Applied Proteogenomics OrganizationaL Learning and Outcomes (APOLLO-5)

Summary

This collection contains subjects from the National Cancer Institute's Applied Proteogenomics OrganizationaL Learning and Outcomes (APOLLO) network. The APOLLO network is a collaboration between NCI, the Department of Defense (DoD), and the Department of Veterans Affairs (VA) to incorporate proteogenomics into patient care as a way of looking beyond the genome, to the activity and expression of the proteins that the genome encodes. The emerging field of proteogenomics aims to better predict how patients will respond to therapy by screening their tumors for both genetic abnormalities and protein information, an approach that has been made possible in recent years due to advances in proteomic technology. Radiology and pathology images from APOLLO patients are being collected and made publicly available by The Cancer Imaging Archive to enable researchers to investigate cancer phenotypes that may correlate to corresponding proteomic, genomic, and clinical data.

Image data are being made available on a release schedule that is coordinated with the APOLLO program releases of proteomic and genomic data. 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 can be heterogeneous in terms of scanner modalities, manufacturers, and acquisition protocols. Pathology imaging is collected as part of the APOLLO qualification and laser capture microdissection (in some cases) workflow. *Li mited Access* data is available only to members of the APOLLO network, and *Publicly Available* data may be under a limited publication embargo.

Collection	Cancer Type	Modalities	Subjects
APOLLO-5-BLCA	Bladder Cancer	CT, MR	1
APOLLO-5-BRCA	Breast Cancer	CT, MG, MR, NM, US	22
APOLLO-5-CCRCC	Kidney Clear Cell Renal Cell Carcinoma	CT. MR, NM, PT	9
APOLLO-5-CHOL	Cholangiocarcinoma	CT, MR, PT	4
APOLLO-5-CM	Cutaneous Melanoma	CT, MR, NM, PT, US	5
APOLLO-5-COAD	Colon adenocarcinoma	CT, MR, PT, US	34
APOLLO-5-ENDOCRINE-MISC	Endocrine	CT, US	1
APOLLO-5-ESCA	Esophageal Carcinoma	CT, PT	3
APOLLO-5-GIST	Gastrointestinal Stromal Tumor	CT, PT	2
APOLLO-5-HNSCC	Head And Neck Squamous Cell Carcinoma	CT, MR, PT	8
APOLLO-5-KICH	Kidney Chromophobe	CT, MR	2
APOLLO-5-KIRP	Kidney Renal Papillary Cell Carcinoma	CT, MR	3
APOLLO-5-LIHC	Liver Hepatocellular Carcinoma	CT, MR, NM, PT, US	6

APOLLO-5-LSCC	Lung Squamous Cell Carcinoma	CT, MR, PT	45
APOLLO-5-LUAD	Lung Adenocarcinoma	CT, MR, PT	67
APOLLO-5-LUNG-MISC	Lung Other	CT, PT	6
APOLLO-5-MISC	Miscellaneous	CT, MR, PT	4
APOLLO-5-MSG	Major Salivary Gland	CT, PT	1
APOLLO-5-NET	Neuroendocrine Tumors (all sites)	CT, MR, NM, PT, US, XA	15
APOLLO-5-NONCANCER	Pathologically Benign	CT, MR, NM, PT	5
APOLLO-5-OV	Ovarian Cancer	CT, PT	23
APOLLO-5-PAAD	Pancreatic Adenocarcinoma	CT, MR, PT, US	5
APOLLO-5-PRAD	Prostate Adenocarcinoma	CT, PT, US, MR	3
APOLLO-5-SAR	Soft Tissue	CT, MR, NM	4
APOLLO-5-THCA	Thyroid Carcinoma	CT, NM, PT, US	2
APOLLO-5-THYM	Thymoma	CT, MR, PT	2
APOLLO-5-UCEC	Uterine Corpus Endometrial Carcinoma	CT	1

Data Access

Data Access

This is a limited access data set that is currently only available to APOLLO investigators. If you are a member of this network and would like to request access, please fill out this PDF form. Please email the completed form to the email address indicated on the form. You will be notified with further information when it is received.

Data Type	Download all or Query/Filter	License
Images (DICOM)	*Search*	APOLLO Restricted

Click the Versions tab for more info about data releases.

Detailed Description

Detailed Description

Collection Statistics	
Modalities	CT, MG, MR, NM, PET, US, XA
Number of Participants	283
Number of Studies	3,586
Number of Series	22,419
Number of Images	2,527,224

Image Size (GB)	1.3 TB

De-identification of DICOM dates

The resulting DICOM dates are meaningless yet preserve the relative temporal distance between studies for a patient

De-identification of dates uses the DICOM standard “Retain Longitudinal With Modified Dates Option” which allows dates to be retained as long as they are modified from the original date. Date and Date-Time fields in TCIA DICOM image headers are de-identified by normalizing to a base date of January 1, 1975 and then shifted by the number of days between the original Study Date and an "anchor date". The anchor date for APOLLO is the Date of Diagnosis. The choice of '1975' was arbitrary, but it allows one to ensure that the dates in de-identified DICOM files have been properly de-identified as anything not around that year would be suspect.

TCIA Study Date = 01/01/1975 + (Original Study Date – Date of Diagnosis).

For example, if the original Study Date was 03/29/2018 and the Date of Diagnosis was 03/27/2018 then the Days from Diagnosis would be +2 and the TCIA Study Date would become 01/03/1975.

This technique de-identifies the dates while preserving the longitudinal relationship between dates. Therefore, a researcher won’t know the precise date the scan occurred, but if a follow up scan was performed 120 days later, that same 120 day difference between scans of a subject will exist in the TCIA images. Dates that occur in DICOM tags other than Date or Date-Time fields are removed. An example of this would be a date entered into the Series Description field. If the date is associated with a library for Code Meaning then that date is preserved as the date would be required to look up the meaning in the correct version of the library. To show that the dates have been modified, the term “MODIFIED” is written into DICOM tag (0028,0303) “LongitudinalTemporalInformationModified”.

Original dates will be first normalized to 01 January, 1975 and then offset relative to the date of diagnosis. The CTP code for shifting the StudyDate is shown below:

<e en="T" t="00080020" n="StudyDate"> @dateinterval(StudyDate,diagnosisdate,PatientID,@NORMDATE)</e>

Insertion of computed "Days from Diagnosis" value

The inserted "Days from Diagnosis" value can be compared with similar values in the APOLLO clinical data to understand the clinical context of the imaging study

The number of days the study occurred relative to the date of diagnosis is calculated by the CTP software (using the diagnosis date in the CTP lookup table at the submission site) and automatically stored in the DICOM tag (0012,0052) Longitudinal Temporal Offset from Event with the associated tag (0012,0053) Longitudinal Temporal Event Type set to “Days from Diagnosis”. The days from diagnosis links the imaging data to the clinical data for a given subject. The CTP code for this is:

<e en="T" t="00120052" n="LongitudinalTemporalOffsetfromEvent">@always()@dateinterval(StudyDate,ddate, PatientID)</e>

<e en="T" t="00120053" n="LongitudinalTemporalEventType">@always()@param(@LTET)</e> (where LTET is defined as DIAGNOSIS)

Insertion of "Diagnosis Year"

It is important for cancer researchers to know the timeframe for which the cancer was diagnosed to relate the prescribed cancer treatment or staging to what was available at that time.

In order to relate the treatments that were available at the time of the diagnosis, the year that the primary diagnosis was made is recorded in a CTP owned group 13 private tag as follows.

<e en="T" t="00131051" n="DiagnosisYear">@always()@lookup(PatientID,diagnosisdate)</e>

In a separate stage of the pipeline the diagnosisdate is truncated to be just the year that the diagnosis was made.

<e en="T" t="00131051" n="DiagnosisYear">@truncate(DiagnosisYear,-4)</e>

The approximate StudyYear can be calculated by adding the days from diagnosis in tag LongitudinalTemporalOffsetfromEvent to the DiagnosisYear.

In order to use a normalized date function the private tags must also be de-identified at the site using a CTP script that encapsulates the TCIA Safe Private Tag Knowledge Base. With this approach, only the Safe Private Tags contained within the TCIA Private Tag Knowledge Base and encoded into the CTP script at the time the CTP script was created will be retained. If there are Private Tags that are known to be important but not part of the current Safe tags of the TCIA Private Tag Knowledge Base, then it is up to the submitting site to submit a Private Tag Dictionary of those tags to TCIA for consideration.

The normalized date workflow described above requires that diagnosis date be present and this workflow does not handle the example where there no diagnosis date is present.

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:

Acknowledgement

Citing APOLLO Data in Publications

The APOLLO Research Network require that publications using data from this program (1) cite all relevant publications and preprints describing the APOLLO data referenced in the manuscript; and (2) cite the relevant DOIs and/or study accession numbers for the data referenced in the manuscript.

Acknowledging the APOLLO Research Network in Publications

The APOLLO Research Network requests that publications using data from this program include the following statement: “Data used in this publication were generated by the Applied Proteogenomics OrganizationaL Learning and Outcomes (APOLLO) Research Network, a Federal Precision Oncology and Cancer Moonshot Program of the Department of Defense, Department of Veterans Affairs, and National Cancer Institute."

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 that leverage our data. At this time, we are not aware of any publications based on this data. If you have a publication you'd like to add, please contact TCIA's Helpdesk.

Versions

Versions

Versions are not being tracked during the project phase while access is limited to APOLLO investigators. Beginning in August 2022 new data are being added on a monthly basis as the program accrues them.