Data Deluge: Challenges and Solutions in Implementing a Precision Medicine Approach to Create the **TRACK-TBI Information Commons**

Tatiana Khasanova[1]; Mary Vassar[2]; Mónica Coelho[1]; Sabrina Taylor[2]; Laura Brovold[1]; Geoffrey Manley[2] and the TRACK-TBI Investigators[3]

[1] Rancho BioSciences, LLC, 16955 Via Del Campo #220, San Diego, CA

[2] University of California San Francisco, Department of Neurosurgery, Brain and Spinal Injury Center

TRACK-TBI Background

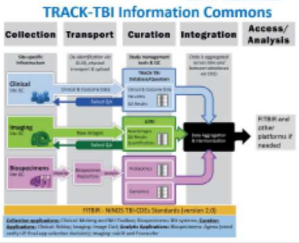
- **Traumatic Brain Injury (TBI)** is a serious condition, affecting an estimated 5.3 million people living with the long-term physical, cognitive, and psychological health disabilities, with annual direct and indirect costs estimated at over $76 billion [1]. Each year in the U.S. there are approximately 2.8 million TBI-related emergency department (ED) visits, hospitalizations, and deaths [2]. TBI is a contributing factor in a third of all injury-related US deaths. Despite promising preclinical research on neuroprotective agents these advances have failed to translate into a single successful clinical trial or treatment [3]. The heterogeneity of TBI has pose challenges to the design of clinical trials and approaches to effective treatment of TBI represents a great unmet need in public health.
- **The NINDS and DoD funded[2] multicenter Transforming Research and Clinical Knowledge in Traumatic Brain Injury (TRACK-TBI)** study aims to improve TBI classification and identify new diagnostic and prognostic markers for future clinical treatment trials.
- Since 2013, in collaboration with expert public-private partners and industry, we are collecting and analyzing comprehensive acute and long-term data on subjects at 19 U.S. trauma centers, across the spectrum of injury. This includes physiology measures, advanced neuroimaging, blood biospecimens, and outcome assessments. TRACK-TBI has currently enrolled 3,264 persons with TBI and will complete follow-ups for the first 2700 TBI and another 300 Orthopedic Control subjects in 2019.
- **TRACK-TBI** network has established a robust research infrastructure involving hundreds of multidisciplinary experts collaborating in a concerted effort to **improve the diagnosis, treatment and outcomes for persons with TBI.**

Precision Medicine

The ultimate goal for this study is to enable development of new diagnostic and prognostic biomarkers for future treatment clinical studies. Ideally, all data types should be analyzed holistically to arrive at the credible biomarker hypotheses

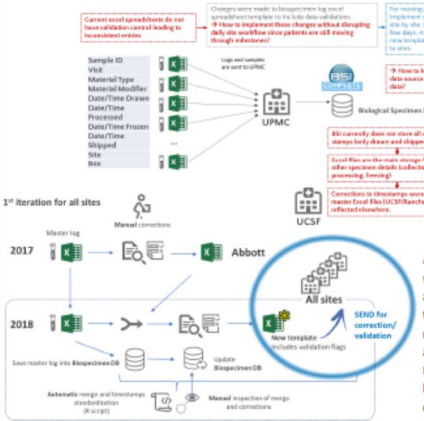
To achieve this goal, the ***TRACK-TBI Information Commons*** is being built to harmonize disparate datasets across multiple domains and applications.

- Our teams have been developing curation workflows in order to create "analysis ready" datasets for analytic teams.
- Data has been harmonized with the NIH TBI Common Data Elements and submitted to the Federated Interagency TBI Repository (FITBIR) for future sharing with the research community
- Examples are presented for the workflows, challenges and solutions



Biospecimen Validation Process

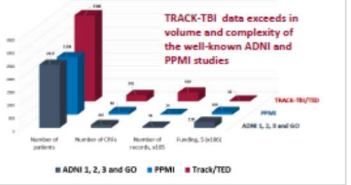
At the time the study was launched biospecimen data collection consisted of Excel sheets logs used by the study sites with no validation controls to flag and prevent errors.



After the curation workflow was developed, a new Biospecimen template with validation rules is used at all sites and a separate Biospecimen DB is used to keep a master track of the data.

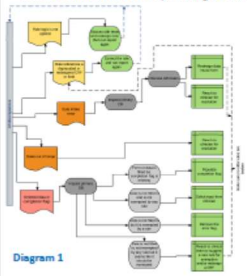
Data Types

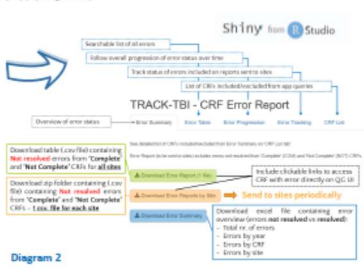
- Clinical, standardized high resolution neuroimaging, proteomic, genomic and outcome biomarkers
- Injury spectrum ranges from concussion to coma for patients enrolled within the first 24 hours following injury
- Subjects are followed during their acute care phase and at four follow-up visits over 12 months
- Clinical and outcomes core data are collected in a web-based electronic case report form that has over 200 forms and 15,000 data fields
- The initial clinical brain scans and high resolution magnetic resonance research imaging sequences obtained at 2-weeks and 6-months are submitted electronically to a central imaging repository
- Biospecimen logs are maintained for the inventory of over 90,000 biosamples collected at multiple time points, shipped to a central repository and processed for distribution to analytic laboratories



Challenges and Solutions

- In examples, we harmonized three primary sources of data: clinical and outcomes (via CRFs), Biospecimen samples, and neuroimaging findings. For clinical data, we started by implementing OpenClinica solution to track data entry errors and omissions. This work allowed us to de-lineate sources of mistakes (**Diagram 1**). However, a more succinct solution was developed using an R Shiny app (**Diagram 2**).

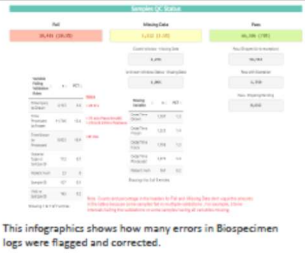




- R Shiny application was developed to improve analytics, provide dashboard reports and facilitate sharing of datasets among the curation team members.
- For Biospecimen data, the main challenges are to: (1) incorporate incremental data updates from multiple sites as sample collection is ongoing; (2) bring timestamp coding to consistency; and (3) identify omissions in collection protocols and request corrective actions.
- We solved these issues by developing a uniform template for biospecimen log entries (with multiple rules for automatic data checking) and a biospecimen validation process.

Process Examples

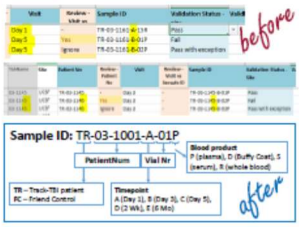
Improved process and data validation allowed TRACK-TBI team to identify and correct multiple omissions and incorrect data



This infographics shows how many errors in Biospecimen logs were flagged and corrected.

Examples of errors are shown below:

- Mismatched suffixes on sample vials resulting in loss of provenance.
- After validation and correction, Sample IDs entries are brought to consistency and correct formats.



Future Directions

- Complete remaining data curation efforts for the first 3000 subjects for harmonization of datasets across the core repositories and creation of locked datasets for sharing.
- Develop analytic pipelines for biomarker-based pathoanatomic classification of TBI and employ this paradigm-shifting approach in an exploratory clinical trial utilizing the established TRACK-TBI infrastructure.

References

1. https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6227a2.htm
2. https://www.cdc.gov/traumaticbraininjury/basics.html
3. Maas AI, Roozenbeek B and Manley GT. *Neurotherapeutics*. Vol. 7, 115–126, January 2010

Acknowledgments

*TRACK-TBI Principal Investigators: Ramon Diaz-Arrastia, Univ. Pennsylvania; Joseph T. Giacino, Harvard Univ.; Pratik Mukherjee, UC San Francisco; David O. Okonkwo, Univ. Pittsburgh; Claudia S. Robertson, Baylor College of Medicine; Nancy Temkin, Univ. Washington.

*Research was supported by National Institutes of Health (grant U01 NS086090) and US Department of Defense (grant W81XWH-14-2-0176). Abbott Laboratories provided funding for add-in TRACK-TBI clinical studies. One Mind provided funding for patient stipends and support to clinical sites. Pfizer and NeuroTrauma Sciences, LLC provided funding for data curation services.