SDTM datasets of clinical data and measurements for selected cancer collections to TCIA (DI-Cubed-Reports)

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 4 TCIA breast cancer collections (Multi-center breast DCE-MRI data and segmentations from patients in the I-SPY 1/ACRIN 6657 trials (ISPY1), BREAST-DIAGNOSIS, Single site breast DCE-MRI data and segmentations from patients undergoing neoadjuvant chemotherapy (Breast-MRI-NACT-Pilot), The Cancer Genome Atlas Breast Invasive Carcinoma Collection (TCGA-BRCA)) and the Ivy Glioblastoma Atlas Project (IvyGAP) 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)

Data Access
Data Access

Data Type	Download all or Query/Filter
SAS Transport Files (XPT)	**⊙ Download**
Image Analysis (CSV)	**⊙ Download**

Please contact help@cancerimagingarchive.net with any questions regarding usage.

Collections Used in this Third Party Analysis

Below is a list of the Collections used in these analyses:

- ISPY1,
- BREAST-DIAGNOSIS,
- Breast-MRI-NACT-Pilot,
- TCGA-BRCA,
- Ivy GAP

Detailed Description
Detailed Description

TCIA breast cancer collections used:

- Multi-center breast DCE-MRI data and segmentations from patients in the I-SPY 1/ACRIN 6657 trials (ISPY1)
- BREAST-DIAGNOSIS
- Single site breast DCE-MRI data and segmentations from patients undergoing neoadjuvant chemotherapy (Breast-MRI-NACT-Pilot)
- The Cancer Genome Atlas Breast Invasive Carcinoma Collection (TCGA-BRCA)

TCIA brain cancer collection used:

- Ivy Glioblastoma Atlas Project (IvyGAP)

Citations & Data Usage Policy

Users of this data must abide by the TCIA Data Usage Policy and the Creative Commons Attribution 3.0 Unported License under which it has been published. Attribution should include references to the following citations:

Data Citation

Hickman H., Ver Hoef W., Hastak S., Neville J., Clunie D., Wagner U., Helton E. (2019). **SDTM datasets of clinical data and measurements for selected cancer collections to TCIA [Dataset]**. The Cancer Imaging Archive. doi: 10.7937/TCIA.2019.zfv154m9

Publication Citation

Clunie, D., Hickman, H., Ver Hoef, W., Hastak, S., Evans, J., Neville, J., & Wagner, U. (2020). **Observations from the Data Integration and Imaging Informatics (DI-Cubed) Project**. MDPI AG. https://doi.org/10.20944/preprints202008.0474.v1

TCIA Citation

Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., & Prior, F. (2013). **The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository**. Journal of Digital Imaging, 26(6), 1045–1057. https://doi.org/10.1007/s10278-013-9622-7

Other Publications Using This Data

TCIA maintains a list of publications that leverage TCIA data. If you have a manuscript you'd like to add please contact TCIA's Helpdesk.

Versions

Version 1 (Current): 2019/06/21

Data Type	Download all or Query/Filter
SAS Transport Files (XPT)	Download
Image Analysis (CSV)	Download