The Clinical Proteomic Tumor Analysis Consortium Colon Adenocarcinoma Collection (CPTAC-COAD)

Summary

Redirection Notice

This page will redirect to https://www.cancerimagingarchive.net/collection/cptac-coad/ in about 5 seconds.

This collection contains subjects from the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium CPTAC Colon Adenocarcinoma cohort. CPTAC is a national effort to accelerate the understanding of the molecular basis of cancer through the application of large-scale proteome and genome analysis, or proteogenomics. Radiology and pathology images from CPTAC patients are being collected and made publicly available by The Cancer Imaging Archive to enable researchers to investigate cancer phenotypes which may correlate to corresponding proteomic, genomic and clinical data.

Imaging from each cancer type will be contained in its own TCIA Collection, with the collection name "CPTAC-*cancertype*". Radiology imaging is collected from standard of care imaging performed on patients immediately before the pathological diagnosis, and from follow-up scans where available. For this reason the radiology image data sets are heterogeneous in terms of scanner modalities, manufacturers and acquisition protocols. Pathology imaging is collected as part of the CPTAC qualification workflow.

All CPTAC cohorts are released as either a single combined cohort, or split into *Discovery* and *Confirmatory* where applicable. There are two main types of proteomic studies: discovery proteomics and targeted proteomics. The term "discovery proteomics" is in reference to "untargeted" identification and quantification of a maximal number of proteins in a biological or clinical sample. The term "targeted proteomics" refers to quantitative measurements on a defined subset of total proteins in a biological or clinical sample, often following the completion of discovery proteomics studies to confirm interesting targets selected. Commonly used proteomic technologies and platforms are different types of mass spectrometry and protein microarrays depending on the needs, throughput and sample input requirement of an analysis, with further development on nanotechnologies and automation in the pipeline in order to improve the detection of low abundance proteins, increase throughput, and selectively reach a target protein *in vivo*. Once the protein targets of interest are identified, high-throughput targeted assays are developed for confirmatory studies: tests to affirm that the initial tests were accurate. A summary of CPTAC imaging efforts can be found on the CPTAC Imaging Proteomics page.

CPTAC Imaging Special Interest Group

You can join the CPTAC Imaging Special Interest Group to be notified of webinars & data releases, collaborate on common data wrangling tasks and seek out partners to explore research hypotheses! Artifacts from previous webinars such as slide decks and video recordings can be found on the CPTAC SIG Webinars page.

Data Access

Data Access

Data Type	Download all or Query/Filter	License
Tissue Slide Images (SVS, 66.69 GB)	*Download Search* (Download and apply the IBM-Aspera-Connect plugin to your browser to retrieve this faspex package)	CC BY 3.0

Click the Versions tab for more info about data releases.

Additional Resources for this Dataset

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)
- Proteomic Data Commons (PDC) (Proteomic & Clinical Data)
- Genomic Data Commons (GDC) (Genomic & Clinical Data)

Detailed Description
Detailed Description

	Pathology Image Statistics
Modalities	Pathology
Number of Participants	106
Number of Images	373
Images Size (GB)	66.69

Accessing the Proteomic & Genomic Clinical Data

To access/download the clinical data on the **Proteomic Data Commons (PDC)** and **Genomic Data Commons (GDC)**, once you have identified the data of your interest, move to the 'Clinical' tab on the browse page. Select the checkbox to select a specific row, all rows on the page or all pages and click the export clinical manifest button in CSV or TSV format on the GDC, or TSV or JSON format on the PDC.

A Note about TCIA and CPTAC Subject Identifiers

A subject with radiology and pathology images stored in TCIA is identified with a de-identified project Patient ID that is identical to the Patient ID of the same subject with clinical, proteomic, and/or genomic data stored in other CPTAC databases and web sites.

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**
>
> National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC). (2020). **The Clinical Proteomic Tumor Analysis Consortium Colon Adenocarcinoma Collection (CPTAC-COAD)** (Version 1) [Data set]. The Cancer Imaging Archive. https://doi.org/10.7937/TCIA.YZWQ-ZZ63

> ⓘ **Acknowledgement**
>
> The CPTAC 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 Clinical Proteomic Tumor Analysis Consortium (CPTAC).**"

> ⓘ **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 TCIA's Helpdesk.

<u>Versions</u>

Version 1 (Current) : 2021/02/02

Data Type	Download all or Query/Filter
Tissue Slide Images (SVS, 66.69 GB)	*Download* *Search* (Download and apply the IBM-Aspera-Connect plugin to your browser to retrieve this faspex package)
Clinical Data (external)	Proteomic Data Commons
Proteomic Data (external)	Proteomic Data Commons
Genomics Data (external)	Genomic Data Commons