathways, can reverse the effect of PMA activation on phosphorylation patterns in healthy donor cells. The team found that RvD1 does not reverse PMA-induced activation globally; instead, RvD1 exposure after PMA activation upregulates pTyr on many peptides, which may work to counteract the signaling pathways activated by PMA. The team is ready to look at what happened in septic samples.

Dr. Voldman gave a summary of his presentation. To date, the team has met all five milestones related to their grant and established the ability to prepare 10 µL samples and analyze them for PMN functions and SPM pathways. These technologies have broad applicability, so Dr. Voldman invited those on the call who are facing challenges in low sample volume—such as those studying pediatric and neonatal populations—to reach out if they are interested in reducing the sample sizes needed for their assays. The team can also assist with the phosphoproteomics aspect from small samples, if that is of interest.

Dr. Files asked if Dr. Voldman had compared capillary samples to conventional venous samples on their assays. Dr. Voldman agrees that differences are expected between these two samples. The team has not used fingerstick blood (capillary) in the R21 phase, so the comparison will be done in the R33 phase.

Dr. Bastarache asked if these assays must be run on fresh samples or if samples can be stored and/or shipped first. Dr. Voldman explained that neutrophils are a sensitive cell type, so they are processed and fractionated immediately. For mass spectrometry, fresh samples are immediately processed and then stored for further processing and analysis in bulk. Dr. Bastarache wonders if the MDDS devices can be shared with other investigators for them to use. Dr. Voldman explained that the MDDS device essentially replaces the MACS procedure of cell isolation. It would be best to ship samples once they are processed into more stable states such as proteins or peptides. Dr. Bruce Levy noted that the team has shipped fixed cells at -80°C.

Dr. Patrie asked if the team has studied the sensitivity of the phosphoproteomic platform to sampling handling or storage conditions. Dr. Voldman said that the phosphoproteomics team may have looked at many variables that have affected their results, including the MACS kit they used for cell isolation and time from collection to processing. He invited Dr. Patrie to reach out to Dr. Forest White (MIT) on the phosphoproteomics team to talk more about this.

Dr. McMahon asked if cells run through the MDDS devices are subject to shear forces. Dr. Voldman explained that there are shear forces in this device because those forces are important for separating cell types. However, the force is not enough to activate PMNs in the cell isolation process.

Dr. Furdui asked how many WBCs were isolated from 10 µL blood for phosphoproteomics analysis and if this was representative of WBC heterogeneity in larger volumes. Dr. Atreya followed up by asking if these approaches could identify neutrophil subsets. Dr. Levy explained that the team is using the phosphoproteome to begin to define neutrophil subsets. He noted that some neutrophils in the IDS analysis appear to be at a resting state and others appear active, so it would be possible to separate the cells based on their IDS profile and study them further.

Dr. Bollyky asked if using small starting sample sizes introduces challenges to sample processing and storage. Dr. Voldman replied that there are technical challenges to using small samples.

Optimizing Methods of Clinical Sample Processing for scRNA-Seq and Mechanistic Studies in Sepsis to Enable Reliable, Reproducible, and High-Yield Multi-Center Collection Efforts

Michael Filbin, M.D., M.S., Harvard Medical School, and Massachusetts General Hospital

Dr. Filbin introduced his study, which is designed to find practical methods for scaling up single-cell RNAseq pipelines to enable future multicenter studies. The team has expertise in emergency medicine, scRNAseq, biorepository, immunology, and infectious diseases. He stated that bulk RNAseq has historically been the method of choice for characterizing the host immune response in sepsis and deriving immune endotypes. However, bulk RNAseq yields gene signatures averaged over all circulating immune cells. Single-cell RNAseq, on the other hand, can measure gene expression in individual cells (more than 1,000 cells per sample) and give a precise readout (quantum leap) of gene expression programs within specific immune cell types and cell substates. An analogy is a fruit smoothie for bulk RNAseq versus a fruit salad bowl for scRNAseq.

Dr. Filbin explained that his team started this line of inquiry with a urosepsis patient cohort with varying degrees of severity, which has relatively uniform source infections. They measured single-cell transcriptomes of these samples and discerned cell lineages on single-cell t-distributed stochastic embedding (tSNE) plots and 2D uniform manifold approximation & projection (UMAP) visualization. Homing into the monocyte linage, expression-based clusters emerged, which were mapped into a clinical "sepsis density" per cell (Reyes M et al. Nat Med. 2020; Reyes M et al. Sci Transl Med. 2021). They identified a substate of monocytes, called MS1 monocytes, with a unique transcriptional profile that is enriched in patients with urosepsis and corresponds to the highest sepsis density. Dr. Filbin noted that MS1 monocytes can also be readily identified in flow cytometry, and they also present in sepsis of other infectious sources. Through mechanistic and functional studies, the team has shown that MS1 monocytes are immunosuppressive.

Dr. Filbin described how the team applied these findings to a larger cohort of acutely ill COVID-19 patients. When the team clustered patient transcriptomes, three patient clusters emerged: those with elevated levels of MS1 monocyte substate, those with high expression of IFNγ, and those with low expression of IFN-α/β/γ. The high MS1 monocyte endotype correlated with COVID-19 severity. These results showed that scRNAseq clustering can not only unravel cell-level heterogeneity but also help resolve patient-level heterogeneity.

Dr. Filbin noted the benefits and drawbacks of scRNAseq, which can help resolve the heterogeneity of host immune responses in sepsis and can provide viable cells for functional studies to understand mechanisms. However, it is difficult to scale, given the current state of sample processing methods. Dr. Filbin's goal is to make scRNAseq scalable to support multicenter investigations.

Dr. Filbin listed the goals for the R21 phase of the grant. First, they will develop an optimized blood processing method to support scRNAseq in multicenter sepsis studies. Next, they will validate a biological readout of the optimized processing method compared to standard methods. Finally, they will assess the functional viability of cells harvested through this method to support mechanistic studies.

Dr. Filbin described the current approach to scRNAseq studies. Whole-blood samples are collected from patients and must be processed at the clinical site by research technicians, a process where PBMCs are isolated from fresh whole blood through Ficoll-gradient centrifugation that takes 2 to 3 hours. PBMCs are frozen at -80°C and then shipped to a central laboratory for batched sequencing.

The proposed new method for blood processing is dubbed whole-blood cryopreservation. In this technique, 10% DMSO is added directly to the whole-blood sample immediately after it is drawn and frozen at -80°C. Frozen samples are then shipped to Broad for post-thaw PBMC isolation and sequencing.

The team is working to optimize and validate the post-thaw processing workflow for scRNAseq following whole-blood cryopreservation. To do so, they are enrolling septic patients in the ED and collecting 2 mL of whole blood to be cryopreserved and 5–8 mL of whole blood to be freshly processed onsite through standard PBMC isolation through Ficoll separation. To date, more than 40 patients have been enrolled, and data have been collected on 6.

Dr. Filbin and his team tested three different methods for post-thaw immune cell isolation. The first method, direct-flow cytometer sorting of thawed blood, was proven not viable because the samples often clogged the cytometer. The second method, standard Ficoll-gradient centrifugation on thawed samples, did not recover high cell numbers. Finally, magnetic cell sorting (MACS) seems to yield the highest cell counts with the greatest cell viability. comparable to those of freshly processed samples. In addition, this method is fast (only requires 20 to 30 minutes to perform) and is compatible with downstream FACS and scRNAseq procedures.

Dr. Filbin presented his preliminary quality metric comparisons of cryopreserved whole blood to the gold standard of fresh Ficoll-separated PBMCs. The two sets of samples displayed similar quality, including by metrics such as percent mitochondrial reads, genes per cell, and unique transcripts per cell. Looking at biological readouts, such as cell substate UMAP visualization, the clustering based on single-cell profiling is very similar across the two methods. The density of the clusters is higher for whole-blood cryopreservation because it actually yielded more cells than standard Ficoll isolation of fresh samples. When plotting the fractional abundance of different cell substates, there was a good correlation between the two methods for all cell substates (R=0.87), and that for the MS1 substate has an R of 0.99. Furthermore, expression levels of the top 20 genes that define the MS1 substate were highly correlated between the two processing methods, and the distribution of expression of each gene among individual cells is also similar by each method.

Dr. Filbin noted the conclusions of the preliminary results. Whole-blood cryopreservation yields scRNAseq results comparable to standard Ficoll-based processing with minimal clinical site labor (5 minutes) and low blood volume requirements (1 mL). This approach centralizes processing to one location, which is in the hands of experienced technicians; this eliminates user/site variability and centralizes resources for processing. Dr. Filbin also gave an update on the status of their specific aims. They are scheduled to complete sample collection in December 2023. Aim 1 is complete, as they have identified MACS sorting as the optimal post-cryopreservation cell separation method. Aim 2 is ongoing,

and the team plans to validate transcriptional signatures between the two methods with a larger sample size. Aim 3 will study the functional responses of cryopreserved PBMCs to stimuli.

The future directions of this project are to develop a multicenter ED enrollment approach at other sites to support scRNA-seq studies, which allows for deeper investigations of scRNA-seq endotyping in sepsis, which will generate more refined biological hypotheses than bulk RNAseq. In the future, these methods could be embedded into randomized controlled clinical trials to harness the power of randomization to study endotype-specific responses to therapy.

In conclusion, Dr. Filbin remarked that given the nature of the disease, sepsis investigators will never be able to reliably control when patient samples are collected and processed, so a streamlined approach like this may offer a practical solution to the arduous task of PBMC isolation for scRNAseq.

Dr. Seymour asked if the sepsis density is a single-cell measure or a clinical measure. Dr. Bhattacharyya explained that the dots represent single cells, which are colored based on retrospective independent clinical adjudication of sepsis severity of the patient from which each cell was derived. In other words, cells with the MS1 signature predominantly came from patients with sepsis as adjudicated clinically. Dr. Bollyky commented that longitudinal scRNAseq studies to see whether heterogeneous responses reflect patient heterogeneity versus different exposures would be desirable. Dr. Atreya commented that the ability to track the temporal dynamics of cell substate changes during the course of crucial illness will be fascinating.

Drs. Files and Burnham asked if the time from collection to processing or time stored in -80°C, affects transcriptional readouts. Dr. Filbin noted that these questions have not been studied yet. Dr. McMahon added that they have been working on site-friendly methods that separate multiple compartments (plasma, PBMCs, PMNs, RBCs) before cryopreservation, and perhaps post-cryopreservation processing is better. Dr. Filbin is happy to share best practices around post-cryopreservation processing methods.

Dr. Levy asked if the whole blood thaw induces platelet activation and if so, if the platelet activation has an impact on leukocyte activation. Dr. Bhattacharyya explained that platelets are sorted out through flow cytometry prior to sequencing, but that they may still have an effect. However, based on similarity of immune signatures in the post-thawing group versus the fresh Ficoll isolation, this impact seems to be minimal, but more work is needed to confirm this.

Dr. Segal asked if the team has tried cell hashing, which is a limiting factor to scaling up and keeping costs low. Dr. Bhattacharyya explained that the team does 8-plex hashing, meaning that there are eight samples loaded per 10x Genomics run. They also do CITEseq to measure surface protein epitopes. He also noted that they had only done antibody-based cell hashing using blood samples, and both hashing and CITEseq are compatible with post-cryopreservation thawing. Dr. Segal added that cell hashing in BAL samples is even more challenging.

Dr. Langelier asked if the team compared different methods (other than DMSO) for whole-blood cryopreservation. Dr. Filbin explained that the team did a literature review and found data showing that DMSO cryopreservation yields high cell viability for fresh tissue. They think cryopreservation of whole blood should be fairly similar, although there is limited literature on the exact topic.

Dr. Atreya pointed out that the current challenges of the bulk RNAseq method are that results are often overwhelmed by neutrophil signatures, and asked if scRNAseq has any means to overcome that, such as

using hashing. Dr. Filbin noted that neutrophils usually don't survive the freeze and thawing process. They have separated neutrophils and have done bulk sequencing on them before, and they are comparing the cell substate signatures derived from scRNAseq to bulk RNAseq on the same samples. Dr. Bhattacharyya noted that bulk deconvolution has worked surprisingly well to overcome this because there are still enough transcripts that are not expressed in neutrophils (despite many being missing), which can be used to infer other cell populations.

Dr. McMahon asked to what extent red blood cells in these samples are lysed and/or recovered. He noted that red blood cell lysis may influence leukocyte activation. Additionally, some researchers are interested in studying red blood cells in sepsis. Dr. Filbin explained that the first step in the post-thawing protocol is to remove the RBCs so that the impact on leukocyte activation would be minimal. They observed a fair amount of RBC lysis, judging by the redness of the samples, but have not tried quantifying it.

Redox Trapping for Biospecimen Preservation and Innovation in Sepsis Care

Cristina M. Furdui, Ph.D., Wake Forest University School of Medicine
D. Clark Files, M.D., Wake Forest University School of Medicine

Dr. Furdui introduced the study team and the project, which is aimed at developing a new formulation for blood collection to enable redox-related discoveries in sepsis. Dr. Furdui noted that this is designed to be an alternative to commonly used blood collection additives, such as EDTA or heparin, because they are not amenable to studying questions of redox metabolism.

Dr. Furdui explained why the team is interested in redox metabolism. First, reactive oxygen species (ROS) and redox mechanisms have well-established associations with infections and sepsis. Redox mechanisms are critical to control bacterial infections (e.g. bacterial killing) and are also fundamental to augment the host responses to infections. Some household disinfectants, such as Lysol cleaner and bleach, are powerful oxidants (H2O2, HOCI, and silver nitrate). Some pharmaceuticals actively modify redox metabolism. For example, acetaminophen can reduce the ferryl-protoporphyrin radical to lower sepsis-induced oxidative injury and was tested in the PETAL trial ASTER, pending results. Selenium may not be viewed as a conventional antioxidant, but it is essential for several antioxidant proteins including glutathione peroxidase and thioredoxin reductase. Selenium showed some benefits in mortality and length of hospital stay in meta-analysis of clinical studies. Ascorbate (vitamin C) was tested in the CITRIS-ALI (septic ARDS) and LOVIT (septic shock) trials with divergent results, which may reflect the differing oxidation states of the patient populations. Because not all antioxidants are equal, it is important to know when and how much of this treatment to administer.

Dr. Furdui's team is focused on better understanding the redox metabolism in sepsis to further understand how therapies may modify the redox state in sepsis. This is particularly relevant for sepsis, since redox metabolism is continuously being remodeled during the disease progression, with an oxidative phase occurring within hours of initial infection and a reductive phase later in the disease course. As a result, clinicians need to time and match redox-based treatments to redox state. Redox-based diagnosis can inform the selection and timing of treatment with redox-based therapeutics. Additionally, they can increase understanding of the heterogeneity of the pathogenesis and resolution of sepsis. Despite this need, the current methods of specimen collection fail to capture the intrinsic and disease-acquired heterogeneity of their redox state. For example, there are reported differences in redox

state based on age (increased ROS with aging, Jones DP. Redox Biology 2015), gender (higher homocysteine in males and with aging), and genetics (G6PD and FADS polymorphisms create decreased capacity to generate NADPH and increase n-6 PUFA utilization leading to synthesis of pro-inflammatory, pro-oxidant oxylipins). Coincidently, age and gender have an impact as susceptibility factors for sepsis as well.

Identifying a new specimen collection technique that can trap the endogenous redox processes and prevent reactions that damage the specimens over time would have several benefits, including the advancement of clinical sepsis research and a reduction in preanalytical variability due to *ex vivo* metabolism and/or degradation of clinical specimens, which is a critical step toward personalized medicine. In addition, this advancement could lead to the discovery of new biomarkers to aid in disease staging and the development of new treatment regimens.

Dr. Furdui explained that the team chose to focus on blood because it is a readily accessible sample type and is a top specimen for diagnostic use and biomarker discovery that is subjected to long-term storage. However, blood is highly unstable and prone to oxidative damage due to its high iron content.

To overcome these challenges, the team adopted a multifactorial approach that would preserve redox during blood, including steps to prevent artifactual ROS formation during collection and storage, stop the propagation of single-electron/radical reactions (e.g., lipid peroxidation), and block reactive centers in proteins and small molecules. Dr. Furdui expects that the final product will meet a few key requirements: quench the redox state, not interfere with or even improve current "gold standard" clinical lab assays for sepsis patients, and be relatively low cost.

Dr. Furdui listed the team's experimental goals for the R21 phase:

1. Achieve efficient trapping of endogenous redox state or under oxidant challenges.
2. Ensure quantitative redox quenching of a variable volume of blood.
3. Produce higher-quality specimens using the redox formulation (RMX) when specimens are exposed to repeated freeze/thaw cycles.
4. Enroll sepsis patients into the single-center REDOX-Sepsis study, and test the RMX for a variety of clinical and research laboratory measurements (compare with EDTA in collection tubes).

In the R33 phase, the team plans to scale up the testing of RMX at multiple sites to create a biorepository for redox-focused investigations.

Dr. Furdui explained that the team uses two redox sensors: 2-Cys peroxiredoxins (Prx, breaks down ROOH) and glutathione (important for suppressing lipid peroxidation through Gpx4, for example). Both of these molecules are highly abundant in cells, recycled by NADPH, critically regulate important signaling pathways, and cycle between a reduced and oxidized state. The oxidized Prx dimer and reduced Prx monomer can be distinguished through a nonreducing protein gel and Western blot, while reduced and oxidized glutathione (GSH vs. GSSG) can be detected by mass spectrometry.

Dr. Furdui outlined the team's progress toward their four goals. They collected healthy donor blood into EDTA tubes and RMX tubes. RMX achieved rapid and stable redox quenching, as seen by more Prx monomers and increased GSH/GSSG, as compared to EDTA/Heparin samples. The team also challenged the blood samples with various concentrations of hydrogen peroxide and LPS. The blood samples

collected into the RMX solution saw less oxidized Prx and GSSG, indicating that the RMX prevented the artifactual oxidation postcollection.

The team has also made progress toward their second goal of ensuring that excess blood volume does not affect redox quenching. Dr. Furdui displayed data showing that there is no difference in Prx oxidation between 1 mL blood samples and 1.25 mL blood samples treated with the same volume of RMX mixture.

For the third goal of determining protein stability after freeze/thaw, the team has collected preliminary data on the stability of vascular endothelial growth factor (VEGF) after one or four freeze/thaw cycles after 1 month at -80°C, which was known to be highly sensitive to freeze/thaw. The RMX significantly improved the stability of VEGF with repeated freeze/thaw cycles. In the future, the team will study the stability of VEGF after freeze/thaw cycles after 3 months, 6 months, and 12 months at -80°C.

Their fourth goal is to enroll sepsis patients under verbal informed consent into the REDOX-Sepsis study and test RMX-treated blood samples on clinical and research measurements. Blood collected from a subset of 50 patients was to be measured for a complete metabolic profile (CMP) plus lactate; blood collected from the remaining 25 patients was to be tested for a complete blood count (CBC) plus lactate. Approximately 10 mL of blood is drawn from each patient. The blood samples were to be distributed between RMX tubes and standard protocol tubes for the intended assays (EDTA, NaF/Kox, or heparin). Enrollment was stratified by age quartiles and race, which is representative of their ICU demographics. Dr. Furdui explained that the team has collected data from a total of 25 patients thus far. She expects to complete enrollment before the end of 2023.

Dr. Furdui highlighted preliminary findings from this study. Lactic acid, an important measurement for the Sepsis-3 definition, is subject to artifacts from many sources, including processing delays and temperature. RMX on ice prevented a significant amount of lactic acid accumulation post-blood collection as compared to NaF/Kox on ice, while the difference was even larger when both reagents were at room temperature. Lactic acid tested with a handheld i-STAT device showed similar improvement when RMX was added to the samples. Some analytes within the CMP and CBC panels showed equal performance to standard collection methods. RMX improved detections of Lactate, Creatinine, and Bilirubin. Since RMX contains Na and K+ counterions, proper detection of affected ions in the CMP panel may need minor adjustments of the RMX ingredients. In addition, RMX performance for the AlkPhos and AST (SGOT) tests needs further investigation.

Using sepsis samples, RMX captures reduced Prx monomer better than the EDTA collection method, which showed heterogeneous expression among patients. Using 4-HNE as a marker, RMX was shown to quench PUFA lipid peroxidation, which is expected to improve specimen stability and quantification of pro- and anti-inflammatory lipid signaling mediators (e.g., oxylipins).

Dr. Furdui concluded that RMX has met the three key requirements of an ideal redox-trapping product for blood preservation, as mentioned above. It has the potential to improve both clinical research and clinical care. She envisions RMX as a universal blood collection tool to replace current standards. Moving forward, the team hopes to recruit 150 additional patients from five clinical sites to build a biorepository capable of supporting redox studies including multi-omics of plasma and cell fractions (225 sepsis and controls with paired samples of standard and redox collection).

Dr. Foster asked what the components of RMX are. Dr. Files explained that the RMX formulation is protected by IP, but interested investigators could contact the team for more information.

Use of a Rapid Index of Endothelial Glycocalyx Degradation to Identify Vascular Endotypes of Sepsis

Eric P. Schmidt, M.D., Massachusetts General Hospital and Harvard Medical School

Dr. Schmidt introduced his study on behalf of MPI Dr. Nathan Shapiro and explained that it is related to rapidly assessing the integrity of the endothelial glycocalyx at the bedside to better understand sepsis. He noted that the study was funded in July 2023, so it is still in its early stages.

Dr. Schmidt explained that it is well known that the endothelial surface is lined by a thick apical glycocalyx (Van der Berg BM et al. *Circ Res*. 2003), which is primarily composed of proteoglycans (proteins anchored to the cell membrane via GPI or transmembrane domains) decorated with glycosaminoglycans (GAGs) projecting into the vascular lumen. The predominant endothelial glycocalyx GAG is heparan sulfate, and to a lesser extent, chondroitin sulfate.

Heparan sulfate is a linear sugar that can be ≥40–60 saccharides long (Schmidt EP et al. Compr Physiol. 2016). It is adorned with sulfation, which gives the polysaccharide a negative charge that allows it to electrostatically bind to proteins. Through these binding events, heparan sulfate can affect the function of these proteins; for example, heparan sulfate can bind to growth factors (or coagulation factors) and either activate or inhibit growth (or coagulation). Heparan sulfate contributes to the housekeeping functions of blood vessels—that is, it maintains the endothelial barrier to fluid and proteins (Curry FR. Microcirculation. 2005), aids mechanotransduction of shear stress that produces NO (Florian JA, *Circ Res*, 2003), regulates leukocyte-endothelial adhesion (Schmidt EP et al., *Nat Med.* 2012 Aug;18(8):1217-23) and prevents intravascular coagulation (Dimitrievska S et al. *Artioscler Thromb Vasc Biol* 2016).

Dr. Schmidt explained that his laboratory has previously used novel in vivo microscopy approaches to observe the endothelial glycocalyx in the pulmonary and glomerular circulations. His laboratory found that during sepsis, induction of mammalian heparinase cleaves heparan sulfate. Heparan sulfate cleavage strips the endothelium of the endothelial glycocalyx, leading to local vascular dysfunction and lung and kidney injury (Schmidt et al. *Nat Med* 2012; Han et al. *Sci Transl Med* 2016; Schmidt et al. *J Biol Chem*. 2014; Nelson A et al *Acta Anesthesiol Scan* 2014; Schmidt et al. *Am J Respir Crit Care Med* 2016). Additionally, glycocalyx fragmentation releases heparan sulfate fragments into the circulation, where they interact with and alter the function of proteins throughout the body to induce an endocrine-like effect (Schmidt et al. *J Biol Chem*. 2014 Mar 21;289(12):8194-202). For example, his laboratory has shown that heparan sulfate fragments can penetrate the brain selectively and inhibit the function of a hippocampal brain-derived neurotrophic factor, contributing to septic neurocognitive dysfunction among survivors (Hippensteel et al. *J Clin Invest* 2019; Zhang X et al. *PNAS* 2019).

Dr. Schmidt explained that these findings were based on animal studies and small human studies, so the team wanted to confirm the findings in a large human trial. They used samples from the CLOVERS trial (Shapiro NI et al., *N Engl J Med*. 2023), which enrolled 1,563 patients in the ED; patients were randomized to receive a fluid-liberal or a fluid-conservative/vasopressor-liberal treatment. The trial was stopped early for futility, but the clinical samples were collected early and available for use.

Dr. Schmidt explained that the team used plasma collected from approximately 600 patients on Day 0, Day 1, and Day 3 of ED admission to study glycocalyx degradation. They performed mass spectrometry to quantify circulating glycans, which is laborious and took approximately 1 year to complete. Analysis of this result is expected to be done soon. The team also conducted syndecan-1 ELISA as a proxy for glycocalyx degradation. Syndecan-1 was shed from the endothelium rapidly when the sugar chains attached to it were cleaved off, perhaps due to the exposure of bare syndecan-1 to proteinase cleavage. The team found no effect of fluid resuscitation on endothelial glycocalyx degradation by syndecan-1 ELISA, which conflicts with some animal studies and retrospective or observational human studies, but nevertheless is in line with the null effects of different fluid resuscitation strategies observed in the CLOVERS trial.

Dr. Schmidt also investigated whether baseline syndecan-1 shedding can predict 90-day mortality during sepsis, and the answer is yes. Patients with the highest tertile of syndecan-1 shedding had the poorest 90-day mortality. After controlling for confounders such as age, SOFA score, race, kidney failure, and treatment allocation, syndecan-1 levels showed a linear relationship to the estimated probability of death for patients in the low to med-tertiles, which does tend to plateau among high shedders. Since 80% of patients are in the low to med range, rapid assessment of the syndecan-1 level could have great clinical utility.

Thus, the next goal for the team is to develop a point-of-care assay to rapidly measure glycocalyx degradation as a prognostic biomarker. This is challenging to measure, especially in humans. As mentioned above, mass spectrometry of GAG fragments is the gold standard as it is ultrasensitive, specific, and yields a wealth of information on the shed fragments; however, it is expensive and laborious. Intravital microscopy, which analyzes the microcirculation under the tongue in humans, is an option but this technique is potentially subject to operator variability and sublingual microcirculation is of unclear importance as a vascular bed in sepsis. The syndecan-1 ELISA is less expensive but not fast enough for clinical decision-making. The dimethyl methylene blue (DMMB) assay is fast and inexpensive and thus promising for point-of-care use. This is an existing assay previously used to detect chondroitin sulfate in synovial fluid (cartilage degradation) and validated on sulfated GAGs in the airspace fluid (Rizzo et al. *JCI Insight* 2022) and urine (Schmidt et al. *AJRCCM* 2016). However, this has never been optimized for use in human blood.

The team wants to optimize and validate the DMMB assay in human blood. It is a colorimetric assay that uses a blue dye that binds to negative charges with a particular avidity for GAGs. The DMMB assay has two potential challenges: First, the color of blood components may interfere with the colorimetric readout of the DMMB assay. Second, cell-free DNA often released from neutrophil NETs or dead cells, which is also negatively charged, interferes with the binding of the dye (Schmidt et al. *AJRCCM* 2016).

The aims of the R21 phase of this grant are to:

1. Spike normal human blood with known concentrations of heparan sulfate, then determine the accuracy of the DMMB assay with optimization steps to minimize the impact of blood color and cell-free DNA
    a. So far, the results of this study have been positive; the team has generated standard curves for sulfated GAGs in both water and pooled healthy plasma
2. Use blood from patients with sepsis (collected as part of a preexisting biorepository) to compare mass spectrometry measures of circulating GAGs to the DMMB assay

    a. This is valuable even when the assays do not agree, since the team can determine which patient subtypes or characteristics the assay will lose its accuracy
3. Determine if GAGs in the urine (collected under waivered consent) can be measured and serve as a noninvasive but accurate measurement of plasma GAGs
    a. The team had shown previously that plasma GAGs are rapidly cleared into the urine of septic mice (Zhang et al. *PNAS* 2019). They will prospectively collect matched urine and blood samples from patients with sepsis and compare DMMB results with or without the presence of kidney failure.

In the R33 phase, the team hopes to study sepsis endotypes focusing on endothelial injury. Dr. Shapiro measured eight circulating vascular injury markers from 2,095 patients with suspected sepsis, and LCA analysis of mean standardized value comparisons revealed two distinct clusters of vascular injury patterns. Dr. Schmidt wants to know if DMMB assay can quickly and sensitively discern these patient subgroups. This is designed to be a multicenter study and will enroll sepsis patients. A question of interest is which patients are likely to have endothelial injury or demonstrate endothelial resistance.

Dr. Sanchez-Pinto referenced data suggesting that saline resuscitation causes more glycocalyx degradation than balanced fluids, at least in pediatric sepsis patients. He wonders if the CLOVERS trial used saline or balanced fluids, and if that can be taken into account for data analysis. Dr. Shapiro explained that in the CLOVERS trial, balanced fluids were recommended but not mandated, so the team has the data to investigate this question and will look into this.

Dr. Atreya asked if Dr. Schmidt can differentiate circulating GAGs by their organ of origin on mass spectrometry given the recognized heterogeneity in endothelial glycocalyx within and between organs. Dr. Schmidt noted that his team has tried to detect organ signatures of enzyme-stimulated glycocalyx degradation using isolated and perfused murine organs, which is a bit easier to do in the lung due to systemic pulmonary circulation. He explained that the mass spectrometry methods would allow for this, since it can detect the ultrastructure of the shed sugar. The challenge of MS is that it cannot be performed at the bedside. He thinks that DMMB can solve this challenge, because it binds more avidly to GAGs with increased sulfation and thus may partially reflect the sulfation characteristics of circulating GAGs. Furthermore, it can potentially detect other GAGs of interest such as chondroitin sulfates. The technology is not there yet, but studying the organ signature of injury will help advance precision medicine.

Dr. Ware asked how AKI would impact DMMB assays on urine. Dr. Schmidt explained that in a previous study (Schmidt et al. *AJRCCM* 2016), he measured GAGs in the urine (collected from a prospective cohort and older samples from the ARMA study) by DMMB assay and found that the presence of GAGs in the urine predicted kidney injury 2 days early than creatinine, which he believes is a maker of glomerular injury at the time. Retrospectively, it is possible that the raised GAGs in these patients may have been due to a systemic endothelial injury endotype that predisposed patients to kidney injury—rather than detecting early kidney injury, perhaps they are detecting an endothelial-injured sepsis endotype prone to kidney injury. Dr. Schmidt also explained that a patient with measurable shedding in the urine but not in the blood may truly have kidney dysfunction, but a patient with both may suggest a systemic phenotype.

Dr. Bollyky noted that hyaluronan is abundant in the glycocalyx and serum. He asked if hyaluronan could interfere with these assays, or if they are highly specific for heparan sulfate. Dr. Schmidt explained that hyaluronan is detected on the mass spectrometry but not the DMMB assay because it is not sulfated.

Dr. Nuala Meyer asked if Dr. Schmidt was using plasma or serum and if the sample was vulnerable to hemolysis or high bilirubin. Dr. Schmidt explained that the samples are EDTA or citrate plasma. The team cannot use heparinized plasma because heparin is a glycosaminoglycan and will bind DMMB.

Dr. Files asked Dr. Schmidt if he plans to study the association of syndecan-1 with delirium or long-term cognitive outcomes with CLOVERS samples. Dr. Shapiro responded that the team plans to study long-term cognitive function with both syndecan-1 ELISA and MS GAGs. Dr. Schmidt added that they had done that in smaller human studies but wanted to confirm the results in CLOVERS. A colleague (Joe Hippensteel, University of Colorado) is studying the direct deliriogenic effect of chondroitin sulfate via activation of AMPA signaling under the support of a K08 award.

Dr. Bollyky asked if studying endothelial glycocalyx shedding could help predict disseminated intravascular coagulation (DIC). Dr. Schmidt believes that glycocalyx shedding may occur prior to DIC as supported by publications in the trauma field. A colleague of his (Kaori Oshima) is exploring the impact of glycocalyx degradation on septic coagulopathy under the support of an AHA career development award, but patient heterogeneity is a major challenge to the study.

Dr. Zhao noted that there is a device in clinical trials that mimics GAGs to bind and eliminate pathogens in circulation (Seraph100 Microbind Affinity Blood Filter) (Chitty SA et al. *Crit Care Explor*. 2022 Mar 25;4(4):e0662). She asked if shed glycocalyx binds to pathogens. Dr. Schmidt explained that his team is very interested in studying how GAGs interact with pathogens. He noted that gram-negative bacteria are associated with higher circulating GAGs than gram-positive bacteria, but both had higher levels of GAGs than culture-negative cases. GAGs also alter pathogen behavior, either by inhibiting cationic antimicrobial properties or by directly interacting with bacterial exotoxins. His Ph.D. student published this looking at alveolar GAGs and MRSA (Langouët-Astrié C, *Cell Rep*. 2022 Nov 29;41(9):111721).

Dr. Bernard asked if syndecan-1 levels correlate with other inflammatory markers such as C-reactive protein, white blood cell count, or fever. Dr. Shapiro confirmed that syndecan-1 tracks with illness severity, and accordingly, tracks with these measures of inflammation. Dr. Schmidt added that he is looking forward to learning if glycocalyx degradation also correlates with other endotyping markers, such as Calfee classes, using samples from the CLOVERS study.

Dr. Bollyky asked how specific circulating syndecan-1 is for endothelial cells, and if it matters what cell types it originated from. Dr. Schmidt explained that syndecan-1 is not specific to endothelial cells, as it can be found in epithelial cells. In the past, the team has found a tight correlation between heparan sulfate and endothelial injury markers and hopes to find the same with syndecan-1 using CLOVERS samples.

Biobank of Small Extracellular Vesicles for Pediatric Sepsis

Basilia Zingarelli, M.D., Ph.D., Cincinnati Children's Hospital Medical Center and University of Cincinnati College of Medicine

Dr. Zingarelli began with an introduction to the project, which is designed to build a biorepository for the study of extracellular vesicles (EVs) in pediatric sepsis. Dr. Zingarelli explained that they just received the award, and she doesn't have a lot of preliminary data, so she will share her vision about the important aspects of building an EV biobank. EVs are nano-sized lipid bilayer membrane vesicles that are released into the extracellular environment by nearly all cell types. EVs are classified by cell type of origin, size

(diameter), and function. Types of EVs include exosomes (known as small EVs, 40–120 nanometers), microvesicles (50–1,000 nanometers), and apoptotic bodies (500–2,000 nanometers).

Dr. Zingarelli noted that her focus is on exosomes, which are homogenous in structure but retain characteristics of the membrane from their cell of origin. For example, they contain transmembrane proteins, clusters of differentiation (CD) markers, integrins, and immune components such as major histocompatibility molecules. sEVs also contain cargo comprised of DNA, RNA, amino acids, proteins, lipids, and metabolites.

sEVs play an important role in cellular communication and regulation. When released from a cell, sEVs can interact with target cells and modify their function through receptors, phagocytosis, or release of proteins or RNA through genetic reprogramming of the target cells. These communication events have been found to play an integral role in physiology with respect to immune surveillance, blood coagulation, stem cell maintenance, and tissue repair.

Dr. Zingarelli explained that sEVs also play an integral role in pathophysiology because their cargo changes during disease to contain DAMPs. This has been demonstrated in cancer metastasis (started the exosome research), autoimmunity, neurodegeneration, sepsis, and infection.

Because sEVs are released into the extracellular environment, they can be extracted from any patient fluid, including blood, urine, and lymph, for analysis. In addition, because sEVs cargos retain the signatures of the cell types they were released by, they can give insight into the mechanism of organ injury.

Dr. Zingarelli presented a study from South Korea with more than 200 patients showing that the concentration of sEVs in plasma correlates with the severity of organ failure and mortality in sepsis patients, which increases linearly with patients went on to septic shock (Im Y et al. *J Cell Mol Med*. 2020). Other sepsis sEV studies have demonstrated that most sEVs in sepsis originate from platelets, monocytes, neutrophils, and lymphocytes, but characterizing cell origin types is difficult.

The two questions this study will work to answer are:

1. Are sEVs released into the circulation from other injured organs such as the liver or kidney?
2. Can sEVs be characterized by their compartment-specific cargo?

To begin to study these questions, the team used an animal model of sepsis to learn how feasible it would be to isolate high-quality sEVs for -omics studies. Through an RNAseq analysis of the protein-coding RNA transcripts in the sEVs, the team found differences in RNA profiles between control and septic mice. The differentially expressed genes were indicative of an increase in acute inflammatory response, cellular response to cytokines, and homeostasis. Additionally, the sepsis group saw increased expression of haptoglobin, suggesting liver damage (organ origin).

Since animal models do not effectively replicate the heterogeneity of human sepsis, Dr. Zingarelli collaborated with Dr. Hongkuan Fan (Medical University of South Carolina), who has access to an existing biorepository of samples from adults with sepsis. They harvested sEVs from frozen serum samples of septic patients and treated endothelial cells in culture. Surprisingly, they found that only the sEVs from ARDS patients can cause endothelial cell damage (Li P et al. *Front Immunol.* 2023; 14:1150564), while there were no correlations with mortality and acute renal failure when cells were treated with septic

serum. The team determined that the majority of circulating sEVs originated from endothelial cells based on surface CD markers and that they were loaded with active caspase-1, a major contributor to endothelial dysfunction. Caspase-1 activity was also correlated with ARDS but not with mortality or acute renal failure.

Dr. Zingarelli explained that the team also studied sEVs harvested from plasma samples from pediatric sepsis patients to further understand the function of sEVs. This time, they incubated the sEVs with human macrophages in culture. The macrophages adopted a pro-inflammatory (M1) phenotype with increased expression of TNFα.

These samples were part of an existing biorepository at Cincinnati Children's Hospital containing more than 22,000 samples (plasma, serum, RNA, and DNA) collected from more than 2,000 patients with associated clinical data. More than 600 of these patients were stratified through the severity risk model associated with the PERSEVERE trial, and more than 400 patients have been analyzed by Nanostring technology to identify their proinflammatory gene expression profiles. Dr. Zingarelli believes the sEVs data would add granularity and value to this biorepository, but a challenge is that there is no standard method for characterizing sEVs (e.g., isolation or direct detection, size, quantity/count, phenotype, morphology, single EV or bulk analysis, variability in the very complex RNA cargo profile, point-of-care potential, etc.).

In the R21 phase of this grant, the team plans to establish a cost-effective, simple, and rapid method for sEV collection. They will determine the recovery and purity of the resulting sEVs and their functional characteristics (physical and physiological properties including their inflammatory profile). As a part of this study, they will compare ultra-centrifugation (gold standard) to a commercial kit based on phosphatidylserine affinity extraction, and process serum versus plasma samples from the same patient (serum contains more platelet-derived sEVs). Additionally, they will demonstrate the suitability of banked sEVs for high-throughput analyses of their RNA cargo profile.

Aim 3 of the R33 phase will characterize functional endotypes and RNA profiles of sEV in critically ill septic patients with specific organ injuries to see if these can be used as classifiers for organ dysfunction, and Aim 4 will use a prospective cohort to confirm this in patients with septic shock, with a special emphasize on cardiovascular organ failure. Dr. Zingarelli finished with a final message that sEVs are useful biomarkers for liquid biopsy and can inform clinicians about the mechanism of injury, and she hopes that they can develop a point-of-care device to detect sEVs.

Dr. Ware asked how Dr. Zingarelli avoids platelet activation and release of EVs into samples *ex vivo* in studies of serum. Dr. Zingarelli noted that this is an ongoing question and part of the reason why she wanted to compare serum to plasma. Another option is platelet-depleted samples. Dr. Ware suggested a comparison to platelet-poor plasma to determine what fraction of EVs are from platelet activation and release *ex vivo*, which is a major issue from her own experience. Dr. Zingarelli noted that they are always very careful to list the characteristics of the samples (e.g., organ origin, therapies received, clinical data, cell count). It will also be possible to determine the cell type of origin through flow cytometry in addition to transcriptomic analysis of the cell origin. Drs. Bollyky and Langelier noted in the chat that it might be important to distinguish between sEV of human origin and outer membrane vesicles of microbial origin, perhaps through testing the exosomes containing microbial nucleic acid.

Session III Panel Discussion

Session Co-Chairs:
Eric P. Schmidt, M.D., Massachusetts General Hospital and Harvard Medical School
Paul Bollyky, M.D., Ph.D., Stanford University Medical Center

Dr. Bollyky began the Session III Panel Discussion by briefly summarizing the presentations in Session III, noting that he is excited about the different concepts, ideas, and biomarkers coming out of this session. He also listed the common themes that arose during the session:

1. New and repurposed technologies (e.g., microfluidics, GAGs, EVs had been used in other contexts) have the potential to improve sepsis diagnostics and biospecimen collection.
2. Obtaining sufficient volumes of clinically relevant biospecimens is as challenging as the technology needed to analyze the samples.
3. Sample processing will need to be tailored to the particular readouts (redox states, cfDNA, etc.); a one-size-fits-all approach will not suffice.
4. New targets are on the horizon, such as bacteriophages, EVs, GAGs, and redox states.

Dr. Bollyky listed some ideas presented so far for sample collection and processing that allow analysis using emerging and future technologies—for example, novel sample processing approaches including whole blood cryo and REDOX-sepsis. He added that these approaches must be assay-appropriate (cfDNA and scDNAseq have different needs). He asked the group to discuss the following questions:

1. Are platforms or synergies between methods possible? How about standardized protocols or repositories amenable to all the research questions being pursued here?
2. How important is speed? Should that impact technology prioritization?

Dr. Schmidt stated his worry that without unified sampling and processing standards, investigators risk "chasing the noise," given how granular the current technologies are. He recommended that the team assess the need to set standards for evaluating sepsis diagnostics and ways to adhere to these across multiple study sites. Dr. Bollyky asked investigators who are setting up biorepositories (these were presented on Day 1), if it is possible that the standards can be coordinated for these different types of new technologies presented on Day 2.

Dr. Segal reiterated that protocol optimization depends on research goals and questions, the timing, and the technologies being used, meaning that what works best for one study may not work for another. However, he thinks that an integrated meeting like this one is very helpful so people can share their experiences. Within a consortium where a group of investigators comes together to achieve the same goal (e.g., APS), they can set a master protocol that is modifiable to allow for the add-on of different approaches. It is important to have different groups contributing to the biorepository effort.

Dr. Atreya remarked that teams should take advantage of the synergy that exists between their studies. For example, he and Drs. Schmidt and Zingarelli are all (in their own way) studying organ injury signatures. Instead of measuring systemic GAGs or biomarkers of specific organ injury separately, they should synergize their different aspects of the study to add more predictive power to these biomarkers.

Dr. Sanchez-Pinto noted that many of the investigators on the call were trained by or collaborated with the late Dr. Hector Wong, who was a pioneer in sepsis research and biorepository development. All these investigators have different expertise, a very diverse group from multiple institutions. Following Dr.

Wong's legacy of collaboration, Dr. Sanchez-Pinto expressed his interest in building a cohort discovery tool that tracks existing sepsis biorepositories and allows investigators to see what types of samples are available for shared use, what metadata are available, how the samples were collected, and whom they should contact to request samples for their particular assay of interest. There are the latest bioinformatic tools available to tackle such a project, which would foster future collaborations.

Dr. Bastarache asked the group on the call which approach they believed would lead to a breakthrough more rapidly: multiple investigators using uniform approaches to sample collection, data collection, and analysis; or investigators all working on different high-risk, novel projects. Dr. Jennifer Kaplan noted the importance of testing procedures at multiple institutions to ensure reproducibility and generalizability. Dr. Voldman expressed that both options are appropriate, as investigators can develop new methods and approaches, and then successful methods are shared and scaled up. Dr. Schmidt expressed his agreement with Dr. Voldman and said that consistency and novelty are not mutually exclusive.

Dr. Files agreed with the discussion and wondered how best to embrace innovative and highly sensitive techniques without "chasing noise." He also explained that the expertise to build biorepositories and phenotyping patients is not a simple task, and there is a need for a network of experienced research centers that are poised to scale up innovative studies across the sepsis continuum. Regarding complementary approaches, the APS consortium uses standard protocol. It may be of value to build a network that could scale up the innovations that are presented here. This would avoid a scenario where novel techniques are abandoned because investigators are unable to scale up to multicenter studies. Dr. Balamuth added in the chat that an additional advantage to developing scalable and standardized protocols is that they could be activated quickly when the next pandemic arrives to develop diagnostics/therapeutics.

Dr. Burnham noted that as investigators build large biorepositories such as the APS consortium, it is important to consider both the sample collection and storage piece. As the funded study may not have the bandwidth to test all innovative technologies being developed, she recommended that the community develop new methods for collection methods that are amenable to a wide range of analytic methods and improve long-term storage that preserves sample integrity, so investigators can validate new results and test new methods using these biorepositories. Dr. Bollyky added that his career was built on the re-analysis of publicly available genomic data, so banking and sharing data are also very important to enable futuristic research projects.

Dr. Bhattacharyya added in the chat that his group stored cryopreserved cells in liquid nitrogen across a range of weeks to more than 1 year, and they have not noticed a difference in scRNAseq assays. They have not looked at the duration of -80°C storage at hospitals before shipment, though, but they have data on collection versus shipment times and can cross-compare QC metrics across samples—that is, doing storage time courses on individual samples.

Dr. Furdui explained that what drove her grant proposal—in addition to elucidating questions about the redox state—was her concern about the quality of samples that have been frozen at -80°C for many years and possibly subjected to freeze/thawing cycles. She referenced the fact that lactate is not stable in EDTA plasma (especially for costly complex analyses such as metabolic and multi-omic) and questioned if any data collected from these previously frozen samples is trustworthy. Frozen samples don't necessarily lock in metabolism, signaling molecules, or epigenetics, in her opinion. Dr. Furdui thinks

that the specimen-preservation methods are as important as the collection method to generate valid analytic results.

Dr. Patrie emphasized the importance of tracking patient and specimen information. Bioinformatics tools are continually improving and can be used to re-analyze existing datasets if they are well annotated. In his opinion, tracking the samples and data is as important as the collection process.

Dr. Zhao emphasized the importance of understanding why a technique works, not simply if it does or does not, which is important for meaningful validation of the technique later. Dr. Voldman thinks that NIH could facilitate the ability to assess the reproducibility and generalizability of promising approaches. Dr. Zhao echoed this statement and noted that rigor and reproducibility are an NIH priority. She also encouraged the wide sharing of data and samples so that investigators could test protocols.

Dr. Schmidt added that for the development of innovative technologies, it may be hard to set a common standard. However, it may be useful to set standards for controls or references. As new and innovative technologies are developed, having a standard approach to assess their utility and improvement over the existing technologies will be important. The challenge is how to determine "comparable" groups. In his case, even though EDTA and heparin sample collections are all commonly available methods, he can only choose EDTA due to the interference of heparin with the DMMB assay. It would be nice to have a common control for all these new technologies, in order to "set a floor."

Dr. Bollyky noted that sepsis is a field of medicine where time is of the essence. He asked the group if eliminating the bottlenecks in timely acquiring and processing samples and analyzing the subsequent data should be a focus. Dr. Voldman discussed that in the case of research for discovery, he thinks sampling frequently is more important (e.g., up to every hour during the disease course), while having a lag time from sample collection to analysis is reasonable. However, for techniques that are to be translated to clinical decision-making, such as the diagnostic method Dr. Bollyky's group is working on, quick assaying is very important. Dr. Voldman is optimistic that a technique that is too time-consuming with current technology may quicken as the technology advances if there is a need in that area. Dr. Lautz agreed that the speed issue is dichotomized between discovery-oriented approaches and therapy-oriented approaches (e.g., endotyping predictive of response to therapy).

Dr. Yang suggested a "working backward approach" where investigators identify high-priority analytes or readouts that are important for the sepsis community and then work backward to identify the ideal sample collection and processing approaches for that analyte.

Facilitated Discussion

Meeting Co-Chairs and All Participants

Dr. Files explained that the purpose of the facilitated discussion is to brainstorm ways to move the field of sepsis research forward to be included in a summary of the 2-day meeting. The investigators on the call agreed that the collection of biospecimens and robust clinical data is critical due to:

1. Diagnosis of what sepsis is, and how it progresses and resolves. It is important to include appropriate controls, including traditionally hard-to-obtain samples (patients in EMS, at home, controls for developing assays, etc.) and understanding the temporal development and heterogeneity of the sepsis syndrome
2. Providing a repository of samples and data that can be used to test emerging and future technologies.

Dr. Files noted that all the above activities are framed in the ultimate goal of finding effective treatments for sepsis. He went on to highlight that the goal of the last session was to leverage group input to summarize what has been learned and which best practices the group has to move the field forward, and he needed everyone to participate to make this as informed as possible. He also highlighted four recurring themes he came up with, with Dr. Bastarache, for the investigators on the call to consider:

1. Tension between feasibility (sample collection, volume, processing, timing) and novelty (iPSCs, BAL, cells); often the more novel a sample type is, the less feasible it is for many investigators
2. Clinical and phenotypic complexity of sepsis: whether to embrace it or focus on enrollment of specific subgroups
3. Optimal timing to study development, course, and resolution of sepsis; which goals of the biorepository—for example, for diagnostic, discovery, and prognostic—will affect the timing of sample/data collection, and how long will we need to follow the patients?
4. How to preserve samples for future studies that will use technologies unknown to us now

Dr. Files summarized the main messages of Session I, which focused on **novel methods in patient recruitment and sample collection**. This included a discussion on the use of remnant biospecimens and clinical phenotyping and enrollment, including discussions on protocol-based enrollment with specific inclusion criteria versus broad patient enrollment. Session I also included discussions on an EHR-based strategy for patient enrollment versus a more traditional “boots on the ground” research coordinator approach where clinical investigators identify patients to enroll. There was discussion about sample timing—that is, how early and how late and who to include. Finally, there were discussions regarding how best to acquire patient consent, an important topic that may be outside of the scope of this meeting.

Dr. Files also summarized the main messages of Session II, which focused on **comprehensive -omics analysis and data integration questions**. In this session, there were discussions on how to balance the ease of sample collection with sample quality for new studies, and how to optimize the timing of sample collection to maximize understanding of the development and trajectory of sepsis. There were also discussions about what the best -omics approaches are and how they should be analyzed, including what control groups to include and which standards to judge an -omics analysis against.

Finally, Dr. Files summarized the main messages of Session III, which focused on expanding the horizon of analytic techniques. The investigators on the call discussed how to embrace innovation and new technologies without getting lost in the noise that high-sensitivity techniques detect. Another main message of Session III was keeping track of data and sharing protocols and using the proper controls when evaluating new technologies and comparing them to "gold standards." There was also a discussion about the speed of data analysis, the importance of which depends on the research question and whether it is a discovery or a practical project.

Dr. Patrie noted the importance of having clearly defined ontologies associated with all aspects of the workflow for a biorepository, such as disease stage, sample types, and endotype definitions. This would eliminate scenarios where investigators use different language to describe the same phenomenon. This would be helpful for researchers to find the differences and similarities between their studies, especially for investigators who are new to the sepsis field.

Dr. Bastarache reopened the discussion about how much investigators should work together versus work on their own projects, and how it would be possible to embrace both approaches at the same time.

Dr. Foster noted that the investigators on the call are competing for funding through the R33 phase of this award, so it will be a challenge to balance collaboration with competition. Dr. Zhao noted that new collaborations will lead to projects bigger than the R33 phase, meaning that the future projects might be supported by funding mechanisms beyond the R21/R33 biospecimens program. Dr. Sanchez-Pinto pitched in and described parallel research projects as Venn diagrams, where projects have overlaps or connections as well as their own unique components. Once the ontologies of these projects are clearly defined, investigators are better positioned to share their data and project outcomes, at the levels of the samples, clinical interests, particular phenotypes, or at the levels of complementary expertise (e.g., technology expertise vs. patient recruitment expertise). It would be helpful if a platform like this would foster transformative collaboration in the future.

Dr. Shapiro noted that as an investigator who has been developing biorepositories for more than 20 years, he noted that the initial phase of biorepository utilization (up to ~60%) is fairly similar across sites and studies—for example, case finding and clinical data abstraction for similar sample/data types. Dr. Shapiro explained that this opens up an opportunity to leverage common infrastructure including CRF and common approaches synergistically, and facilitate the hypothesis-generating phase beyond that. He also explained that his group is currently operating under a model in which multiple hypotheses are pursued under the same infrastructure (e.g., parallel use of RBC, white cells, etc., from the same biorepository). This way, when an investigator makes a promising discovery, it can be validated through a second preexisting cohort. Dr. Bastarache agreed, and commented that these collaborations would need funding and support. Dr. Sanchez-Pinto added in the chat that it is also important to harmonize biorepository metadata. Dr. Kaplan commented that for some investigators, it's premature to consolidate at this point, as they are at different phases of the work and have different aims/goals. It is more important to learn what works/doesn't work before being able to generalize or speak to common practices. Dr. Zhao confirmed that the sepsis biospecimen program is to support the early phase of hypothesis-generating, and investigators can apply for research project awards when a hypothesis is generated.

Next, Dr. Files highlighted the opportunities and challenges of the different blood-related sample collection methods presented in this meeting—that is, remnant, whole-blood dried, prospective

research blood draw, and novel technologies such as microfluidics. Dr. Files iterated that remnant samples have broad implications even though there is only one focused talk on that topic. It is appealing because they are widely available, and no additional blood volume is needed outside what is drawn for clinical use. The challenges with using remnant biospecimens are that it is difficult to standardize clinical laboratory procedures across sites, and investigators are not in control of when samples are collected or the various times to processing. Dr. Files then opened a discussion about what the community needs to do to move remnant biospecimen collection forward.

Dr. Bastarache responded that the investigators obtained so much information from prospectively collected samples, and they are not confident about how reliable or comparable the remnant samples are to this gold standard of biorepositories plus the less innovative and exciting aspect of collecting remnant samples, so more comparisons need to be done. Dr. Seymour agreed with the need to validate the value of remnant samples and stated that peer-reviewed publications from multiple sites demonstrating stability and comparable results between these two sample collection types would be beneficial. Dr. Seymour noted that using remnant samples is a lower cost than prospectively collected samples, making them accessible to more studies and sites. Dr. Kaplan noted in the chat that stability is an important aspect of remnant sampling.

Dr. Langelier suggested comparing these collection methods in an assay-specific manner because different analytes or measurements (such as human or microbial DNAs, host RNAs, or protein) may have different sensitivities to storage conditions, and assaying using samples collected at 4 hours versus 4 days in storage may give different results. He suggested the development of assay-specific cut-offs for how long a sample can sit before assaying; Dr. Files noted that it can be difficult to get studies like this published, so an alternative is to deposit and share findings to avoid waste, perhaps in a public repository of methods. Dr. Seymour thinks that a lot of investigators in the call have access to fresh and remnant samples to compare them, and the results of such may be publishable in journals such as *Clinical Chemistry*.

Dr. Files transitioned the discussion to the merits of studying dried whole blood. The benefits of using dried blood are that it requires no additional blood volume and reflects the whole-blood compartment. However, limited published data using dried blood exists. Dr. Files welcomed the investigators on the call to share their opinions on how to move this technique forward. Dr. Foster explained that dried blood is routinely used for clinical laboratory assays but not for -omic types of studies. It is easy to add whole-blood Mitra tips microsampling to an existing protocol because the barrier and cost are low. Dr. Files agreed that layering a novel technique on top of standard prospective sampling—at least in a subset of patients—is a great way to move the technique forward. Dr. McMahon added that Mitra fingerstick sampling may allow for home and remote sampling after patients have been discharged from the hospital.

Dr. Files then reiterated some of the known benefits of prospective research blood draws, such that the investigators can control the kinetics of sampling the processing, but it is expensive—requiring frequent coordinator efforts—and not available in remote and resource-limited settings.

He then discussed some other novel technologies such as microfluidics. The benefit of these techniques is that they use very small volumes, but they are still being developed, like dried blood. Dr. Bastarache advocates for an easier sample collecting procedure, and she thinks that the sample processing and

analysis could be a more centralized endeavor. Dr. Levy added that an opportunity to operationalize this idea is to collect remnant blood at the time of intravenous line establishment.

Dr. Segal noted that suggestions should be sample-specific and approach-specific. Data from studying different compartments will be valuable to add to the whole picture. He is interested in the lower airway and collects respiratory samples based on his own expertise. He noted the importance of studying the site of disease. Because of this, he believes that any roadmaps for sample collection that the community sets should not hamper innovation. Consortium is good at establishing baseline standards, but it should be amenable to further advances in technology and modifications.

Dr. Bernard noted the importance of annotating samples in a biorepository with the patient's current clinical state, including clinical laboratory tests or physical examination. These clinical markers could add credence to research findings. Dr. Bastarache added that her team always collects a set of basic EHR data in the data collection form but has not thought about any additional clinical data that would be useful to link to biospecimens. This process built a connection between the clinical data and the biospecimens through the EHR number so they can go back to collect more clinical data if necessary. However, it may be helpful to build a list of bedside assessments for targeted areas to add to what is already collected into EHR. Dr. Files agreed, and gave plateau pressure as an important metric that is not collected in EHR. Dr. Sanchez-Pinto noted that there is an opportunity to improve EHR data collection if those added variables provide high clinical value but do not add a significant burden to bedside care (e.g., plateau pressure in ARDS patients, standardized capillary refill in septic shock, etc.).

Dr. Bastarache went on to discuss other sample types—mainly liquid samples such as urine, plasma, and HME filter fluid. They are easy to collect, feasible at all sites, have straightforward processing and storage, and are amenable to protein studies. However, these offer no cellular information and have some limitations on assays. Liquid samples can be collected from the respiratory compartment by BAL, HME, and trachea aspirate. However, BALs are invasive, time-consuming, and expensive, and HME filter fluid does not collect cells. Dr. Burnham noted that exhaled breath condensate could be parallel samples to HME filter samples because many sepsis patients are not intubated, and researchers want to capture samples in the early stage of the disease, while HME samples could be captured as follow-up samples in intubated patients. Dr. Bastarache agreed that the major limitation of the HME filter is that it can only be collected from intubated patients.

Cells collected from the blood are available anytime blood is drawn, can facilitate greater biological understanding, and can be used to generate patient-specific iPSCs. However, this collection method is limited to circulating cells and requires challenging sample processing. Dr. Segal voiced his enthusiasm for collecting BAL samples but noted that the viability of cells from BALs is variable, so BAL cells are not amenable to all assay types traditionally done on PBMCs. Nevertheless, lower airway samples are not replaceable when studying local diseases. There is a tension between simplifying the method to allow broad patient recruitment or pursuing a more in-depth study on a more refined patient population. The latter is more important for the pursuit of the biological truth in longitudinal studies.

Dr. Ware noted that the APS Consortium decided to collect mini BALs on intubated patients. She is eager to see the results of this study and learn how feasible it will be to collect mini BALs in a big multicenter study when some sites do not have experience collecting them. She noted that she is hopeful that the APS Consortium will collect cells to study the alveolar compartment that the HME filter misses.

Dr. Langelier noted that BALs and mini BALs are often challenging to collect. Tracheal aspirates can effectively capture host transcriptomic biomarkers of infection in patients with pneumonia and sample the lung microbiome, making them an attractive and feasible alternative to BALs in intubated patients.

Dr. Prescott remarked that she has easy access to neonatal patient cohorts but does not necessarily have the expertise to process and analyze specialized samples like BALs. She hopes that this group can work together and leverage each other's strengths to make use of specialized samples. Dr. Bastarache noted that her team can collect HME filter fluid from intubated neonates (roughly 100 µL).

Dr. Schmidt echoed Dr. Langelier's point about the importance of studying pathogens because sepsis is often driven by host/pathogen interactions. He noted that investigators should collect blood and urine to look for pathogens. As for ways to study pathogens, some new methods are on the horizon, such as the phage study by Dr. Bollyky, and there are limitations on the blood-culture results. He asked the group about the best way to analyze the pathogens in the samples. Dr. Files agreed, and emphasized that techniques for capturing the types of infections or discerning sample contamination from true infections are highly limited given the fact that sepsis is driven by known or suspected infections. Dr. Bastarache referenced Dr. Esper's presentation on Day 1, which described collecting environmental samples of the patients. Dr. Bastarache noted that these samples of environmental exposure may offer complimentary information to patients' clinical data.

Drs. Langelier, Seymour, and Bollyky agreed that studying pathogens in sepsis is important. Dr. Bollyky said that paired microbiological and patient samples would be beneficial. Dr. Seymour noted that all conceptual models of sepsis consider the pathogen aspect. Dr. Schmidt also noted that treatments against the pathogen (e.g., antibiotics) are often the most effective for sepsis. Dr. Levy remarked that for studies investigating pathogen sources, patients with bacteremia but not sepsis may be an important control group to include.

Dr. Bastarache opened a new discussion about urine and referenced Dr. Schmidt's work on studying the glycocalyx in urine. She explained that it is easy to collect but wondered what it can be used for. Dr. Schmidt discussed that the appeal of collecting urine is that it is easy to collect with a waiver of consent and can be collected longitudinally from a broad patient group. He noted the great potential of urine in a possible dipstick test (point-of-care) to assess organ injury. Dr. Sanchez-Pinto agreed that the generalizability of the dipstick approach is enticing. Dr. Kaplan echoed that her institution allows investigators to collect urine from catheterized patients through a waiver of consent because urine is considered a discarded sample. Dr. Seymour noted that the AKI community including his colleague Dr. Kellum had positive feedback about a urine biomarker test Nephrocheck based on TIMP-2 and IGFBP-7.

Dr. Burnham noted that urine can be used for toxicology and environmental exposure testing (or prior drugs) that may have altered patient phenotypes. She thinks that this kind of urine test can improve the clinical care of patients, although it does not do much to advance science directly. Dr. Ware agreed on the utility of urine and noted that she had measured many upcoming AKI biomarkers in a large cohort, and most of these biomarkers are measured in urine. She also tested urine for clinical care when part of the patient history was missing, such as cigarette smoke exposure. She encouraged investigators to continue collecting urine.

Dr. Foster noted that he has had success measuring proteomics in urine. Urine has advantages in collecting samples from remote locations, and it is stable at room temperature. In his experience,

bacteria may grow in samples that sit for long stretches of time, but host protein information is still retained. Dry urine spot samples can also be used to test biomarkers.

Dr. Shapiro noted that collecting excess urine with no predefined reason to study it may not be time- or cost effective. Because it is so easy to collect, it may be easier to start prospectively collecting when a use arises. Additionally, Dr. Shapiro highlighted the need for a hypothesis-driven study design instead of merely collecting a large amount of data with predictable results.

Dr. Zhao noted in the chat that the Technology Development R21 NOFO PAR-22-126 is a potential funding mechanism for high-risk projects such as those developing new urine assays for organ injury.

Dr. Bollyky stressed the point about including noninfectious controls for sepsis studies, such as SIRS patients with hemostatic complications but no identifiable pathogens. This type of control would add more information to diagnostic studies to help distinguish host immune responses to commensal flora or cardiogenic shock versus microbial-mediated responses. However, Dr, Bollyky noted that these controls are difficult to enroll, as they are sorted into a different group during triage, and they are highly heterogeneous. He thinks that infectious cases are not hard to find in publicly available datasets, so it is these proper controls that make study results more meaningful—that is, whether a leukocyte marker is truly an infection marker or not. Dr. Bastarache pointed out that they have always been enrolling patients broadly to get controls. Dr. Files added that a clear definition of the proper control is important since sepsis is defined as a dysregulated host response to infection.

Dr. Bastarache noted that the lines that distinguish these critical illness syndromes are blurred. Subphenotyping of sepsis and ARDS is based on probabilities, making yes/no distinctions clinically convenient but likely not biologically definitive. For this reason, Dr. Bastarache agrees with Dr. Bollyky on the need to include different control groups. In response, Dr. Filbin highlighted the need to better characterize the cohort with suspected infections. He explained that his team enrolls patients broadly by first capturing a suspicion of infection in the ED, and in his experience, the cohort with confirmed sepsis (e.g., based on bacteremia and vasopressors, etc.) is better studied. Retrospective analysis of the suspected infection cohort suggests that they might be aligned with the sepsis definition, but he didn't see many clinical reports about this group in the literature. This is a gap in the field, especially when investigators try to link clinical data with high-fidelity analytic results using samples from these patients.

Dr. Seymour agreed with the blurred boundaries of critical syndromes and added in the chat that treatment effects for marginal patients in a subgroup are quite different from the core cohort (e.g., more clearly defined). He mentioned a few relevant publications: Dr. Angus Derek's opinion paper regarding the "fuzzy logic" of the Sepsis-3 definition (Angus D.C. et al. *Am J Respir Crit Care Med*. 2016 Jul 1; 194(1): 14–15.), Dr. Chanu's work on the "infection suspicion" dilemma (Shappell CN et al. *Crit Care Med*. 2021 Nov 1; 49(11): e1144–e1150.), and Dr. Churpek's paper looking at different implementations of "suspicion" in various EMR parameterizations (Bashiri FS. *J Am Med Inform Assoc*. 2022 Oct; 29(10): 1696–1704.).

Dr. Segal noted previous work from Dr. Langelier (Langelier C et al. *Proc Natl Acad Sci* U S A. 2018 Dec 26; 115(52): E12353–E12362) showed that culture-independent methods of pathogen identification combined with host responses and microbiome can detect signatures of an infection process and hard-to-detect pathogens. He also noted that commensal flora was also shown to contribute to the host immune tone and pathogenesis, making the definition between commensal versus pathogenic less clear.

He recommends a more generous definition of pathogens and various novel methods for broad characterization of the microbial world. Dr. Langelier expressed that investigators cannot underestimate the power of integrated analysis on the host and the microbe, especially in syndromes like sepsis or pneumonia, which are driven by both factors, which is the key that drives pathology.

Dr. Levy emphasized the importance of longitudinal sampling to increase understanding of disease trajectory during sepsis. This would aid in identifying predictive signals for patients who are not responding to treatment before they develop organ failure, especially if host-directed therapies are in future management strategies. Dr. Lautz agreed, and noted that longitudinal sampling is also important given that patients present to the ED at different time points in their disease course. Dr. Voldman added that sampling should be commensurate with the dynamics of the disease whenever possible.

Dr. Files opened a new discussion about study enrollment and phenotyping. He welcomed those on the call to discuss the merits of enrolling under strict inclusion and exclusion criteria versus enrolling broad populations of acute and critically ill patients. The benefit of strict enrollment is that it generates a more clinically homogenous group, making it easier to report; however, this method may exclude appropriate patient controls and is more challenging to do. Dr. Files also encouraged discussions about EHR data extraction versus the use of traditional case report forms and coordinator adjudication. EHR data extraction is faster and more cost effective but may not include key variables that are important for the proposed studies. Case report forms allow for prospective data but require more time. Dr. Files also opened a discussion on how early (prehospitalization) and how late (posthospitalization) investigators should collect biospecimens. One concern with posthospitalization samples is that they may reflect comorbidities, not the hospital event related to sepsis.

Dr. Bernard noted that the APS Consortium has been thinking about the same questions. They ultimately decided on broad inclusion criteria but within the boundaries of critical illness based on the idea that historical definitions of sepsis and ARDS are not clearly defined but still enrich the cohorts of interests. He also noted that the APS Consortium decided to collect posthospitalization samples on only a subset of the cohort because posthospitalization samples are difficult to acquire. Dr. Bernard noted that studies of acutely ill patients have high rates of dropout at both ends of the spectrum, either due to death or fast recovery. Thus, the current patient enrollment strategies are already biased against these otherwise very useful groups, and there is selection bias for which kinds of patients will likely come back for follow-up studies. Dr. Bernard reminded investigators to keep this bias in mind when analyzing the results. He also added in the chat that the APS Consortium will ultimately collect more than 500,000 sample aliquots during the study period. Dr. Seymour thanked the APS Consortium members on the call for sharing their perspectives. He noted that the biology of postsepsis organ recovery has historically gone understudied, and the opportunity in this area of research was discussed at the 2023 Banbury Sepsis conference.

Dr. Bhattacharyya noted that a broad enrollment approach makes sense for a syndrome like sepsis where the clinical definitions and subtype definitions are blurred. However, the high cost and effort associated with a broad enrollment approach would justify some streamlining and simplification of the process. Dr. Bhattacharyya also noted that what he did for the scRNA analysis was to enroll broad cohorts but start by analyzing the well-defined cases, and then analyzing the less defined groups with unclear expectations of the results; in this way, the well-defined cases could serve as a reference to sort the more heterogenous cases. Dr. Esper noted that with the potential for disease definitions to change over

time, having broad enrollment criteria is the better choice. Dr. Bastarache emphasized that adopting broad enrollment criteria with few exclusion criteria is easier for study nurses than adopting strict inclusion and exclusion criteria. However, she has not performed a cost analysis, so she cannot speak to cost.

Dr. Lautz pointed out that broad enrollment may not be cost effective in pediatrics. He explained that there are many reasons why children are in the pediatric ICU beyond critical illness, including potentially being critically ill, making them flawed controls for a study of septic children.

Dr. Seymour noted that there is precedence for having study nurses visit patients at home for posthospitalization phlebotomy. This is cost effective and improves study adherence. Dr. Files noted that during the COVID-19 pandemic, he partnered with a company in North Carolina that made in-home visits for follow-up. Dr. Filbin noted that he has experience with in-home sample collection as well. Dr. Bernard noted that obtaining IRB approval for in-home collection is not straightforward and is even more complicated when a company is involved in the process.

Day 2 Closing Remarks

Meeting Co-Chairs:

Julie A. Bastarache, M.D., Vanderbilt University Medical Center

D. Clark Files, M.D., Wake Forest University School of Medicine

Drs. Bastarache and Files thanked those on the call for a great meeting and their contributions to the discussion. Dr. Zhao expressed her hope that all the investigators on the call have new ideas or will form new collaborations based on the discussions. She noted that sharing ontology, sample collection methods, best practices, and biospecimens will be a key part of advancing the field of sepsis research.

References

Angus D.C. et al. *Am J Respir Crit Care Med*. 2016 Jul 1; 194(1): 14–15
Argelaguet R et al. *Mol Syst Biol*. 2018 Jun 20;14(6):e8124.
Atreya MR et al. *Crit Care*. 2022 Dec 14;26(1):387
Bashiri FS. *J Am Med Inform Assoc*. 2022 Oct; 29(10): 1696–1704.).
Bos LD. *Lancet Respir Med*. 2021 Dec;9(12):1377-1386
Brant EB et al. *NPJ Digit Med*. 2022; 5: 44
Brant EB et al. *Intensive Care Med*. 2020 Apr; 46(4): 823–824
Caval T et al. *Front Immunol*. 2021 Jan 14:11:608466.
Chitty SA et al. *Crit Care Explor*. 2022 Mar 25;4(4):e0662.
Cline EN et al. *Anal Chem*. 2021 Dec 21;93(50):16741-16750
Curry FR. *Microcirculation*. 2005 Jan-Feb;12(1):17-31
DeMerle K et al. *JAMA*. 2021;325(20):2041-2042.
Dimitrievska S et al. *Artioscler Thromb Vasc Biol* 2016
Engoren M et al. *PLoS One*. 2022 Mar 11;17(3):e0265052.
Florian JA. *Circ Res*. 2003 Nov 14;93(10):e136-42.
Haddock NL. *Nat Microbiol*. 2023 Aug;8(8):1495-1507
Haddock NL et al. *mSystems*. 2023 Aug 31;8(4):e0049723
Han S et al. *Sci Transl Med*. 2016 Apr 20;8(335):335ra55
Hasin Y et al. *Genome Biol*. 2017 May 5;18(1):83.
Hippensteel JA et al. *J Clin Invest*. 2019 Apr 1;129(4):1779-1784.
Hossler P, et al. *Biotechnol Bioeng*. 2006 Dec 5;95(5):946-60.
Im Y et al. *J Cell Mol Med*. 2020;24:9439–9445
Jeon H et al. *Lab Chip*, 2020, 20, 3612
Jones DP. *Redox Biol*. 2015 Aug;5:71-79.
Jundi B et al. *CI Insight*. 2021 Aug 9;6(15):e148866
Jundi B et al. *Nat Biomed Eng*. 2019 Dec;3(12):961-973.
Langelier C et al. *Proc Natl Acad Sci U S A*. 2018 Dec 26; 115(52): E12353–E12362
Langouët-Astrié C, *Cell Rep*. 2022 Nov 29;41(9):111721.
Li P et al. *Front Immunol.* 2023; 14:1150564
Lieuwe DJ Bos. *Am J Respir Crit Care Med*. 2019 Jul 1;200(1):42-50.
Maslove DM and Wong HR. *Trends Mol Med*. 2014 Apr;20(4):204-13.
Nelson A et al. Acta Anaesthesiol Scand. 2014;58:36–43
Reyes M et al. *Nat Med*. 2020 Mar;26(3):333-340.
Reyes M et al. *Sci Transl Med*. 2021 Jun 16;13(598):eabe9599.
Rizzo AN et al. *JCI Insight*. 2022. PMID: 34874923
Sanchez-Pinto LN et al. *JAMA Netw Open*. 2020 Aug 3;3(8):e209271.
Sanchez-Pinto LN et al. *Pediatr Crit Care Med*. 2023 Jun 2;24(10):795-806
Schaller TH. *J Proteome Res*. 2019 Aug 2;18(8):3032-3041
Schmidt EP et al. *Am J Respir Crit Care Med*. 2016 Aug 15;194(4):439-49
Schmidt EP et al., *Nat Med.* 2012 Aug;18(8):1217-23
Schmidt EP et al. *J Biol Chem*. 2014 Mar 21;289(12):8194-202
Schmidt EP et al. *Compr Physiol*. 2016 Mar 15;6(2):945-73
Seymour CW et al. *Crit Care Explor*. 2023 Oct 27;5(11):e0974
Shapiro NI et al. *N Engl J Med*. 2023 Feb 9;388(6):499-510
Shappell CN et al. *Crit Care Med*. 2021 Nov 1; 49(11): e1144–e1150.
Sharma R et al. *JAMA Netw Open*. 2023;6(2):e230982.

Shen X et al. *Nat Biomed Eng*. 2024 Jan;8(1):11-29
Singh S et al. *Mucosal Immunol*. 2022 May;15(5):837-845.
Sinha P. *Curr Opin Crit Car*e. 2019 Feb;25(1):12-20
Smith LM et al. *Nat Methods*. 2013. PMID: 23443629
Van der Berg BM et al. *Circ Res*. 2003 Apr 4;92(6):592-4.
Volani C et al. *Metabolites*. 2023 Jan 18;13(2):146.
Whelan SA. *Anal. Chem*. 2023, 95, 11007-11018
Wong et al, *Am J Respir Crit Care Med*, 2015
Wong HR. *Sci Transl Med*. 2019 Nov 13;11(518):eaax9000
Zhang X et al. *Proc Natl Acad Sci U S A*. 2019 May 7;116(19):9208-9213