



# Research Week 2023

## Development and validation of harmonization strategy for identifying sociodemographic characteristics across studies in the Federal Interagency Traumatic Brain Injury Research informatic system (FITBIR)

Sara Hannon (co-first author), Will Baker-Robinson (co-first author), Maya O'Neil, David Cameron

David Cameron, MPH, Oregon Health and Science University, Department of Psychology, School of Medicine  
Clinical Psychology PhD candidate

Portland VA Medical Center  
Center to Improve Veteran Involvement in Care  
Research Health Science Specialist

### Keywords

Traumatic brain injury, FAIR data, harmonization, sociodemographic

### Abstract

Existing traumatic brain injury (TBI) literature supports associations between TBI exposure poor health and functional outcomes. However, associations of patient and injury characteristics with the prognosis of psychological health, functioning, and quality of life outcomes are still poorly understood. These gaps can be addressed by harmonizing and pooling data from existing studies to address relevant research questions. The Federal Interagency Traumatic Brain Injury Research informatic system (FITBIR) is the largest existing repository of TBI study data. While the size of FITBIR presents unique research opportunities, there are many practical challenges to harmonizing these data due to heterogeneity in study design, data collection, and data contribution. We developed a data model that accommodates heterogeneity in measurements and procedures for identifying and harmonizing sociodemographic characteristics across studies using study metadata, including variable title and definition text. Overall, 44 studies have made their data available. Means, proportions, and frequency of missingness are presented for age, gender, race, ethnicity, education, employment, income, and socioeconomic status overall and by study. Performance characteristics were calculated for keyword, algorithmic, and machine learning classification approaches. We used manual review of a stratified random sample as the gold standard. This work was motivated by the need to meet Findable, Accessible, Interoperable, Reusable (FAIR) data principles. Harmonizing the overlapping elements of FITBIR studies is a first step towards producing more impactful and clinically meaningful findings and addressing present knowledge gaps in the TBI literature.