

Title: Advancing Multi-center Clinical Research in Critically Ill Patients through the Development of a Common Longitudinal ICU Format (CLIF)

Authors: Juan C. Rojas MD MS¹, Kaveri Chhikara MS², Patrick G Lyons MD MS³, Vaishvik Chaudhari MS¹, Nicholas E. Ingraham MD⁴, Brenna Park-Egan MS³, Catherine A Gao MD⁵, Yuan Luo, PhD⁵, Rachel Baccile, MPP², Kevin D Smith, MD², Kevin Buell, MBBS², Jay Koyner, MD², Xuan Han, MD, MS⁶, Chad Hochberg MD MHS⁷, Sivasubramaniam Bhavani MD⁸, William Parker, MD, PhD²

Affiliations:

1. Rush University, Department of Internal Medicine
2. University of Chicago, Pritzker School of Medicine
3. Oregon Health & Science University
4. University of Minnesota, Department of Medicine
5. Northwestern University Feinberg School of Medicine
6. Tufts University School of Medicine
7. John Hopkins University School of Medicine
8. Emory University School of Medicine

RATIONALE:

Multicenter critical care research often relies on sharing sensitive patient data across sites, requiring complex data use agreements (DUAs) and yielding redundant work to account for diverse data infrastructures. Common data models (CDMs) like the Observational Medical Outcomes Partnership can allow for standardized analysis scripts, mitigating these challenges but requiring advanced data engineering skills and risking a loss of crucial granular clinical information. To overcome these barriers, we present the Common Longitudinal ICU Format (CLIF), designed specifically for observational studies of critically ill patients across multiple centres. Through CLIF, we aim to streamline data organization into a longitudinal format and establish standard vocabularies to facilitate standardized analysis scripts and improve data readability for researchers and clinicians.

METHODS:

The CLIF consortium, comprising critical care researchers from seven US academic health systems, collaboratively developed CLIF's schema, clinical vocabularies, and "proof of concept" datasets. CLIF's tables emphasize care processes, clinical outcomes, and granular clinical physiology measures. One of CLIF's key contributions is an open-source web application that enables users to convert a relational database into a longitudinal dataset with custom time intervals, select study-specific variables, and choose a preferred programming language. This facilitates straightforward data processing and enables effortless cross-center comparisons and integrations, bypassing the need for DUAs when analytic queries do not need pooled patient-level data.

RESULTS:

CLIF's deployment across four of the planned eight health systems has successfully compiled a robust ICU encounter-centric relational database, documenting 87,120 ICU admissions and capturing data from 71,190 unique patients. The data across the four sites indicate a majority of male patients, with an average male percentage of approximately 53.9%. Racial and ethnic diversity were captured effectively, with White patients constituting 20.9% to 80.3%, Black patients up to 70.1%, and Hispanic or Latino patients between 5.9% and 19.8% across the medical centers.

CONCLUSIONS:

By standardizing and harmonizing data, CLIF streamlines single- and multi-center clinical research on critically ill patients. CLIF establishes agreed-upon standards for data elements and formatting, simplifying cohort identification and data cleaning. This report on CLIF's design and development serves as a conceptual roadmap and practical resource for observational healthcare data studies in critical illness. CLIF will underpin various critical care epidemiology, outcomes, and informatics publications, supporting prediction models, clinical sub-phenotyping projects, and comparative effectiveness research based on observable practice variation.