

CPTAC-LSCC



Summary

This collection contains subjects from the National Cancer Institute's [Clinical Proteomic Tumor Analysis Consortium](#) Lung Squamous Cell Carcinoma (CPTAC-LSCC) 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*", and is being made available on a release schedule that is coordinated with the CPTAC program releases of proteomic and genomic data. A summary of CPTAC imaging efforts can be found on the [CPTAC Imaging Proteomics](#) page.

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.

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.

Acknowledgements






We would like to acknowledge the individuals and institutions that have provided data for this collection:

- Beaumont Health System, Royal Oak, MI - Special thanks to **George D. Wilson, PhD** from the Department of Radiation Oncology Research, **Barbara Pruetz** of the Biobank, **Debra Kapczynski, MHSA, CIIP, RT(R) (CT)** and **Rachel Deyer** from the Department of Diagnostic Radiology.
- Cureline, Inc. team and clinical network, Brisbane, CA - Special thanks to **Olga Potapova, PhD**, **Vladislav Golubkov, PhD**, **Victoria Fulidou, MD**, **Alexander Sviridov**, **Dmitry Belyaev, MD**, **Oxana Paklina, MD**, **Dr. Sc.**, **Galiya Setdikova, MD, PhD**, and **Denis Golbin, MD, PhD**.
- St. Joseph's Hospital and Medical Center, Phoenix, AZ - Special thanks to **Jennifer Eschbacher, MD** from the Department of Neuropathology, **Catherine Seiler, PhD**, **Rosy Singh** and **Beth Hermes** from the Biobank Core Facility, and **Victor Sisneros, RT(R)(CT)**, CPSA.

- International Institute for Molecular Oncology, Pozna, Poland - Special thanks to **Maciej Wiznerowicz MD, PhD** and **Jan Lubiski MD PhD**; **Wojciech Dyszkiewicz MD, PhD.**, and **ukasz Gsiorowski MD** from Greater Poland Center of Pulmonology and Thoracic Surgery, Pozna, Poland; **Bartosz Kubisa MD, PhD.**, **Malgorzata Wojtys MD and PhD.**, and **Tomasz Grodzki MD, PhD** of Regional Hospital Complex, Szczecin, Poland
- BioPartners, CA - Special thanks to **Alexander Gasparian, PhD.** from the Department of Drug Discovery and Biomedical Sciences, University of South Carolina College of Pharmacy, **Kakhaber Zaalishvili, MD** Medical Advisor and Staff Pathologist at BioPartners, LLC, **Milla Gorodnia**, President of BioPartners, Inc., **Victoria Christensen**, Global Business Development/Project Coordination Manager, **Oksana Havryliuk, MD.** Chief of Research department of radiodiagnostics of NCI (Ukraine), **Marianna Gredil'**, Director of BioPartners, LLC, and **Anna Legenka** Chief of the Data Department at BioPartners, LLC
- Boston Medical Center, Boston, MA - Special thanks to **Chris D. Andry M.Phil, PhD** from the Department of Pathology and Laboratory Medicine, **Margaret Lavoye**, **Artem Kaliev**, **Wilson Chavez**, **Stephan Anderson**, **Jorge Soto**, and **Mitchell Horn** from the Department of Radiology, **Elizabeth Duffy, MA** and **Cheryl Spencer, MA** of the Biobank.

Data Access

Data Access

Data Type	Download all or Query/Filter
Images (DICOM, 28.8 GB)	  (Download requires the NBIA Data Retriever)
Tissue Slide Images (SVS, 414 GB)	 
Clinical Data API (JSON - more info)	
Discovery Study (CPTAC Data Portal)	<ul style="list-style-type: none"> • Proteomic Data Commons • Genomic Data Commons • Pathology • Radiology

Click the Versions tab for more info about data releases.

Detailed Description

Detailed Description

	Radiology Image Statistics	Pathology Image Statistics
--	-----------------------------------	-----------------------------------

Modalities	CT, CR, DX, NM, PT	Pathology
Number of Participants	39	212
Number of Studies	63	N/A
Number of Series	264	N/A
Number of Images	52,072	1,081
Images Size (GB)	28.8	414

A Note about TCIA and CPTAC Subject Identifiers and Dates

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.

Dates:

The radiology imaging data is in DICOM format. To provide temporal context information aligned with events in the clinical data set for each patient, TCIA has inserted information in DICOM tag (0012,0050) *Clinical Trial Time Point ID*. This DICOM tag contains the number of days from the *date the patient was initially diagnosed pathologically with the disease* to the date of the scan. E.g. a scan acquired 3 days before the diagnosis would contain the value -3. A follow up scan acquired 90 days after diagnosis would contain the value 90.

The DICOM date tags (i.e. birth dates, imaging study dates, etc.) are modified per TCIA's standard process which offsets them 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.

Citations & Data Usage Policy

Citations & Data Usage Policy

Users of this data must abide by the [Creative Commons Attribution 3.0 Unported License](#) under which it has been published. **CPTAC proteomic and genomic data use must also comply with the [CPTAC Data Use Agreement](#).** Questions may be directed to help@cancerimagingarchive.net. Attribution should include references to the following citations:

Data Citation

National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC). (2018). **Radiology Data from the Clinical Proteomic Tumor Analysis Consortium Lung Squamous Cell Carcinoma [CPTAC-LSCC] Collection [Data set]**. The Cancer Imaging Archive. <https://doi.org/10.7937/k9/tcia.2018.6emub512>

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. **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 TCIA data. If you have a manuscript you'd like to add please [contact the TCIA Helpdesk](#).







Versions

Version 13 (Current): 2021/01/27

Data Type	Download all or Query/Filter
Images (DICOM, 28.8 GB)	  (Download requires the NBIA Data Retriever)
Tissue Slide Images (SVS, 414 GB)	 
Clinical Data API (JSON - more info)	
Discovery Study (CPTAC Data Portal)	<ul style="list-style-type: none"> Proteomic Data Commons Genomic Data Commons Pathology Radiology






New Genomic and Proteomic data added

Version 12: Updated 2020/06/25

Data Type	Download all or Query/Filter
Images (DICOM, 28.8 GB)	  (Requires NBIA Data Retriever)
Tissue Slide Images (SVS, 414 GB)	 
Clinical Data API (JSON - more info)	
Proteomics (web)	


Added 12 participants' radiology imaging from two sites.



Version 11: Updated 2020/03/31

Data Type	Download all or Query/Filter
Images (DICOM, 18.7 GB)	 (Requires NBIA Data Retriever .)
Tissue Slide Images (SVS, 414 GB)	 
Clinical Data API (JSON - more info)	
Proteomics (web)	

Added 1 radiology subject





Version 10 : 2020/02/14

Data Type	Download all or Query/Filter
Images (DICOM, 18 GB)	 (Requires NBIA Data Retriever .)

Tissue Slide Images (SVS, 414 GB)	
Clinical Data (CSV)	Coming soon
Proteomics (web)	





Added 6 new pathology subjects and 74 new pathology images.

Version 9: 2020/01/24

Data Type	Download all or Query/Filter
Images (DICOM, 18 GB)	 (Requires NBIA Data Retriever)
Tissue Slide Images (SVS, 341 GB)	 
Clinical Data (CSV)	Coming soon
Proteomics (web)	

Added 51 new pathology patients and 367 new WSIs.




Version 8: 2019/12/19

Data Type	Download all or Query/Filter
Images (DICOM, 18 GB)	 (Requires NBIA Data Retriever)
Tissue Slide Images (SVS, 148 GB)	 
Clinical Data (CSV)	Coming soon
Proteomics (web)	

Added 3 new radiological subjects.




Transferred two subjects radiology imaging (C3N-00738 C3N-02290) to CPTAC-LUAD as a correction.

Version 7: 2019/09/30

Data Type	Download all or Query/Filter
Images (DICOM, 19.5 GB)	 (Requires NBIA Data Retriever .)
Tissue Slide Images (SVS, 147.6 GB)	
Clinical Data (CSV)	Coming soon
Proteomics (web)	




Added New Subjects.

Version 6 : Updated 2019/06/30

Data Type	Download all or Query/Filter
Images (DICOM, 13.9 GB)	 (Requires NBIA Data Retriever)
Tissue Slide Images (SVS, 147.6 GB)	
Clinical Data (CSV)	Coming soon
Proteomics (web)	




Added Subjects.

Version 5: Updated 2019/03/31

Data Type	Download all or Query/Filter
Images (DICOM, 12.6 GB)	 (Requires NBIA Data Retriever)
Tissue Slide Images (web)	
Clinical Data (CSV)	Coming soon
Proteomics (web)	




Added new subjects

Version 4 : Updated 2018/12/31

Data Type	Download all or Query/Filter
Images (DICOM, 2.5 GB)	 (Requires NBIA Data Retriever .)
Tissue Slide Images (web)	
Proteomics (web)	




Added new subjects.

Version 3: Updated 2018/06/30

Data Type	Download all or Query/Filter
Images (DICOM, 0.6 GB)	 (Requires NBIA Data Retriever)
Tissue Slide Images (web)	
Proteomics (web)	



Added new subjects.

Version 2: Updated 2018/01/10

Data Type	Download all or Query/Filter
Images (DICOM, 0.4 GB)	 (Requires NBIA Data Retriever)
Tissue Slide Images (web)	
Proteomics (web)	

Added new subjects.

Version 1: Updated 2018/01/10

Data Type	Download all or Query/Filter
Images (DICOM, 0.132 GB)	 (Requires NBIA Data Retriever)
Tissue Slide Images (web)	
Proteomics (web)	