SDTM datasets of clinical data and measurements for selected cancer collections to TCIA

Description

The Data Integration & Imaging Informatics (DI-Cubed) project explored the issue of lack of standardized data capture at the point of data creation, as reflected in the non-image data accompanying various TCIA breast cancer collections (ISPY1, BREAST-DIAGNOSIS, Breast-MRI-NACT-Pilot, TCGA-BRCA) and the Ivy Glioblastoma Atlas Project (Ivy GAP) brain cancer collection. The work addressed the desire for semantic interoperability between various NCI initiatives by aligning on common clinical metadata elements and supporting use cases that connect clinical, imaging, and genomics data. Accordingly, clinical and measurement data imported into I2B2 were cross-mapped to industry standard concepts for names and values including those derived from BRIDG, CDISC SDTM, DICOM Structured Reporting models and using NCI Thesaurus, SNOMED CT and LOINC controlled terminology.

A subset of the standardized data was then exported from I2B2 in SDTM compliant SAS transport files. The SDTM data was derived from data taken from both the curated TCIA spreadsheets as well as tumor measurements and dates from the TCIA Restful API. Due to the nature of the available data not all SDTM conformance rules were applicable or adhered to.

These Study Data Tabulation Model format (SDTM) datasets were validated using Pinnacle 21 CDISC validation software. The validation software reviews datasets according to their degree of conformance to rules developed for the purposes of FDA submissions of electronic data. Iterative refinements were made to the datasets based upon group discussions and feedback from the validation tool.

Export datasets for the following SDTM domains were generated:

- DM (Demographics)
- DS (Disposition)
- MI (Microscopic Findings)
- PR (Procedures)
- SS (Subject Status)
- TU (Tumor/Lesion Identification)
- TR (Tumor/Lesion Results)

Download

- Source DICOM images (require the NBIA Data Retriever to open)
    - NBIA-manifest-SDTM-ISPY1.tcia
    - NBIA-manifest-STDM-Breast-MRI-NACT-Pilot.tcia
    - NBIA-manifest-SDTM-Ivy-GAP-TCGA-BRCA.tcia
- You can also download zipped accompanying data:
    - SAS transport files (.xpt) and
    - corresponding CSV files for each domain.