

CPTAC Clinical Data API

Clinical data from CPTAC-3 are now being made available through an API hosted by ESAC. This is documented in detail at <https://clinicalapi-cptac.esacinc.com/api/tcia/>. Data are provided in **JSON**, which is a file format commonly used by software developers for data interchange. Most data scientists and statisticians should already be familiar with this type of file, but for users who are not familiar working with JSON, we are providing a few examples of ways to access this data.

1. Your Browser will display the interface differently:
 - a. [Firefox](#) (recommended)
 - b. [Chrome](#)
 - c. [Edge](#) (and notepad++)
 - d. Safari (and sublime text or oxygenxml)
2. You may
 - a. view the data online (Firefox) in human-readable format
 - b. download to your local machine for viewing and analysis


View JSON files in Firefox

Open Firefox and navigate to one of our CPTAC collection pages (e.g. [The Clinical Proteomic Tumor Analysis Consortium Clear Cell Renal Cell Carcinoma Collection \(CPTAC-CCRCC\)](#)) and click the blue download button for the Clinical Data API:

[Data Access](#) [Detailed Description](#) [Citations & Data Usage Policy](#) [Versions](#)

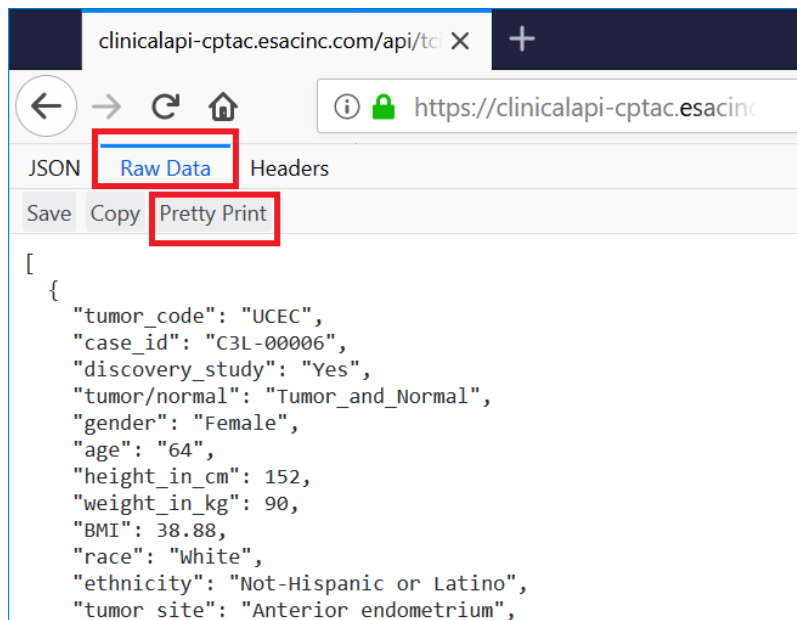
Data Access

Click 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, 31.9 GB)	Download Search
Tissue Slide Images (SVS, 151.6 GB)	Download Search
Clinical Data API (JSON - more info)	Download 
Discovery Study Proteomics/Clinical Data (external)	<ul style="list-style-type: none"> CPTAC Data Portal (Georgetown) Proteomics Data Commons
Genomics/Clinical Data (external)	Genomic Data Commons

Click the Versions tab for more info about data releases.

Click the **Raw Data** button at the top of the page, and then click **Pretty Print**. This will arrange the data in a way that is more human-readable.



The screenshot shows a web browser window with the address bar displaying "clinicalapi-cptac.esacinc.com/api/tci". The page has a dark header with a "Raw Data" button highlighted by a red box. Below the header, there are tabs for "JSON", "Raw Data", and "Headers". The "Raw Data" tab is selected. Below the tabs, there are buttons for "Save", "Copy", and "Pretty Print", with the "Pretty Print" button also highlighted by a red box. The main content area displays a JSON object in a human-readable format:

```
[
  {
    "tumor_code": "UCEC",
    "case_id": "C3L-00006",
    "discovery_study": "Yes",
    "tumor/normal": "Tumor_and_Normal",
    "gender": "Female",
    "age": "64",
    "height_in_cm": 152,
    "weight_in_kg": 90,
    "BMI": 38.88,
    "race": "White",
    "ethnicity": "Not-Hispanic or Latino",
    "tumor_site": "Anterior endometrium",
```

Online JSON to CSV conversion tools

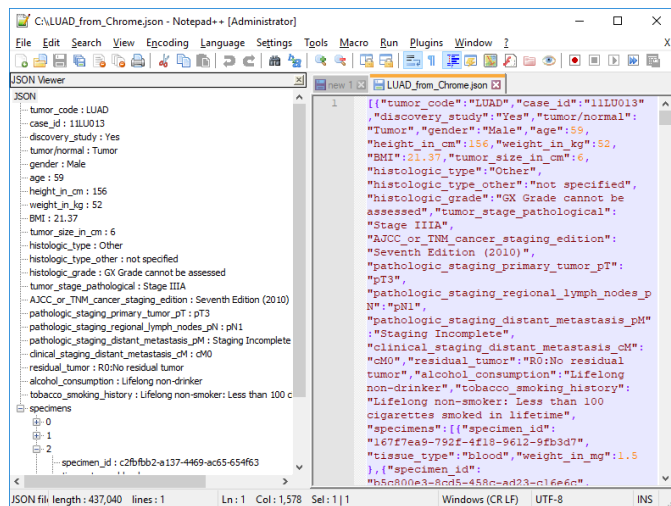
There are a number of free online tools such as <https://json-csv.com/> which will convert between JSON and CSV, which can be viewed in most spreadsheet applications like Excel. After the page loads simply paste in the URL for the cancer type you were trying to access, e.g. https://clinicalapi-cptac.esacinc.com/api/tcia/clinical_data/tumor_code/UCEC.

For Edge browser, and Notepad++ editor:

blocked URL

Open **Notepad++**. From the Plugins menu click "**Plugins Admin**" then type JSON into the search bar, you should see "**JSON viewer**" as an option. Click install.

Go to the Plugins Menu. Click JSON Viewer, and open your file. There is a tree-view at left and colored view in the main window.



Loading JSON files into Excel

The following guide has been prepared for researchers who are not used to working with JSON data or APIs. By following these steps you will be able to view the data in newer versions of Excel via the "Power Query" or "Get and Transform" feature.

Excel 2010 and Excel 2013 for Windows

Power Query is a free add-in that can be downloaded here: [Download the Power Query add-in](#). Once enabled, Power Query functionality is available from the Power Query tab on the ribbon.


To generate the JSON file:

- First access the desired CPTAC collection page (e.g. <https://wiki.cancerimagingarchive.net/display/Public/CPTAC-CCRCC>)
- Click on the blue download button next to “**Clinical Data API...**) data type, in the Data Access Box

[Data Access](#) [Detailed Description](#) [Citations & Data Usage Policy](#) [Versions](#)

Data Access

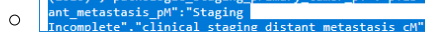
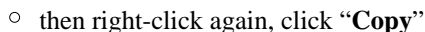
Click the **Download** button to save a ".tcia" manifest file to your computer, which you must open with the [NBIA Data Retriever](#). Click the **Search** button to open our Data Portal, where you can browse the data collection and/or download a subset of its contents.

Data Type	Download all or Query/Filter
Images (DICOM, 31.9 GB)	Download Search
Tissue Slide Images (SVS, 151.6 GB)	Download Search
Clinical Data API (JSON - more info)	Download 
Discovery Study Proteomics/Clinical Data (external)	<ul style="list-style-type: none"> • CPTAC Data Portal (Georgetown) • Proteomics Data Commons
Genomics/Clinical Data (external)	Genomic Data Commons

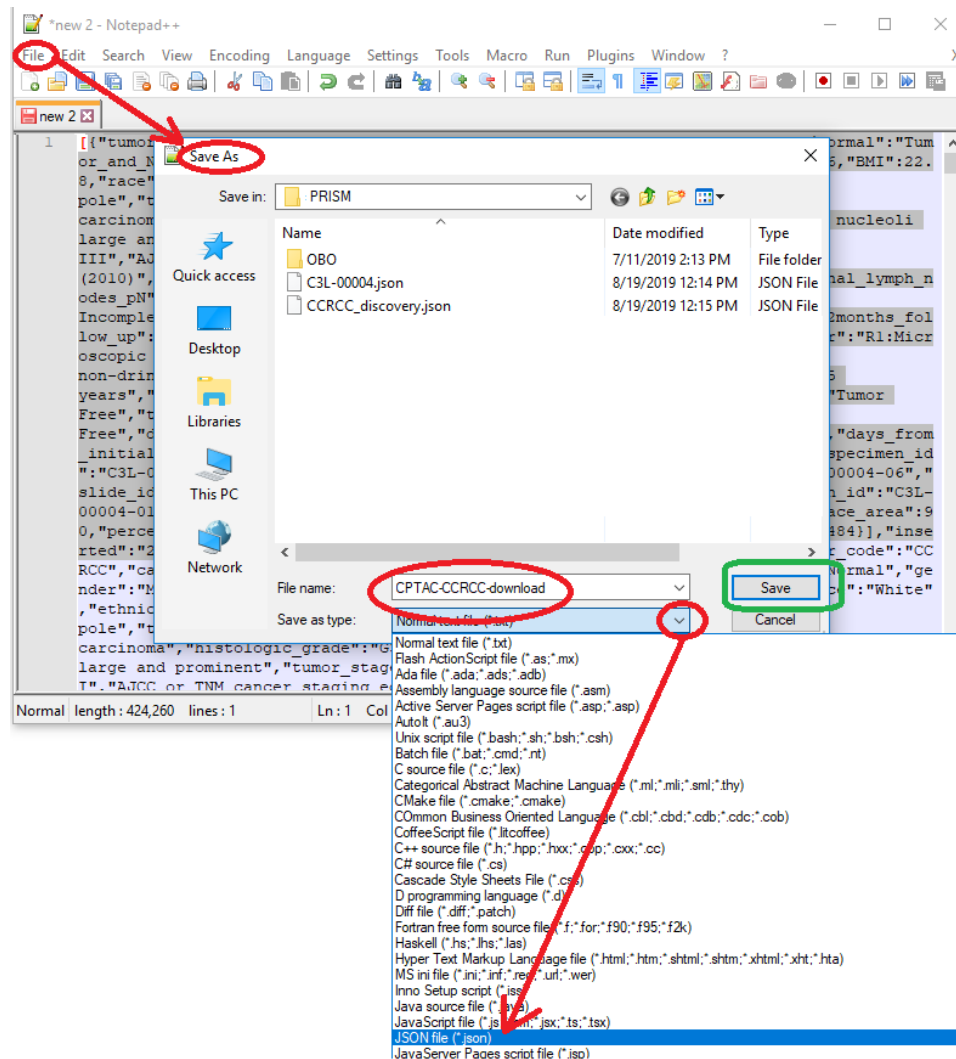
Click the Versions tab for more info about data releases.

- For **Chrome** Browser:
 - Right-click anywhere in the in the middle of the page and select “**Save as...**” and save the file in your desired location

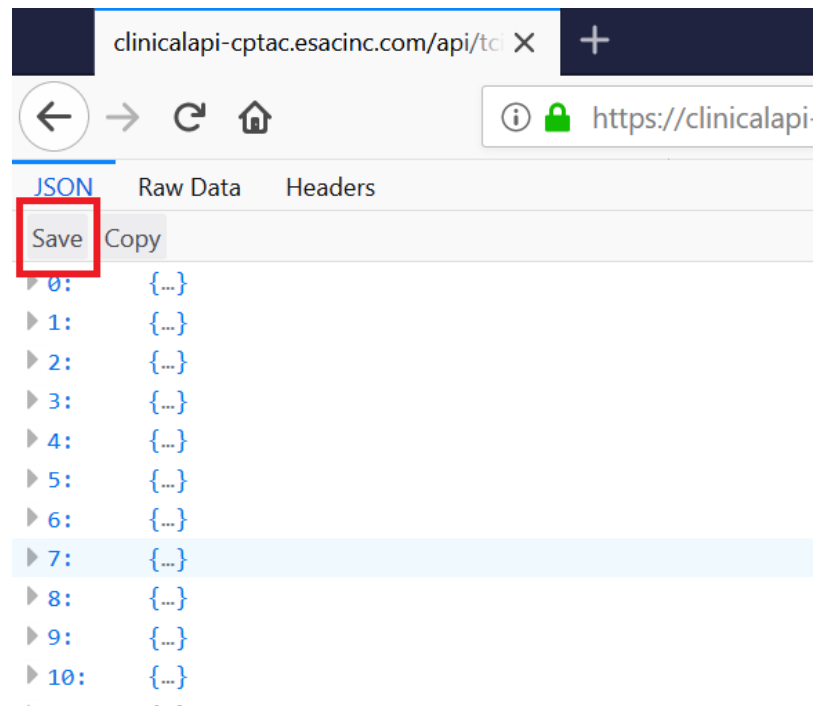
- Right-click anywhere in the in the middle of the page and select "**Select all**"



- "Paste" the contents of the file into an editor, for example "Notepad++",
 - Give your new file a **new name**
 - as file type **JSON**
 - Click **Save**.

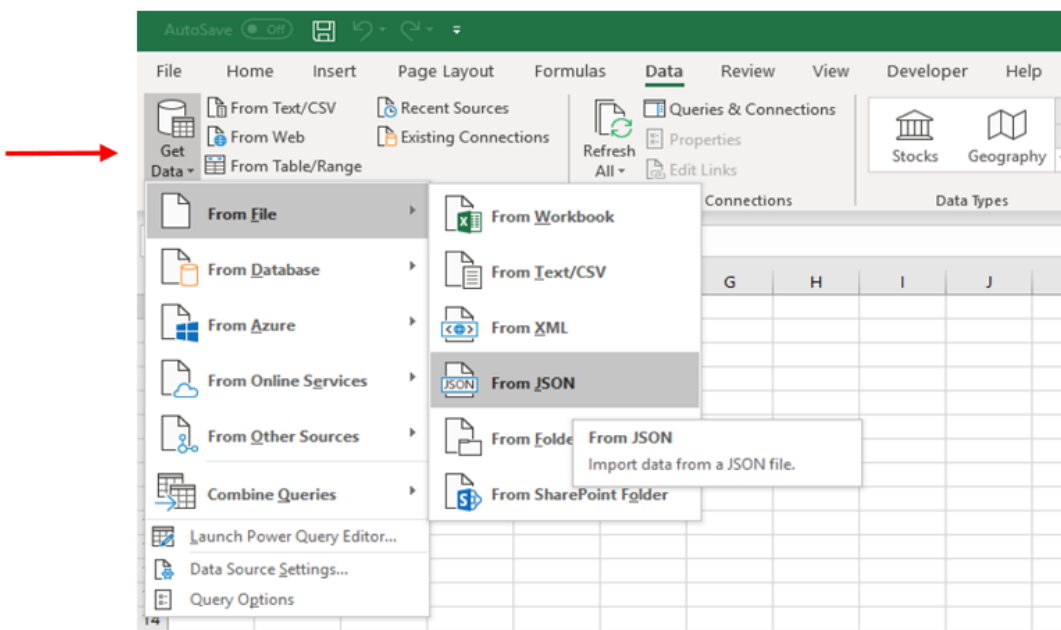


- For **Firefox** browser:
 - Click on "Save" near the top left corner of the page and save the file in your desired location

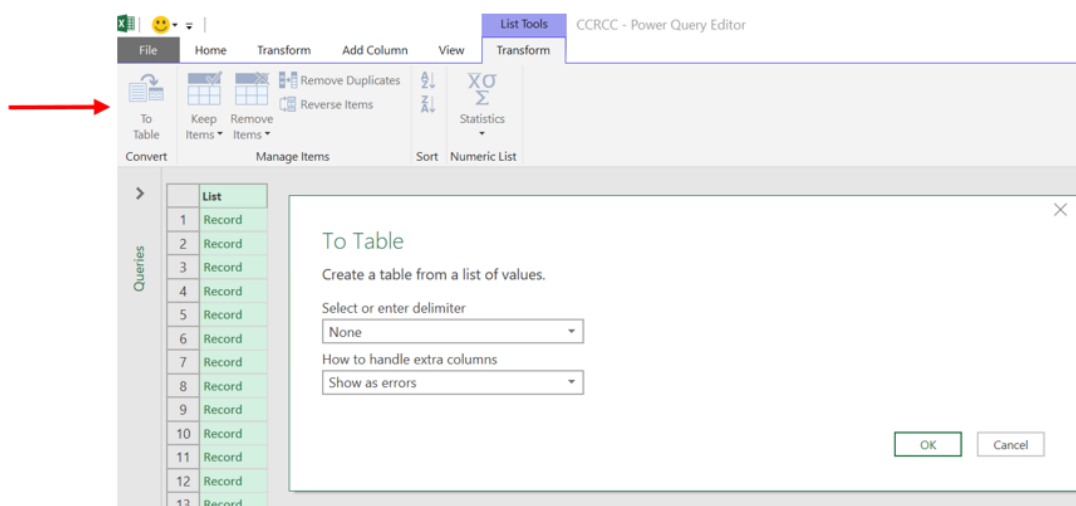


Load and Convert the JSON Clinical Data file into an Excel Sheet (Office 365)

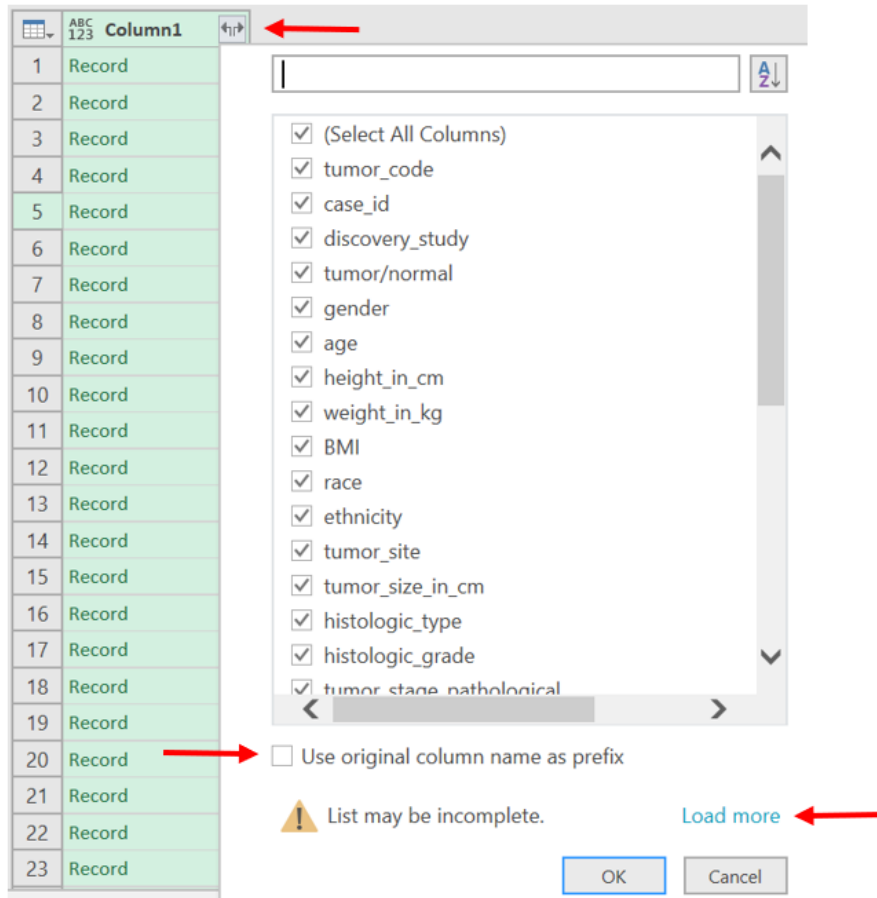
- Open a blank Excel sheet in Office 365
- From the main menu, select “**Data**”
- From the “**Get Data**” drop down menu, select “**From File**”, then “**From JSON**” ****if you can't, see below for custom query**
- Select your CPTAC JSON file from the directory where you saved it previously.



- From the power query editor select “**To Table**” from the main menu.
- In the table window, keep the delimiter and extra column selections as shown below and click on “**ok.**”



- Click on the expansion arrow icon next to the “**Column1**” header, then click on “**load more**” at the bottom of the column header list to ensure all columns are selected.
- Uncheck the box next to “**Use original column name as prefix**”
- Click on “**ok**” after the full list has loaded and “**ok**” again to load all columns into the table.

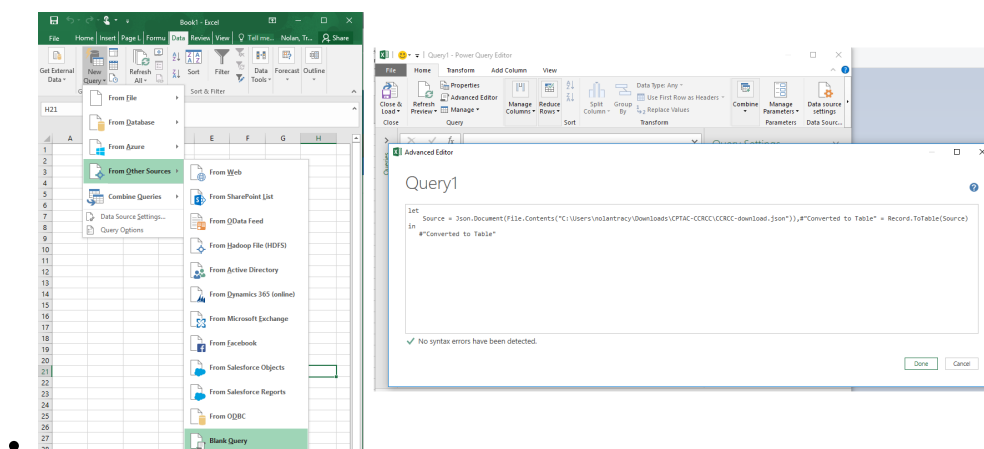


- **Note:** The “specimens” column has a sub-level with specimen-related column headings (specimen id, slide id, etc) which prevent it from showing up without additional processing. We recommend simply deleting this column.
- **Note:** You can also use the Microsoft Excel **Query Editor** to create formulas to connect to JSON files. If you don't see the **Get Data From JSON** button, you can go to **New Query > From Other Sources > Blank Query**, then go to **Advanced Editor** and input your custom query string, then follow the other instructions above to convert from the json file to the tabular format.

Blank Query to import JSON

```
let
    Source = Json.Document(File.Contents("C:\Users\Name\Desktop\CPTAC-AML-download.json")), #"Converted to Table" = Record.ToTable(Source)

in
    #"Converted to Table"
```



Save As Excel Worksheet

- After you have customized the table, select **"Close and Load"**
- Then click on **"File"**, and **"Save"** to save as an Excel document

tumor_code	case_id	discovery_study	tumor/normal	gender	age	height_in_cm	weight_in_kg	BMI
CCRCC	C3L-00004	Yes	Tumor_and_Normal	Male	72	170	66	22.1
CCRCC	C3L-00010	Yes	Tumor_and_Normal	Male	30	177	107	34.1
CCRCC	C3L-00011	Yes	Tumor_and_Normal	Female	63	180	89	27.4
CCRCC	C3L-00026	Yes	Tumor_and_Normal	Female	65	163	75	28.2
CCRCC	C3L-00079	Yes	Tumor_and_Normal	Male	49	175	116	37.8
CCRCC	C3L-00088	Yes	Tumor_and_Normal	Male	72	178	104	32.8
CCRCC	C3L-00096	Yes	Tumor_and_Normal	Male	52	185	94	27.4
CCRCC	C3L-00097	Yes	Tumor_and_Normal	Male	59	175	113	36.8
CCRCC	C3L-00103	Yes	Tumor_and_Normal	Male	56	180	93	28.1
CCRCC	C3L-00165	No		Male	47	168	61	21.6

