

# CPTAC-CCRCC

# Summary

This collection contains subjects from the National Cancer Institute's [Clinical Proteomic Tumor Analysis Consortium](#) Clear Cell Renal Cell Carcinoma (CPTAC-CCRCC) 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.

On February 4, 2020 Dr. David Clarke presented the consortium's proteogenomic analyses of the CPTAC-CCRCC discovery cohort. This deep dive into the CCRCC genomic and proteomic datasets will help researchers better understand how these can be correlated with features derived from the imaging data. ([Download the slides](#))

## 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.
- International Institute for Molecular Oncology, Pozna, Poland - Special thanks to **Maciej Wiznerowicz MD, PhD** and **Jan Lubiski MD PhD, Andrzej Antczak MD, PhD**, and **Gabriel Bromiski MD**, from Józef Stru City Hospital and Polyclinic in Pozna, Poland.
- 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**.
- 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 Kaliaev, Wilson Chavez, Stephan Anderson, J**

**orge Soto**, and **Mitchell Horn** from the Department of Radiology, **Elizabeth Duffy, MA** and **Cheryl Spencer, MA** of the Biobank.

- 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
- University of Pittsburgh/UPMC, Pittsburgh, PA - Special thanks to **Scott Beasley (MD, FACR)** and **Rose Jarosz** in the Department of Radiology; **Rajiv Dhir (MBBS, MBA)** and **Tony Green (HT (ASCP), AS)** in the Department of Pathology (PBC).

### Data Access

#### Data Access

Data Type	Download all or Query/Filter
Images (DICOM, 54.7 GB)	<a href="#">Download</a> <a href="#">Search</a> (Download requires the <a href="#">NBIA Data Retriever</a> )
Tissue Slide Images (SVS, 190 GB)	<a href="#">Download</a> <a href="#">Search</a>
Clinical Data API (JSON - <a href="#">more info</a> )	<a href="#">Download</a>
<a href="#">Discovery Study</a> (CPTAC Data Portal)	<ul style="list-style-type: none"> <li>• <a href="#">Proteomic Data Commons</a></li> <li>• <a href="#">Genomic Data Commons</a></li> <li>• <a href="#">Pathology</a></li> <li>• <a href="#">Radiology</a></li> </ul>

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:

- [Crowds Cure Cancer: Data collected at the RSNA 2018 annual meeting \(Crowds-Cure-2018\)](#)

### Detailed Description

#### Detailed Description

	Radiology Image Statistics	Pathology Image Statistics
Modalities	CT, MR, CR, SR, DX	Pathology
Number of Participants	63	222

Number of Studies	82	N/A
Number of Series	670	N/A
Number of Images	94,500	783
Images Size (GB)	54.7	190

## 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](mailto: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 Clear Cell Renal Cell Carcinoma [CPTAC-CCRCC] collection [Data set]. The Cancer Imaging Archive. <https://doi.org/10.7937/k9/tcia.2018.oblamn27>

#### 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 9 (Current): 2020/09/03

Data Type	Download all or Query/Filter
Images (DICOM, 54.7 GB)	  (Requires <a href="#">NBIA Data Retriever</a> )
Tissue Slide Images (SVS, 190 GB)	 
Clinical Data (web)	<a href="#">CPTAC Discovery Study (Clinical)</a>
Proteomics (web)	<a href="#">CPTAC Discovery Study (Proteomics)</a>




Changed to new Aspera download link for histopathology slides.

#### Version 8: Updated 2019/09/30

Data Type	Download all or Query/Filter
Images (DICOM, 54.7 GB)	  (Requires <a href="#">NBIA Data Retriever</a> )
Tissue Slide Images (SVS, 190 GB)	 
Clinical Data (web)	<a href="#">CPTAC Discovery Study (Clinical)</a>
Proteomics (web)	<a href="#">CPTAC Discovery Study (Proteomics)</a>



Added new subjects.

### Version 7: Updated 2019/06/30

Data Type	Download all or Query/Filter
Images (DICOM, 35.8 GB)	 (Requires <a href="#">NBIA Data Retriever</a> )
Tissue Slide Images (SVS, 190 GB)	 
Clinical Data (web)	<a href="#">CPTAC Discovery Study (Clinical)</a>
Proteomics (web)	<a href="#">CPTAC Discovery Study (Proteomics)</a>


Added new subjects.

### Version 6: Updated 2019/03/30

Data Type	Download all or Query/Filter
Images (DICOM, 37.1 GB)	 (Requires <a href="#">NBIA Data Retriever</a> )
Tissue Slide Images (web)	
Clinical Data (web)	<a href="#">CPTAC Discovery Study (Clinical)</a>
Proteomics (web)	<a href="#">CPTAC Discovery Study (Proteomics)</a>

Added new subjects




### Version 5: Updated 2019/01/23

Data Type	Download all or Query/Filter
Images (DICOM, 19.9 GB)	 (Requires <a href="#">NBIA Data Retriever</a> )
Tissue Slide Images (web)	
Clinical Data (web)	<a href="#">CPTAC Discovery Study (Clinical)</a>

Proteomics (web)	<a href="#">CPTAC Discovery Study (Proteomics)</a>
------------------	--




Added links to clinical and proteomic data from CPTAC Discovery Study (first 110 patients)

## Version 4: Updated 2018/10/24

Data Type	Download all or Query/Filter
Images (DICOM, 19.9 GB)	 (Requires <a href="#">NBIA Data Retriever.</a> )
Tissue Slide Images (web)	
Proteomics (web)	


Added new subjects.



## Version 3: Updated 2018/06/30

Data Type	Download all or Query/Filter
Images (DICOM, 13.5 GB)	 (Requires <a href="#">NBIA Data Retriever.</a> )
Tissue Slide Images (web)	
Proteomics (web)	

Added new subjects.



## Version 2: Updated 2018/04/26

Data Type	Download all or Query/Filter
Images (DICOM, 11.8 GB)	 (Requires <a href="#">NBIA Data Retriever.</a> )

Tissue Slide Images (web)	
Proteomics (web)	

Added new subjects.

### Version 1: Updated 2018/01/10

Data Type	Download all or Query/Filter
Images (DICOM, 2.96 GB)	 (Requires <a href="#">NBIA Data Retriever</a> .)
Tissue Slide Images (web)	
Proteomics (web)	