TCGA-OV

Summary

The Cancer Genome Atlas Ovarian Cancer (TCGA-OV) data collection is part of a larger effort to build a research community focused on connecting cancer phenotypes to genotypes by providing clinical images matched to subjects from The Cancer Genome Atlas (TCGA). Clinical, genetic, and pathological data resides in the Genomic Data Commons (GDC) Data Portal while the radiological data is stored on The Cancer Imaging Archive (TCIA).

Matched TCGA patient identifiers allow researchers to explore the TCGA/TCIA databases for correlations between tissue genotype, radiological phenotype and patient outcomes. Tissues for TCGA were collected from many sites all over the world in order to reach their accrual targets, usually around 500 specimens per cancer type. For this reason the image data sets are also extremely heterogeneous in terms of scanner modalities, manufacturers and acquisition protocols. In most cases the images were acquired as part of routine care and not as part of a controlled research study or clinical trial.

CIP TCGA Radiology Initiative

Imaging Source Site (ISS) Groups are being populated and governed by participants from institutions that have provided imaging data to the archive for a given cancer type. Modeled after TCGA analysis groups, ISS groups are given the opportunity to publish a marker paper for a given cancer type per the guidelines in the table above. This opportunity will generate increased participation in building these multi-institutional data sets as they become an open community resource. Learn more about the TCGA Ovarian Phenotype Research Group.

Acknowledgements

- Mayo Clinic, Rochester, MN - Special thanks to **Brad Erickson, MD, PhD. Ca**

Data Access

C lick the **Download** button to save a ".tcia" manifest file to your computer, which you must open with the NBIA Data Retriever . Click the **Search** button to open our Data Portal, where you can browse the data collection and/or download a subset of its contents.

Data Type	Download all or Query/Filter
Images (DICOM, 28.3GB)	Download Search
Tissue Slide Images (web)	Search
Clinical Data (TXT)	Download
Biomedical Data (TXT)	Download
Genomics (web)	Search

Click the Versions tab for more info about data releases.

Third Party Analyses of this Dataset

TCIA encourages the community to publish your analyses of our datasets. Below is a list of such third party analyses published using this Collection:

- Integration of CT-based Qualitative and Radiomic Features with Proteomic Variables in Patients with High-Grade Serous Ovarian Cancer: An Exploratory Analysis (TCGA-OV-Proteogenomics)
- Imaging Features, and Correlations with Genomic and Clinical Data from the TCGA Ovarian Radiology Research Group (TCGA-OV-Radiogenomics)
- Crowds Cure Cancer: Data collected at the RSNA 2017 annual meeting (Crowds-Cure-2017)

Detailed Description

Image Statistics	
Modalities	CT
Number of Participants	143

Number of Studies	322
Number of Series	844
Number of Images	53,662
Images Size (GB)	28.3

GDC Data Portal - Clinical and Genomic Data

- The GDC Data Porta l has extensive clinical and genomic data, which can be matched to the patient identifiers on the images here in TCIA. Below is a snapshot of clinical data available for the cases with imaging on the TCIA (extracted on 1/5/2016):

  - TCGA-OV Clinical Data.zip

  Case report forms (with verbose explanations) used to collect the clinical data can be found on the Biospecimen Core Resource Clinical Data Forms linked below:

  - Ovary Case Quality Control Form
  - Ovary Enrollment Form
  - Ovary Follow-Up Form

  Additional Documentation

  - OV Clinical Follow Up
  - O V Clinical Follow up Notes

  Cases by Site:

  - O V TCIA Cases - listed by site
  - OV TCIA Cases - listed by source site code

Shared List: **Ovarian Cases with Early Post-Operative Scans**

This shared list is tor researchers wanting to evaluate the completeness of OV tumor surgical de-bulking, this group of cases has sequential CT scans that have both pre-op (baseline) and early post-op (less than ~ 2 months).

A Note about TCIA and TCGA Subject Identifiers and Dates

Subject Identifiers: a subject with radiology images stored in TCIA is identified with a Patient ID that is identical to the Patient ID of the same subject with demographic, clinical, pathological, and/or genomic data stored in TCGA. For each TCGA case, the baseline TCGA imaging studies found on TCIA are pre-surgical.

Dates: TCIA and TCGA handle dates differently, and there are no immediate plans to reconcile:

- <u>TCIA Dates:</u> dates (be they birth dates, imaging study dates, etc.) in the Digital Imaging and Communications in Medicine (DICOM) headers of TCIA radiology images have been offset by a random number of days. The offset is a number of days between 3 and 10 years prior to the real date that is consistent for each TCIA image-submitting site and collection, but that varies among sites and among collections from the same site. Thus, the number of days between a subject's longitudinal imaging studies are accurately preserved when more than one study has been archived while still meeting HIPAA requirements.

- <u>TCGA Dates:</u> the patient demographic and clinical event dates are all the number of days from the index date, which is the actual date of pathologic diagnosis. So all the dates in the data are relative negative or positive integers, except for the "days_to_pathologic_diagnosis" value, which is 0 – the index date. The years of birth and diagnosis are maintained in the distributed clinical data file. The NCI retains a copy of the data with complete dates, but those data are not made available. With regard to other TCGA dates, if a date comes from a HIPAA "covered entity's" medical record, it is turned into the relative day count from the index date. Dates like the date TCGA received the specimen or when the TCGA case report form was filled out are not such covered dates, and they will appear as real dates (month, day, and year).

<u>Citations & Data Usage Policy</u>

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:

ⓘ **TCGA Attribution**

"The results <published or shown> here are in whole or part based upon data generated by the TCGA Research Network: http://cancergenome.nih.gov/."

ⓘ **Data Citation**

Holback, C., Jarosz, R., Prior, F., Mutch, D. G., Bhosale, P., Garcia, K., … Erickson, B. J. (2016). Radiology Data from The Cancer Genome Atlas Ovarian Cancer [TCGA-OV] collection. The Cancer Imaging Archive. http://doi.org/10.7937/K9/TCIA.2016.NDO1MDFQ

ⓘ **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. **The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository**, Journal of Digital Imaging, Volume 26, Number 6, December, 2013, pp 1045-1057. (paper)

Other Publications Using This Data

TCIA maintains a list of publications which leverage our data. At this time we are not aware of any manuscripts based on this data. If you have a manuscript you'd like to add please contact the TCIA Helpdesk.

Versions

Version 4 (Current): Updated 2020/05/29

Data Type	Download all or Query/Filter
Images (DICOM, 28.3GB)	Download Search
Tissue Slide Images (web)	Search
Clinical Data (TXT)	Download
Biomedical Data (TXT)	Download
Genomics (web)	Search

Updated clinical data link with latest spreadsheets from GDC. Added new biomedical spreadsheets from GDC.

Version 3: Updated 2020/04/03

On 2020-04-03 DICOM tags associated with TCGA-30-1891 were updated with a corrected version of the file.

Version 2 : Updated 2016/01/05

Data Type	Download all or Query/Filter
Images (DICOM, 28.3GB)	Download Search
Clinical Data (TXT)	Download
Genomics (web)	Search

Extracted latest release of clinical data (TXT) from the GDC Data Portal.

Version 1: Updated 2014/11/05

Data Type	Download all or Query/Filter
Images (DICOM, 28.3GB)	⊘ Download Q Search
Clinical Data (TXT)	⊘ Download
Genomics (web)	Q Search