Cancer Moonshot Biobank - Gastroesophageal Cancer Collection (CMB-GEC)

Reproduced from https://wiki.cancerimagingarchive.net

Summary

The Cancer Moonshot Biobank is a National Cancer Institute initiative to support current and future investigations into drug resistance and sensitivity and other NCI-sponsored cancer research initiatives, with an aim of improving researchers' understanding of cancer and how to intervene in cancer initiation and progression. During the course of this study, biospecimens (blood and tissue removed during medical procedures) and associated data will be collected longitudinally from at least 1000 patients across at least 10 cancer types, who represent the demographic diversity of the U.S. and receiving standard of care cancer treatment at multiple NCI Community Oncology Research Program (NCORP) sites.

This collection contains de-identified radiology and histopathology imaging procured from subjects in NCI's Cancer Moonshot Biobank - Gastroesophageal Cancer (CMB-GEC) cohort. Associated genomic, phenotypic and clinical data will be hosted by The Database of Genotypes and Phenotypes (dbGaP) and other NCI databases. A summary of Cancer Moonshot Biobank imaging efforts can be found on the Cancer Moonshot Biobank Imaging page.

Data Access

Data Access

Data Type	Download all or Query/Filter	License
Images (DICOM, 839 MB)	*Download* *Search* (Download requires the NBIA Data Retriever)	CC BY 4.0
Tissue Slide Images, Pathology Metadata (SVS, JSON, 103 MB)	*Download* *Search* (Download requires Aspera)	CC BY 4.0

Click the Versions tab for more info about data releases.

Additional Resources for this Dataset

The database of Genotypes and Phenotypes (dbGaP) hosts genomic, phenotypic, and clinical data for NCI's Cancer Moonshot Biobank (CMB) project. Information and access to the data can be found at:

- dbGaP - Cancer Moonshot Biobank (Genomic, Phenotypic & Clinical Data)

The NCI Cancer Research Data Commons (CRDC) provides access to additional data and a cloud-based data science infrastructure that connects data sets with analytics tools to allow users to share, integrate, analyze, and visualize cancer research data.

- Imaging Data Commons (IDC) (Imaging Data)

Reproduced from https://wiki.cancerimagingarchive.net

Detailed Description

Detailed Description

Image Statistics	Radiology Image Statistics	Pathology Image Statistics
Modalities	CT, PT	WSI
Number of Patients	1	1
Number of Studies	4	N/A
Number of Series	14	N/A
Number of Images	1293	3
Images Size (MB)	839	103

Reproduced from https://wiki.cancerimagingarchive.net

Introduction

Biobank radiology imaging data on TCIA contains the “days from enrollment (registration)” for each scan, embedded in the DICOM files (DICOM tag (0012,0053)). This allows for temporal alignment between the imaging on TCIA and clinical events data found on the Biobank Catalog.
Note: In order that the images display properly in DICOM readers, the radiology imaging data also contains de-identified dates that preserve the temporal sequence relationship between scans in a given study.

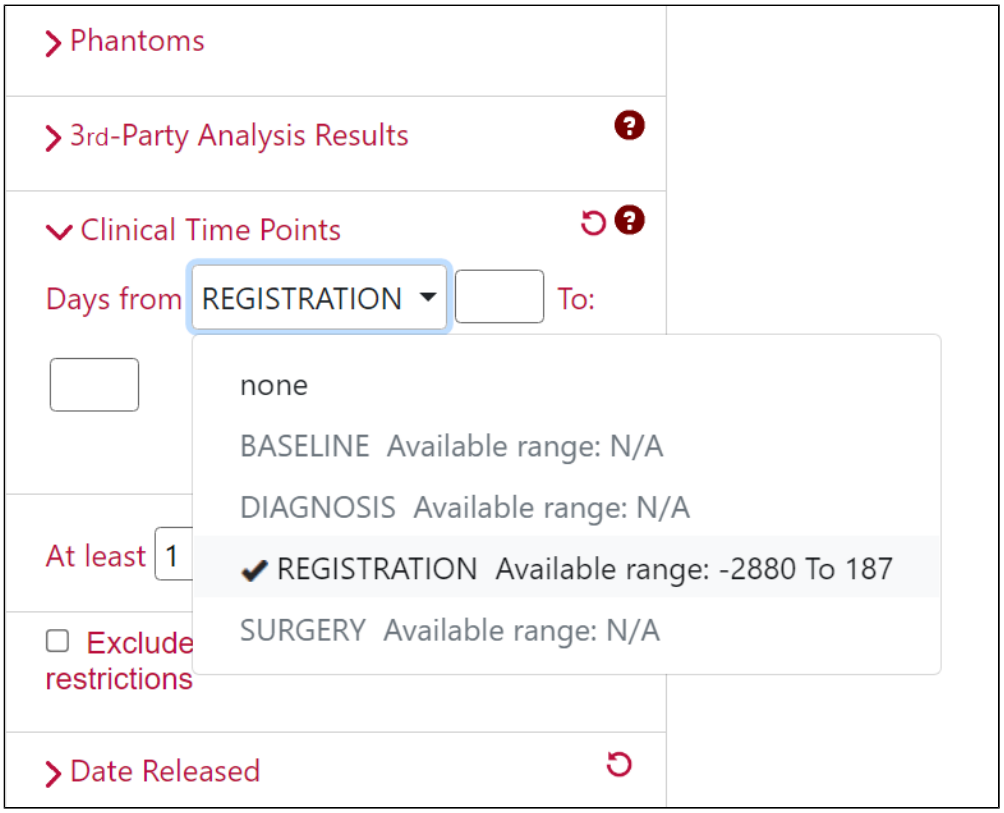
Reproduced from https://wiki.cancerimagingarchive.net

Days from enrollment (registration)

In addition to modifying the actual date fields in the DICOM header, the "days from registration" values are calculated and stored in the DICOM tag **(0012,0052) Longitudinal Temporal Offset from Event** with the associated tag **(0012,0053) Longitudinal Temporal Event Type** set to “REGISTRATION”. Here is an example DICOM header from a scan where the patient's imaging was performed 2 days before the registration, resulting in a negative offset value.

(0012,0052)	Longitudinal Temporal Offset from Event	-2.0
(0012,0053)	Longitudinal Temporal Event Type	REGISTRATION

If you would like to filter your search results using this information, you can leverage the "Clinical Trial Time Points" filter via our data portal at https://nbia.cancerimagingarchive.net/nbia-search/.



De-identification of DICOM dates

De-identification of dates for this dataset uses the DICOM Part 3.15 Annex E standard “Retain Longitudinal With Modified Dates Option” which allows dates to be retained as long as they are modified from the original date. TCIA implements this using a technique which de-identifies the dates while preserving the longitudinal relationship between them. Original dates will be first normalized to January 1, 1960 and then offset relative to the date of registration for each patient. This normalized date system was chosen in order to make it obvious that the dates are not real, and to make it easy to quickly determine how much time has passed between the date of registration and the patients' related imaging studies.

For example, if the real date of a patient's registration was 03/27/2018 and the original imaging Study Date was 03/29 /2018 then the anonymized TCIA Study Date would become 01/03/1960 (two days after the base date of 1/1/1960).

Citations & Data Usage Policy

Citations & Data Usage Policy

Users must abide by the TCIA Data Usage Policy and Restrictions. Attribution should include references to the following citations:

Data Citation

Cancer Moonshot Biobank. (2022). Cancer Moonshot Biobank - Gastroesophageal Cancer Collection (CMB-GEC) (Version 1) [Data set]. The Cancer Imaging Archive. https://doi.org/10.7937/E7KH-R486

Acknowledgement

The Cancer Moonshot Biobank program requests that publications using data from this program include the following statement: “**Data used in this publication were generated by the National Cancer Institute Cancer Moonshot Biobank**.”

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. In Journal of Digital Imaging (Vol. 26, Issue 6, pp. 1045–1057). Springer Science and Business Media LLC. https://doi.org/10.1007/s10278-013-9622-7

Other Publications Using This Data

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

Versions

Version 1 (Current): Updated 2022/08/12

Data Type	Download all or Query/Filter	License
Images (DICOM, 839 MB)	*Download Search* (Requires NBIA Data Retriever.)	CC BY 4.0

Reproduced from https://wiki.cancerimagingarchive.net

Tissue Slide Images (SVS, 103 MB)	*Download Search* (Download requires Aspera)	CC BY 4.0