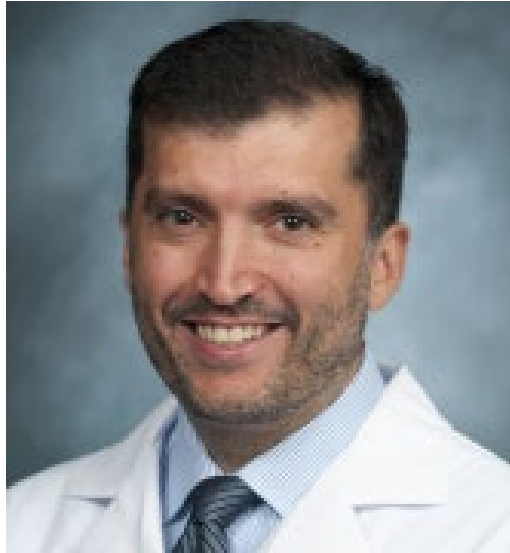


# ATS Highlights 2022: Critical Care Assembly Early Career Professionals



## Majid Afshar, MD, MSCR

Assistant Professor

Division of Pulmonary and Critical Care

University of Wisconsin

Website: <https://www.medicine.wisc.edu/apcc/icu-data-science-lab>

Twitter: @Majeans2011

LinkedIn: Majid Afshar, MD, MSCR

### *Tell us about yourself.*

I am a critical care physician with a focus in both clinical and research informatics. I spend a quarter of my time practicing in the medical ICU, and the remainder spans both health operations and research development. I currently have a NIH grant focused in data-driven methods for screening hospitalized patients with unhealthy substance use.

### *Is your research clinical, basic science, or translational?*

Clinical Informatics

### *Tell us about your research.*

My research lab employs methods in natural language processing (NLP) and machine learning using electronic health record data for high throughput computable phenotypes and applied predictive modeling. I currently work on automated methods in substance misuse screening and designing quasi-experimental studies in a learning health system.

### *Where do you see yourself in 5 years?*

Building an institute of data science researchers at a major academic center.

### *How has the Critical Care Assembly contributed to your career?*

Introduced me to great collaborators and exposed me to compelling science and research developments.

### *Published manuscripts in ATS journals.*

Research needs for inpatient management of severe alcohol withdrawal syndrome: An Official American Thoracic Society Research Statement: PMID 34609257



# ATS Highlights 2022: Critical Care Assembly Early Career Professionals

## Majid Afshar, MD, MSCR

Assistant Professor

Division of Pulmonary and Critical Care

University of Wisconsin

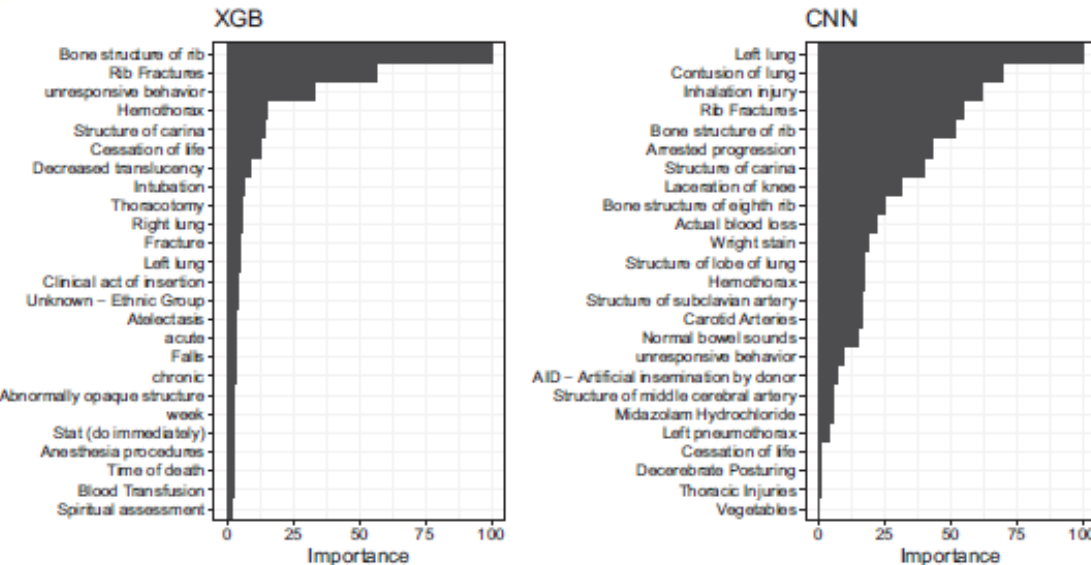
## Comparison and interpretability of machine learning models to predict severity of chest injury

**Rationale:** Trauma quality improvement programs and registries improve care and outcomes for injured patients. Designated trauma centers calculate injury scores using dedicated trauma registrars; however, many injuries arrive at nontrauma centers, leaving a substantial amount of data uncaptured. We propose automated methods to identify severe chest injury using machine learning (ML) and natural language processing (NLP) methods from the electronic health record (EHR) for quality reporting.

**Methods:** A level I trauma center was queried for patients presenting after injury between 2014 and 2018. Prediction modeling was performed to classify severe chest injury using a reference dataset labeled by certified registrars. Clinical documents from trauma encounters were processed into concept unique identifiers for inputs to ML models: logistic regression with elastic net (EN) regularization, extreme gradient boosted (XGB) machines, and convolutional neural networks (CNN). The optimal model was identified by examining predictive and face validity metrics using global explanations.

**Results:** Of 8952 encounters, 542 (6.1%) had a severe chest injury. CNN and EN had the highest discrimination, with an area under the receiver operating characteristic curve of 0.93 and calibration slopes between 0.88 and 0.97. CNN had better performance across risk thresholds with fewer discordant cases. Examination of global explanations demonstrated the CNN model had better face validity, with top features including "contusion of lung" and "hemopneumothorax."

**Discussion:** The CNN model featured optimal discrimination, calibration, and clinically relevant features selected. NLP and ML methods to populate trauma registries for quality analyses are feasible.



**Figure.** Top global model explanations from machine learning models to predict severe chest injury. X-axis represents rescaled variable importance from (1) XGBoost permuted feature importance, and (2) CNN training dataset averaged LIME explanations. Y-axis represents the preferred text definition of concepts; CNN LIME interpretation median  $r^2$  0.69 (IQR 0.46–0.93).