GDC API User's Guide

NCI Genomic Data Commons (GDC)

Contents

Chapter 1

Getting Started

Getting Started

The GDC Application Programming Interface (API): An Overview

The GDC API drives the GDC Data and Submission Portals and provides programmatic access to GDC functionality. This includes searching for, downloading, and submitting data and metadata. The GDC API uses JSON as its communication format, and standard HTTP methods like `GET`, `PUT`, `POST` and `DELETE`.

This guide explains how to construct and execute API requests and interpret API responses.

Tools for communicating with the GDC API

Many third-party tools can be used for communicating with the GDC API and for preparing and visualizing API calls.

Examples of tools for communicating with the GDC API:

Tool	Type
Curl	Command line tool
HTTPie	Command line tool
Postman REST Client	App for Google Chrome and OS X
DHC REST Client	Google Chrome extension
Google Chrome	Google Chrome web browser

Examples of tools that can help build GDC API calls:

Tool	Description
JSONLint	Validate JSON
JSON Formatter	Format, validate, and convert JSON to other formats
Percent-(URL)-encoding tool	Tool for percent-encoding strings
JSON escape tool	Tool for escaping strings using JSON string rules

API Endpoints

Communicating with the GDC API involves making calls to API endpoints. Each GDC API endpoint represents specific API functionality, as summarized in the following table:

Endpoint	Type	Description
status	Status	Get the API status and version information
projects	Search & Retrieval	Search all data generated by a project
cases	Search & Retrieval	Find all files related to a specific case, or sample donor.
files	Search & Retrieval	Find all files with specific characteristics such as file_name, md5sum, data_format and others.
annotations	Search & Retrieval	Search annotations added to data after curation
data	Download	Used to download GDC data
manifest	Download	Generates manifests for use with GDC Data Transfer Tool
slicing	BAM Slicing	Allows remote slicing of BAM format objects
submission	Submission	Returns the available resources at the top level above programs i.e., registered programs

The HTTP URL that corresponds to the latest version of a GDC API endpoint is `https://gdc-api.nci.nih.gov/<endpoint>`, where `<endpoint>` is the name of the endpoint.

The HTTP URL of an endpoint corresponding to a specific major version of the GDC API is `https://gdc-api.nci.nih.gov/<version>/` where `<endpoint>` is the name of the endpoint and `<version>` is the GDC API version.

For example, the address of the latest version of the `status` endpoint is `https://gdc-api.nci.nih.gov/status`, whereas the address of the `status` endpoint corresponding to version 0 of GDC API is `https://gdc-api.nci.nih.gov/v0/status`.

GDC Legacy Archive

To interact with data in the GDC Legacy Archive, add `legacy` to the endpoint URL:

```
1 https://gdc-api.nci.nih.gov/<version>/legacy/<endpoint>
```

Entity UUIDs

All objects (*entities*) in the GDC are assigned a unique identifier in the form of a version 4 universally unique identifier (UUID). The UUID uniquely identifies the entity in the GDC, and is stored in the entity's `id` property.

UUIDs are frequently used in GDC API requests and responses to identify specific entities like files, cases, and samples.

See GDC Data Model for details.

Sample Request

The following is an example of a request to the `files` endpoint, which retrieves information about a BAM file stored in the GDC.

```
1 curl https://gdc-api.nci.nih.gov/files/d853e541-f16a-4345-9f00-88e03c2dc0bc?pretty=true
```

"' python import requests import json

file_endpt = 'https://gdc-api.nci.nih.gov/files/' file_uuid = 'd853e541-f16a-4345-9f00-88e03c2dc0bc' response = requests.get(file_endpt + file_uuid) print json.dumps(response.json(), indent=2)

```
 1  {
 2    "data": {
 3      "data_type": "Aligned Reads",
 4      "updated_datetime": "2016-05-26T17:06:40.003624-05:00",
 5      "created_datetime": "2016-05-26T17:06:40.003624-05:00",
 6      "file_name": "0017ba4c33a07ba807b29140b0662cb1_gdc_realn.bam",
 7      "md5sum": "a08304b120c5df76b6532da0e9a35ced",
 8      "data_format": "BAM",
 9      "acl": [
10        "phs000178"
11      ],
12      "access": "controlled",
13      "platform": "Illumina",
14      "state": "submitted",
15      "file_id": "d853e541-f16a-4345-9f00-88e03c2dc0bc",
16      "data_category": "Raw Sequencing Data",
17      "file_size": 23650901931,
18      "submitter_id": "c30188d7-be1a-4b43-9a17-e19ccd71792e",
19      "type": "aligned_reads",
20      "file_state": "processed",
21      "experimental_strategy": "WXS"
22    },
23    "warnings": {}
24  }
```

Authentication

Authentication is required for downloading controlled-access data, and for all data submission functionality. The GDC API uses tokens for authentication.

Users can obtain authentication tokens from the GDC Data Portal and the GDC Data Submission Portal. See the GDC Data Portal User's Guide and the GDC Data Submission Portal User's Guide for instructions.

Using Authentication Tokens

All API requests that require authentication must include a token as an `X-Auth-Token` custom HTTP header.

In the following example, an authentication token is saved as an environment variable and passed to `curl` to download a controlled-access file:

```
 1  export
      token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
 2
 3  curl -O -J -H "X-Auth-Token: $token"
      'https://gdc-api.nci.nih.gov/data/a1c1b23b-cc41-4e85-b1b7-62a42873c5af'
```

```
 1    % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
 2                                   Dload  Upload   Total   Spent    Left  Speed
 3  100 31.4M  100 31.4M    0     0   290k      0  0:01:50  0:01:50 --:--:--   172k
 4  curl: Saved to filename
      'ACOLD_p_TCGA_Batch17_SNP_N_GenomeWideSNP_6_A03_466078.tangent.copynumber.data.txt'
```

For more information about authentication tokens, including token expiration and rotation, see Data Security.

NOTE: The authentication token should be kept in a secure location, as it allows access to all data accessible by the associated user account.

Chapter 2

Search and Retrieval

Search and Retrieval

Introducing Search and Retrieval Requests

The GDC API provides endpoints that search and retrieve information stored in the GDC according to the GDC Data Model. The general format of requests to search & retrieval endpoints is described below.

Note: Queries described in this section work for datasets that have been released to the GDC Data Portal. Unreleased data that is in the process of being submitted to GDC cannot be queried using these methods. See Submission to learn how to query unreleased data using GraphQL.

Components of a Request

A typical search and retrieval API request specifies the following parameters:

- a `filters` parameter, that specifies the search terms for the query
- several parameters that specify the API response, such as:
    - `format` — specifies response format (JSON, TSV, XML)
    - `fields` — specifies the which data elements should be returned in the response, if available
    - `size` — specifies the the maximum number of results to include in the response
    - other parameters are described below.

Requests can be executed using HTTP GET or HTTP POST. GET requests are limited by maximum URL length, so the POST method is recommended for large queries.

Note: Requests for information stored in the GDC Legacy Archive must be directed to `legacy/` endpoints. See Getting Started for details.

POST Example

The following is an example of an HTTP POST request to the `files` endpoint of the GDC API. It looks for Gene Expression Quantification files associated with specific TCGA cases (represented by TCGA barcodes) and retrieves the associated biospecimen metadata in TSV format.

Request

```
1  curl --request POST --header "Content-Type: application/json" --data @Payload
       'https://gdc-api.nci.nih.gov/files' > response.tsv
```

Payload

```
1  {
2      "filters":{
3          "op":"and",
4          "content":[
5              {
6                  "op":"in",
7                  "content":{
8                      "field":"cases.submitter_id",
9                      "value":[
10                         "TCGA-CK-4948",
11                         "TCGA-D1-A17N",
12                         "TCGA-4V-A9QX",
13                         "TCGA-4V-A9QM"
14                     ]
15                 }
16             },
17             {
18                 "op":"=",
19                 "content":{
20                     "field":"files.data_type",
21                     "value":"Gene Expression Quantification"
22                 }
23             }
24         ]
25     },
26     "format":"tsv",
27     "fields":"file_id,file_name,cases.submitter_id,cases.case_id,data_category,data_type,cases.samples.tumor_des
28     "size":"1000"
29 }
```

Each component of the request is explained below.

GET Example

The above request can be executed as an HTTP GET:

```
1  https://gdc-api.nci.nih.gov/files?filters=%7B%22op%22%3A%22and%22%2C%22content%22%3A%5B%7B%22op%22%3A%22in%22%2
```

Each component of the request is explained below.

Endpoints

The following search and retrieval endpoints are available in the GDC API:

Endpoints	Description
files	Information about files stored in the GDC
cases	Information related to cases, or sample donors.
projects	Information about projects

Endpoints	Description
annotations	Information about annotations to GDC data
_mapping	Information about elements that can be used to query other endpoints

The choice of endpoint determines what is listed in the search results. The `files` endpoint will generate a list of files, whereas the `cases` endpoint will generate a list of cases. Each of the above endpoints, other than `_mapping`, can query and return any of the related fields in the GDC Data Model. So the `cases` endpoint can be queried for file fields (e.g. to look for cases that have certain types of experimental data), and the `files` endpoint can be queried for clinical metadata associated with a case (e.g. to look for files from cases diagnosed with a specific cancer type).

Project Endpoint

The `projects` endpoint provides access to project records, the highest level of data organization in the GDC.

Example

This example is a query for projects contained in the GDC. It uses the from, size, sort, and pretty parameters, and returns the first two projects sorted by project id.

```
1 curl 'https://gdc-api.nci.nih.gov/projects?from=1&size=2&sort=project.project_id:asc&pretty=true'
```

```
1
2    {
3      "data": {
4        "hits": [
5          {
6            "state": "legacy",
7            "project_id": "TCGA-ACC",
8            "primary_site": "Adrenal Gland",
9            "disease_type": "Adrenocortical Carcinoma",
10           "name": "Adrenocortical Carcinoma"
11         },
12         {
13           "dbgap_accession_number": "phs000464",
14           "disease_type": "Acute Lymphoblastic Leukemia",
15           "state": "legacy",
16           "primary_site": "Blood",
17           "project_id": "TARGET-ALL-P2",
18           "name": "Acute Lymphoblastic Leukemia - Phase II"
19         }
20       ],
21       "pagination": {
22         "count": 2,
23         "sort": "project.project_id:asc",
24         "from": 1,
25         "pages": 22,
26         "total": 44,
27         "page": 1,
28         "size": 2
29       }
30     },
31     "warnings": {}
32   }
```

Retrieval of project metadata using project_id

The `project` endpoint supports a simple query format that retrieves the metadata of a single project using its `project_id`:

```
1 curl
      'https://gdc-api.nci.nih.gov/projects/TARGET-NBL?expand=summary,summary.experimental_strategies,summary.data
```

```
1  {
2    "data": {
3      "dbgap_accession_number": "phs000467",
4      "name": "Neuroblastoma",
5      "summary": {
6        "data_categories": [
7          {
8            "case_count": 151,
9            "file_count": 471,
10           "data_category": "Transcriptome Profiling"
11         },
12         {
13           "case_count": 216,
14           "file_count": 1728,
15           "data_category": "Simple Nucleotide Variation"
16         },
17         {
18           "case_count": 1120,
19           "file_count": 1,
20           "data_category": "Clinical"
21         },
22         {
23           "case_count": 270,
24           "file_count": 599,
25           "data_category": "Raw Sequencing Data"
26         }
27       ],
28       "case_count": 1120,
29       "file_count": 2799,
30       "experimental_strategies": [
31         {
32           "case_count": 221,
33           "file_count": 2170,
34           "experimental_strategy": "WXS"
35         },
36         {
37           "case_count": 151,
38           "file_count": 628,
39           "experimental_strategy": "RNA-Seq"
40         }
41       ],
42       "file_size": 8157089415961
43     },
44     "released": true,
45     "state": "legacy",
46     "primary_site": "Nervous System",
47     "project_id": "TARGET-NBL",
48     "disease_type": "Neuroblastoma"
49   },
50   "warnings": {}
51 }
```

Files Endpoint

The GDC Files Endpoint `https://gdc-api.nci.nih.gov/files` enables search and retrieval of information relating to files stored in the GDC, including file properties such as `file_name`, `md5sum`, `data_format`, and others.

Example

This example is a query for files contained in the GDC. It uses the from, size, sort, and pretty parameters, and returns only the first two files, sorted by file size, from smallest to largest.

```
1  curl 'https://gdc-api.nci.nih.gov/files?from=1&size=2&sort=file_size:asc&pretty=true'
```

```
1      {
2        "data": {
3          "hits": [
4            {
5              "origin": "migrated",
6              "data_type": "Raw microarray data",
7              "platform": "HG-U133_Plus_2",
8              "file_name": "TCGA-AB-2842-03A-01R-0757-21.CEL.README",
9              "md5sum": "56f9a6d58b450bf7e9f6431a86220b9d",
10             "data_format": "CEL",
11             "acl": "open",
12             "access": "open",
13             "uploaded_datetime": 1425340539,
14             "state": "live",
15             "data_subtype": "Raw intensities",
16             "file_id": "ca13321c-02aa-4141-bdb6-84d31e3c5711",
17             "file_size": 43,
18             "experimental_strategy": "Gene expression array"
19           },
20           {
21             "origin": "migrated",
22             "data_type": "Raw microarray data",
23             "platform": "HG-U133_Plus_2",
24             "file_name": "TCGA-AB-2809-03A-01R-0757-21.CEL.README",
25             "md5sum": "56f9a6d58b450bf7e9f6431a86220b9d",
26             "data_format": "CEL",
27             "acl": "open",
28             "access": "open",
29             "uploaded_datetime": 1425340539,
30             "state": "live",
31             "data_subtype": "Raw intensities",
32             "file_id": "299d500b-49e2-4c62-9111-c0691592dce1",
33             "file_size": 43,
34             "experimental_strategy": "Gene expression array"
35           }
36         ],
37         "pagination": {
38           "count": 2,
39           "sort": "file_size:asc",
40           "from": 1,
41           "pages": 300936,
42           "total": 601872,
43           "page": 1,
44           "size": 2
45         }
46       },
```

```
47        "warnings": {}
48      }
```

Retrieval of file metadata using individual UUIDs:

The `files` endpoint supports a simple query format that retrieves the metadata of a single file using its UUID:

```
1 curl 'https://gdc-api.nci.nih.gov/files/000225ad-497b-4a8c-967e-a72159c9b3c9?pretty=true'
```

```
1  {
2    "data": {
3      "data_type": "Raw Simple Somatic Mutation",
4      "updated_datetime": "2016-06-04T23:42:25.428738-05:00",
5      "created_datetime": "2016-06-03T19:04:32.950673-05:00",
6      "file_name": "000225ad-497b-4a8c-967e-a72159c9b3c9.snp.Somatic.hc.vcf.gz",
7      "md5sum": "bbe8a7157acbfc9133e47898650b5437",
8      "data_format": "VCF",
9      "acl": [
10        "phs000178"
11     ],
12     "access": "controlled",
13     "state": "submitted",
14     "file_id": "000225ad-497b-4a8c-967e-a72159c9b3c9",
15     "data_category": "Simple Nucleotide Variation",
16     "file_size": 19690,
17     "submitter_id": "TCGA-VR-A8ET-01A-11D-A403-09_TCGA-VR-A8ET-10B-01D-A403-09_varscan",
18     "type": "simple_somatic_mutation",
19     "file_state": "processed",
20     "experimental_strategy": "WXS"
21    },
22    "warnings": {}
23 }
```

files/ids Endpoint

The `files/ids` endpoint corresponds to the "Quick Search" functionality of the GDC Data Portal. The API response includes all files for which the query matches the beginning (or entirety) of any of the following fields:

```
1  project.project_id
2  project.name
3  project.disease_type.analyzed
4  project.primary_site.analyzed
5  case.aliquot_ids
6  case.submitter_aliquot_ids
7  case.analyte_ids
8  case.submitter_analyte_ids
9  case.case_id.raw
10 case.submitter_id.raw
11 case.portion_ids
12 case.submitter_portion_ids
13 case.sample_ids
14 case.slide_ids
15 case.submitter_slide_ids
16 case.submitter_sample_ids
17 file.file_id.raw
18 file.file_name.raw
19 file.submitter_id
```

Requests to this endpoint must be in the format `files/ids?query=`, as provided in the example below. The endpoint returns up to 500 results, which cannot be adjusted with the `size` parameter.

```
1 curl 'https://gdc-api.nci.nih.gov/files/ids?query=nationwidechildrens.org_clinical.TCGA-EM&pretty=true'
```

```
1  {
2    "data": {
3      "pagination": {
4        "count": 5,
5        "sort": "",
6        "from": 1,
7        "page": 1,
8        "total": 81,
9        "pages": 17,
10       "size": 5
11     },
12     "hits": [
13       {
14         "_type": "file",
15         "file_name": "nationwidechildrens.org_clinical.TCGA-EM-A3FQ.xml",
16         "file_id": "efac6904-ac9f-4a44-bf9c-f7d9a822c127",
17         "_score": 4.644438,
18         "cases": [
19           {
20             "case_id": "fef9c64f-5959-4da0-aaa2-66b56fc7b4c3",
21             "submitter_id": "TCGA-EM-A3FQ"
22           }
23         ],
24         "_id": "efac6904-ac9f-4a44-bf9c-f7d9a822c127"
25       },
26       {
27         "_type": "file",
28         "file_name": "nationwidechildrens.org_clinical.TCGA-EM-A4FN.xml",
29         "file_id": "07add35d-66f0-4384-bb2c-9d86661f4073",
30         "_score": 4.644438,
31         "cases": [
32           {
33             "case_id": "f854ce67-c586-4424-a674-2dd67ad0ed7f",
34             "submitter_id": "TCGA-EM-A4FN"
35           }
36         ],
37         "_id": "07add35d-66f0-4384-bb2c-9d86661f4073"
38       },
39       {
40         "_type": "file",
41         "file_name": "nationwidechildrens.org_clinical.TCGA-EM-A3AN.xml",
42         "file_id": "889f222e-09d1-477c-a7b9-a514b65f322b",
43         "_score": 4.644438,
44         "cases": [
45           {
46             "case_id": "6491d025-c061-4180-bfd4-d7c6e6e55f66",
47             "submitter_id": "TCGA-EM-A3AN"
48           }
49         ],
50         "_id": "889f222e-09d1-477c-a7b9-a514b65f322b"
51       },
52       {
53         "_type": "file",
54         "file_name": "nationwidechildrens.org_clinical.TCGA-EM-A4FO.xml",
```

```
55        "file_id": "37db38d4-64ca-4cd4-9adc-d504b812997b",
56        "_score": 4.644438,
57        "cases": [
58          {
59            "case_id": "e0ab95ef-5c96-4d00-8950-04e26e3b4672",
60            "submitter_id": "TCGA-EM-A4FO"
61          }
62        ],
63        "_id": "37db38d4-64ca-4cd4-9adc-d504b812997b"
64      },
65      {
66        "_type": "file",
67        "file_name": "nationwidechildrens.org_clinical.TCGA-EM-A3FN.xml",
68        "file_id": "e2d13acf-3121-4c7b-b2ec-28db49eff699",
69        "_score": 4.644438,
70        "cases": [
71          {
72            "case_id": "59cd0969-a798-4d19-95ed-a311f39d2f38",
73            "submitter_id": "TCGA-EM-A3FN"
74          }
75        ],
76        "_id": "e2d13acf-3121-4c7b-b2ec-28db49eff699"
77      }
78    ],
79    "_shards": {
80      "successful": 5,
81      "failed": 0,
82      "total": 5
83    },
84    "took": 9,
85    "timed_out": false
86  },
87  "warnings": {}
88 }
```

Cases Endpoint

The GDC Cases Endpoint `https://gdc-api.nci.nih.gov/cases` enables search and retrieval of information related to a specific case.

Example

This example is a query for files contained in GDC. It returns case where submitter id is `TCGA-BH-A0EA`, using the pretty and filters parameters and the following filtering operators:

```
1 {"op":"and","content":[{"op":"in","content":{"field":"submitter_id","value":["TCGA-BH-A0EA"]}}]}
```

Command:

```
1 curl
     'https://gdc-api.nci.nih.gov/cases?filters=%7B%22op%22%3A%22and%22%2C%22content%22%3A%5B%7B%22op%22%3A%22in%
```

```
1 {
2 "data": {
3     "hits": [
4         {
5             "sample_ids": [
```

```
 6                    "7f791228-dd77-4ab0-8227-d784a4c7fea1",
 7                    "9a6c71a6-82cd-42b1-a93f-f569370848d6"
 8                ],
 9                "portion_ids": [
10                    "cb6086d1-3416-4310-b109-e8fa6e8b72d4",
11                    "8629bf5a-cdaf-4f6a-90bb-27dd4a7565c5",
12                    "ae4f5816-f97a-4605-9b05-9ab820467dee"
13                ],
14                "submitter_portion_ids": [
15                    "TCGA-BH-A0EA-01A-11",
16                    "TCGA-BH-A0EA-01A-21-A13C-20",
17                    "TCGA-BH-A0EA-10A-01"
18                ],
19                "created_datetime": null,
20                "submitter_aliquot_ids": [
21                    "TCGA-BH-A0EA-01A-11R-A114-13",
22                    "TCGA-BH-A0EA-01A-11D-A111-01",
23                    "TCGA-BH-A0EA-01A-11W-A12T-09",
24                    "TCGA-BH-A0EA-01A-11R-A114-13",
25                    "TCGA-BH-A0EA-01A-11R-A115-07",
26                    "TCGA-BH-A0EA-01A-11D-A111-01",
27                    "TCGA-BH-A0EA-01A-11D-A314-09",
28                    "TCGA-BH-A0EA-01A-11D-A112-05",
29                    "TCGA-BH-A0EA-01A-11D-A10Y-09",
30                    "TCGA-BH-A0EA-01A-11D-A10X-02",
31                    "TCGA-BH-A0EA-01A-11W-A12T-09",
32                    "TCGA-BH-A0EA-01A-11D-A10X-02",
33                    "TCGA-BH-A0EA-01A-11D-A10Y-09",
34                    "TCGA-BH-A0EA-01A-11D-A314-09",
35                    "TCGA-BH-A0EA-01A-11R-A115-07",
36                    "TCGA-BH-A0EA-01A-11D-A112-05",
37                    "TCGA-BH-A0EA-10A-01D-A110-09",
38                    "TCGA-BH-A0EA-10A-01D-A113-01",
39                    "TCGA-BH-A0EA-10A-01W-A12U-09",
40                    "TCGA-BH-A0EA-10A-01D-A10Z-02",
41                    "TCGA-BH-A0EA-10A-01D-A113-01",
42                    "TCGA-BH-A0EA-10A-01D-A110-09",
43                    "TCGA-BH-A0EA-10A-01W-A12U-09",
44                    "TCGA-BH-A0EA-10A-01D-A10Z-02"
45                ],
46                "updated_datetime": "2016-05-02T14:37:43.619198-05:00",
47                "submitter_analyte_ids": [
48                    "TCGA-BH-A0EA-01A-11R",
49                    "TCGA-BH-A0EA-01A-11D",
50                    "TCGA-BH-A0EA-01A-11W",
51                    "TCGA-BH-A0EA-10A-01W",
52                    "TCGA-BH-A0EA-10A-01D"
53                ],
54                "analyte_ids": [
55                    "30cb470f-66d4-4085-8c30-83a42e8453d4",
56                    "66ed0f86-5ca5-4dec-ba76-7ee4dcf31831",
57                    "f19f408a-815f-43d9-8032-e9482b796371",
58                    "69ddc092-88a0-4839-a2bb-9f1c9e760409",
59                    "fe678556-acf4-4bde-a95e-860bb0150a95"
60                ],
61                "submitter_id": "TCGA-BH-A0EA",
62                "case_id": "1f601832-eee3-48fb-acf5-80c4a454f26e",
63                "state": null,
```

```
64              "aliquot_ids": [
65                  "bcb7fc6d-60a0-48b7-aa81-14c0dda72d76",
66                  "97c64d6a-7dce-4d0f-9cb3-b3e4eb4719c5",
67                  "edad5bd3-efe0-4c5f-b05c-2c0c2951c45a",
68                  "bcb7fc6d-60a0-48b7-aa81-14c0dda72d76",
69                  "ca71ca96-cbb7-4eab-9487-251dda34e107",
70                  "97c64d6a-7dce-4d0f-9cb3-b3e4eb4719c5",
71                  "eef9dce1-6ba6-432b-bbe2-53c7dbe64fe7",
72                  "42d050e4-e8ee-4442-b9c0-0ee14706b138",
73                  "561b8777-801a-49ed-a306-e7dafeb044b6",
74                  "262715e1-835c-4f16-8ee7-6900e26f7cf5",
75                  "edad5bd3-efe0-4c5f-b05c-2c0c2951c45a",
76                  "262715e1-835c-4f16-8ee7-6900e26f7cf5",
77                  "561b8777-801a-49ed-a306-e7dafeb044b6",
78                  "eef9dce1-6ba6-432b-bbe2-53c7dbe64fe7",
79                  "ca71ca96-cbb7-4eab-9487-251dda34e107",
80                  "42d050e4-e8ee-4442-b9c0-0ee14706b138",
81                  "cfbd5476-e83a-401d-9f9a-639c73a0e35b",
82                  "2beb34c4-d493-4a73-b21e-de77d43251ff",
83                  "b1a3739d-d554-4202-b96f-f25a444e2042",
84                  "cde982b7-3b0a-49eb-8710-a599cb0e44c1",
85                  "2beb34c4-d493-4a73-b21e-de77d43251ff",
86                  "cfbd5476-e83a-401d-9f9a-639c73a0e35b",
87                  "b1a3739d-d554-4202-b96f-f25a444e2042",
88                  "cde982b7-3b0a-49eb-8710-a599cb0e44c1"
89              ],
90              "slide_ids": [
91                  "90154ea1-6b76-4445-870e-d531d6fa1239",
92                  "a0826f0d-986a-491b-8c6f-b34f8929f3ee"
93              ],
94              "submitter_sample_ids": [
95                  "TCGA-BH-A0EA-01A",
96                  "TCGA-BH-A0EA-10A"
97              ]
98          }
99      ],
100     "pagination": {
101         "count": 1,
102         "sort": "",
103         "from": 1,
104         "page": 1,
105         "total": 1,
106         "pages": 1,
107         "size": 10
108     }
109 },
110 "warnings": {}
111 }
```

Retrieval of case metadata using individual UUIDs:

The `cases` endpoint supports a simple query format that retrieves the metadata of a single case using its UUID:

```
1 curl 'https://gdc-api.nci.nih.gov/cases/1f601832-eee3-48fb-acf5-80c4a454f26e?pretty=true&expand=diagnoses'
```

```
1 {
2   "data": {
```

```
3      "diagnoses": [
4        {
5          "classification_of_tumor": "not reported",
6          "last_known_disease_status": "not reported",
7          "updated_datetime": "2016-05-16T10:59:16.740358-05:00",
8          "primary_diagnosis": "c50.9",
9          "submitter_id": "TCGA-BH-A0EA_diagnosis",
10         "tumor_stage": "stage iia",
11         "age_at_diagnosis": 26548.0,
12         "vital_status": "dead",
13         "morphology": "8500/3",
14         "days_to_death": 991.0,
15         "days_to_last_known_disease_status": null,
16         "days_to_last_follow_up": null,
17         "state": null,
18         "days_to_recurrence": null,
19         "diagnosis_id": "84654ad5-2a2c-5c3b-8340-ecac6a5550fe",
20         "tumor_grade": "not reported",
21         "tissue_or_organ_of_origin": "c50.9",
22         "days_to_birth": -26548.0,
23         "progression_or_recurrence": "not reported",
24         "prior_malignancy": "not reported",
25         "site_of_resection_or_biopsy": "c50.9",
26         "created_datetime": null
27       }
28     ],
29     "sample_ids": [
30       "7f791228-dd77-4ab0-8227-d784a4c7fea1",
31       "9a6c71a6-82cd-42b1-a93f-f569370848d6"
32     ],
33     "portion_ids": [
34       "cb6086d1-3416-4310-b109-e8fa6e8b72d4",
35       "8629bf5a-cdaf-4f6a-90bb-27dd4a7565c5",
36       "ae4f5816-f97a-4605-9b05-9ab820467dee"
37     ],
38     "submitter_portion_ids": [
39       "TCGA-BH-A0EA-01A-11",
40       "TCGA-BH-A0EA-01A-21-A13C-20",
41       "TCGA-BH-A0EA-10A-01"
42     ],
43     "created_datetime": null,
44     "submitter_aliquot_ids": [
45       "TCGA-BH-A0EA-01A-11R-A114-13",
46       "TCGA-BH-A0EA-01A-11D-A111-01",
47       "TCGA-BH-A0EA-01A-11W-A12T-09",
48       "TCGA-BH-A0EA-01A-11R-A114-13",
49       "TCGA-BH-A0EA-01A-11R-A115-07",
50       "TCGA-BH-A0EA-01A-11D-A111-01",
51       "TCGA-BH-A0EA-01A-11D-A314-09",
52       "TCGA-BH-A0EA-01A-11D-A112-05",
53       "TCGA-BH-A0EA-01A-11D-A10Y-09",
54       "TCGA-BH-A0EA-01A-11D-A10X-02",
55       "TCGA-BH-A0EA-01A-11W-A12T-09",
56       "TCGA-BH-A0EA-01A-11D-A10X-02",
57       "TCGA-BH-A0EA-01A-11D-A10Y-09",
58       "TCGA-BH-A0EA-01A-11D-A314-09",
59       "TCGA-BH-A0EA-01A-11R-A115-07",
60       "TCGA-BH-A0EA-01A-11D-A112-05",
```

```
61        "TCGA-BH-A0EA-10A-01D-A110-09",
62        "TCGA-BH-A0EA-10A-01D-A113-01",
63        "TCGA-BH-A0EA-10A-01W-A12U-09",
64        "TCGA-BH-A0EA-10A-01D-A10Z-02",
65        "TCGA-BH-A0EA-10A-01D-A113-01",
66        "TCGA-BH-A0EA-10A-01D-A110-09",
67        "TCGA-BH-A0EA-10A-01W-A12U-09",
68        "TCGA-BH-A0EA-10A-01D-A10Z-02"
69      ],
70      "updated_datetime": "2016-05-02T14:37:43.619198-05:00",
71      "submitter_analyte_ids": [
72        "TCGA-BH-A0EA-01A-11R",
73        "TCGA-BH-A0EA-01A-11D",
74        "TCGA-BH-A0EA-01A-11W",
75        "TCGA-BH-A0EA-10A-01W",
76        "TCGA-BH-A0EA-10A-01D"
77      ],
78      "analyte_ids": [
79        "30cb470f-66d4-4085-8c30-83a42e8453d4",
80        "66ed0f86-5ca5-4dec-ba76-7ee4dcf31831",
81        "f19f408a-815f-43d9-8032-e9482b796371",
82        "69ddc092-88a0-4839-a2bb-9f1c9e760409",
83        "fe678556-acf4-4bde-a95e-860bb0150a95"
84      ],
85      "submitter_id": "TCGA-BH-A0EA",
86      "case_id": "1f601832-eee3-48fb-acf5-80c4a454f26e",
87      "state": null,
88      "aliquot_ids": [
89        "bcb7fc6d-60a0-48b7-aa81-14c0dda72d76",
90        "97c64d6a-7dce-4d0f-9cb3-b3e4eb4719c5",
91        "edad5bd3-efe0-4c5f-b05c-2c0c2951c45a",
92        "bcb7fc6d-60a0-48b7-aa81-14c0dda72d76",
93        "ca71ca96-cbb7-4eab-9487-251dda34e107",
94        "97c64d6a-7dce-4d0f-9cb3-b3e4eb4719c5",
95        "eef9dce1-6ba6-432b-bbe2-53c7dbe64fe7",
96        "42d050e4-e8ee-4442-b9c0-0ee14706b138",
97        "561b8777-801a-49ed-a306-e7dafeb044b6",
98        "262715e1-835c-4f16-8ee7-6900e26f7cf5",
99        "edad5bd3-efe0-4c5f-b05c-2c0c2951c45a",
100       "262715e1-835c-4f16-8ee7-6900e26f7cf5",
101       "561b8777-801a-49ed-a306-e7dafeb044b6",
102       "eef9dce1-6ba6-432b-bbe2-53c7dbe64fe7",
103       "ca71ca96-cbb7-4eab-9487-251dda34e107",
104       "42d050e4-e8ee-4442-b9c0-0ee14706b138",
105       "cfbd5476-e83a-401d-9f9a-639c73a0e35b",
106       "2beb34c4-d493-4a73-b21e-de77d43251ff",
107       "b1a3739d-d554-4202-b96f-f25a444e2042",
108       "cde982b7-3b0a-49eb-8710-a599cb0e44c1",
109       "2beb34c4-d493-4a73-b21e-de77d43251ff",
110       "cfbd5476-e83a-401d-9f9a-639c73a0e35b",
111       "b1a3739d-d554-4202-b96f-f25a444e2042",
112       "cde982b7-3b0a-49eb-8710-a599cb0e44c1"
113     ],
114     "slide_ids": [
115       "90154ea1-6b76-4445-870e-d531d6fa1239",
116       "a0826f0d-986a-491b-8c6f-b34f8929f3ee"
117     ],
118     "submitter_sample_ids": [
```

```
119        "TCGA-BH-A0EA-01A",
120        "TCGA-BH-A0EA-10A"
121      ]
122    },
123    "warnings": {}
124 }
```

Annotations Endpoint

The GDC Annotation Endpoint `https://gdc-api.nci.nih.gov/annotations` enables search and retrieval of annotations stored in the GDC.

Example

This example is a query for any annotations **directly** associated with the following GDC entities:

- the case with UUID e0d36cc0-652c-4224-bb10-09d15c7bd8f1
- the sample with UUID 25ebc29a-7598-4ae4-ba7f-618d448882cc
- the aliquot with UUID fe660d7c-2746-4b50-ab93-b2ed99960553

The query uses the filters parameter to specify entity UUIDs. Code samples below include the bare and percent-encoded filter JSON.

```
1  {
2      "op":"in",
3      "content":{
4          "field":"entity_id",
5          "value":[
6              "e0d36cc0-652c-4224-bb10-09d15c7bd8f1",
7              "25ebc29a-7598-4ae4-ba7f-618d448882cc",
8              "fe660d7c-2746-4b50-ab93-b2ed99960553"
9          ]
10      }
11 }
```

```
1  %7B%22op%22%3A%22in%22%2C%22content%22%3A%7B%22field%22%3A%22entity_id%22%2C%22value%22%3A%5B%22e0d36cc0-652c-42
```

```
1  curl
       'https://gdc-api.nci.nih.gov/annotations?filters=%7B%22op%22%3A%22in%22%2C%22content%22%3A%7B%22field%22%3A%
```

```
1  {
2    "data": {
3      "hits": [
4        {
5          "status": "Approved",
6          "category": "Item flagged DNU",
7          "entity_id": "fe660d7c-2746-4b50-ab93-b2ed99960553",
8          "classification": "CenterNotification",
9          "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
10         "created_datetime": "2015-09-28T13:39:13-04:00",
11         "annotation_id": "5ddadefe-8b57-5ce2-b8b2-918d63d99a59",
12         "notes": "The aliquot failed Broad pipeline QC and not all files are suitable for use. Consult
                the SDRF file to determine which files are usable.",
13         "entity_type": "aliquot",
14         "submitter_id": "29087",
15         "case_id": "41b59716-116f-4942-8b63-409870a87e26",
```

```
16          "entity_submitter_id": "TCGA-DK-A3IM-10A-01D-A20B-01",
17          "case_submitter_id": "TCGA-DK-A3IM"
18        },
19        {
20          "status": "Approved",
21          "category": "Item is noncanonical",
22          "entity_id": "25ebc29a-7598-4ae4-ba7f-618d448882cc",
23          "classification": "Notification",
24          "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
25          "created_datetime": "2012-07-12T15:00:15-04:00",
26          "annotation_id": "d6500f94-618f-5334-a810-ade76b887ec9",
27          "notes": "No Matching Normal",
28          "entity_type": "sample",
29          "submitter_id": "8009",
30          "case_id": "bd114e05-5a97-41e2-a0d5-5d39a1e9d461",
31          "entity_submitter_id": "TCGA-08-0514-01A",
32          "case_submitter_id": "TCGA-08-0514"
33        },
34        {
35          "status": "Approved",
36          "category": "Prior malignancy",
37          "entity_id": "e0d36cc0-652c-4224-bb10-09d15c7bd8f1",
38          "classification": "Notification",
39          "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
40          "created_datetime": "2013-03-12T10:05:14-04:00",
41          "annotation_id": "33336cdf-2cf0-5af2-bb52-fecd3427f180",
42          "notes": "Patient had a prior lymphoma. Unknown radiation or systemic chemotherapy.",
43          "entity_type": "case",
44          "submitter_id": "15630",
45          "case_id": "e0d36cc0-652c-4224-bb10-09d15c7bd8f1",
46          "entity_submitter_id": "TCGA-FS-A1ZF",
47          "case_submitter_id": "TCGA-FS-A1ZF"
48        }
49      ],
50      "pagination": {
51        "count": 3,
52        "sort": "",
53        "from": 1,
54        "page": 1,
55        "total": 3,
56        "pages": 1,
57        "size": 10
58      }
59    },
60    "warnings": {}
61 }
```

Example

This example is a query for any annotations that are associated with the following *cases*, **directly or via a child entity**:

- the case with UUID 513c5f34-dc6e-4caa-81cc-907fd6a825b1
- the case with UUID 942c0088-c9a0-428c-a879-e16f8c5bfdb8

The query uses the filters parameter to specify entity UUIDs. Code samples below include the bare and percent-encoded filter JSON.

```
1 {
2     "op":"in",
3     "content":{
4         "field":"annotation.case_id",
5         "value":[
6             "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
7             "942c0088-c9a0-428c-a879-e16f8c5bfdb8"
8         ]
9     }
10 }
```

```
1 %7B%22op%22%3A%22in%22%2C%22content%22%3A%7B%22field%22%3A%22annotation.case_id%22%2C%22value%22%3A%5B%22513c5f3
```

```
1 curl
     'https://gdc-api.nci.nih.gov/annotations?filters=%7B%22op%22%3A%22in%22%2C%22content%22%3A%7B%22field%22%3A%
```

```
1 {
2   "data": {
3     "hits": [
4       {
5         "status": "Approved",
6         "category": "BCR Notification",
7         "entity_id": "39c2e6c5-379e-4ffa-b02c-e1db298b34f7",
8         "classification": "Notification",
9         "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
10        "created_datetime": "2014-09-05T09:02:25-04:00",
11        "annotation_id": "8bed6748-11dd-5767-9a9c-577712f9b616",
12        "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
              The sequencing and characterization centers have reported that molecular signature of the
              tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
              of the tumor is very low. Therefore, the data from this patient should be used with caution.",
13        "entity_type": "sample",
14        "submitter_id": "22104",
15        "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
16        "entity_submitter_id": "TCGA-56-8623-10A",
17        "case_submitter_id": "TCGA-56-8623"
18      },
19      {
20        "status": "Approved",
21        "category": "BCR Notification",
22        "entity_id": "e2bca154-9ff3-42a6-a5bc-0daa14080c92",
23        "classification": "Notification",
24        "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
25        "created_datetime": "2014-09-05T09:02:32-04:00",
26        "annotation_id": "a6997495-8544-58b1-81aa-7e628e46af7e",
27        "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
              The sequencing and characterization centers have reported that molecular signature of the
              tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
              of the tumor is very low. Therefore, the data from this patient should be used with caution.",
28        "entity_type": "aliquot",
29        "submitter_id": "22118",
30        "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
31        "entity_submitter_id": "TCGA-56-8623-01A-11D-2395-08",
32        "case_submitter_id": "TCGA-56-8623"
33      },
34      {
35        "status": "Approved",
36        "category": "History of unacceptable prior treatment related to a prior/other malignancy",
```

```
37        "entity_id": "942c0088-c9a0-428c-a879-e16f8c5bfdb8",
38        "classification": "Notification",
39        "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
40        "created_datetime": "2012-11-10T05:30:37-05:00",
41        "annotation_id": "6d57211a-402b-52c9-b857-665164c63339",
42        "notes": "Pt with synchronous B-cell lymphoma treated with cyclophosphamide, Adriamycin,
              vincristine, prednisone, and Rituxin prior to procurment of TCGA tumor.",
43        "entity_type": "case",
44        "submitter_id": "12063",
45        "case_id": "942c0088-c9a0-428c-a879-e16f8c5bfdb8",
46        "entity_submitter_id": "TCGA-CJ-4642",
47        "case_submitter_id": "TCGA-CJ-4642"
48      },
49      {
50        "status": "Approved",
51        "category": "BCR Notification",
52        "entity_id": "265c0257-54b6-4b79-8c9a-d56ca8bbdb48",
53        "classification": "Notification",
54        "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
55        "created_datetime": "2014-09-05T09:02:41-04:00",
56        "annotation_id": "0fe698fa-169a-51e2-b41b-6452d0b5cefc",
57        "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
              The sequencing and characterization centers have reported that molecular signature of the
              tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
              of the tumor is very low. Therefore, the data from this patient should be used with caution.",
58        "entity_type": "aliquot",
59        "submitter_id": "22122",
60        "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
61        "entity_submitter_id": "TCGA-56-8623-01A-11R-A28V-07",
62        "case_submitter_id": "TCGA-56-8623"
63      },
64      {
65        "status": "Approved",
66        "category": "BCR Notification",
67        "entity_id": "a4d9f761-ece9-4a6b-8818-ecf4a8f7d380",
68        "classification": "Notification",
69        "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
70        "created_datetime": "2014-09-05T09:02:28-04:00",
71        "annotation_id": "6dd575b6-78a7-55ca-9dbe-654079eeca0f",
72        "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
              The sequencing and characterization centers have reported that molecular signature of the
              tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
              of the tumor is very low. Therefore, the data from this patient should be used with caution.",
73        "entity_type": "analyte",
74        "submitter_id": "22110",
75        "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
76        "entity_submitter_id": "TCGA-56-8623-10A-01D",
77        "case_submitter_id": "TCGA-56-8623"
78      },
79      {
80        "status": "Approved",
81        "category": "BCR Notification",
82        "entity_id": "a5370483-6bdd-4e2a-9b4b-ec1eb381052b",
83        "classification": "Notification",
84        "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
85        "created_datetime": "2014-09-05T09:02:25-04:00",
86        "annotation_id": "3d268354-9193-50a9-a8b5-79f00d36d3c6",
```

```
87          "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
               The sequencing and characterization centers have reported that molecular signature of the
               tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
               of the tumor is very low. Therefore, the data from this patient should be used with caution.",
88          "entity_type": "sample",
89          "submitter_id": "22103",
90          "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
91          "entity_submitter_id": "TCGA-56-8623-01A",
92          "case_submitter_id": "TCGA-56-8623"
93        },
94        {
95          "status": "Approved",
96          "category": "BCR Notification",
97          "entity_id": "2aa6f51a-e0fa-4bdb-830c-80b79c59cfc9",
98          "classification": "Notification",
99          "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
100         "created_datetime": "2014-09-05T09:02:32-04:00",
101         "annotation_id": "d4f40e94-f701-58d1-aa0c-b3d687844c7b",
102         "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
               The sequencing and characterization centers have reported that molecular signature of the
               tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
               of the tumor is very low. Therefore, the data from this patient should be used with caution.",
103         "entity_type": "aliquot",
104         "submitter_id": "22117",
105         "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
106         "entity_submitter_id": "TCGA-56-8623-01A-11D-2391-01",
107         "case_submitter_id": "TCGA-56-8623"
108       },
109       {
110         "status": "Approved",
111         "category": "BCR Notification",
112         "entity_id": "d9fffb10-ae12-43ad-983b-1e2336f915a3",
113         "classification": "Notification",
114         "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
115         "created_datetime": "2014-09-05T09:02:43-04:00",
116         "annotation_id": "ab91b0d6-28fd-5600-8933-3457c0939687",
117         "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
               The sequencing and characterization centers have reported that molecular signature of the
               tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
               of the tumor is very low. Therefore, the data from this patient should be used with caution.",
118         "entity_type": "aliquot",
119         "submitter_id": "22126",
120         "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
121         "entity_submitter_id": "TCGA-56-8623-10A-01D-2395-08",
122         "case_submitter_id": "TCGA-56-8623"
123       },
124       {
125         "status": "Approved",
126         "category": "BCR Notification",
127         "entity_id": "7402c87f-c9b7-451f-9b8d-a8979bf0d98b",
128         "classification": "Notification",
129         "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
130         "created_datetime": "2014-09-05T09:02:30-04:00",
131         "annotation_id": "b16b9f8f-adfd-5ab8-b10e-16b0ed277dbc",
132         "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
               The sequencing and characterization centers have reported that molecular signature of the
               tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
               of the tumor is very low. Therefore, the data from this patient should be used with caution.",
```

```
133          "entity_type": "analyte",
134          "submitter_id": "22114",
135          "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
136          "entity_submitter_id": "TCGA-56-8623-11A-01R",
137          "case_submitter_id": "TCGA-56-8623"
138        },
139        {
140          "status": "Approved",
141          "category": "Synchronous malignancy",
142          "entity_id": "942c0088-c9a0-428c-a879-e16f8c5bfdb8",
143          "classification": "Notification",
144          "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
145          "created_datetime": "2012-11-10T05:30:35-05:00",
146          "annotation_id": "b02b2a31-7c5a-5ff0-a2ea-6ab54b721a86",
147          "notes": "Pt with synchronous B-cell lymphoma treated with cyclophosphamide, Adriamycin,
                 vincristine, prednisone, and Rituxin prior to procurment of TCGA tumor.",
148          "entity_type": "case",
149          "submitter_id": "12062",
150          "case_id": "942c0088-c9a0-428c-a879-e16f8c5bfdb8",
151          "entity_submitter_id": "TCGA-CJ-4642",
152          "case_submitter_id": "TCGA-CJ-4642"
153        },
154        {
155          "status": "Approved",
156          "category": "BCR Notification",
157          "entity_id": "5a676d7d-0b8d-48fb-b898-0eb023a871e1",
158          "classification": "Notification",
159          "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
160          "created_datetime": "2014-09-05T09:02:27-04:00",
161          "annotation_id": "96d82ca1-5e27-5048-82d5-ee5886437580",
162          "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
                 The sequencing and characterization centers have reported that molecular signature of the
                 tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
                 of the tumor is very low. Therefore, the data from this patient should be used with caution.",
163          "entity_type": "analyte",
164          "submitter_id": "22107",
165          "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
166          "entity_submitter_id": "TCGA-56-8623-01A-11H",
167          "case_submitter_id": "TCGA-56-8623"
168        },
169        {
170          "status": "Approved",
171          "category": "BCR Notification",
172          "entity_id": "2327bab1-e352-4c43-ab40-6162e8632c26",
173          "classification": "Notification",
174          "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
175          "created_datetime": "2014-09-05T09:02:26-04:00",
176          "annotation_id": "95873790-c215-5185-babe-0228e32d5689",
177          "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
                 The sequencing and characterization centers have reported that molecular signature of the
                 tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
                 of the tumor is very low. Therefore, the data from this patient should be used with caution.",
178          "entity_type": "analyte",
179          "submitter_id": "22106",
180          "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
181          "entity_submitter_id": "TCGA-56-8623-01A-11D",
182          "case_submitter_id": "TCGA-56-8623"
183        },
```

```json
184          {
185            "status": "Approved",
186            "category": "BCR Notification",
187            "entity_id": "258a054e-b0ec-4215-81c9-8e1c715abc9b",
188            "classification": "Notification",
189            "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
190            "created_datetime": "2014-09-05T09:02:30-04:00",
191            "annotation_id": "f0a7ee98-09aa-5c46-b88a-66a09176c408",
192            "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
                  The sequencing and characterization centers have reported that molecular signature of the
                  tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
                  of the tumor is very low. Therefore, the data from this patient should be used with caution.",
193            "entity_type": "analyte",
194            "submitter_id": "22113",
195            "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
196            "entity_submitter_id": "TCGA-56-8623-11A-01H",
197            "case_submitter_id": "TCGA-56-8623"
198          },
199          {
200            "status": "Approved",
201            "category": "BCR Notification",
202            "entity_id": "7c9e5386-1540-40f2-96ed-c06d752598df",
203            "classification": "Notification",
204            "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
205            "created_datetime": "2014-09-05T09:02:45-04:00",
206            "annotation_id": "f34eb057-6f81-5d37-858c-9ed5cc8a2052",
207            "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
                  The sequencing and characterization centers have reported that molecular signature of the
                  tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
                  of the tumor is very low. Therefore, the data from this patient should be used with caution.",
208            "entity_type": "aliquot",
209            "submitter_id": "22129",
210            "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
211            "entity_submitter_id": "TCGA-56-8623-11A-01D-2391-01",
212            "case_submitter_id": "TCGA-56-8623"
213          },
214          {
215            "status": "Approved",
216            "category": "BCR Notification",
217            "entity_id": "77eccef5-254d-4ddc-a70f-576436278e0a",
218            "classification": "Notification",
219            "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
220            "created_datetime": "2014-09-05T09:02:46-04:00",
221            "annotation_id": "9bf047b4-bcfc-5bfc-bc42-584a1efd09b4",
222            "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
                  The sequencing and characterization centers have reported that molecular signature of the
                  tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
                  of the tumor is very low. Therefore, the data from this patient should be used with caution.",
223            "entity_type": "aliquot",
224            "submitter_id": "22131",
225            "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
226            "entity_submitter_id": "TCGA-56-8623-11A-01H-2402-13",
227            "case_submitter_id": "TCGA-56-8623"
228          },
229          {
230            "status": "Approved",
231            "category": "BCR Notification",
232            "entity_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
```

```
233        "classification": "Notification",
234        "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
235        "created_datetime": "2014-09-05T08:54:02-04:00",
236        "annotation_id": "53ccabc5-0c89-5e55-a267-130dd021d9ff",
237        "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
              The sequencing and characterization centers have reported that molecular signature of the
              tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
              of the tumor is very low. Therefore, the data from this patient should be used with caution.",
238        "entity_type": "case",
239        "submitter_id": "22102",
240        "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
241        "entity_submitter_id": "TCGA-56-8623",
242        "case_submitter_id": "TCGA-56-8623"
243      },
244      {
245        "status": "Approved",
246        "category": "BCR Notification",
247        "entity_id": "9fe04d07-57ba-434f-8254-184fe00ab23e",
248        "classification": "Notification",
249        "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
250        "created_datetime": "2014-09-05T09:02:46-04:00",
251        "annotation_id": "100caf14-4fd6-51ef-8f3c-8e7ca62d0233",
252        "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
              The sequencing and characterization centers have reported that molecular signature of the
              tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
              of the tumor is very low. Therefore, the data from this patient should be used with caution.",
253        "entity_type": "aliquot",
254        "submitter_id": "22132",
255        "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
256        "entity_submitter_id": "TCGA-56-8623-11A-01R-A28V-07",
257        "case_submitter_id": "TCGA-56-8623"
258      },
259      {
260        "status": "Approved",
261        "category": "BCR Notification",
262        "entity_id": "26368154-1ed7-4471-87b9-d50d5fe8046b",
263        "classification": "Notification",
264        "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
265        "created_datetime": "2014-09-05T09:02:26-04:00",
266        "annotation_id": "75e180f1-4216-591f-8bd4-426e923d0778",
267        "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
              The sequencing and characterization centers have reported that molecular signature of the
              tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
              of the tumor is very low. Therefore, the data from this patient should be used with caution.",
268        "entity_type": "sample",
269        "submitter_id": "22105",
270        "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
271        "entity_submitter_id": "TCGA-56-8623-11A",
272        "case_submitter_id": "TCGA-56-8623"
273      },
274      {
275        "status": "Approved",
276        "category": "Prior malignancy",
277        "entity_id": "942c0088-c9a0-428c-a879-e16f8c5bfdb8",
278        "classification": "Notification",
279        "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
280        "created_datetime": "2011-02-05T13:15:23-05:00",
281        "annotation_id": "59f7f353-2119-58f2-a340-34bc39c3d7ef",
```

```
282          "notes": "[intgen.org]: Prior Malignancy",
283          "entity_type": "case",
284          "submitter_id": "1272",
285          "case_id": "942c0088-c9a0-428c-a879-e16f8c5bfdb8",
286          "entity_submitter_id": "TCGA-CJ-4642",
287          "case_submitter_id": "TCGA-CJ-4642"
288        },
289        {
290          "status": "Approved",
291          "category": "BCR Notification",
292          "entity_id": "51ded92f-852f-4445-96b2-dd9db1544a9d",
293          "classification": "Notification",
294          "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
295          "created_datetime": "2014-09-05T09:02:27-04:00",
296          "annotation_id": "8ac6de9f-6b86-5421-a4be-e1fe8a4a8a38",
297          "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
                  The sequencing and characterization centers have reported that molecular signature of the
                  tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
                  of the tumor is very low. Therefore, the data from this patient should be used with caution.",
298          "entity_type": "analyte",
299          "submitter_id": "22108",
300          "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
301          "entity_submitter_id": "TCGA-56-8623-01A-11R",
302          "case_submitter_id": "TCGA-56-8623"
303        },
304        {
305          "status": "Approved",
306          "category": "Molecular analysis outside specification",
307          "entity_id": "942c0088-c9a0-428c-a879-e16f8c5bfdb8",
308          "classification": "Notification",
309          "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
310          "created_datetime": "2010-10-29T00:00:00-04:00",
311          "annotation_id": "e0fcd5e3-02f0-52b4-b949-95ed4de324f6",
312          "notes": "Molecular results off spec",
313          "entity_type": "case",
314          "submitter_id": "831",
315          "case_id": "942c0088-c9a0-428c-a879-e16f8c5bfdb8",
316          "entity_submitter_id": "TCGA-CJ-4642",
317          "case_submitter_id": "TCGA-CJ-4642"
318        },
319        {
320          "status": "Approved",
321          "category": "BCR Notification",
322          "entity_id": "3cffe6b0-3aef-413a-a9a9-4a2fcefff6ad",
323          "classification": "Notification",
324          "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
325          "created_datetime": "2014-09-05T09:02:45-04:00",
326          "annotation_id": "dfad3f3e-9d62-51e0-a433-77f89b1c24ca",
327          "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
                  The sequencing and characterization centers have reported that molecular signature of the
                  tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
                  of the tumor is very low. Therefore, the data from this patient should be used with caution.",
328          "entity_type": "aliquot",
329          "submitter_id": "22130",
330          "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
331          "entity_submitter_id": "TCGA-56-8623-11A-01D-2395-08",
332          "case_submitter_id": "TCGA-56-8623"
333        },
```

```
334         {
335           "status": "Approved",
336           "category": "BCR Notification",
337           "entity_id": "6f688bcb-0534-4a50-a39d-a3a149165694",
338           "classification": "Notification",
339           "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
340           "created_datetime": "2014-09-05T09:02:33-04:00",
341           "annotation_id": "03716d9e-c294-570b-89ab-53af108850f3",
342           "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
                  The sequencing and characterization centers have reported that molecular signature of the
                  tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
                  of the tumor is very low. Therefore, the data from this patient should be used with caution.",
343           "entity_type": "aliquot",
344           "submitter_id": "22120",
345           "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
346           "entity_submitter_id": "TCGA-56-8623-01A-11H-2402-13",
347           "case_submitter_id": "TCGA-56-8623"
348         },
349         {
350           "status": "Approved",
351           "category": "BCR Notification",
352           "entity_id": "f1090757-a85d-4702-a16f-0ee35284c45d",
353           "classification": "Notification",
354           "updated_datetime": "2016-05-01T15:00:21.638875-05:00",
355           "created_datetime": "2014-09-05T09:02:29-04:00",
356           "annotation_id": "84236594-5c53-504d-bd27-1e8415cb768a",
357           "notes": "Possible tumor/normal sample swap, cross-contamination, and/or sample purity issues.
                  The sequencing and characterization centers have reported that molecular signature of the
                  tumor DNA and RNA cluster with the normal controls. Their analysis suggests that the purity
                  of the tumor is very low. Therefore, the data from this patient should be used with caution.",
358           "entity_type": "analyte",
359           "submitter_id": "22112",
360           "case_id": "513c5f34-dc6e-4caa-81cc-907fd6a825b1",
361           "entity_submitter_id": "TCGA-56-8623-11A-01D",
362           "case_submitter_id": "TCGA-56-8623"
363         }
364       ],
365       "pagination": {
366         "count": 24,
367         "sort": "",
368         "from": 1,
369         "page": 1,
370         "total": 24,
371         "pages": 1,
372         "size": 30
373       }
374     },
375     "warnings": {}
376 }
```

__mapping Endpoint

Each search and retrieval endpoint is equipped with a `_mapping` endpoint that provides information about available fields. For example, `files/_mapping` endpoint provides information about fields and field groups available at the `files` endpoint: https://gdc-api.nci.nih.gov/files/_mapping.

The high-level structure of a response to a `_mapping` query is as follows:

```
1 "_mapping": {}
2 , "defaults": []
3 , "expand": []
4 , "fields": []
5 , "multi": []
6 , "nested": []
```

Each part of the response is described below:

Part	Description
_mapping	All available fields and their descriptions. The endpoint-agnostic field names provided here are compatible with the `filters` parameter but are not always compatible with the `fields` parameter
defaults	The default set of fields included in the API response when the `fields` parameter is not used in the request
expand	Field group names for use with the `expand` parameter
fields	All available fields in an endpoint-specific format that is compatible with both the `filters` and `fields` parameters
multi	GDC internal use
nested	Nested fields

Example

```
1 curl 'https://gdc-api.nci.nih.gov/projects/_mapping'
```

```
1 {
2     ...
3
4       "_mapping": {
5         "projects.disease_type": {
6           "doc_type": "projects",
7           "field": "disease_type",
8           "type": "id"
9         },
10        "projects.name": {
11          "doc_type": "projects",
12          "field": "name",
13          "type": "id"
14        }
15      }
16
17    ...
18
19 }
```

Similar information can be obtained using the `fields` parameter; `fields` queries provide additional information in the response, such as the name of the Elastic Search document (`doc_type`), the field name and the type of value. A list of supported types (such as `string`, `long`, `float`, ...) can be obtained from Elastic Search Documentation.

Request Parameters

The GDC API supports the following search & retrieval request parameters:

Parameter	Default	Description
filters	null	Specifies search parameters

Parameter	Default	Description
format	JSON	Specifies the API response format: JSON, XML, or TSV
pretty	false	Returns response with indentations and line breaks in a human-readable format
fields	null	Specifies which fields to include in the response
size	10	Specifies the number of results to return
from	1	Specifies the first record to return from a set of search results
sort	null	Specifies sorting for the search results
facets	null	Provides all existing values for a given field and the number of records having this value.

Filters: Specifying the Query

The `filters` parameter enables passing of complex search queries to the GDC API. The parameter carries a query in the form of a JSON object.

Query Format

A `filters` query consists of an operator (or a nested set of operators) with a set of `field` and `value` operands.

The following `filters` query operators are supported by the GDC API:

Operator	Description	Number of Operands	Logic example
=	equals (string or number)	one	gender = "female"
!=	does not equal (string or number)	one	project_id != "TARGET-AML"
<	less than (number)	one	age at diagnosis < 90y
<=	less than or equal (number)	one	age at diagnosis <= 17
	greater than (number)	one	age at diagnosis > 50
=	greater than or equal (number)	one	age at diagnosis >= 18
is	is (missing)	one	gender is missing
not	not (missing)	one	race not missing
in	matches a string or number in (a list)	multiple	primary_site in [Brain, Lung]
exclude	does not match any strings or values in (a list)	multiple	experimental_strategy exclude [WXS, WGS, "Genotyping array"]
and	(operation1) and (operation2)	multiple	{primary_site in [Brain, Lung]} and {gender = "female"}
or	(operation1) or (operation2)	multiple	{project_id != "TARGET-AML"} or {age at diagnosis < 90y}

The `field` operand specifies a field that corresponds to a property defined in the GDC Data Dictionary. A list of supported fields is provided in Appendix A; the list can also be accessed programmatically at the _mapping endpoint.

The `value` operand specifies the search terms. Users can get a list of available values for a specific property by making a call to the appropriate API endpoint using the `facets` parameter, e.g. `https://gdc-api.nci.nih.gov/v0/cases?facets=demographic.gender&siz` See Facets for details.

A simple query with a single operator looks like this:

```
1 {
2     "op":"=",
3     "content":{
4         "field":"cases.demographic.gender",
5         "value":[
6             "male"
7         ]
8     }
9 }
```

A more complex query with multiple operators looks like this:

```
1  {
2      "op":"and",
3      "content":[
4          {
5              "op":"in",
6              "content":{
7                  "field":"cases.submitter_id",
8                  "value":[
9                      "TCGA-CK-4948",
10                     "TCGA-D1-A17N",
11                     "TCGA-4V-A9QX",
12                     "TCGA-4V-A9QM"
13                 ]
14             }
15         },
16         {
17             "op":"=",
18             "content":{
19                 "field":"files.data_type",
20                 "value":"Gene Expression Quantification"
21             }
22         }
23     ]
24 }
```

Example: HTTP GET Request

This example requests `male` cases using HTTP GET.

The JSON object to be passed to the GDC API looks like:

```
1 {"op": "=",
2     "content": {
3         "field": "cases.demographic.gender",
4         "value": ["male"]
5     }
6 }
```

URL-encoding the above JSON object using Percent-(URL)-encoding tool results in the following string:

```
1 %7b%22op%22%3a+%22%3d%22%2c%0d%0a++++++%22content%22%3a+%7b%0d%0a++++++++++%22field%22%3a+%22cases.clinical.gend
```

The above string can now be passed to the GDC API using the `filters` parameter:

```
1  curl
2      'https://gdc-api.nci.nih.gov/cases?filters=%7b%22op%22%3a+%22%3d%22%2c%0d%0a++++++%22content%22%3a+%7b%0d%0
```

```
1 import requests
2 import json
3 cases_endpt = 'https://gdc-api.nci.nih.gov/cases'
4 filt = {"op":"=",
5         "content":{
6             "field": "cases.demographic.gender",
7             "value": ["male"]
8         }
9 }
10 params = {'filters':json.dumps(filt), 'sort':'demographic.gender:asc'}
11 ## requests URL-encodes automatically
12 response = requests.get(cases_endpt, params = params)
13 print json.dumps(response.json(), indent=2)
```

```
1 {
2   "data": {
3     "hits": [
4       {
5         "sample_ids": [
6           "1d014bf1-95ae-42e3-ae39-97ff4841d8ca",
7           "6b685bfc-651b-48d1-8e68-32c8096ea205"
8         ],
9         "portion_ids": [
10          "c061217a-266a-496d-8a96-3489191afa87",
11          "0d3a6a58-0e00-4889-bc73-5ddb5a387738",
12          "e858ee92-0438-48e9-a70d-80ef2c0ad539"
13        ],
14        "submitter_portion_ids": [
15          "TCGA-66-2770-01A-21-2193-20",
16          "TCGA-66-2770-01A-01",
17          "TCGA-66-2770-11A-01"
18        ],
19        "created_datetime": null,
20        "submitter_aliquot_ids": [
21          "TCGA-66-2770-01A-01D-1522-08",
22          "TCGA-66-2770-01A-01D-0848-05",
23          "TCGA-66-2770-01A-01W-0879-09",
24          "TCGA-66-2770-11A-01W-0878-08",
25          "TCGA-66-2770-01A-01R-0849-01",
26          "TCGA-66-2770-01A-01W-0877-08",
27          "TCGA-66-2770-01A-01D-0846-06",
28          "TCGA-66-2770-11A-01W-0880-09",
29          "TCGA-66-2770-01A-01D-0964-09",
30          "TCGA-66-2770-11A-01D-0846-06",
31          "TCGA-66-2770-01A-01D-0845-04",
32          "TCGA-66-2770-01A-01W-0881-10",
33          "TCGA-66-2770-11A-01D-0963-08",
34          "TCGA-66-2770-11A-01D-0844-01",
35          "TCGA-66-2770-01A-01R-0851-07",
36          "TCGA-66-2770-11A-01W-0882-10",
37          "TCGA-66-2770-11A-01D-1522-08",
38          "TCGA-66-2770-01A-01T-1557-13",
39          "TCGA-66-2770-01A-01D-0847-02",
40          "TCGA-66-2770-01A-01D-0844-01",
41          "TCGA-66-2770-11A-01D-0847-02",
42          "TCGA-66-2770-11A-01D-0964-09",
43          "TCGA-66-2770-01A-01D-0963-08",
44          "TCGA-66-2770-01A-01R-0850-03",
```

```
45          "TCGA-66-2770-11A-01D-0845-04",
46          "TCGA-66-2770-01A-01T-0852-07"
47        ],
48        "updated_datetime": "2016-05-02T15:57:03.730994-05:00",
49        "submitter_analyte_ids": [
50          "TCGA-66-2770-01A-01D",
51          "TCGA-66-2770-11A-01W",
52          "TCGA-66-2770-01A-01T",
53          "TCGA-66-2770-01A-01W",
54          "TCGA-66-2770-01A-01R",
55          "TCGA-66-2770-11A-01D"
56        ],
57        "analyte_ids": [
58          "385807d3-78de-4558-8d93-702d93fc835a",
59          "247acc7a-b4f5-47e9-86da-5ea9b04ad444",
60          "151b8cb9-6b0a-4db9-9b0e-62aa501b35d9",
61          "e549aebd-4dda-4ea8-8ccf-56c03bc8b2be",
62          "631ad4eb-845a-4e70-96ad-4b40157218a8",
63          "9a75640e-09d4-42b7-8cb4-75d62b39e98a"
64        ],
65        "submitter_id": "TCGA-66-2770",
66        "case_id": "f1b357e4-d67a-42c9-b0b7-12f69fa3da58",
67        "state": null,
68        "aliquot_ids": [
69          "a2d10f8e-6b27-4df0-bd25-ac24992d0bb4",
70          "8c1c733a-abed-468f-b4d0-d1ac34ba6d8b",
71          "cad8d384-3b7a-4f70-89c2-5584ae75c5eb",
72          "42e774cf-3c4a-4efd-9665-378cb6b4afac",
73          "3755168b-f5da-422d-847a-566cb112a8d7",
74          "cae4d249-ba67-4316-8761-7e71e3813182",
75          "aa6e700c-ce01-4cc9-87de-8bf615a8aa1a",
76          "ad5c4069-e616-4ab4-9b03-b196f9189b20",
77          "07c26ea4-0584-4cb0-8e5a-d057b8fe6c14",
78          "f95c2cb5-d20a-4f1f-8f2a-95a2d37fbdc4",
79          "817bf327-e583-4704-b294-c3645dcc4adf",
80          "2246cb75-38bd-491f-b6ee-99f4781f2564",
81          "a81b9090-626d-492d-9baf-7fa3ef70111c",
82          "5cd6f026-894e-45f6-bc59-d6f056e63846",
83          "e417903d-ab76-44f0-aae9-3a91fa9a8d3c",
84          "1d809a56-31ca-49d8-a57b-e773236b24de",
85          "df60a743-ef4b-43ea-bc5a-4d75e8befb8a",
86          "871350e2-958f-401c-ae86-6bc880a01942",
87          "3dc4207d-5671-4c3d-b75a-d39ef69b564c",
88          "69b77cc0-d00a-4ea3-9b39-3e3019d9e292",
89          "3d035ee8-9523-4771-8738-c8a5a2f91403",
90          "775e46bd-e56f-40fa-9891-aaedc1d49395",
91          "d1c60049-922a-42d4-bd7e-8cf4ace47f05",
92          "5220a53f-f3fc-476c-aa72-65a038eb2fd8",
93          "b7e44e6e-ccf9-4b75-a258-159912ab51ca",
94          "42750622-28d7-4d32-9262-b139fe77bc01"
95        ],
96        "slide_ids": [
97          "a10196d2-7a81-4e1e-a9a7-62d123c30875",
98          "72edc1ba-916d-42a2-9f22-6254c6e54c5c",
99          "ff15eeb9-550e-4c78-90cc-a6cce8ccc3df",
100         "71ccfb52-169d-4176-94d6-fff5b75f853d"
101       ],
102       "submitter_sample_ids": [
```

```
103          "TCGA-66-2770-11A",
104          "TCGA-66-2770-01A"
105        ]
106      },
107      {
108        "sample_ids": [
109          "06889714-2a40-4248-98ee-f690b301e36a",
110          "9f43a0c6-ea19-4021-b0ed-026f33ce1c33"
111        ],
112        "portion_ids": [
113          "3a001d28-7cf9-4c61-b155-73938aebaa25",
114          "79554cfd-e853-481e-8e37-1e296034094e"
115        ],
116        "submitter_portion_ids": [
117          "TCGA-02-0075-01A-01",
118          "TCGA-02-0075-10A-01"
119        ],
120        "created_datetime": null,
121        "submitter_aliquot_ids": [
122          "TCGA-02-0075-01A-01W-0204-02",
123          "TCGA-02-0075-01A-01R-0194-03",
124          "TCGA-02-0075-01A-01D-0198-02",
125          "TCGA-02-0075-01A-01R-0202-01",
126          "TCGA-02-0075-10A-01W-0207-09",
127          "TCGA-02-0075-01A-01R-0676-04",
128          "TCGA-02-0075-10A-01D-0198-02",
129          "TCGA-02-0075-10A-01D-0197-06",
130          "TCGA-02-0075-10A-01D-0193-01",
131          "TCGA-02-0075-01A-01W-0207-09",
132          "TCGA-02-0075-01A-01W-0206-08",
133          "TCGA-02-0075-01A-01D-0193-01",
134          "TCGA-02-0075-10A-01W-0205-10",
135          "TCGA-02-0075-01A-01R-0201-02",
136          "TCGA-02-0075-10A-01W-0204-02",
137          "TCGA-02-0075-01A-01D-0199-05",
138          "TCGA-02-0075-10A-01W-0206-08",
139          "TCGA-02-0075-01A-01D-0196-04",
140          "TCGA-02-0075-01A-01T-0195-07",
141          "TCGA-02-0075-10A-01D-0196-04",
142          "TCGA-02-0075-01A-01D-0197-06",
143          "TCGA-02-0075-01A-01D-0888-01",
144          "TCGA-02-0075-01A-01R-0195-07",
145          "TCGA-02-0075-01A-01W-0205-10"
146        ],
147        "updated_datetime": "2016-05-02T15:00:01.972331-05:00",
148        "submitter_analyte_ids": [
149          "TCGA-02-0075-01A-01R",
150          "TCGA-02-0075-10A-01D",
151          "TCGA-02-0075-01A-01W",
152          "TCGA-02-0075-01A-01T",
153          "TCGA-02-0075-01A-01D",
154          "TCGA-02-0075-10A-01W"
155        ],
156        "analyte_ids": [
157          "fec22de0-a2b9-45df-9854-1ebe76cee84e",
158          "b4d11c50-61f1-4d4a-815f-1c0413018d7f",
159          "c48673d0-a38d-44e1-8cfd-e91cb23ea2d5",
160          "24f1852c-999a-4ea8-917c-fcfd683e2aca",
```

```
161            "aa431260-a0fc-4924-80ce-61cab8b5e83e",
162            "11f21140-d761-44ca-a9b2-b24099df3b15"
163          ],
164          "submitter_id": "TCGA-02-0075",
165          "case_id": "b196f82b-ef3f-4e05-99f7-da5df65e691e",
166          "state": null,
167          "aliquot_ids": [
168            "75531fe0-101e-4220-bd47-98892c90ee70",
169            "e5ea38d4-f47c-4c8a-8bab-13631e0a9a7b",
170            "d48b7c2c-daac-4496-af8f-1f45ca43f627",
171            "bbba08fc-2514-4e15-afb7-41eecc7e876f",
172            "0685b37f-a47c-4222-a846-bf9f3c000de3",
173            "683986da-3cee-446d-9b7a-83bef25815c9",
174            "e6ffdb20-a1be-4664-bcd3-cc7a4de6f40b",
175            "5d1f25c0-9e1a-41ad-9735-134f39dbf70e",
176            "528b40b9-246f-4ba3-8209-777136638e62",
177            "33131479-5d69-4262-a549-ba8864320f3b",
178            "5c7822fc-cf4f-4f62-8482-7c0ce1b7ab9a",
179            "b95e7659-e3a4-4e96-b98c-f67d26b85322",
180            "30c84aca-f9db-4e07-ac34-1a92b1652ca1",
181            "d5e3b5cc-06e0-4294-9d3c-8f3b63acae3d",
182            "b14b3d09-3a7f-41a6-81df-2757efa67906",
183            "513040e2-dc29-4e2c-86fb-57371eede17a",
184            "21c3be1b-7c1e-4864-99d1-486cfe5d8f1d",
185            "5e28e5dc-6dfa-44a9-8793-9134cb4cdda5",
186            "b8c25892-4773-428f-a02c-f930931268e8",
187            "266d5260-08e4-4cec-87f3-ca415bd98575",
188            "8859a3ae-f85d-4ef2-830b-80f42f98d53e",
189            "ac018a8c-a6e2-4291-a4bf-a330ae9c441e",
190            "4b022f7f-7549-4d97-9d41-4e5f2e9ec74c",
191            "caad3dfa-74a9-4ecc-95c1-86f6fbfd4ab5"
192          ],
193          "slide_ids": [
194            "39f547cd-5dc3-4bf4-99ea-073bb161c23c",
195            "5f096267-0cc2-4cc5-a206-7357159633d7"
196          ],
197          "submitter_sample_ids": [
198            "TCGA-02-0075-10A",
199            "TCGA-02-0075-01A"
200          ]
201        },
202        {
203          "sample_ids": [
204            "ba08195b-31cf-4bb1-a470-23740225c99d",
205            "929889c4-e474-4104-b69b-fac7e414a59e"
206          ],
207          "portion_ids": [
208            "48a36eb4-79fb-45e7-8bb1-0fa1d5fcda2c",
209            "1de5e67a-ac3f-4c18-92c4-27ba1868c7ac",
210            "e09fc5e7-e8d2-4bf9-b12b-17b22e0387e4"
211          ],
212          "submitter_portion_ids": [
213            "TCGA-EJ-A8FU-10A-01",
214            "TCGA-EJ-A8FU-01A-21-A43L-20",
215            "TCGA-EJ-A8FU-01A-11"
216          ],
217          "created_datetime": null,
218          "submitter_aliquot_ids": [
```

```
219            "TCGA-EJ-A8FU-01A-11R-A36B-13",
220            "TCGA-EJ-A8FU-01A-11R-A36G-07",
221            "TCGA-EJ-A8FU-01A-11D-A363-01",
222            "TCGA-EJ-A8FU-10A-01D-A361-01",
223            "TCGA-EJ-A8FU-10A-01D-A362-08",
224            "TCGA-EJ-A8FU-01A-11W-A447-08",
225            "TCGA-EJ-A8FU-01A-11D-A365-05",
226            "TCGA-EJ-A8FU-01A-11D-A364-08",
227            "TCGA-EJ-A8FU-10A-01W-A446-08"
228          ],
229          "updated_datetime": "2016-05-02T15:57:04.948573-05:00",
230          "submitter_analyte_ids": [
231            "TCGA-EJ-A8FU-01A-11W",
232            "TCGA-EJ-A8FU-01A-11D",
233            "TCGA-EJ-A8FU-01A-11R",
234            "TCGA-EJ-A8FU-10A-01W",
235            "TCGA-EJ-A8FU-10A-01D"
236          ],
237          "analyte_ids": [
238            "2d4e4925-6ac8-498f-882b-4bbf319f6b7b",
239            "8d09b982-1256-4674-b383-d6ca4b4bb3c8",
240            "c74495d9-63bf-4ac0-b10e-04b3b06103c1",
241            "b9884d98-af57-4901-8b9d-4fdbf73d2c5a",
242            "2f16ac02-13bf-44fd-bbd7-658c1c384928"
243          ],
244          "submitter_id": "TCGA-EJ-A8FU",
245          "case_id": "23e56e08-e11d-4e83-88a8-1254675b3af8",
246          "state": null,
247          "aliquot_ids": [
248            "e77da017-5dc6-4e32-9568-755e4ee9b533",
249            "c9b286d1-d500-4bb3-bb3d-5bf40b1b1265",
250            "b7867d52-7987-46d4-a595-0ff5b5375a58",
251            "5586ad35-94b7-459e-8982-8e7fb25697a1",
252            "162a63f7-594f-4669-a06d-b4899c7fe86a",
253            "b8b1ab44-ee6e-4ac5-9efd-d5bd07e67b9c",
254            "7adcdf73-3ad3-4da7-ab27-2888f1d4f53a",
255            "eb498e52-3eae-402f-8cac-ec930f8d938d",
256            "293f781c-c2c7-479b-b1a6-5f951a2c5e5a"
257          ],
258          "slide_ids": [
259            "454a95d5-d084-4f36-b1f1-32c6c23ab46e"
260          ],
261          "submitter_sample_ids": [
262            "TCGA-EJ-A8FU-01A",
263            "TCGA-EJ-A8FU-10A"
264          ]
265        },
266        {
267          "sample_ids": [
268            "d43f0112-fe59-4842-9fda-1189e5fb7248",
269            "213cbbe5-c382-47a1-b936-bf40c2c99091"
270          ],
271          "portion_ids": [
272            "26441aae-22e5-4e69-b3f5-34ccde356c93",
273            "60d7a93c-0634-438e-a72a-ce63630bb890",
274            "246a8f01-7ef2-4737-a984-49aa0b41c089"
275          ],
276          "submitter_portion_ids": [
```

```
277        "TCGA-F2-6879-10A-01",
278        "TCGA-F2-6879-01A-21-A39M-20",
279        "TCGA-F2-6879-01A-11"
280      ],
281      "created_datetime": "2016-05-02T16:23:44.347995-05:00",
282      "submitter_aliquot_ids": [
283        "TCGA-F2-6879-01A-11R-2155-13",
284        "TCGA-F2-6879-10A-01D-2153-01",
285        "TCGA-F2-6879-10A-01D-2152-26",
286        "TCGA-F2-6879-01A-11D-2157-05",
287        "TCGA-F2-6879-10A-01D-2154-08",
288        "TCGA-F2-6879-01A-11D-A45X-08",
289        "TCGA-F2-6879-01A-11D-2154-08",
290        "TCGA-F2-6879-01A-11W-2179-08",
291        "TCGA-F2-6879-01A-11D-2153-01",
292        "TCGA-F2-6879-01A-11R-2156-07",
293        "TCGA-F2-6879-01A-11D-2152-26",
294        "TCGA-F2-6879-10A-01D-A45X-08",
295        "TCGA-F2-6879-10A-01W-2179-08",
296        "TCGA-F2-6879-01A-01D-YYYY-23"
297      ],
298      "updated_datetime": "2016-05-02T16:23:44.347995-05:00",
299      "submitter_analyte_ids": [
300        "TCGA-F2-6879-10A-01D",
301        "TCGA-F2-6879-01A-11R",
302        "TCGA-F2-6879-10A-01W",
303        "TCGA-F2-6879-01A-11W",
304        "TCGA-F2-6879-01A-11D"
305      ],
306      "analyte_ids": [
307        "e87dde8d-3bf5-42d8-9a77-620d5c4943e0",
308        "30ade77d-996b-4031-93ab-6b341d49eb0a",
309        "1d94bd70-6621-4a94-8102-d673663e6665",
310        "ea65d92e-1597-410d-84d8-abb2a6235b3e",
311        "79697034-1cec-4d92-8195-8a35258ab477"
312      ],
313      "submitter_id": "TCGA-F2-6879",
314      "case_id": "8d9bd437-8b4b-4da5-87ba-6b5790f05022",
315      "state": null,
316      "aliquot_ids": [
317        "e7533585-b062-4d74-b511-05dc806a1357",
318        "e107952a-cc2b-4410-b0f9-62e7115430a0",
319        "61f1c8b1-986a-485a-9d96-4e4285b6425a",
320        "c043e276-fece-4cb9-a848-a0b16e6099b6",
321        "e5d110e1-63ad-49ce-b9b7-22bbd7ef8a88",
322        "7accb08d-acdb-46bc-bf7f-b9f678193115",
323        "a52cd04b-41d6-40db-b050-00ef3a143f7e",
324        "207fcf5e-c422-4333-9ec2-5dab38d240c7",
325        "5ddd3f83-28a8-4b7f-9aec-203a3c2efbe5",
326        "ccd4dd70-c0e4-42cf-870e-33d1013b201a",
327        "e12314fe-f16a-4d85-95b4-e712ede450f6",
328        "695461e3-283c-4b5b-9325-6b2588b67fd8",
329        "8481be1e-0993-487d-8d73-b0eb72b304ee",
330        "d7200791-4f1c-418f-8744-91b793486d9f"
331      ],
332      "slide_ids": [
333        "bcbcc947-cab1-4400-aebc-1d9e251a3ce8",
334        "cae8d0b9-3605-40af-bf99-7c23df8110a9"
```

```
335            ],
336            "submitter_sample_ids": [
337              "TCGA-F2-6879-10A",
338              "TCGA-F2-6879-01A"
339            ]
340          },
341          {
342            "sample_ids": [
343              "3a66b5bd-7037-463c-9f8d-2ba3de9d5571",
344              "84f603d6-9f71-48fb-b2e3-190424407452"
345            ],
346            "portion_ids": [
347              "fe90de9f-8ee3-4d55-834f-a90538958cb7",
348              "7a0042fd-07f0-4894-adb0-03cebce8aa02"
349            ],
350            "submitter_portion_ids": [
351              "TCGA-VQ-A922-01A-11",
352              "TCGA-VQ-A922-10A-01"
353            ],
354            "created_datetime": "2016-05-02T16:26:23.121974-05:00",
355            "submitter_aliquot_ids": [
356              "TCGA-VQ-A922-10A-01D-A412-01",
357              "TCGA-VQ-A922-01A-11D-A40Z-01",
358              "TCGA-VQ-A922-10A-01D-A413-08",
359              "TCGA-VQ-A922-01A-01D-YYYY-23",
360              "TCGA-VQ-A922-01A-11R-A414-31",
361              "TCGA-VQ-A922-01A-11D-A410-08",
362              "TCGA-VQ-A922-01A-11R-A415-13",
363              "TCGA-VQ-A922-01A-11D-A411-05"
364            ],
365            "updated_datetime": "2016-05-02T16:26:23.121974-05:00",
366            "submitter_analyte_ids": [
367              "TCGA-VQ-A922-01A-11R",
368              "TCGA-VQ-A922-10A-01D",
369              "TCGA-VQ-A922-01A-11D"
370            ],
371            "analyte_ids": [
372              "15bec495-04c7-412b-ad69-26b1f9274ccf",
373              "26a24673-04a1-4837-b888-702b0578aef2",
374              "2c0ecd67-b9ff-4e60-8d2f-7744c79a13aa"
375            ],
376            "submitter_id": "TCGA-VQ-A922",
377            "case_id": "8bd783a3-d6c9-4c87-a2a1-09f903b9c7ca",
378            "state": null,
379            "aliquot_ids": [
380              "58a121b4-265c-44ae-b6a9-79d087ee8b34",
381              "76fbba49-0123-4524-89aa-a1818c5507cb",
382              "0b0805bb-edaa-400f-ae9f-effed3dbb605",
383              "3370d626-d572-4d13-9cd3-1823a5df3d34",
384              "60934993-a9df-4389-b64d-da6844ef22df",
385              "243f24ba-bb0f-44e0-bcb1-69a97b395981",
386              "6cae9f2a-1c6c-4645-98b6-20719aec1413",
387              "44d020d1-c516-4a15-94e8-bcf0cb9c2683"
388            ],
389            "slide_ids": [
390              "0ff02899-57f8-419e-8872-c6ede53f4d3c"
391            ],
392            "submitter_sample_ids": [
```

```
393              "TCGA-VQ-A922-10A",
394              "TCGA-VQ-A922-01A"
395            ]
396          },
397          {
398            "sample_ids": [
399              "5bb5bd60-cf47-413b-88fa-f14977e24035",
400              "82fcf670-1646-4a28-9578-f7e5b2f426e5",
401              "3b87fed0-cfbd-4ee3-b71d-ab595853e836"
402            ],
403            "portion_ids": [
404              "18bf160e-702a-464a-9920-f115024b5484",
405              "10a9c093-009d-4bc0-a344-2afd3f0f9b9f",
406              "8ebd06e1-5eda-47ec-8888-61965ecf005e"
407            ],
408            "submitter_portion_ids": [
409              "TCGA-HU-8243-11A-01",
410              "TCGA-HU-8243-01A-11",
411              "TCGA-HU-8243-10A-01"
412            ],
413            "created_datetime": "2016-05-02T16:17:09.754748-05:00",
414            "submitter_aliquot_ids": [
415              "TCGA-HU-8243-01A-01D-YYYY-23",
416              "TCGA-HU-8243-01A-11D-2340-08",
417              "TCGA-HU-8243-01A-11D-2338-01",
418              "TCGA-HU-8243-01A-11D-2342-05",
419              "TCGA-HU-8243-11A-01D-2338-01",
420              "TCGA-HU-8243-11A-01D-2340-08",
421              "TCGA-HU-8243-10A-01D-2339-01",
422              "TCGA-HU-8243-01A-11R-2343-13",
423              "TCGA-HU-8243-10A-01D-2341-08"
424            ],
425            "updated_datetime": "2016-05-02T16:17:09.754748-05:00",
426            "submitter_analyte_ids": [
427              "TCGA-HU-8243-11A-01D",
428              "TCGA-HU-8243-10A-01D",
429              "TCGA-HU-8243-01A-11R",
430              "TCGA-HU-8243-01A-11D"
431            ],
432            "analyte_ids": [
433              "89c9094d-5cf6-4c7d-ad24-41b7ad9427cc",
434              "2c413e60-0122-426b-afb3-ae94810e2513",
435              "57d41760-0fed-49d2-8606-48231cb244ea",
436              "37ed51fd-b540-408e-8bd6-4447ae4aa84a"
437            ],
438            "submitter_id": "TCGA-HU-8243",
439            "case_id": "77a8eab6-f6a1-4739-9031-75ead40d68cb",
440            "state": null,
441            "aliquot_ids": [
442              "ace3edd6-14a9-42cc-84f3-6127237f2913",
443              "a711abd1-f1c2-4e42-8b66-79b4514ac1c4",
444              "6af7ba34-58f7-4472-8c7e-89fc91ad5ac1",
445              "558ff67a-a584-46f8-9089-8f4a08015294",
446              "71c0a224-5953-4b59-a49c-b7aa1e959f1e",
447              "a460c222-bcac-4959-961f-4dbd73e1ce13",
448              "6e5789d7-4988-457a-86eb-e618c7ab06eb",
449              "ff31f56b-398c-45ee-b122-f10027774527",
450              "9635cfd4-3d26-4fc6-846c-fd74d5b60098"
```

```
451          ],
452          "slide_ids": [
453            "60b7c6b8-594a-40c3-9341-a0902e4e6938",
454            "e55e00a0-2048-404a-b83a-f34106468694"
455          ],
456          "submitter_sample_ids": [
457            "TCGA-HU-8243-10A",
458            "TCGA-HU-8243-01A",
459            "TCGA-HU-8243-11A"
460          ]
461        },
462        {
463          "sample_ids": [
464            "2f5cc9c9-31a9-5eb3-952a-b21e7cef50ca",
465            "4f3f4fc8-4465-5230-83ec-c0ef6aceb2ea"
466          ],
467          "updated_datetime": "2016-05-25T19:12:45.610324-05:00",
468          "submitter_aliquot_ids": [
469            "TARGET-30-PAUXFZ-01A-01D",
470            "TARGET-30-PAUXFZ-10A-01D"
471          ],
472          "submitter_id": "TARGET-30-PAUXFZ",
473          "case_id": "a7ccef7c-14c0-5232-b647-58b4a54fb343",
474          "aliquot_ids": [
475            "9e1e30a8-7607-5b7e-b33c-9a6c5828d5fb",
476            "c56898f9-c394-516a-bdbb-bf32a5af9d3f"
477          ],
478          "submitter_sample_ids": [
479            "TARGET-30-PAUXFZ-01A",
480            "TARGET-30-PAUXFZ-10A"
481          ]
482        },
483        {
484          "sample_ids": [
485            "c1bcb8d1-e13d-4af4-93f4-02d5f7f616a2",
486            "52fcf737-cdcc-43ea-b33c-4018039b42dd"
487          ],
488          "portion_ids": [
489            "e0e97a05-656a-468e-8418-0d08c38e76ab",
490            "3e2a0eab-7d89-4f3c-9c0e-8942e53d3c45"
491          ],
492          "submitter_portion_ids": [
493            "TCGA-KK-A8I9-01A-11",
494            "TCGA-KK-A8I9-11A-11"
495          ],
496          "created_datetime": null,
497          "submitter_aliquot_ids": [
498            "TCGA-KK-A8I9-11A-11D-A361-01",
499            "TCGA-KK-A8I9-11A-11D-A362-08",
500            "TCGA-KK-A8I9-11A-11W-A446-08",
501            "TCGA-KK-A8I9-01A-11R-A36G-07",
502            "TCGA-KK-A8I9-11A-11D-A40C-01",
503            "TCGA-KK-A8I9-01A-11D-A363-01",
504            "TCGA-KK-A8I9-01A-11W-A447-08",
505            "TCGA-KK-A8I9-01A-11D-A365-05",
506            "TCGA-KK-A8I9-01A-11D-A364-08",
507            "TCGA-KK-A8I9-01A-11R-A36B-13"
508          ],
```

```
509        "updated_datetime": "2016-05-02T15:57:29.451686-05:00",
510        "submitter_analyte_ids": [
511          "TCGA-KK-A8I9-11A-11W",
512          "TCGA-KK-A8I9-01A-11R",
513          "TCGA-KK-A8I9-11A-11D",
514          "TCGA-KK-A8I9-01A-11W",
515          "TCGA-KK-A8I9-01A-11D"
516        ],
517        "analyte_ids": [
518          "ddec19cb-5e4c-4151-8b6d-741044abff1e",
519          "96c5b539-8eb7-4156-81d0-7b7fecd68900",
520          "ced38a45-7610-49d4-8bf9-d53a1fc2d489",
521          "476f5deb-1b3f-4a35-8a31-f27763ba8d8a",
522          "c284f2af-1e9b-40cc-8936-b61cfd251d62"
523        ],
524        "submitter_id": "TCGA-KK-A8I9",
525        "case_id": "261c3d74-706e-4751-bd15-8f3c1a402ff0",
526        "state": null,
527        "aliquot_ids": [
528          "4f76de2d-e07a-402b-9818-7f04d3704a43",
529          "96802a73-b1db-47d7-8f5f-4504f3ece5ad",
530          "f376fc45-370a-4d96-833b-9a1322e32a42",
531          "d3e88dd3-66d7-40d4-978a-4ddab868373a",
532          "06f1d087-75c9-4da8-8339-80aff3bfaa12",
533          "50b1e243-b45a-42a1-8692-b7ae5d51250f",
534          "0f1c00d3-f3dc-4d2b-bd8a-ecc31e4f4089",
535          "986a3ed6-ba56-4025-a2bd-9909648e703a",
536          "bebc84b6-9179-420b-8207-858b999e8c0c",
537          "239d5e7e-5fb5-4df3-ae6b-a5a06ee296ae"
538        ],
539        "slide_ids": [
540          "1e174ca5-9298-41b6-a705-728f111a3e7b",
541          "a3e31324-9e06-4799-85b4-4f6236848009"
542        ],
543        "submitter_sample_ids": [
544          "TCGA-KK-A8I9-11A",
545          "TCGA-KK-A8I9-01A"
546        ]
547      },
548      {
549        "sample_ids": [
550          "d43f727a-96d6-40b8-86ae-7a3e0aa46853",
551          "b8329a6d-a87b-47f4-ad00-9e979e62647b"
552        ],
553        "portion_ids": [
554          "8960ddcc-0950-4d6e-a557-8727b652c93b",
555          "e36bfd07-c911-4a98-8424-e58e5e9aaa68"
556        ],
557        "submitter_portion_ids": [
558          "TCGA-QR-A70H-10A-01",
559          "TCGA-QR-A70H-01A-12"
560        ],
561        "created_datetime": null,
562        "submitter_aliquot_ids": [
563          "TCGA-QR-A70H-01A-12R-A35K-07",
564          "TCGA-QR-A70H-01A-12R-A35M-13",
565          "TCGA-QR-A70H-01A-12D-A35E-05",
566          "TCGA-QR-A70H-10A-01D-A35A-01",
```

```
567          "TCGA-QR-A70H-01A-12D-A35C-01",
568          "TCGA-QR-A70H-01A-12W-A43Z-08",
569          "TCGA-QR-A70H-10A-01D-A35B-08",
570          "TCGA-QR-A70H-10A-01W-A441-08",
571          "TCGA-QR-A70H-01A-12D-A35D-08"
572        ],
573        "updated_datetime": "2016-05-02T15:37:31.996088-05:00",
574        "submitter_analyte_ids": [
575          "TCGA-QR-A70H-10A-01D",
576          "TCGA-QR-A70H-10A-01W",
577          "TCGA-QR-A70H-01A-12D",
578          "TCGA-QR-A70H-01A-12W",
579          "TCGA-QR-A70H-01A-12R"
580        ],
581        "analyte_ids": [
582          "c4a41555-dd45-4e10-a3be-50d49a1121a3",
583          "957e01f6-eb3f-446e-9f45-b50c66337e2d",
584          "1acde950-2e0c-4586-852b-b4ac4e1ea4a4",
585          "67c033c0-9fe8-4004-967e-c605e1890f4d",
586          "b0873010-5d60-4691-b700-e172950f1d7c"
587        ],
588        "submitter_id": "TCGA-QR-A70H",
589        "case_id": "13b41b15-a785-4ab7-b864-ffff6d35dd45",
590        "state": null,
591        "aliquot_ids": [
592          "d9120f00-7f10-49d5-ae84-6177e9424c7c",
593          "31c6fa50-200a-46c1-a546-61b52592fd8f",
594          "ab50f38c-2e7d-4d75-a216-27aeaa4d9305",
595          "382d5e31-6c66-4df3-a695-6b8c29cfc681",
596          "51d1fb14-c918-4439-b816-ef6cd3253c64",
597          "f586d8d5-d0c6-4979-aaa7-10217a88fa4c",
598          "2f9a60eb-602e-44bb-bc57-87e20d946f76",
599          "fbafc85e-deff-46cd-a40f-479b9dc92a60",
600          "cacbc8a6-0eb0-4277-931f-d0075c9b1de9"
601        ],
602        "slide_ids": [
603          "2310e34c-0ea5-4876-9f87-bad0b7a44513"
604        ],
605        "submitter_sample_ids": [
606          "TCGA-QR-A70H-01A",
607          "TCGA-QR-A70H-10A"
608        ]
609      },
610      {
611        "sample_ids": [
612          "19dee039-9c98-4d4a-8baf-eea1b6dda8eb",
613          "fdf1e501-f34f-450c-9a5c-611157079a86"
614        ],
615        "portion_ids": [
616          "10b6ccb4-3637-4769-8988-417c0306eaef",
617          "92f8cd48-451d-4ed6-8e60-b15aa93d2c09",
618          "d0d55efa-c91d-45de-92bf-cf6f0d263b21"
619        ],
620        "submitter_portion_ids": [
621          "TCGA-BJ-A18Z-01A-21",
622          "TCGA-BJ-A18Z-01A-11-A21L-20",
623          "TCGA-BJ-A18Z-10A-01"
624        ],
```

```
625          "created_datetime": null,
626          "submitter_aliquot_ids": [
627            "TCGA-BJ-A18Z-01A-21D-A13U-02",
628            "TCGA-BJ-A18Z-10A-01D-A13V-01",
629            "TCGA-BJ-A18Z-01A-21R-A13Y-07",
630            "TCGA-BJ-A18Z-01A-21W-A14T-08",
631            "TCGA-BJ-A18Z-01A-21D-A13Z-05",
632            "TCGA-BJ-A18Z-01A-21D-A37T-08",
633            "TCGA-BJ-A18Z-10A-01D-A13W-08",
634            "TCGA-BJ-A18Z-01A-21R-A13X-13",
635            "TCGA-BJ-A18Z-01A-21D-A13W-08",
636            "TCGA-BJ-A18Z-10A-01D-A13U-02",
637            "TCGA-BJ-A18Z-10A-01W-A14T-08",
638            "TCGA-BJ-A18Z-01A-21D-A13V-01"
639          ],
640          "updated_datetime": "2016-05-02T16:18:19.199189-05:00",
641          "submitter_analyte_ids": [
642            "TCGA-BJ-A18Z-01A-21W",
643            "TCGA-BJ-A18Z-01A-21D",
644            "TCGA-BJ-A18Z-01A-21R",
645            "TCGA-BJ-A18Z-10A-01D",
646            "TCGA-BJ-A18Z-10A-01W"
647          ],
648          "analyte_ids": [
649            "119ebfa1-75b2-4f24-816a-4e9a5061f6b5",
650            "f86759fd-ecc5-4f42-b5fe-b9f079d23968",
651            "39691042-bd28-40ed-b66b-26414ecf1ba0",
652            "76ea5056-d7fa-49fb-94bf-11171ca7c100",
653            "71a822c9-b510-4a4c-8c30-18b8083acc2d"
654          ],
655          "submitter_id": "TCGA-BJ-A18Z",
656          "case_id": "0d497faf-2c1c-4173-a5fe-770cca73323c",
657          "state": null,
658          "aliquot_ids": [
659            "fa580596-e70f-4ed0-85a2-6fb594ca679a",
660            "776cb4b1-8efd-4ea2-b53f-9dff7dd94b10",
661            "85a7922f-0327-437c-bdf5-1bb67a1e932f",
662            "6d532180-0175-4610-8bfa-cca3a7c3697a",
663            "b5977e73-49d8-4e99-9e97-993cc44dad17",
664            "918793fa-b35e-4745-ac75-4d1c868089f8",
665            "ba9479a1-929f-4e4e-8bf5-e23cb280dfcf",
666            "e9776ff5-69b9-4669-ab33-e4bb030461ec",
667            "8ba98907-ab03-4c9e-a900-e31aa16ff810",
668            "35e18649-183e-4223-b2f6-d812bdd9becd",
669            "4aa17671-4420-4989-a6dd-379250f4aeda",
670            "815c53c3-8add-4612-b93c-3ed4bfa530aa"
671          ],
672          "slide_ids": [
673            "7c5b5c77-9fbc-4b48-81f5-48b5ede7c436"
674          ],
675          "submitter_sample_ids": [
676            "TCGA-BJ-A18Z-01A",
677            "TCGA-BJ-A18Z-10A"
678          ]
679        }
680      ],
681      "pagination": {
682        "count": 10,
```

```
683        "sort": "",
684        "from": 1,
685        "page": 1,
686        "total": 6340,
687        "pages": 634,
688        "size": 10
689      }
690    },
691    "warnings": {}
692 }
```

Example: HTTP POST Request

This example demonstrates how to obtain metadata in TSV format for a set of files using their UUIDs (e.g. UUIDs obtained from a download manifest file generated by the GDC Data Portal).

The first step is to construct a JSON query object, including `filters`, `fields`, `format`, and `size` parameters. The object is then submitted as HTTP POST payload to the GDC API using curl, in order to retrieve a TSV file with the requested metadata.

```
 1 {
 2      "filters":{
 3          "op":"in",
 4          "content":{
 5              "field":"files.file_id",
 6              "value":[
 7                  "0001801b-54b0-4551-8d7a-d66fb59429bf",
 8                  "002c67f2-ff52-4246-9d65-a3f69df6789e",
 9                  "003143c8-bbbf-46b9-a96f-f58530f4bb82",
10                  "0043d981-3c6b-463f-b512-ab1d076d3e62",
11                  "004e2a2c-1acc-4873-9379-ef1aa12283b6",
12                  "005239a8-2e63-4ff1-9cd4-714f81837a61",
13                  "006b8839-31e5-4697-b912-8e3f4124dd15",
14                  "006ce9a8-cf38-462e-bb99-7f08499244ab",
15                  "007ce9b5-3268-441e-9ffd-b40d1127a319",
16                  "0084a614-780b-42ec-b85f-7a1b83128cd3",
17                  "00a5e471-a79f-4d56-8a4c-4847ac037400",
18                  "00ab2b5a-b59e-4ec9-b297-76f74ff1d3fb",
19                  "00c5f14e-a398-4076-95d1-25f320ee3a37",
20                  "00c74a8b-10aa-40cc-991e-3365ea1f3fce",
21                  "00df5a50-bce3-4edf-a078-641e54800dcb"
22              ]
23          }
24      },
25      "format":"TSV",
26      "fields":"file_id,file_name,cases.submitter_id,cases.case_id,data_category,data_type,cases.samples.tumor_des
27      "size":"100"
28 }
```

```
 1 curl --request POST --header "Content-Type: application/json" --data @Payload.txt
       'https://gdc-api.nci.nih.gov/files' > File_metadata.txt
```

```
 1 cases_0_submitter_id    cases_0_case_id data_type    cases_0_samples_0_sample_type
       cases_0_samples_0_tissue_type    file_name    cases_0_samples_0_submitter_id
       cases_0_samples_0_portions_0_analytes_0_aliquots_0_aliquot_id    cases_0_samples_0_sample_id file_id
       data_category    cases_0_samples_0_tumor_descriptor
       cases_0_samples_0_portions_0_analytes_0_aliquots_0_submitter_id
```

```
 2 TCGA-B0-5094    8aaa4e25-5c12-4ace-96dc-91aaa0c4457c    Aligned Reads    Solid Tissue Normal
      C345.TCGA-B0-5094-11A-01D-1421-08.5_gdc_realn.bam    TCGA-B0-5094-11A
      b4e4630a-b38c-4b62-b0e8-d73f0e3b4e47    7519d7a8-c3ee-417b-9cfc-111bc5ad0637
      0001801b-54b0-4551-8d7a-d66fb59429bf    Raw Sequencing Data    TCGA-B0-5094-11A-01D-1421-08
 3 TCGA-B0-5117    ae55b2d3-62a1-419e-9f9a-5ddfac356db4    Aligned Reads    Solid Tissue Normal
      C345.TCGA-B0-5117-11A-01D-1421-08.5_gdc_realn.bam    TCGA-B0-5117-11A
      45c68b6b-0bed-424d-9a77-4f87bbaa3649    b1116541-bece-4df3-b3dd-cec50aeb277b
      003143c8-bbbf-46b9-a96f-f58530f4bb82    Raw Sequencing Data    TCGA-B0-5117-11A-01D-1421-08
 4 TCGA-G7-6790    e7a1cbe2-793c-4747-8412-8be794f2382b    Aligned Reads    Blood Derived Normal
      C489.TCGA-G7-6790-10A-01D-1962-08.2_gdc_realn.bam    TCGA-G7-6790-10A
      66cbb40f-14b3-40c0-a332-e8a8e21bca11    4be83d0f-8b09-4e9e-8318-358371d34332
      004e2a2c-1acc-4873-9379-ef1aa12283b6    Raw Sequencing Data    TCGA-G7-6790-10A-01D-1962-08
 5 TCGA-B9-A69E    a4225cb2-7b4b-4122-b6b9-629c26e3ea56    Aligned Reads    Blood Derived Normal
      TCGA-B9-A69E-10A-01D-A31X-10_Illumina_gdc_realn.bam TCGA-B9-A69E-10A
      f4799bdc-b207-4053-9a4b-5a26ebf8ab91    5d6d6cd4-6a7b-499d-936a-1be9bf74b07f
      0084a614-780b-42ec-b85f-7a1b83128cd3    Raw Sequencing Data    TCGA-B9-A69E-10A-01D-A31X-10
 6 TCGA-EE-A2GU    24faa36a-268d-4a13-b3ae-eacd431a2bcc    Aligned Reads    Blood Derived Normal
      C828.TCGA-EE-A2GU-10A-01D-A198-08.2_gdc_realn.bam    TCGA-EE-A2GU-10A
      c3feacc2-5a26-4bb2-a312-8b2ee53ccad1    cc4a5ed8-376a-4842-a25d-ffb07d8e1ca0
      00c74a8b-10aa-40cc-991e-3365ea1f3fce    Raw Sequencing Data    TCGA-EE-A2GU-10A-01D-A198-08
 7 TCGA-CE-A484    e62a728d-390f-428a-bea1-fc8c9814fb11    Aligned Reads    Blood Derived Normal
      C499.TCGA-CE-A484-10A-01D-A23U-08.3_gdc_realn.bam    TCGA-CE-A484-10A
      641a0220-6eec-434a-b606-e256113b65da    27a8008e-044a-4966-b518-cc6905e292ca
      00df5a50-bce3-4edf-a078-641e54800dcb    Raw Sequencing Data    TCGA-CE-A484-10A-01D-A23U-08
 8 TCGA-DA-A1IB    8fc9cc74-f388-49f0-b957-debb62638634    Aligned Reads    Blood Derived Normal
      C828.TCGA-DA-A1IB-10A-01D-A198-08.2_gdc_realn.bam    TCGA-DA-A1IB-10A
      30919a1a-df9f-4604-835e-f66ac7bcacdf    432952c5-6505-4220-a581-f65270a45281
      00ab2b5a-b59e-4ec9-b297-76f74ff1d3fb    Raw Sequencing Data    TCGA-DA-A1IB-10A-01D-A198-08
 9 TCGA-AX-2HG    7a2cf5ce-8317-4fff-946e-b9937afab815    Aligned Reads    Blood Derived Normal
      6c2a8ea343da8d6cc0fd2043492f16df_gdc_realn.bam    TCGA-AX-2HG-10A
      8c34ffe2-9012-4b4a-b610-a42a9c6a9780    ef4b80ec-b453-48ec-8ad8-ccac83e1e4db
      00c5f14e-a398-4076-95d1-25f320ee3a37    Raw Sequencing Data    TCGA-AX-2HG-10A-01D-A17D-09
10 TCGA-EC-A24G    b5c1e511-baf2-45b3-9919-110e8941e3c2    Aligned Reads    Blood Derived Normal
      671333b193812fc2bd2744053b383459_gdc_realn.bam    TCGA-EC-A24G-10A
      2a8cb8fe-b64f-453e-8139-7ede12f3fc51    61cf2e54-1b8d-40a0-9c73-a7449cbd570a
      00a5e471-a79f-4d56-8a4c-4847ac037400    Raw Sequencing Data    TCGA-EC-A24G-10A-01D-A16D-09
11 TCGA-B5-A0K0    29c8f468-5ac1-4d6c-8376-e36e6d246926    Aligned Reads    Blood Derived Normal
      TCGA-B5-A0K0-10A-01W-A062-09_IlluminaGA-DNASeq_exome_gdc_realn.bam    TCGA-B5-A0K0-10A
      02e65074-ffda-4795-b8f5-1bfd20bd1019    1df69e2e-f392-465f-8e61-4671ba2fcd35
      007ce9b5-3268-441e-9ffd-b40d1127a319    Raw Sequencing Data    TCGA-B5-A0K0-10A-01W-A062-09
12 TCGA-C8-A27B    f0d8a1fe-e313-44f1-99cc-b965cbeeff0e    Aligned Reads    Blood Derived Normal
      3c99d98ea8eb6acbf819e67fc77623d9_gdc_realn.bam    TCGA-C8-A27B-10A
      922226ba-6244-4953-ad42-f4daa474c288    31139082-7978-45aa-9d8f-ac4789ac5cec
      006b8839-31e5-4697-b912-8e3f4124dd15    Raw Sequencing Data    TCGA-C8-A27B-10A-01D-A167-09
13 TCGA-E9-A295    fec0da58-1047-44d2-b6d1-c18cceed43dc    Aligned Reads    Blood Derived Normal
      fd4421a6bbf3efd4e3d5c17fdd610314_gdc_realn.bam    TCGA-E9-A295-10A
      cd761feb-9a20-4495-8943-c6243532a5cf    e74183e1-f0b4-412a-8dac-a62d404add78
      002c67f2-ff52-4246-9d65-a3f69df6789e    Raw Sequencing Data    TCGA-E9-A295-10A-01D-A16D-09
14 TCGA-EB-A440    c787c4da-c564-44f1-89eb-dd9da107acb1    Aligned Reads    Blood Derived Normal
      C828.TCGA-EB-A440-10A-01D-A250-08.3_gdc_realn.bam    TCGA-EB-A440-10A
      c723584a-c404-4c88-bfea-e40f5dbba542    5b738547-1825-4684-81bd-864bf2eb43ef
      006ce9a8-cf38-462e-bb99-7f08499244ab    Raw Sequencing Data    TCGA-EB-A440-10A-01D-A250-08
15 TCGA-A2-A3XX    53886143-c1c6-40e9-88e6-e4e5e0271fc8    Aligned Reads    Blood Derived Normal
      b40998d4778f18ed80d6dd8bff0eb761_gdc_realn.bam    TCGA-A2-A3XX-10A
      e96d5811-4736-40dd-966d-e0e172aeb0af    c6eb6218-ad71-40a6-88b7-a4f1a015b816
      0043d981-3c6b-463f-b512-ab1d076d3e62    Raw Sequencing Data    TCGA-A2-A3XX-10A-01D-A23C-09
16 TCGA-EB-A3XB    a9255dcb-b236-4777-ac43-555e3a5386c3    Aligned Reads    Blood Derived Normal
      C828.TCGA-EB-A3XB-10B-01D-A23B-08.1_gdc_realn.bam    TCGA-EB-A3XB-10B
```

```
9f4ffc2f-d006-4d86-b3b1-b25020481893    0e1d4c7c-204d-4765-b090-68ed4cd83835
005239a8-2e63-4ff1-9cd4-714f81837a61    Raw Sequencing Data    TCGA-EB-A3XB-10B-01D-A23B-08
```

Format

Specifies the format of the API response: JSON (default), `TSV` or `XML`.

Examples

```
1 curl  'https://gdc-api.nci.nih.gov/cases?fields=submitter_id&size=5&format=TSV'
```

```
1 import requests
2
3 cases_endpt = 'https://gdc-api.nci.nih.gov/cases'
4 params = {'fields':'submitter_id',
5           'format':'TSV'}
6 response = requests.get(cases_endpt, params = params)
7 print response.content
```

```
1 submitter_id
2 TCGA-RC-A6M6
3 TCGA-B6-A0RV
4 TCGA-MB-A5Y8
5 TCGA-BQ-5876
6 TCGA-Z6-A9VB
```

```
1 curl  'https://gdc-api.nci.nih.gov/cases?fields=submitter_id&size=5&format=XML&pretty=true'
```

```
1 import requests
2
3 cases_endpt = 'https://gdc-api.nci.nih.gov/cases'
4 params = {'fields':'submitter_id',
5           'format':'XML',
6           'pretty':'true'}
7 response = requests.get(cases_endpt, params = params)
8 print response.content
```

```
1 <?xml version="1.0" ?>
2 <response>
3     <data>
4         <hits>
5             <item>
6                 <submitter_id>TCGA-MQ-A4LV</submitter_id>
7             </item>
8             <item>
9                 <submitter_id>TCGA-N9-A4Q1</submitter_id>
10            </item>
11            <item>
12                <submitter_id>TCGA-78-7154</submitter_id>
13            </item>
14            <item>
15                <submitter_id>TCGA-S7-A7WX</submitter_id>
16            </item>
17            <item>
18                <submitter_id>TCGA-XF-AAML</submitter_id>
19            </item>
```

```
20          </hits>
21          <pagination>
22              <count>5</count>
23              <sort/>
24              <from>1</from>
25              <pages>2811</pages>
26              <total>14052</total>
27              <page>1</page>
28              <size>5</size>
29          </pagination>
30      </data>
31      <warnings/>
32 </response>
```

Pretty

Returns when the `pretty` parameter is set to `true`, the API response is formatted with additional whitespace to improve legibility.

Example

```
1 curl  'https://gdc-api.nci.nih.gov/cases?fields=submitter_id&sort=submitter_id:asc&size=5'
```

```
1 {"data": {"hits": [{"submitter_id": "TARGET-20-PABGKN"}, {"submitter_id": "TARGET-20-PABHET"},
      {"submitter_id": "TARGET-20-PABHKY"}, {"submitter_id": "TARGET-20-PABLDZ"}, {"submitter_id":
      "TARGET-20-PACDZR"}], "pagination": {"count": 5, "sort": "submitter_id.raw:asc", "from": 1, "pages":
      2811, "total": 14052, "page": 1, "size": 5}}, "warnings": {}}
```

```
1 curl  'https://gdc-api.nci.nih.gov/cases?fields=submitter_id&sort=submitter_id:asc&size=5&pretty=true'
```

```
1  {
2    "data": {
3      "hits": [
4        {
5          "submitter_id": "TARGET-20-PABGKN"
6        },
7        {
8          "submitter_id": "TARGET-20-PABHET"
9        },
10        {
11          "submitter_id": "TARGET-20-PABHKY"
12        },
13        {
14          "submitter_id": "TARGET-20-PABLDZ"
15        },
16        {
17          "submitter_id": "TARGET-20-PACDZR"
18        }
19      ],
20      "pagination": {
21        "count": 5,
22        "sort": "submitter_id.raw:asc",
23        "from": 1,
24        "pages": 2811,
25        "total": 14052,
26        "page": 1,
27        "size": 5
28      }
```

```
29     },
30     "warnings": {}
31 }
```

Fields

This query parameter specifies which fields are to be included in the API response. A listing of available fields for each endpoint is provided in Appendix A.

Example

The following example requests case submitter ID, file UUID, file name and file size from the `files` endpoint.

```
1 curl 'https://gdc-api.nci.nih.gov/files?fields=cases.submitter_id,file_id,file_name,file_size&pretty=true'
```

```
1 import requests
2 import json
3
4 files_endpt = 'https://gdc-api.nci.nih.gov/files'
5 params = {'fields':'cases.submitter_id,file_id,file_name,file_size'}
6 response = requests.get(files_endpt, params = params)
7 print json.dumps(response.json(), indent=2)
```

```
1 {
2   "data": {
3     "hits": [
4       {
5         "file_name": "NARKY_p_TCGAb69_SNP_N_GenomeWideSNP_6_H03_697832.grch38.seg.txt",
6         "cases": [
7           {
8             "submitter_id": "TCGA-BP-4989"
9           }
10        ],
11        "file_id": "3bd4d5dc-563a-481c-87a6-ec0017d0d58a",
12        "file_size": 54200
13      },
14      {
15        "file_name": "652ecf99-1af9-41fc-b0a5-d3e5c07a7b5d.FPKM.txt.gz",
16        "cases": [
17          {
18            "submitter_id": "TCGA-60-2709"
19          }
20        ],
21        "file_id": "b3286166-01f9-4149-81b5-a2ea5f27c50e",
22        "file_size": 530665
23      },
24      {
25        "file_name": "CUSKS_p_TCGAb47_SNP_1N_GenomeWideSNP_6_D05_628212.nocnv_grch38.seg.txt",
26        "cases": [
27          {
28            "submitter_id": "TCGA-A8-A07Z"
29          }
30        ],
31        "file_id": "282cc9d1-c5e9-49ff-b27b-e00c1e5529c6",
32        "file_size": 15806
33      },
34      {
35        "file_name": "REEDY_p_TCGAb65_SNP_N_GenomeWideSNP_6_F01_697686.nocnv_grch38.seg.txt",
```

```
36        "cases": [
37          {
38            "submitter_id": "TCGA-CJ-4871"
39          }
40        ],
41        "file_id": "fe44a644-eefc-42c5-aac7-a216bc1e88e1",
42        "file_size": 6179
43      },
44      {
45        "file_name": "84df7a8fee9fedb5e8e22849ec66d294_gdc_realn.bam",
46        "cases": [
47          {
48            "submitter_id": "TCGA-A2-A0CO"
49          }
50        ],
51        "file_id": "acd0ec73-c1fe-463e-912c-84e8416510e5",
52        "file_size": 15545555724
53      },
54      {
55        "file_name": "ed8c4bb6-891a-4cf2-80ba-42c5594760d0.vcf",
56        "cases": [
57          {
58            "submitter_id": "TCGA-BQ-7059"
59          }
60        ],
61        "file_id": "ed8c4bb6-891a-4cf2-80ba-42c5594760d0",
62        "file_size": 264694
63      },
64      {
65        "file_name": "nationwidechildrens.org_clinical.TCGA-IG-A6QS.xml",
66        "cases": [
67          {
68            "submitter_id": "TCGA-IG-A6QS"
69          }
70        ],
71        "file_id": "fe8cf009-f033-4536-95c7-836adcba5bf3",
72        "file_size": 36996
73      },
74      {
75        "file_name": "05f6f9f7-6fb7-4c95-b79c-fdfaba16539d.vep.reheader.vcf.gz",
76        "cases": [
77          {
78            "submitter_id": "TCGA-DK-A3IV"
79          }
80        ],
81        "file_id": "05f6f9f7-6fb7-4c95-b79c-fdfaba16539d",
82        "file_size": 415044
83      },
84      {
85        "file_name": "C484.TCGA-12-5301-01A-01D-1486-08.7_gdc_realn.bam",
86        "cases": [
87          {
88            "submitter_id": "TCGA-12-5301"
89          }
90        ],
91        "file_id": "3b0293c2-4a26-428c-b097-9489f23a2a2d",
92        "file_size": 23661175335
93      },
```

```
94          {
95            "file_name": "75a36e71-400d-46a5-93b0-7813cf0595ea.FPKM.txt.gz",
96            "cases": [
97              {
98                "submitter_id": "TCGA-BF-A5EO"
99              }
100           ],
101           "file_id": "28f763c7-8064-4151-ae0e-31e70cd9bfe8",
102           "file_size": 488422
103         }
104       ],
105       "pagination": {
106         "count": 10,
107         "sort": "",
108         "from": 1,
109         "page": 1,
110         "total": 216435,
111         "pages": 21644,
112         "size": 10
113       }
114     },
115     "warnings": {}
116 }
```

Expand

The `expand` parameter provides a shortcut to request multiple related fields (field groups) in the response. Instead of specifying each field using the `fields` parameter, users can specify a field group name using the `expand` parameter to request all fields in the group. Available field groups are listed in Appendix A; the list can also be accessed programmatically at the _mapping endpoint. The `fields` and `expand` parameters can be used together to request custom combinations of field groups and individual fields.

Example

```
1 curl
      'https://gdc-api.nci.nih.gov/files/ac2ddebd-5e5e-4aea-a430-5a87c6d9c878?expand=cases.samples&pretty=true'
```

```
1 {
2   "data": {
3     "data_type": "Aligned Reads",
4     "updated_datetime": "2016-09-18T04:25:13.163601-05:00",
5     "created_datetime": "2016-05-26T18:55:53.506549-05:00",
6     "file_name": "000aa811c15656604161e8f0e3a0aae4_gdc_realn.bam",
7     "md5sum": "200475f5f6e42520204e5f6aadfe954f",
8     "data_format": "BAM",
9     "acl": [
10      "phs000178"
11     ],
12     "access": "controlled",
13     "platform": "Illumina",
14     "state": "submitted",
15     "file_id": "ac2ddebd-5e5e-4aea-a430-5a87c6d9c878",
16     "data_category": "Raw Sequencing Data",
17     "file_size": 12667634731,
18     "cases": [
19       {
20         "samples": [
```

```
21              {
22                "sample_type_id": "11",
23                "updated_datetime": "2016-09-08T11:00:45.021005-05:00",
24                "time_between_excision_and_freezing": null,
25                "oct_embedded": "false",
26                "tumor_code_id": null,
27                "submitter_id": "TCGA-QQ-A5VA-11A",
28                "intermediate_dimension": null,
29                "sample_id": "b4e7558d-898e-4d68-a897-381edde0bbcc",
30                "is_ffpe": false,
31                "pathology_report_uuid": null,
32                "created_datetime": null,
33                "tumor_descriptor": null,
34                "sample_type": "Solid Tissue Normal",
35                "state": null,
36                "current_weight": null,
37                "composition": null,
38                "time_between_clamping_and_freezing": null,
39                "shortest_dimension": null,
40                "tumor_code": null,
41                "tissue_type": null,
42                "days_to_sample_procurement": null,
43                "freezing_method": null,
44                "preservation_method": null,
45                "days_to_collection": 5980,
46                "initial_weight": 810.0,
47                "longest_dimension": null
48              }
49            ]
50          }
51        ],
52        "submitter_id": "32872121-d38a-4128-b96a-698a6f18f29d",
53        "type": "aligned_reads",
54        "file_state": "processed",
55        "experimental_strategy": "WXS"
56      },
57      "warnings": {}
58 }
```

Size and From

GDC API provides a pagination feature that limits the number of results returned by the API. It is implemented using `size` and `from` query parameters.

The `size` query parameter specifies the maximum number of results to return. Default `size` is 10. If the number of query results is greater than `size`, only some of the results will be returned.

The `from` query parameter specifies the first record to return out of the set of results. For example, if there are 20 cases returned from the `cases` endpoint, then setting `from` to `11` will return results 11 to 20. The `from` parameter can be used in conjunction with the `size` parameter to return a specific subset of results.

Example

```
1 curl 'https://gdc-api.nci.nih.gov/files?fields=file_name&from=0&size=2&pretty=true'
```

```
1 import requests
2 import json
3
```

```
4 files_endpt = 'https://gdc-api.nci.nih.gov/files'
5 params = {'fields':'file_name',
6           'from':0, 'size':2}
7 response = requests.get(files_endpt, params = params)
8 print json.dumps(response.json(), indent=2)
```

```
1 {
2   "data": {
3     "hits": [
4       {
5         "file_name":
              "unc.edu.276a1e00-cf3a-4463-a97b-d544381219ea.2363081.rsem.isoforms.normalized_results"
6       },
7       {
8         "file_name": "nationwidechildrens.org_clinical.TCGA-EY-A5W2.xml"
9       }
10    ],
11    "pagination": {
12      "count": 2,
13      "sort": "",
14      "from": 1,
15      "pages": 300936,
16      "total": 601872,
17      "page": 1,
18      "size": 2
19    }
20  },
21  "warnings": {}
22 }
```

```
1 curl 'https://gdc-api.nci.nih.gov/files?fields=file_name&from=101&size=5&pretty=true'
```

```
1 import requests
2 import json
3
4 files_endpt = 'https://gdc-api.nci.nih.gov/files'
5 params = {'fields':'file_name',
6           'from':101, 'size':5}
7 response = requests.get(files_endpt, params = params)
8 print json.dumps(response.json(), indent=2)
```

```
1 {
2   "data": {
3     "hits": [
4       {
5         "file_name": "unc.edu.956590e9-4962-497b-a59f-81ee0a1c0caf.1379536.junction_quantification.txt"
6       },
7       {
8         "file_name": "MATZO_p_TCGAb40_SNP_1N_GenomeWideSNP_6_G09_667760.ismpolish.data.txt"
9       },
10      {
11        "file_name": "GIRTH_p_TCGA_b108_137_SNP_N_GenomeWideSNP_6_D06_787864.hg18.seg.txt"
12      },
13      {
14        "file_name": "PLENA_p_TCGAb63and64_SNP_N_GenomeWideSNP_6_B12_697382.CEL"
15      },
16      {
17        "file_name": "TCGA-HU-8604-01A-11R-2402-13.isoform.quantification.txt"
```

```
18          }
19      ],
20      "pagination": {
21        "count": 5,
22        "sort": "",
23        "from": 100,
24        "pages": 109553,
25        "total": 547761,
26        "page": 21,
27        "size": 5
28      }
29    },
30    "warnings": {}
31 }
```

Sort

The `sort` query parameter sorts the results by a specific field, and with the sort direction specified using the `:asc` (ascending) or `:desc` (descending) prefix, e.g. `sort=field:desc`. A list of all valid *field* names is available in Appendix A; the list can also be accessed programmatically at the _mapping endpoint.

Example

Sort cases by `submitter_id` in ascending order:

```
1 curl  'https://gdc-api.nci.nih.gov/cases?fields=submitter_id&sort=submitter_id:asc&pretty=true'
```

```
1 import requests
2 import json
3
4 cases_endpt = 'https://gdc-api.nci.nih.gov/cases'
5 params = {'fields':'submitter_id',
6           'sort':'submitter_id:asc'}
7 response = requests.get(cases_endpt, params = params)
8 print json.dumps(response.json(), indent=2)
```

```
1 {
2   "data": {
3     "hits": [
4       {
5         "submitter_id": "TARGET-20-PABGKN"
6       },
7       {
8         "submitter_id": "TARGET-20-PABHET"
9       },
10      {
11        "submitter_id": "TARGET-20-PABHKY"
12      },
13      {
14        "submitter_id": "TARGET-20-PABLDZ"
15      },
16      {
17        "submitter_id": "TARGET-20-PACDZR"
18      },
19      {
20        "submitter_id": "TARGET-20-PACEGD"
21      },
22      {
```

```
23         "submitter_id": "TARGET-20-PADDXZ"
24       },
25       {
26         "submitter_id": "TARGET-20-PADYIR"
27       },
28       {
29         "submitter_id": "TARGET-20-PADZCG"
30       },
31       {
32         "submitter_id": "TARGET-20-PADZKD"
33       }
34     ],
35     "pagination": {
36       "count": 10,
37       "sort": "submitter_id.raw:asc",
38       "from": 1,
39       "pages": 1406,
40       "total": 14052,
41       "page": 1,
42       "size": 10
43     }
44   },
45   "warnings": {}
46 }
```

Facets

The `facets` parameter provides aggregate information for a specified field. It provides all values that exist for that field, and the number of entities (cases, projects, files, or annotations) that this value. The primary intended use of this parameter is for displaying aggregate information in the GDC Data Portal.

The `facets` parameter can be used in conjunction with the `filters` parameter to get aggregate information for a set of search results. The following limitations apply when using `facets` and `filters` together:

1. The `filters` object's top level operator must be `and`, and the internal filters must be limited to: `=`, `!=`, `in`, `exclude`, `is`, and `not`.
2. The information provided by `facets` for a given field will disregard any filters applied to that same field.

Example

This is an example of a request for a count of projects in each program.

```
1 curl
      'https://gdc-api.nci.nih.gov/projects?facets=program.name&from=1&size=0&sort=program.name:asc&pretty=true'
```

```python
1 import requests
2 import json
3
4 projects_endpt = 'https://gdc-api.nci.nih.gov/projects'
5 params = {'facets':'program.name',
6           'from':1, 'size':0,
7           'sort':'program.name:asc'}
8 response = requests.get(projects_endpt, params = params)
9 print json.dumps(response.json(), indent=2)
```

```
 1
 2      {
 3        "data": {
 4          "pagination": {
 5            "count": 0,
 6            "sort": "program.name:asc",
 7            "from": 1,
 8            "pages": 46,
 9            "total": 46,
10            "page": 1,
11            "size": 0
12          },
13          "hits": [],
14          "aggregations": {
15            "program.name": {
16              "buckets": [
17                {
18                  "key": "TCGA",
19                  "doc_count": 37
20                },
21                {
22                  "key": "TARGET",
23                  "doc_count": 9
24                }
25              ]
26            }
27          }
28        },
29        "warnings": {}
30      }
```

Example

In this sample POST request, both `filters` and `facets` parameters are used. Note that `facets` ignores the `primary_site` filter.

```
 1 {
 2      "filters":{
 3          "op":"and",
 4          "content":[
 5              {
 6                  "op":"=",
 7                  "content":{
 8                      "field":"cases.project.primary_site",
 9                      "value":"Kidney"
10                  }
11              },
12              {
13                  "op":"=",
14                  "content":{
15                      "field":"project.program.name",
16                      "value":"TCGA"
17                  }
18              }
19          ]
20      },
21      "size":"0",
22      "facets":"project.primary_site",
```

```
23      "pretty":"true"
24 }
```

```
1 curl --request POST --header "Content-Type: application/json" --data @Payload
      'https://gdc-api.nci.nih.gov/v0/cases'
```

```
 1 {
 2   "data": {
 3     "pagination": {
 4       "count": 0,
 5       "sort": "",
 6       "from": 1,
 7       "page": 1,
 8       "total": 941,
 9       "pages": 941,
10       "size": 0
11     },
12     "hits": [],
13     "aggregations": {
14       "project.primary_site": {
15         "buckets": [
16           {
17             "key": "Brain",
18             "doc_count": 1133
19           },
20           {
21             "key": "Breast",
22             "doc_count": 1098
23           },
24           {
25             "key": "Lung",
26             "doc_count": 1089
27           },
28           {
29             "key": "Kidney",
30             "doc_count": 941
31           },
32           {
33             "key": "Colorectal",
34             "doc_count": 635
35           },
36           {
37             "key": "Uterus",
38             "doc_count": 617
39           },
40           {
41             "key": "Ovary",
42             "doc_count": 608
43           },
44           {
45             "key": "Head and Neck",
46             "doc_count": 528
47           },
48           {
49             "key": "Thyroid",
50             "doc_count": 507
51           },
52           {
```

```
53            "key": "Prostate",
54            "doc_count": 500
55          },
56          {
57            "key": "Stomach",
58            "doc_count": 478
59          },
60          {
61            "key": "Skin",
62            "doc_count": 470
63          },
64          {
65            "key": "Bladder",
66            "doc_count": 412
67          },
68          {
69            "key": "Liver",
70            "doc_count": 377
71          },
72          {
73            "key": "Cervix",
74            "doc_count": 308
75          },
76          {
77            "key": "Adrenal Gland",
78            "doc_count": 271
79          },
80          {
81            "key": "Soft Tissue",
82            "doc_count": 261
83          },
84          {
85            "key": "Bone Marrow",
86            "doc_count": 200
87          },
88          {
89            "key": "Esophagus",
90            "doc_count": 185
91          },
92          {
93            "key": "Pancreas",
94            "doc_count": 185
95          },
96          {
97            "key": "Testis",
98            "doc_count": 150
99          },
100         {
101           "key": "Thymus",
102           "doc_count": 124
103         },
104         {
105           "key": "Pleura",
106           "doc_count": 87
107         },
108         {
109           "key": "Eye",
110           "doc_count": 80
```

```
111              },
112              {
113                "key": "Lymph Nodes",
114                "doc_count": 58
115              },
116              {
117                "key": "Bile Duct",
118                "doc_count": 51
119              }
120            ]
121          }
122        }
123      },
124      "warnings": {}
125    }
```

Alternative Request Format

The GDC API also supports POST requests with `Content-Type: application/x-www-form-urlencoded` (curl default), which require payloads in the following format:

```
1 filters=%7B%0A%20%20%20%20%22op%22%3A%22in%22%2C%0A%20%20%20%20%22content%22%3A%7B%0A%20%20%20%20%20%20%20%20%22
```

Additional Examples

More examples of API functionality described in this section are provided in Additional Examples.

Chapter 3

Downloading Files

Downloading Files

The GDC API implements file download functionality using `data` and `manifest` endpoints. The `data` endpoint allows users to download files stored in the GDC by specifying file UUID(s). The `manifest` endpoint generates a download manifest file that can be used with the GDC Data Transfer Tool to transfer large volumes of data.

Note: Downloading controlled access data requires the use of an authentication token. See Getting Started: Authentication for details.

Note: Requests to download data from the GDC Legacy Archive must be directed to `legacy/data`. See Getting Started: Legacy Archive for details.

Data endpoint

To download a file, users can pass UUID(s) to the `data` endpoint. If a single UUID is provided, the API will return the associated file. If a comma-separated list of UUIDs is provided, the API will return an archive file containing the requested files.

The `data` endpoint supports GET and POST requests as demonstrated in the following examples.

Related Files

If the `related_files=true` parameter is specified, the following related files, if available, will be included in the download package by the GDC API:

- BAM index files (BAI files)
- Metadata files (such as SRA XML or MAGE-TAB files)

For example, this request will download a legacy copy number segmentation file and its associated MAGE-TAB metadata file:

```
1 https://gdc-api.nci.nih.gov/legacy/data/7efc039a-fde3-4bc1-9433-2fc6b5e3ffa5?related_files=true
```

Downloading a Single File using GET

This example demonstrates downloading a single file from the GDC. Here we pass the file's UUID to the `data` endpoint with a GET request.

```
1 curl --remote-name --remote-header-name
    'https://gdc-api.nci.nih.gov/data/5b2974ad-f932-499b-90a3-93577a9f0573'
```

```
1   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
2                                  Dload  Upload   Total   Spent    Left  Speed
3 100 6111k  100 6111k    0     0   414k      0  0:00:14  0:00:14 --:--:--  412k
4 curl: Saved to filename '14-3-3_beta-R-V_GBL1112940.tif'
```

Downloading Multiple Files using GET

This example demonstrates downloading multiple files from the GDC using a GET request. The GDC API returns a `.tar.gz` archive containing the downloaded files.

```
1 curl --remote-name --remote-header-name
      'https://gdc-api.nci.nih.gov/data/e3228020-1c54-4521-9182-1ea14c5dc0f7,18e1e38e-0f0a-4a0e-918f-08e6201ea140'
```

```
1 % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
2                                Dload  Upload   Total   Spent    Left  Speed
3 100   287k    0   287k    0     0  30131      0 --:--:--  0:00:09 --:--:-- 42759
4 curl: Saved to filename 'gdc_download_064d1aa8cc8cbab33e93979bebbf7d6af2d6a802.tar.gz'
```

Note: This method supports downloading a limited number of files at one time. To download a large number of files, please use POST.

Downloading Multiple Files using POST

The following two examples demonstrate downloading multiple files from the GDC using a POST request that contains a payload in one of two formats: percent-encoded form data or JSON. The GDC API returns a `.tar.gz` archive containing the downloaded files.

As noted above, both the `data` and `legacy/data` endpoints accept POST requests.

POST request with form data payload

POST requests that carry a payload of percent-encoded form data must include the HTTP header `Content-Type: application/x-www-form-urlencoded`.

The payload is a string in the following format:

```
1 ids=UUID1&ids=UUID2&ids=UUID3...
```

where UUID corresponds to the UUIDs of the files to be downloaded.

In this example we use `curl` to download a set of files from the GDC Legacy Archive. The payload is stored in a plain text file named `Payload`; `curl` includes the `Content-Type: application/x-www-form-urlencoded` header by default.

```
1 ids=556e5e3f-0ab9-4b6c-aa62-c42f6a6cf20c&ids=e0de63e2-02f3-4309-9b24-69f4c24e85fc&ids=f1a06178-2ec2-4b06-83f3-3a
```

```
1 curl --remote-name --remote-header-name --request POST 'https://gdc-api.nci.nih.gov/legacy/data' --data
      @Payload
```

```
1   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
2                                  Dload  Upload   Total   Spent    Left  Speed
3 100 2562k    0 2561k  100   983   880k    337  0:00:02  0:00:02 --:--:--  880k
4 curl: Saved to filename 'gdc_download_20160701_011153.tar.gz'
```

POST request with JSON payload

POST requests that carry a JSON payload must include the HTTP header `Content-Type: application/json`.

The payload is a string in the following format:

```
1 {
2     "ids":[
3         "UUID1",
4         "UUID2",
5                 ...
6         "UUID3"
7     ]
8 }
```

where UUID corresponds to the UUIDs of the files to be downloaded.

In this example we use `curl` to download a set of files from the GDC Legacy Archive; the payload is stored in a plain text file named `Payload`.

```
1 {
2     "ids":[
3         "556e5e3f-0ab9-4b6c-aa62-c42f6a6cf20c",
4         "e0de63e2-02f3-4309-9b24-69f4c24e85fc",
5         "f1a06178-2ec2-4b06-83f3-3aedac332cfe",
6         "11a8aca0-c8e6-4ff8-8ab6-fe18a1b8ba82",
7         "69a69c84-00de-45ff-b397-fd2b6713ed4f",
8         "9ec48233-395d-401e-b205-951c971f8dd4",
9         "93129547-378c-4b69-b858-532abfff678e",
10        "8d4277e9-a472-4590-886d-24dc2538ea65",
11        "6733b412-56da-4f1c-a12b-ff804cb656d7",
12        "a72eec98-c5e0-4866-8953-765780acb6c1",
13        "e77b2294-1bdd-4fba-928a-d81d2622312f",
14        "965e01fc-318e-4c02-a801-d6fad60bfae4",
15        "21ad5409-fe0b-4728-97e4-15520b9fc287",
16        "1a777521-277c-4aeb-baf1-66871a7c2d2a",
17        "c13a3449-9e0d-45a9-bcc0-518f55e45c8a",
18        "5f2d329b-d59d-4112-b490-5114b830e34d",
19        "bb966617-6c1f-4bb0-a1ed-ceb37ecade67",
20        "05d11519-2b33-4742-aa87-3934632f2f2b",
21        "39bfafe2-9628-434e-bd72-148051a47477",
22        "481bea69-3cd5-45f3-8a52-2d4cc8fc8df7",
23        "f95e407b-de69-416c-920c-6be8c9414862",
24        "75940293-8fa6-47f9-ad5d-155b61933fdc",
25        "e8e84ccf-f8a8-4551-9257-ef731d02116f",
26        "e4991159-f088-4a2a-88b7-38d6ac47c6bc"
27    ]
28 }
```

```
1 curl --remote-name --remote-header-name --request POST --header 'Content-Type: application/json' --data
     @request.txt 'https://gdc-api.nci.nih.gov/legacy/data'
```

```
1   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
2                                  Dload  Upload   Total   Spent    Left  Speed
3 100 2562k    0 2561k  100  1145   788k    352  0:00:03  0:00:03 --:--:--  788k
4 curl: Saved to filename 'gdc_download_20160701_011007.tar.gz'
```

Downloading Controlled-access Files

To download controlled-access files, a valid authentication token must be passed to the GDC API using the `X-Auth-Token` HTTP header:

```
1 export
    token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3 curl --remote-name --remote-header-name --header "X-Auth-Token: $token"
    'https://gdc-api.nci.nih.gov/data/0eccf79d-1f1e-4205-910f-8e126b08276e'
```

```
1   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
2                                   Dload  Upload   Total   Spent    Left  Speed
3 100 31.4M  100 31.4M    0     0   290k      0  0:01:50  0:01:50 --:--:--  172k
4 curl: Saved to filename
    'ACOLD_p_TCGA_Batch17_SNP_N_GenomeWideSNP_6_A03_466078.tangent.copynumber.data.txt'
```

Manifest endpoint

The `manifest` endpoint generates a download manifest file that can be used with the GDC Data Transfer Tool. The Data Transfer Tool is recommended for transferring large volumes of data. The GDC API can also generate a download manifest from a list of results that match a Search and Retrieval query. To do this, append `&return_type=manifest` to the end of the query.

Using the manifest endpoint

The `manifest` endpoint allows users to create a download manifest, which can be used with the GDC Data Transfer Tool to download a large volume of data. The `manifest` endpoint generates a manifest file from a comma-separated list of UUIDs.

```
1 curl --remote-name --remote-header-name
    'https://gdc-api.nci.nih.gov/v0/manifest/ae9db773-78ab-48d0-972d-debe1bedd37d,3d815e6e-db97-419d-ad7f-dba4e4
```

```
1   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
2                                   Dload  Upload   Total   Spent    Left  Speed
3 100   274  100   274    0     0   1042      0 --:--:-- --:--:-- --:--:--  1041
4 curl: Saved to filename 'gdc_manifest_20160428_234614.txt'
```

The `manifest` endpoint also supports HTTP POST requests in the same format as the `data` endpoint; see above for details.

Using return_type=manifest

Alternatively, users can create a manifest by appending `&return_type=manifest` to a Search and Retrieval query. In this example, we generate a download manifest for RNA-seq data files from solid tissue normal samples, that are part of the TCGA-KIRC project:

```
1 curl --remote-name --remote-header-name
    'https://gdc-api.nci.nih.gov/files?filters=%7B%0A%20%20%20%20%22op%22%3A%22and%22%2C%0A%20%20%20%20%22conten
```

```
1   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
2                                   Dload  Upload   Total   Spent    Left  Speed
3 100 40663    0 40663    0     0  77109      0 --:--:-- --:--:-- --:--:-- 77306
4 curl: Saved to filename 'gdc_manifest.2016-06-28T13:26:33.850459.tsv'
```

Chapter 4

BAM Slicing

BAM Slicing

The GDC API provides remote BAM slicing functionality that enables downloading of specific parts of a BAM file instead of the whole file. This functionality can be accessed at the `slicing` endpoint, using a syntax similar to that of widely used bioinformatics tools such as `samtools`.

About the slicing endpoint

The `slicing` endpoint accepts HTTP GET requests in the form of a URL, and HTTP POST requests that carry a JSON payload. POST requests are more appropriate in cases where query parameters make the GET URL very long.

The response will be a BAM-formatted file containing the header of the source BAM file, as well as any alignment records that are found to overlap the specified regions, sorted by chromosomal coordinate.

Please note the following:

- The functionality of this API differs from the usual functionality of `samtools` in that alignment records that overlap multiple regions will not be returned multiple times.
- A request with no region or gene specified will return the BAM header, which makes it easy to inspect the references to which the alignment records were aligned.
- A request for regions that are not included in the source BAM is not considered an error, and is treated the same as if no records existed for the region.
- Examples provided for BAM slicing functionality are intended for use with GDC harmonized data (i.e. BAM files available in the GDC Data Portal). Slicing of unharmonized BAM files (i.e. BAM files in the GDC Legacy Portal) is not supported.
- Bam slicing does not create an associated bam index (.bai) file. For applications requiring a .bai file users will need to generate this file from the bam slice using a tool and command such as `samtools index`.

Query Parameters

The following query parameters and JSON fields are supported:

Description	Query Parameter	JSON Field	Query format
entire chromosome, or a position or region on the chromosome, specified using chromosomal coordinates	region	regions	region=(:(-)?)?
region specified using a HGNC / GENCODE v22 gene name	gencode	gencode	gencode=

NOTE: The successfully sliced BAM will contain all reads that overlap (entirely or partially) with the specified region or gene. It is possible to specify an open-ended region, e.g. `chr2:10000`, which would return all reads that (completely or partially) overlap with the region of chromosome 2 from position 10,000 to the end of the chromosome.

JSON Schema

JSON payloads can be syntactically verified using the following JSON schema:

```
1  {
2    "$schema": "http://json-schema.org/schema#",
3    "type": "object",
4    "properties": {
5      "regions": {
6        "type": "array",
7        "items": {
8          "type": "string",
9          "pattern": "^[a-zA-Z0-9]+(:([0-9]+)?(-[0-9]+)?)?$"
10       }
11     },
12     "gencode": {
13       "type": "array",
14       "items": {
15         "type": "string"
16       }
17     }
18   }
19 }
```

Examples: Specifying a region

The following two requests are examples of BAM slicing using region(s).

```
1  export
       token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3  curl --header "X-Auth-Token: $token"
       'https://gdc-api.nci.nih.gov/slicing/view/df80679e-c4d3-487b-934c-fcc782e5d46e?region=chr1&region=chr2:10000
       --output get_regions_slice.bam
```

```
1  {
2      "regions": [
3          "chr1",
4          "chr2:10000",
5          "chr3:10000-20000"
6      ]
7  }
```

```
1  export
       token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3  curl --header "X-Auth-Token: $token" --request POST
       https://gdc-api.nci.nih.gov/slicing/view/9ca90dfa-e62f-4f9c-9946-dfcecfd3ca4d --header "Content-Type:
       application/json" -d@Payload --output post_regions_slice.bam
```

```
1  Response:
2  HTTP/1.1 206
```

```
3
4 <bam_data_stream>
```

Examples: Specifying a gene

The following two requests are examples of BAM slicing using HGNC / GENCODE v22 gene name(s).

```
1 export
      token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3 curl --header "X-Auth-Token: $token"
      'https://gdc-api.nci.nih.gov/slicing/view/df80679e-c4d3-487b-934c-fcc782e5d46e?gencode=BRCA1'
      --output get_brca1_slice.bam
```

```
1 {
2     "gencode": [
3         "BRCA1",
4         "BRCA2"
5     ]
6 }
```

```
1 curl --header "X-Auth-Token: $token" --request POST
      https://gdc-api.nci.nih.gov/slicing/view/df80679e-c4d3-487b-934c-fcc782e5d46e --header "Content-Type:
      application/json" -d@Payload --output post_brca12_slice.bam
```

```
1 Response:
2 HTTP/1.1 206
3
4 <bam_data_stream>
```

After downloading, the sliced BAM file can be converted to SAM using the following command if `samtools` is installed on the user's system:

```
1 samtools view -h brca1_slice.bam -o brca1_slice.sam
```

Errors

When slicing cannot be performed, the GDC API will provide JSON error responses and HTTP error codes.

JSON Error Responses

JSON error responses have the following structure:

```
1 {
2     "error": "<error-message>"
3 }
```

For example, when making a request for a protected BAM without supplying a GDC authentication token:

```
1 curl https://gdc-api.nci.nih.gov/v0/slicing/view/15b0bf8e-ff20-41ab-8366-a495c11b30be
```

```
1 HTTP/1.1 403 FORBIDDEN
2 {
3     "error": "Please specify a X-Auth-Token"
4 }
```

HTTP error codes

Potential HTTP error codes include:

Error Code	Description
400	Bad Request – The regions specified are malformed
403	Unauthorized – The user could not be authenticated
403	Unauthorized – The user is not authorized for access to the source BAM
404	Not Found – No BAM is specified
404	Not Found – No BAM can be found for the specified GDC BAM ID
504	BAI Not Found – No BAI can be found for the BAM

Transfer Errors

In the case that an error occurs during transfer of the resulting BAM, the BGZF EOF marker will not be present. This early truncation of the BAM file will cause errors if the file is used as input to other programs. For example, `samtools` will provide the error "EOF marker is absent".

Early truncation can arise when connection is interrupted or when slicing fails due to BAM corruption.

Chapter 5

Submission

Submission

Overview

The GDC Submission API uses methods and endpoints that are distinct from those that drive the functionality of the GDC Data Portal. In particular, data and metadata that are in the process of being submitted can only be queried using GraphQL.

This section describes the GDC API's submission functionality, including methods for submitting, deleting, updating, searching, and retrieving data and metadata.

Submission endpoint

Constructing the endpoint URL

The endpoint for submitting data to a specific project in the GDC is constructed as follows:

```
1  https://gdc-api.nci.nih.gov/[API_version/]submission/Program.name/Project.code
```

where **API_version/]** is the optional API version component (see [Getting Started).

The values of `Program.name` and `Project.code` can be obtained from the project URL on the GDC Data Submission Portal:

```
1  https://gdc-portal.nci.nih.gov/submission/Program.name/Project.code/dashboard
```

For more information about program name and project code see The GDC Data Model section.

Example

The following are URL examples for a project with `Program.name` "TCGA" and `Project.code` "ALCH":

- Submission Portal URL: `https://gdc-portal.nci.nih.gov/submission/TCGA/ALCH/dashboard`
- API submission endpoint (versioned): `https://gdc-api.nci.nih.gov/v0/submission/TCGA/ALCH`
- API submission endpoint (unversioned): `https://gdc-api.nci.nih.gov/submission/TCGA/ALCH`

Submission Formats

Metadata Formats

JSON and TSV

The GDC API accepts project metadata in JSON and TSV formats for the purpose of creating entities in the GDC Data Model. This includes clinical and biospecimen metadata such as disease name and stage, patient age, sample type, and certain details about the types of data collected. Upon successful data submission and project release, this metadata is indexed and becomes available for queries by data users via the GDC Data Portal and the GDC API. See GDC Data Model (below) for information on accepted metadata elements and instructions for obtaining templates for metadata submission.

Content-Type Header   JSON is the default format for metadata submission. Submission API calls with JSON payloads should include the HTTP header `Content-Type: application/json`. Requests with TSV payloads must instead include the header `Content-Type: text/tsv`.

Binary Mode   Metadata files must be uploaded in raw, unencoded form. Binary mode should be used, if available, to ensure that file contents are not encoded by the upload tool before transmission. For example, when using the `curl` command-line tool, the `--data-binary` switch should be used instead of `--data`. The `--data-binary` switch is required for uploading TSV files.

BCR XML

While JSON and TSV are the recommended formats for submitting metadata, the GDC API can also extract metadata elements from BCR XML files. Users wishing to submit metadata as BCR XML must contact GDC User Services and ensure that appropriate element mapping is in place before initiating XML submission.

To submit BCR XML, make `PUT` requests with the `Content-Type: application/xml` header to the following URLs, replacing Program.name and Project.code as desribed in Submission Endpoint (above):

0. For Biospecimen BCR XML: `https://gdc-api.nci.nih.gov/v0/submission/Program.name/Project.code/xml/biospecimen`
1. For Clinical BCR XML: `https://gdc-api.nci.nih.gov/v0/submission/Program.name/Project.code/xml/clinical/bcr/`

Biospecimen BCR XML creates Case entities in the GDC Data Model, whereas Clinical BCR XML does not. Unless the associated cases already exist in the GDC, Biospecimen BCR XML must be uploaded before Clinical BCR XML.

BCR XML files can be submitted in dry run mode, described below, by appending `_dry_run` to the above URLs.

The following is a sample shell command for submitting an XML file:

```
1 curl --request PUT --header "X-Auth-Token: $token"  --header 'Content-Type: application/xml'
    --data-binary @biospecimen.xml
    'https://gdc-api.nci.nih.gov/v0/submission/GDC/INTERNAL/xml/biospecimen/bcr/_dry_run'
```

NOTE: A typical BCR XML file contains more information than what is extracted and indexed by the GDC. XML files submitted to the above endpoints are not retained or distributed to GDC data users, so the same files should also be submitted as data files (i.e. as clinical or biospecimen supplements).

Data File Formats

The GDC API accepts a variety of data files after their metadata has been registered: BAM and FASTQ files, clinical and biospecimen supplements, slide images, and other file types. Supported data file formats are listed on the GDC website.

GDC Data Model

Submitters should review the GDC Data Model documentation and the GDC Data Dictionary before initiating submission.

UUIDs

Submitters can assign UUIDs to all submittable entities other than those that correspond to files (entities in categories `data_file` or `metadata_file`). If the submitter does not provide a UUID, it will be assigned by the GDC and returned in the API response upon successful completion of the submission transaction. See Appendix C for details of the API response format. To learn more about UUIDs see the GDC Data Model documentation.

Submitter IDs

In addition to `id`, many entities also include a `submitter_id` field. This field can contain any string (e.g. a "barcode") that the submitter wishes to use to identify the entity. Typically this string identifies a corresponding entry in submitter's records. The GDC's only requirement with respect to `submitter_id` is that it be a string that is unique for all entities within a project. The GDC Submission API requires a `submitter_id` for most entities.

Note: For `case` entities, `submitter_id` must correspond to a `submitted_subject_id` of a study participant registered with the project in dbGaP.

GDC Data Dictionary Endpoints

Information in the GDC Data Dictionary can be accessed programmatically as described below.

Submission Templates

Submission templates are accessible programmatically at the `templates` endpoint. Template format (`json`, `tsv` or `csv`) is specified using the `format` parameter.

For example, the JSON template for `case` entities can be obtained from:

```
1 https://gdc-api.nci.nih.gov/v0/submission/template/case?format=json
```

A set of templates for all entities in the GDC Data Model can be downloaded from:

```
1 https://gdc-api.nci.nih.gov/v0/submission/template/?format=json
```

Entity JSON Schemas

The entire collection of GDC entity schemas can be downloaded from the `dictionary` endpoint:

```
1 https://gdc-api.nci.nih.gov/v0/submission/_dictionary/_all
```

Individual schemas can be downloaded by specifying entity type. For example, the JSON schema for `case` entities can be found at:

```
1 https://gdc-api.nci.nih.gov/v0/submission/_dictionary/case
```

Making Requests to the Submission API

Requests to create or update entities in the GDC must specify the entity `type`, the entity `id` or `submitter_id`, relationships (links) that the entity has to existing entities in the GDC Data Model, and entity properties as defined by the GDC Data Dictionary. To delete entities, only the `id` property is required. The general format of GDC API submission requests and responses is provided in Appendix C.

Submission Transactions

Submission of data to the GDC involves a series of transactions initiated by the submitter, that create and link entities according to the GDC Data Model. With the exception of `program`, which is an administrative entity created by the GDC, all new entities must be linked, at creation, to existing entities or to new entities being created in the same transaction. For example, a submitter cannot create a `portion` entity unless the submitter either (1) has previously created the corresponding `case` and `sample` entities, or (2) is creating those entities in the same transaction. This also means that entities cannot be deleted if they have "child" entities attached to them.

If multiple entities are being created and/or updated in a transaction, and an error is encountered for one of the entities, then the transaction will fail and no changes will be made to the GDC.

Dry Run Transactions

The `submission` endpoint provides a `_dry_run` mode that simulates submission transactions without making changes to the GDC. This mode is activated by appending `/_dry_run` to the end of a submission endpoint.

The following is an example of a POST request, that simulates creating an entity in dry run mode:

```
1 {
2    "project_id": "TCGA-ALCH",
3    "type": "case",
4    "submitter_id": "TCGA-ALCH-000001",
5    "projects": {
6      "code": "ALCH"
7    }
8 }
```

```
1 export
    token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3 curl --header "X-Auth-Token: $token" --request POST --data-binary @Request --header 'Content-Type:
    application/json' https://gdc-api.nci.nih.gov/v0/submission/TCGA/ALCH/_dry_run
```

```
1 {
2    "cases_related_to_created_entities_count": 0,
3    "cases_related_to_updated_entities_count": 0,
4    "code": 200,
5    "created_entity_count": 1,
6    "entities": [
7      {
8        "action": "create",
9        "errors": [],
10       "id": "61f48d1c-9439-448c-a90c-d6dbe76b3654",
11       "related_cases": [],
12       "type": "case",
13       "unique_keys": [
14         {
15           "project_id": "TCGA-ALCH",
16           "submitter_id": "TCGA-ALCH-000001"
17         }
18       ],
19       "valid": true,
20       "warnings": []
21     }
22   ],
23   "entity_error_count": 0,
24   "message": "Transaction would have been successful. User selected dry run option, transaction aborted,
        no data written to database.",
```

```
25    "success": true,
26    "transaction_id": null,
27    "transactional_error_count": 0,
28    "transactional_errors": [],
29    "updated_entity_count": 0
30 }
```

Dry Run Commit

For convenience, the GDC enables users to commit earlier `_dry_run` transactions instead of uploading the same data again to execute the changes. This `commit` action is allowed on transactions that (1) have not been previously committed and (2) were successful `dry_run` transactions.

Note that the `commit` action is a separate transaction with its own transaction id, and it can be executed asynchronously. If the state of the submission project has changed in a way that would make the original `_dry_run` transaction invalid if it were run again (e.g. entities with the same `submitter_id` have since been created in another transaction), then then `commit` action will fail.

To commit a transaction, submit a POST or PUT request to `/submission/Program.name/Project.code/transactions/transaction_`, replacing `Program.name`, `Project.code`, and `transaction_id` with values associated with the transaction.

```
1 export
      token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3 curl --header "X-Auth-Token: $token" --request POST
      https://gdc-api.nci.nih.gov/v0/submission/TCGA/ALCH/transactions/467/commit?async=true
```

```
1 {
2   "code": 200,
3   "message": "Transaction submitted.",
4   "transaction_id": 468,
5 }
```

Dry Run Close

The GDC Submission API also provides a `close` action on `_dry_run` transactions. This `close` action is allowed on `_dry_run` transactions that have not been previously closed. Closing a `_dry_run` transaction prevents it from being committed in the future.

To close a transaction, submit a POST or PUT request to `/submission/Program.name/Project.code/transactions/transaction_id/`, replacing `Program.name`, `Project.code`, and `transaction_id` with values associated with the transaction.

```
1 export
      token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3 curl --header "X-Auth-Token: $token" --request POST
      https://gdc-api.nci.nih.gov/v0/submission/TCGA/ALCH/transactions/467/close
```

```
1 {
2     "code": 200,
3     "message": "Closed transaction.",
4     "transaction_id": 467
5 }
```

Asynchronous Transactions

The `submission` endpoint provides an asynchronous mode that provides immediate response and executes submission transactions in the background. This mode is activated by appending `?async=true` to the end of a submission endpoint. The API will respond with the `transaction_id` which can be used to look up the result of the transaction at a later time via the GraphQL endpoint. If the server has too many asynchronous jobs scheduled already, your request to schedule a transaction may fail.

Example

The following is an example of a PUT request, that creates a case asynchronously:

```
1 {
2   "project_id": "TCGA-ALCH",
3   "type": "case",
4   "submitter_id": "TCGA-ALCH-000001",
5   "projects": {
6     "code": "ALCH"
7   }
8 }
```

```
1 export
      token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3 curl --header "X-Auth-Token: $token" --request POST --data-binary @Request --header 'Content-Type:
      application/json' https://gdc-api.nci.nih.gov/v0/submission/TCGA/ALCH?async=true
```

```
1 {
2   "code": 200,
3   "message": "Transaction submitted.",
4   "transaction_id": 467,
5 }
```

The following is a GraphQL request that looks up the state of the above transaction:

```
1 query {
2   transaction_log(id: 467) {
3     is_dry_run
4     committed_by
5     state
6   }
7 }
```

```
1 {
2   "data": {
3     "transaction_log": [
4       {
5         "committed_by": null,
6         "is_dry_run": false,
7         "state": "FAILED"
8       }
9     ]
10   }
11 }
```

Transaction Status

The following transaction fields can be queried using GraphQL and are helpful in determining the status of a transaction:

Field	Type	Description
id	ID	Transaction identifier
is_dry_run	Boolean	Indicates whether the transaction is a dry run
closed	Boolean	For dry run transactions, indicates whether the transaction has been closed to prevent it from being committed in the future.

Field	Type	Description
committable	Boolean	Indicates whether the transaction can be committed (i.e. it is a successful dry run transaction that has not been committed previously and has not been closed)
state	String	Indicates the state of the transaction: PENDING, SUCCEEDED, FAILED (due to user error), or ERRORED (due to system error)
committed_by	ID	The ID of the transaction that committed this transaction

Note: To check whether a dry run transaction was committed successfully, check the state of the transaction that executed the commit. The state of the dry run transaction itself does not represent the status of a subsequent commit.

Creating and Updating Entities

The GDC Submission API supports HTTP POST and HTTP PUT methods for creating entities:

- **POST** will create entities that do not exist, and will fail if any of the entities in the transaction already exist in the GDC.

- **PUT** will create new entities and update existing entities, and identify which entities were created or updated in the API response.

The GDC suggests using POST for creating new entities, and using PUT only for updating entities. This helps to avoid inadvertent entity updates that can occur when using PUT for creating entities.

Note: Once a relationship has been created between two entities, it cannot be removed by updating an entity. To remove a relationship, the child entity must be deleted.

Example: Creating and Updating Case Entities (JSON)

In this example, a case entity is created using POST. Then an attempt is made to create the same entity again using POST, resulting in an error. Then the originally created entity is updated (with the same information) using PUT.

The JSON in the request was generated using the case JSON template that can be obtained from the GDC Data Dictionary Viewer and from `https://gdc-api.nci.nih.gov/v0/submission/template/case?format=json`.

Note: For case entities, submitter_id must correspond to a submitted_subject_id of a study participant registered with the project in dbGaP.

```
1 {
2   "type": "case",
3   "submitter_id": "TCGA-ALCH-000001",
4   "projects": {
5     "code": "ALCH"
6   }
7
8 }
```

```
1 export
     token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3 curl --header "X-Auth-Token: $token" --request POST --data-binary @Request --header 'Content-Type:
     application/json' https://gdc-api.nci.nih.gov/v0/submission/TCGA/ALCH
```

```
1  {
2    "cases_related_to_created_entities_count": 0,
3    "cases_related_to_updated_entities_count": 0,
4    "code": 201,
5    "created_entity_count": 1,
6    "entities": [
7      {
8        "action": "create",
9        "errors": [],
10        "id": "fbf69646-5904-4f95-92d6-692bde658f05",
11        "related_cases": [],
12        "type": "case",
13        "unique_keys": [
14          {
15            "project_id": "TCGA-ALCH",
16            "submitter_id": "TCGA-ALCH-000001"
17          }
18        ],
19        "valid": true,
20        "warnings": []
21      }
22    ],
23    "entity_error_count": 0,
24    "message": "Transaction successful.",
25    "success": true,
26    "transaction_id": 215,
27    "transactional_error_count": 0,
28    "transactional_errors": [],
29    "updated_entity_count": 0
30  }
```

```
1  curl --header "X-Auth-Token: $token" --request POST --data-binary @Request --header 'Content-Type:
      application/json' https://gdc-api.nci.nih.gov/v0/submission/TCGA/ALCH
```

```
1  {
2    "cases_related_to_created_entities_count": 0,
3    "cases_related_to_updated_entities_count": 0,
4    "code": 400,
5    "created_entity_count": 0,
6    "entities": [
7      {
8        "action": null,
9        "errors": [
10          {
11            "keys": [
12              "id"
13            ],
14            "message": "Cannot create entity that already exists. Try updating entity (PUT instead of
                POST)",
15            "type": "NOT_UNIQUE"
16          }
17        ],
18        "id": null,
19        "related_cases": [],
20        "type": "case",
21        "unique_keys": [
22          {
23            "project_id": "TCGA-ALCH",
```

```
24          "submitter_id": "TCGA-ALCH-000001"
25        }
26      ],
27      "valid": false,
28      "warnings": []
29    }
30  ],
31  "entity_error_count": 1,
32  "message": "Transaction aborted due to 1 invalid entity.",
33  "success": false,
34  "transaction_id": null,
35  "transactional_error_count": 0,
36  "transactional_errors": [],
37  "updated_entity_count": 0
38 }
```

```
1 curl --header "X-Auth-Token: $token" --request PUT --data-binary @Request --header 'Content-Type:
     application/json' https://gdc-api.nci.nih.gov/v0/submission/TCGA/ALCH
```

```
1 {
2   "cases_related_to_created_entities_count": 0,
3   "cases_related_to_updated_entities_count": 0,
4   "code": 200,
5   "created_entity_count": 0,
6   "entities": [
7     {
8       "action": "update",
9       "errors": [],
10      "id": "fbf69646-5904-4f95-92d6-692bde658f05",
11      "related_cases": [],
12      "type": "case",
13      "unique_keys": [
14        {
15          "project_id": "TCGA-ALCH",
16          "submitter_id": "TCGA-ALCH-000001"
17        }
18      ],
19      "valid": true,
20      "warnings": []
21    }
22  ],
23  "entity_error_count": 0,
24  "message": "Transaction successful.",
25  "success": true,
26  "transaction_id": 216,
27  "transactional_error_count": 0,
28  "transactional_errors": [],
29  "updated_entity_count": 1
30 }
```

Example: Creating an Aliquot Entity (JSON)

In this example, an `aliquot` entity and a `sample` entity are created in a single transaction. The `aliquot` is linked to `sample` which is linked to `case`. The first request is an example of using `submitter_id` properties to link entities together. The second request is an example of using UUIDs for creating the links.

Request 1: Creating Links Using submitter_id

```
 1 [
 2   {
 3     "type": "sample",
 4     "submitter_id": "TCGA-ALCH-000001-SAMPLE000001",
 5     "sample_type": "Primary Tumor",
 6     "sample_type_id": "01",
 7     "cases": {
 8       "submitter_id": "TCGA-ALCH-000001"
 9     }
10   },
11   {
12     "type": "aliquot",
13     "submitter_id": "TCGA-ALCH-000001-SAMPLE000001-ALIQUOT000001",
14     "samples": {
15       "submitter_id": "TCGA-ALCH-000001-SAMPLE000001"
16     }
17   }
18 ]
```

```
 1 export
       token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
 2
 3 curl --header "X-Auth-Token: $token" --request POST --data-binary @Request --header 'Content-Type:
       application/json' https://gdc-api.nci.nih.gov/v0/submission/TCGA/ALCH
```

```
 1 {
 2   "cases_related_to_created_entities_count": 1,
 3   "cases_related_to_updated_entities_count": 0,
 4   "code": 201,
 5   "created_entity_count": 2,
 6   "entities": [
 7     {
 8       "action": "create",
 9       "errors": [],
10       "id": "48270338-6464-448f-bbef-b09d4f80b11b",
11       "related_cases": [
12         {
13           "id": "fbf69646-5904-4f95-92d6-692bde658f05",
14           "submitter_id": "TCGA-ALCH-000001"
15         }
16       ],
17       "type": "sample",
18       "unique_keys": [
19         {
20           "project_id": "TCGA-ALCH",
21           "submitter_id": "TCGA-ALCH-000001-SAMPLE000001"
22         }
23       ],
24       "valid": true,
25       "warnings": []
26     },
27     {
28       "action": "create",
29       "errors": [],
30       "id": "7af58da0-cb3e-43e2-a074-4bd8f27565ba",
31       "related_cases": [
```

```
32        {
33           "id": "fbf69646-5904-4f95-92d6-692bde658f05",
34           "submitter_id": "TCGA-ALCH-000001"
35        }
36      ],
37      "type": "aliquot",
38      "unique_keys": [
39        {
40           "project_id": "TCGA-ALCH",
41           "submitter_id": "TCGA-ALCH-000001-SAMPLE000001-ALIQUOT000001"
42        }
43      ],
44      "valid": true,
45      "warnings": []
46    }
47  ],
48  "entity_error_count": 0,
49  "message": "Transaction successful.",
50  "success": true,
51  "transaction_id": 222,
52  "transactional_error_count": 0,
53  "transactional_errors": [],
54  "updated_entity_count": 0
55 }
```

Request 2: Creating Links Using UUID

```
1 [
2   {
3     "type": "sample",
4     "submitter_id": "TCGA-ALCH-000001-SAMPLE000001",
5     "id": "2aa7a07b-e706-4eef-aeba-b849972423a0",
6     "sample_type": "Primary Tumor",
7     "sample_type_id": "01",
8     "cases": {
9       "id": "fbf69646-5904-4f95-92d6-692bde658f05"
10    }
11  },
12  {
13    "type": "aliquot",
14    "submitter_id": "TCGA-ALCH-000001-SAMPLE000001-ALIQUOT000001",
15    "samples": {
16      "id": "2aa7a07b-e706-4eef-aeba-b849972423a0"
17    }
18  }
19 ]
```

```
1 export
     token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3 curl --header "X-Auth-Token: $token" --request POST --data-binary @Request --header 'Content-Type:
     application/json' https://gdc-api.nci.nih.gov/v0/submission/TCGA/ALCH
```

```
1 {
2   "cases_related_to_created_entities_count": 1,
3   "cases_related_to_updated_entities_count": 0,
4   "code": 201,
5   "created_entity_count": 2,
```

```
 6    "entities": [
 7      {
 8        "action": "create",
 9        "errors": [],
10        "id": "2aa7a07b-e706-4eef-aeba-b849972423a0",
11        "related_cases": [
12          {
13            "id": "fbf69646-5904-4f95-92d6-692bde658f05",
14            "submitter_id": "TCGA-ALCH-000001"
15          }
16        ],
17        "type": "sample",
18        "unique_keys": [
19          {
20            "project_id": "TCGA-ALCH",
21            "submitter_id": "TCGA-ALCH-000001-SAMPLE000001"
22          }
23        ],
24        "valid": true,
25        "warnings": []
26      },
27      {
28        "action": "create",
29        "errors": [],
30        "id": "545096d5-ce1c-433f-80f0-fd0b04b56cb6",
31        "related_cases": [
32          {
33            "id": "fbf69646-5904-4f95-92d6-692bde658f05",
34            "submitter_id": "TCGA-ALCH-000001"
35          }
36        ],
37        "type": "aliquot",
38        "unique_keys": [
39          {
40            "project_id": "TCGA-ALCH",
41            "submitter_id": "TCGA-ALCH-000001-SAMPLE000001-ALIQUOT000001"
42          }
43        ],
44        "valid": true,
45        "warnings": []
46      }
47    ],
48    "entity_error_count": 0,
49    "message": "Transaction successful.",
50    "success": true,
51    "transaction_id": 219,
52    "transactional_error_count": 0,
53    "transactional_errors": [],
54    "updated_entity_count": 0
55 }
```

Example: Creating Two Samples (TSV)

In this example, a TSV file containing metadata for two samples is uploaded to the GDC in dry run mode.

```
1 type    project_id  submitter_id      cases.submitter_id  sample_type sample_type_id  tumor_descriptor
2 sample  GDC-INTERNAL    GDC-INTERNAL-000022-sampleA GDC-INTERNAL-000022 Additional Metastatic   01
3 sample  GDC-INTERNAL    GDC-INTERNAL-000022-sampleB GDC-INTERNAL-000022 Solid Tissue Normal 02
```

```
 1 curl --header "X-Auth-Token: $token" --header 'Content-Type: text/tsv' --request PUT --data-binary
     @Samples.tsv 'https://gdc-api.nci.nih.gov/submission/GDC/INTERNAL/_dry_run'
```

```
 1 {
 2   "cases_related_to_created_entities_count": 1,
 3   "cases_related_to_updated_entities_count": 0,
 4   "code": 200,
 5   "created_entity_count": 2,
 6   "entities": [
 7     {
 8       "action": "create",
 9       "errors": [],
10       "id": "b55e10af-5b7f-48f1-b230-0f8e6b7a7afe",
11       "related_cases": [
12         {
13           "id": "6e2e3b31-c5d2-45df-a911-eb3577640b70",
14           "submitter_id": "GDC-INTERNAL-000022"
15         }
16       ],
17       "type": "sample",
18       "unique_keys": [
19         {
20           "project_id": "GDC-INTERNAL",
21           "submitter_id": "GDC-INTERNAL-000022-sampleA"
22         }
23       ],
24       "valid": true,
25       "warnings": []
26     },
27     {
28       "action": "create",
29       "errors": [],
30       "id": "15076660-fccc-4406-b981-c745eb992034",
31       "related_cases": [
32         {
33           "id": "6e2e3b31-c5d2-45df-a911-eb3577640b70",
34           "submitter_id": "GDC-INTERNAL-000022"
35         }
36       ],
37       "type": "sample",
38       "unique_keys": [
39         {
40           "project_id": "GDC-INTERNAL",
41           "submitter_id": "GDC-INTERNAL-000022-sampleB"
42         }
43       ],
44       "valid": true,
45       "warnings": []
46     }
47   ],
48   "entity_error_count": 0,
49   "message": "Transaction would have been successful. User selected dry run option, transaction aborted,
       no data written to database.",
50   "success": true,
51   "transaction_id": 51284,
52   "transactional_error_count": 0,
53   "transactional_errors": [],
54   "updated_entity_count": 0
```

```
55 }
```

Retrieving Entities

Entities Endpoint

JSON objects representing submitted entities can be retrieved with a GET request to the `entities` endpoint. This endpoint retrieves entities by UUID. A single UUID or a comma-separated list of UUIDs can be passed to this endpoint as a query.

```
1 export
      token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3 curl --header "X-Auth-Token: $token"
      https://gdc-api.nci.nih.gov/v0/submission/TCGA/ALCH/entities/fbf69646-5904-4f95-92d6-692bde658f05
```

```
1 {
2   "entities": [
3     {
4       "program": "TCGA",
5       "project": "ALCH",
6       "properties": {
7         "created_datetime": "2016-04-14T08:44:43.361800-05:00",
8         "id": "fbf69646-5904-4f95-92d6-692bde658f05",
9         "project_id": "TCGA-ALCH",
10        "projects": [
11          {
12            "id": "d9906779-f1da-5d9f-9caa-6d5ecb2e3cd6",
13            "submitter_id": null
14          }
15        ],
16        "state": "validated",
17        "submitter_id": "TCGA-ALCH-000001",
18        "type": "case",
19        "updated_datetime": "2016-04-14T21:29:28.401212-05:00"
20      }
21    }
22  ]
23 }
```

Export Endpoint

The `export` endpoint provides additional functionality for exporting entities from the GDC submission system. The `ids` parameter accepts a UUID or a comma-separated list of UUIDs. The `format` parameter allows the user to specify the preferred format of the API response: JSON, TSV, or CSV. When the `with_children` parameter is set to `with_children`, the response includes the metadata stored in all "child" entities of the entity being requested. The `export` endpoint accepts GET requests.

```
1 export
      token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3
4 curl --header "X-Auth-Token: $token"
      'https://gdc-api.nci.nih.gov/v0/submission/TCGA/ALCH/export?ids=11f8321-832f-4a8b-8384-a2f6256557e0&format=j
```

```
1 {
2   "case": [
```

```json
   3        {
   4          "tissue_source_sites": [],
   5          "submitter_id": "TCGA-ALCH-000026",
   6          "project_id": "TCGA-ALCH",
   7          "type": "case",
   8          "id": "11f83251-832f-4a8b-8384-a2f6256557e0",
   9          "projects": [
  10            {
  11              "code": "ALCH",
  12              "id": "d9906779-f1da-5d9f-9caa-6d5ecb2e3cd6"
  13            }
  14          ]
  15        }
  16      ],
  17      "sample": [
  18        {
  19          "sample_type_id": "10",
  20          "time_between_excision_and_freezing": null,
  21          "oct_embedded": "false",
  22          "tumor_code_id": null,
  23          "submitter_id": "Blood-00001_api26",
  24          "intermediate_dimension": null,
  25          "id": "23308708-6a63-471e-947c-6a93c6e85983",
  26          "time_between_clamping_and_freezing": null,
  27          "pathology_report_uuid": null,
  28          "tumor_descriptor": null,
  29          "sample_type": "Blood Derived Normal",
  30          "project_id": "TCGA-ALCH",
  31          "current_weight": null,
  32          "composition": null,
  33          "is_ffpe": null,
  34          "shortest_dimension": null,
  35          "tumor_code": null,
  36          "tissue_type": null,
  37          "days_to_sample_procurement": null,
  38          "cases": [
  39            {
  40              "id": "11f83251-832f-4a8b-8384-a2f6256557e0",
  41              "submitter_id": "TCGA-ALCH-000026"
  42            }
  43          ],
  44          "freezing_method": null,
  45          "type": "sample",
  46          "preservation_method": null,
  47          "days_to_collection": null,
  48          "initial_weight": null,
  49          "longest_dimension": null
  50        }
  51      ],
  52      "read_group": [
  53        {
  54          "library_name": "Solexa-34688",
  55          "is_paired_end": true,
  56          "size_selection_range": null,
  57          "adapter_sequence": null,
  58          "library_strand": null,
  59          "submitter_id": "Blood-00001-aliquot_lane1_barcode26",
  60          "library_preparation_kit_name": null,
```

```
 61        "adapter_name": null,
 62        "target_capture_kit_name": null,
 63        "includes_spike_ins": null,
 64        "library_preparation_kit_version": null,
 65        "id": "90163202-cfd7-4f6a-8214-e7e4e924d3a6",
 66        "spike_ins_concentration": null,
 67        "target_capture_kit_vendor": null,
 68        "read_length": 75,
 69        "sequencing_date": "2010-08-04",
 70        "spike_ins_fasta": null,
 71        "to_trim_adapter_sequence": null,
 72        "RIN": null,
 73        "platform": "Illumina",
 74        "library_selection": "Hybrid_Selection",
 75        "library_strategy": "WXS",
 76        "library_preparation_kit_catalog_number": null,
 77        "target_capture_kit_target_region": null,
 78        "fastq_name": null,
 79        "target_capture_kit_version": null,
 80        "aliquots": [
 81          {
 82            "id": "e66dee54-5f4c-4471-9e08-dba0f6cdaaa4",
 83            "submitter_id": "Blood-00001-aliquot26"
 84          }
 85        ],
 86        "read_group_name": "205DD.3-2",
 87        "library_preparation_kit_vendor": null,
 88        "project_id": "TCGA-ALCH",
 89        "type": "read_group",
 90        "target_capture_kit_catalog_number": null,
 91        "instrument_model": "Illumina HiSeq 2000",
 92        "base_caller_name": null,
 93        "experiment_name": "Resequencing",
 94        "flow_cell_barcode": "205DDABXX",
 95        "sequencing_center": "BI",
 96        "base_caller_version": null
 97      }
 98    ],
 99    "aliquot": [
100      {
101        "source_center": "23",
102        "centers": [],
103        "analytes": [],
104        "submitter_id": "Blood-00001-aliquot26",
105        "amount": 10,
106        "samples": [
107          {
108            "id": "23308708-6a63-471e-947c-6a93c6e85983",
109            "submitter_id": "Blood-00001_api26"
110          }
111        ],
112        "concentration": 0.07,
113        "project_id": "TCGA-ALCH",
114        "type": "aliquot",
115        "id": "e66dee54-5f4c-4471-9e08-dba0f6cdaaa4"
116      }
117    ],
118    "submitted_unaligned_reads": [
```

```
119    {
120      "read_groups": [
121        {
122          "id": "90163202-cfd7-4f6a-8214-e7e4e924d3a6",
123          "submitter_id": "Blood-00001-aliquot_lane1_barcode26"
124        }
125      ],
126      "data_type": "Unaligned Reads",
127      "file_name": "dummy.fastq",
128      "md5sum": "70c48a8a670ed2a02327601a10038d06",
129      "data_format": "FASTQ",
130      "submitter_id": "Blood-00001-aliquot_lane1_barcode26.fastq",
131      "state_comment": null,
132      "data_category": "Sequencing Data",
133      "file_size": 38,
134      "project_id": "TCGA-ALCH",
135      "type": "submitted_unaligned_reads",
136      "id": "6d45f2a0-8161-42e3-97e6-e058ac18f3f3",
137      "experimental_strategy": "WGS"
138    },
139    {
140      "read_groups": [
141        {
142          "id": "90163202-cfd7-4f6a-8214-e7e4e924d3a6",
143          "submitter_id": "Blood-00001-aliquot_lane1_barcode26"
144        }
145      ],
146      "data_type": "Unaligned Reads",
147      "file_name": "dummy.fastq",
148      "md5sum": "70c48a8a670ed2a02327601a10038d06",
149      "data_format": "FASTQ",
150      "submitter_id": "Blood-00001-aliquot_lane1_barcode27.fastq",
151      "state_comment": null,
152      "data_category": "Sequencing Data",
153      "file_size": 38,
154      "project_id": "TCGA-ALCH",
155      "type": "submitted_unaligned_reads",
156      "id": "4faabdd6-45bb-4259-8868-13d5b1149748",
157      "experimental_strategy": "WGS"
158    }
159  ]
160 }
```

GraphQL

Submitters can use the GraphQL query language for advanced search and retrieval of data from the GDC Submission Portal. See GraphQL for more information.

Deleting Entities

The `entities` endpoint can also be used to delete entities. This is accomplished using a DELETE request to the endpoint, specifying the entity's UUID. If an entity cannot be deleted because it is linked to child entities, the GDC Submission API will respond with an error providing a list of entities that must be deleted prior to deleting the subject entity.

A subgraph (a parent along with all of its child entities) can be deleted in a single transaction by passing a comma-separated list of UUIDs to the `entities` endpoint.

```
1 export
     token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3 curl --header "X-Auth-Token: $token" --request DELETE
     https://gdc-api.nci.nih.gov/v0/submission/TCGA/ALCH/entities/67782964-0065-491d-b051-2ae404bb734d
```

```
1 {
2    "code": 200,
3    "deleted_entity_count": 1,
4    "dependent_ids": "",
5    "entities": [
6      {
7        "action": "delete",
8        "errors": [],
9        "id": "67782964-0065-491d-b051-2ae404bb734d",
10       "related_cases": [],
11       "type": "case",
12       "valid": true,
13       "warnings": []
14     }
15   ],
16   "entity_error_count": 0,
17   "message": "Successfully deleted 1 entities",
18   "success": true,
19   "transaction_id": 192,
20   "transactional_error_count": 0,
21   "transactional_errors": []
22 }
```

Working With Files

Uploading Data Files

Experimental data files like BAM and FASTQ can be uploaded directly to the API using the `files` endpoint, by specifying the UUID of the corresponding `data_file` entity. Binary upload mode must be used if available. Uploading large files may be more efficiently performed using the GDC Data Transfer Tool.

```
1 export
     token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3 curl --header "X-Auth-Token: $token" --output needed_to_show_progress_bar.log --request PUT --data-binary
     @GDC-INTERNAL-000084-S1-Q1-RG1.fastq.zip
     https://gdc-api.nci.nih.gov/v0/submission/GDC/INTERNAL/files/c414a205-376e-4993-af48-2a4689eb433e &&
     rm needed_to_show_progress_bar.log
4
5  "&& rm needed_to_show_progress_bar.log" at the end of the command above
6  removes the temporary file required to show upload progress bar. This
7  will not work on Windows platforms. Windows users must remove this
8  string and can delete the file manually.
```

Upload Manifest

The `manifest` endpoint generates a manifest for uploading files using the GDC Data Transfer Tool. It requires a comma-separated list of file UUIDs to generate a manifest.

```
1 https://gdc-api.nci.nih.gov/v0/submission/PROGRAM/PROJECT/manifest?ids=bf0751ca-fc3b-4760-b876-0fefce040be5,9016
```

Downloading Files

Unreleased files that have been uploaded to the GDC can be downloaded by submitters using the `data` endpoint and an appropriate authentication token. See Downloading Files for details.

Deleting Files

Uploaded files can be deleted by deleting the entity that corresponds to the file. See Deleting Entities for details.

Querying Submitted Data Using GraphQL

GraphQL Overview

GraphQL is a query language that makes it easy to search and retrieve data from graph data structures such as the GDC Data Model.

Unlike the methods outlined in Search and Retrieval, which provide access to public releases (or snapshots) of GDC data, the `/graphql` endpoint of GDC Submission API makes it possible for submitters to access "live" data, which provides a real-time view of the state of entities in a project.

NOTE: Access to GDC Submission API GraphQL service is limited to authorized and authenticated submitters. Submitters may only access data in their own project using GraphQL.

GraphQL IDE

The GDC GraphQL IDE is an instance of GraphiQL, an in-browser GraphQL IDE that facilitates construction and execution of GraphQL queries. The GDC GraphQL IDE provides tab-completion and syntax checking using schema from the GDC Data Dictionary. It can be found at https://gdc-portal.nci.nih.gov/submission/graphiql.

Before interacting directly with the GDC Submission API's GraphQL endpoint, users are encouraged to become familiar with executing queries using the GDC GraphQL IDE.

GraphQL Endpoint

GDC data submitters can access the GDC Submission API GraphQL endpoint at:

```
1 https://gdc-api.nci.nih.gov/[API_version/]submission/graphql
```

where **API_version/]** is the optional API version component (see [Getting Started).

NOTE: An authentication token is required for all requests to the `graphql` endpoint. Queries are restricted to those projects for which the submitter has obtained authorization.

Constructing a Query

When sending GraphQL requests to the API directly, the bare GraphQL query must be wrapped in a "query" JSON object as shown below:

When using the GDC GraphQL IDE, the bare JSON query must be used without a JSON wrapper.

Bare GraphQL query

In its simplest form, a GraphQL query is a **selection set** (curly brackets) that encloses a set of **fields**. The selection set defines the set of information that is to be retrieved. Furthermore, in GraphQL fields are conceptually equivalent to functions that retrieve additional fields and, in some cases, can take arguments. So each field in a selection set can have its own selection set, thereby creating a nested query structure that can navigate complex data relationships. See GraphQL Specification for further details.

In GDC GraphQL IDE, a root field (field within the outermost/umbrella selection set) typically corresponds to an entity, whereas fields inside nested selection sets are typically a combination of entities and entity properties.

The "Docs" panel on the right-hand side of the GDC GraphQL IDE allows users to discover the fields that can be queried with GraphQL. Note that the panel contains a lot of information and users may experience a delay before it is displayed.

A simple GraphQL query looks like this:

```
1 {
2   case (project_id: "TCGA-ALCH", first: 0) {
3     id
4     submitter_id
5
6   }
7   _case_count (project_id: "TCGA-ALCH")
8 }
```

The query above has two root fields: `case` and `_case_count`. The `case` field corresponds to the `case` entity in the GDC Data Model. The query supplies two arguments to the field:

1. `project_id: "TCGA-ALCH"`, which requests only cases in the TCGA-ALCH project.
2. `first: 0`, which requests that the API provide all results in the response, without pagination (a nonzero positive integer value of `first` specifies the number of results to return, 10 by default; "pages" are selected using `offset`).

The `_case_count` field is a special field that returns the number of cases that match the supplied argument.

The bare query above can be used as is in the GraphQL IDE. In order to pass this query to the GDC API directly, it needs to be further processed as described below.

Passing GraphQL queries to GDC API directly

Before a bare GraphQL query is passed to the GDC API, it must be processed as follows:

1. Escape the query using JSON string rules
2. Wrap the query in a "query" JSON object.
3. Pass the query to the `graphql` endpoint in an HTTP POST request.

Using the `case` and `_case_count` example above as the starting point, the results are as follows:

```
1 {
2     case (project_id: "TCGA-ALCH", first: 0) {
3         id
4         submitter_id
5
6     }
7     _case_count (project_id: "TCGA-ALCH")
8 }
```

```
1 {\n\tcase (project_id: \"TCGA-ALCH\", first: 0) {\n\t\tid\n\t\tsubmitter_id\n\n\t}\n\t_case_count
      (project_id: \"TCGA-ALCH\")\n}
```

```
1 {
2     "query": "{\n\tcase (project_id: \"TCGA-ALCH\", first: 0)
          {\n\t\tid\n\t\tsubmitter_id\n\n\t}\n\t_case_count (project_id: \"TCGA-ALCH\")\n}",
3     "variables": null
4 }
```

```
1 export
     token=ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcToKeN=0123456789-ALPHANUMERICTOKEN-01234567890+AlPhAnUmErIcT
2
3 curl --request POST --header "X-Auth-Token: $token" 'https://gdc-api.nci.nih.gov/v0/submission/graphql'
     --data-binary @Query_json
```

```
1  {
2    "data": {
3      "_case_count": 20,
4      "case": [
5        {
6          "id": "700d1110-b6b4-4251-89d4-fa6f0698e3f8",
7          "submitter_id": "TCGA-ALCH-000004"
8        },
9        {
10          "id": "be01357d-7348-40b4-a997-8a61ae7af17d",
11          "submitter_id": "TCGA-ALCH-000005"
12        },
13        {
14          "id": "e5638697-6ef3-4bf8-a373-102519093f33",
15          "submitter_id": "TCGA-ALCH-000008"
16        },
17        {
18          "id": "4871d41a-680e-4fd0-901c-b06f06ecae33",
19          "submitter_id": "TCGA-ALCH-000007"
20        },
21        {
22          "id": "2f18c2c1-bff2-43b6-9702-e138c72d8c6b",
23          "submitter_id": "TCGA-ALCH-000009"
24        },
25        {
26          "id": "ec83e038-4f01-47a6-bc69-47fb297d0282",
27          "submitter_id": "TCGA-ALCH-000006"
28        },
29        {
30          "id": "e4642952-d259-4be1-9c53-ed95aa1fc50b",
31          "submitter_id": "TCGA-ALCH-000011"
32        },
33        {
34          "id": "8bcaf0b3-21d0-45c6-87ee-c997efb417dc",
35          "submitter_id": "TCGA-ALCH-000010"
36        },
37        {
38          "id": "83de027e-bcbf-4239-975b-7e8ced82448e",
39          "submitter_id": "TCGA-ALCH-000013"
40        },
41        {
42          "id": "bbd91cc1-06e2-4e60-8b93-e09c3b16f00c",
43          "submitter_id": "TCGA-ALCH-000014"
44        },
45        {
46          "id": "574fd163-4368-440c-9548-d76a0fbc9056",
```

```
47          "submitter_id": "TCGA-ALCH-000015"
48        },
49        {
50          "id": "47c92cdd-ff11-4c25-b0f0-0f7671144271",
51          "submitter_id": "TCGA-ALCH-000016"
52        },
53        {
54          "id": "9f13caab-1fda-4b2a-b500-f79dc978c6c1",
55          "submitter_id": "TCGA-ALCH-000017"
56        },
57        {
58          "id": "9418f194-8741-44db-bd8f-36f4fd8c3bf2",
59          "submitter_id": "TCGA-ALCH-000018"
60        },
61        {
62          "id": "6fb2a018-c5f3-45e5-81d3-e58e7e4bf921",
63          "submitter_id": "TCGA-ALCH-000019"
64        },
65        {
66          "id": "70236972-e796-414a-9b7a-3b29b849ba7c",
67          "submitter_id": "TCGA-ALCH-000020"
68        },
69        {
70          "id": "6f78e86f-9e31-4af5-a0d9-b8970ece476d",
71          "submitter_id": "TCGA-ALCH-000021"
72        },
73        {
74          "id": "c6fcb2f0-c6bb-4b40-a761-bae3e63869cb",
75          "submitter_id": "TCGA-ALCH-000002"
76        },
77        {
78          "id": "67782964-0065-491d-b051-2ae404bb734d",
79          "submitter_id": "TCGA-ALCH-000001"
80        },
81        {
82          "id": "b45d2891-ba81-4ecc-a250-c58060934227",
83          "submitter_id": "TCGA-ALCH-000012"
84        }
85      ]
86    }
87 }
```

Additional Examples

Example: File UUID

GraphQL query to find the file UUID based on file `submitter_id`:

```
1 {
2
3   submitted_unaligned_reads (project_id: "GDC-INTERNAL", submitter_id:
        "Blood-00001-aliquot_lane1_barcode23.fastq") {
4     id
5     submitter_id
6     file_name
7     project_id
8 }
```

```
 9 }
```

```
 1 {
 2     "query": "{\n \n  submitted_unaligned_reads (project_id: \"GDC-INTERNAL\", submitter_id:
            \"Blood-00001-aliquot_lane1_barcode23.fastq\") {\n    id\n    submitter_id\n    file_name\n
            project_id\n}\n}",
 3     "variables": null
 4 }
```

```
 1 curl --request POST --header "X-Auth-Token: $token" 'https://gdc-api.nci.nih.gov/v0/submission/graphql'
      --data-binary @escaped_GraphQL
```

```
 1 {
 2   "data": {
 3     "submitted_unaligned_reads": [
 4       {
 5         "file_name": "dummy.fastq",
 6         "id": "616eab2f-791a-4641-8cd6-ee195a10a201",
 7         "project_id": "GDC-INTERNAL",
 8         "submitter_id": "Blood-00001-aliquot_lane1_barcode23.fastq"
 9       }
10     ]
11   }
12 }
```

Example: Case Without Diagnosis

GraphQL query for any one case in 'TCGA-LUAD' without Diagnosis information:

```
 1 {
 2   case (project_id: "TCGA-LUAD", without_links: ["diagnoses"], first: 1) {
 3       submitter_id
 4   }
 5 }
```

```
 1 {
 2   "data": {
 3     "case": [
 4       {
 5         "submitter_id": "TCGA-17-Z050"
 6       }
 7     ]
 8   }
 9 }
```

Example: Number of Cases Without Diagnosis

GraphQL query for the number of cases in 'TCGA-LUAD' without Diagnosis information:

```
 1 {
 2   _case_count (project_id: "TCGA-LUAD", without_links: ["diagnoses"])
 3 }
```

```
 1 {
 2   "data": {
 3     "_case_count": 5
 4   }
 5 }
```

Example: Aliquot State

Query for the `state` of aliquots belonging to case with `submitter_id: "TCGA-ALCH-000001"`:

```
1 {
2   aliquot(with_path_to: {type: "case", submitter_id:"TCGA-ALCH-000001"}) {
3     id release_state
4   }
5 }
```

```
1 {
2   "data": {
3     "aliquot": [
4       {
5         "id": "7af58da0-cb3e-43e2-a074-4bd8f27565ba",
6         "state": "validated"
7       }
8     ]
9   }
10 }
```

Example: Aliases

GraphQL query that uses a GraphQL fragment to get specific properties from two portions and give them aliases in the response:

```
1 {
2   some_portion: portion (first: 1) {
3     ...portionProperties
4   }
5   specific_portion: portion(submitter_id: "TCGA-67-6217-01A-13-2191-20") {
6     ...portionProperties
7   }
8 }
9
10 fragment portionProperties on portion {
11   submitter_id
12   is_ffpe
13 }
```

```
1 {
2   "data": {
3     "some_portion": [
4       {
5         "is_ffpe": false,
6         "submitter_id": "TCGA-62-A471-10A-01"
7       }
8     ],
9     "specific_portion": [
10       {
11         "is_ffpe": false,
12         "submitter_id": "TCGA-67-6217-01A-13-2191-20"
13       }
14     ]
15   }
16 }
```

Example: Biospecimen Tree

GraphQL Query for a case in "TCGA-LUAD" and return a biospecimen tree:

```
1  {
2    case(project_id: "TCGA-LUAD", first: 1) {
3      id
4      samples(first: 1) {
5        id
6        portions(first: 1) {
7          id
8          analytes(first: 1) {
9            id
10           aliquots(first: 1) {
11             id
12           }
13         }
14       }
15     }
16   }
17 }
```

```
1  {
2    "data": {
3      "case": [
4        {
5          "id": "19ca36e6-2154-4224-89b1-117a4a4407f6",
6          "samples": [
7            {
8              "id": "5e2625d2-290d-48cd-af5c-27dc8e3c8b6a",
9              "portions": [
10               {
11                 "analytes": [
12                   {
13                     "aliquots": [
14                       {
15                         "id": "8e1820d5-dcd8-4760-9962-221e2b71d4b9"
16                       }
17                     ],
18                     "id": "6449533c-e52a-4e58-bae7-0732f48153ef"
19                   }
20                 ],
21                 "id": "26b75643-8fcd-445e-a0e0-9868cac589ea"
22               }
23             ]
24           }
25         ]
26       }
27     ]
28   }
29 }
```

Chapter 6

System Information

System Information

Overview

The GDC API offers endpoints that provide information about the system. These endpoints are described below.

GDC Notifications Endpoint

The `notifications` endpoint provides current user-facing notifications.

GDC notifications have a corresponding `level` with the following meanings:

Level	Meaning
INFO	Non-essential information, e.g. regarding a new dataset
WARNING	Important user information, e.g. regarding a dataset to be removed
ERROR	Important system information, e.g. regarding a GDC component
DEBUG	Unimportant system information, e.g. testing the notification system

Notifications will indicate the GDC `components` to which they apply:

Component	Description
PORTAL	The GDC Data Portal
LEGACY	The GDC Legacy Archive
SUBMISSION	The GDC Data Submission Portal
DOCUMENTATION	The GDC documentation site that contains GDC user guides, release notes, and the GDC Data Dictionary
WEBSITE	The GDC project website that includes information about the system. This does not include any of the above-listed GDC components.

Sample Request

```
1  curl --request GET https://gdc-api.nci.nih.gov/v0/notifications
```

```
1  {
2    "data": [
3      {
4        "level": "INFO",
5        "components": [
6          "SUBMISSION_API",
7          "LEGACY_API"
8        ],
9        "message": "The system is up!"
10     }
11   ]
12 }
```

API Status Endpoint

The `status` endpoint provides information about the current status and version of the GDC API.

Sample Request

```
1  curl https://gdc-api.nci.nih.gov/status
```

```
1  import requests
2  import json
3
4  status_endpt = 'https://gdc-api.nci.nih.gov/status'
5  response = requests.get(status_endpt)
6  print json.dumps(response.json(), indent=2)
```

```
1  {
2    "commit": "74e1e3583c0f39fbf2149322addb7378206be3b9",
3    "status": "OK",
4    "tag": "1.2.0",
5    "version": 1
6  }
```

Chapter 7

Additional Examples

Additional Examples

Data Search and Retrieval

Endpoint Examples

This section contains additional examples for using endpoints.

Project Endpoint Example

This example is a query for Projects contained in GDC. It returns only the first five projects sorted by project name.

```
1 curl 'https://gdc-api.nci.nih.gov/projects?from=1&size=5&sort=project.name:asc&pretty=true'
```

```
1  {
2    "data": {
3      "hits": [
4        {
5          "state": "legacy",
6          "project_id": "TARGET-AML",
7          "primary_site": "Blood",
8          "disease_type": "Acute Myeloid Leukemia",
9          "name": "Acute Myeloid Leukemia"
10       },
11       {
12         "state": "legacy",
13         "project_id": "TCGA-LAML",
14         "primary_site": "Blood",
15         "disease_type": "Acute Myeloid Leukemia",
16         "name": "Acute Myeloid Leukemia"
17       },
18       {
19         "state": "legacy",
20         "project_id": "TARGET-AML-IF",
21         "primary_site": "Blood",
22         "disease_type": "Acute Myeloid Leukemia Induction Failure",
23         "name": "Acute Myeloid Leukemia Induction Failure"
24       },
25       {
26         "state": "legacy",
```

```json
27          "project_id": "TARGET-ALL-P2",
28          "primary_site": "Blood",
29          "disease_type": "Acute Lymphoblastic Leukemia",
30          "name": "Acute Lymphoblastic Leukemia - Phase II"
31        },
32        {
33          "state": "legacy",
34          "project_id": "TARGET-ALL-P1",
35          "primary_site": "Blood",
36          "disease_type": "Acute Lymphoblastic Leukemia",
37          "name": "Acute Lymphoblastic Leukemia - Phase I"
38        }
39      ],
40      "pagination": {
41        "count": 5,
42        "sort": "project.name:asc",
43        "from": 1,
44        "pages": 10,
45        "total": 46,
46        "page": 1,
47        "size": 5
48      }
49    },
50    "warnings": {}
51 }
```

Files Endpoint Example

This example is a query for files contained in GDC. It returns only the first two files, sorted by file size, from smallest to largest.

```
1 curl 'https://gdc-api.nci.nih.gov/files?from=1&size=2&sort=file_size:asc&pretty=true'
```

```json
1 {
2    "data": {
3      "hits": [
4        {
5          "state": "legacy",
6          "project_id": "TARGET-AML",
7          "primary_site": "Blood",
8          "disease_type": "Acute Myeloid Leukemia",
9          "name": "Acute Myeloid Leukemia"
10       },
11       {
12         "state": "legacy",
13         "project_id": "TCGA-LAML",
14         "primary_site": "Blood",
15         "disease_type": "Acute Myeloid Leukemia",
16         "name": "Acute Myeloid Leukemia"
17       },
18       {
19         "state": "legacy",
20         "project_id": "TARGET-AML-IF",
21         "primary_site": "Blood",
22         "disease_type": "Acute Myeloid Leukemia Induction Failure",
23         "name": "Acute Myeloid Leukemia Induction Failure"
24       },
25       {
```

```
26          "state": "legacy",
27          "project_id": "TARGET-ALL-P2",
28          "primary_site": "Blood",
29          "disease_type": "Acute Lymphoblastic Leukemia",
30          "name": "Acute Lymphoblastic Leukemia - Phase II"
31        },
32        {
33          "state": "legacy",
34          "project_id": "TARGET-ALL-P1",
35          "primary_site": "Blood",
36          "disease_type": "Acute Lymphoblastic Leukemia",
37          "name": "Acute Lymphoblastic Leukemia - Phase I"
38        }
39      ],
40      "pagination": {
41        "count": 5,
42        "sort": "project.name:asc",
43        "from": 1,
44        "pages": 10,
45        "total": 46,
46        "page": 1,
47        "size": 5
48      }
49    },
50    "warnings": {}
51 }
```

Cases Endpoint Example

This example is a query for cases contained in GDC. It returns only the first five files.

```
1 curl 'https://gdc-api.nci.nih.gov/cases?from=1&size=5&pretty=true'
```

```
1 {
2    "data": {
3      "hits": [
4        {
5          "sample_ids": "fae164e6-16ed-4547-9872-15d53c79bb45",
6          "portion_ids": "0a5fa1fd-aa9b-49d1-8a32-3522271a56e8",
7          "submitter_portion_ids": "TCGA-78-7535-10A-01",
8          "submitter_aliquot_ids": "TCGA-78-7535-10A-01W-2107-08",
9          "days_to_index": 0,
10         "submitter_analyte_ids": "TCGA-78-7535-10A-01W",
11         "analyte_ids": "14081a57-a8ee-497d-a944-3f24ef8efddb",
12         "submitter_id": "TCGA-78-7535",
13         "case_id": "46592b7b-6968-42a6-83af-0917c9f4a9a5",
14         "submitter_sample_ids": "TCGA-78-7535-10A",
15         "aliquot_ids": "22036caf-c6c9-4ad4-8a69-912b8e56aace"
16       },
17       {
18         "sample_ids": "094cf919-3e36-4d9e-9d37-a00ae04736ee",
19         "portion_ids": "1a723c9e-ac2e-40fd-b342-6f9fe7795681",
20         "submitter_portion_ids": "TCGA-DJ-A2Q9-01A-11-A21M-20",
21         "submitter_aliquot_ids": "TCGA-DJ-A2Q9-01A-21R-A18B-13",
22         "days_to_index": 0,
23         "submitter_analyte_ids": "TCGA-DJ-A2Q9-01A-21R",
24         "analyte_ids": "f888f0c4-7f33-4a64-8975-316d88e214b3",
```

```
25          "submitter_id": "TCGA-DJ-A2Q9",
26          "case_id": "061fab24-727a-4551-a205-89eeb9f530ea",
27          "submitter_sample_ids": "TCGA-DJ-A2Q9-01A",
28          "aliquot_ids": "29c9b306-d3e9-4d09-bc54-21e46f92ad8a"
29        },
30        {
31          "sample_ids": "8a59b137-8e5e-4484-a9a9-65a596a47ef8",
32          "portion_ids": "9f590b62-a6ab-489d-92b9-6e7802812a15",
33          "submitter_portion_ids": "TCGA-J4-A83I-01A-11",
34          "submitter_aliquot_ids": "TCGA-J4-A83I-01A-11W-A447-08",
35          "days_to_index": 0,
36          "submitter_analyte_ids": "TCGA-J4-A83I-01A-11W",
37          "analyte_ids": "f12a0133-c4c3-4240-a013-1ce159cc08f6",
38          "submitter_id": "TCGA-J4-A83I",
39          "case_id": "5dc7e186-7e01-4a54-8ae8-350dace2297b",
40          "submitter_sample_ids": "TCGA-J4-A83I-01A",
41          "aliquot_ids": "dffd09c5-965d-45ac-93ca-f4a547d78684"
42        },
43        {
44          "sample_ids": "587bf402-b61b-444d-af40-6f67bb04c323",
45          "portion_ids": "b2e05da1-d75d-4075-9704-2efbd7ed51f3",
46          "submitter_portion_ids": "TCGA-BG-A0VW-10A-01",
47          "submitter_aliquot_ids": "TCGA-BG-A0VW-10A-01D-A122-09",
48          "days_to_index": 0,
49          "submitter_analyte_ids": "TCGA-BG-A0VW-10A-01W",
50          "analyte_ids": "2f3e28c3-60ec-433a-8f13-6466dd68c5ac",
51          "submitter_id": "TCGA-BG-A0VW",
52          "case_id": "62d71839-4fba-42e9-9929-8d937f0fe287",
53          "submitter_sample_ids": "TCGA-BG-A0VW-10A",
54          "aliquot_ids": "a9cd1596-89b6-46fd-9864-8a62f47d1f8b"
55        },
56        {
57          "sample_ids": "26a0b6bc-f8aa-45f1-a215-c90e4e840607",
58          "portion_ids": "0dfe9858-d9bf-4e75-84f5-cae466ece831",
59          "submitter_portion_ids": "TCGA-P4-AAVL-11A-11",
60          "submitter_aliquot_ids": "TCGA-P4-AAVL-11A-11D-A42M-10",
61          "days_to_index": 0,
62          "submitter_analyte_ids": "TCGA-P4-AAVL-11A-11D",
63          "analyte_ids": "f310ac58-97b8-4bba-9dab-ae7768185375",
64          "submitter_id": "TCGA-P4-AAVL",
65          "case_id": "39847790-c951-4f9d-b23c-88c7f44d20a0",
66          "submitter_sample_ids": "TCGA-P4-AAVL-11A",
67          "aliquot_ids": "fb226698-844d-4a24-86c2-29571549b9bb"
68        }
69      ],
70      "pagination": {
71        "count": 5,
72        "sort": "",
73        "from": 1,
74        "pages": 2822,
75        "total": 14108,
76        "page": 1,
77        "size": 5
78      }
79    },
80    "warnings": {}
81 }
```

Annotations Endpoint Example

This example is a query for annotations contained in the GDC. It returns only the first two annotations.

```
1 curl 'https://gdc-api.nci.nih.gov/annotations?from=1&size=2&pretty=true'
```

```
1  {
2    "data": {
3      "hits": [
4        {
5          "category": "Item flagged DNU",
6          "status": "Approved",
7          "entity_id": "2b61b856-b988-43ca-8dc5-9f97600118ec",
8          "classification": "CenterNotification",
9          "entity_type": "aliquot",
10         "created_datetime": 1294525038,
11         "annotation_id": "7d01080f-e82d-5e58-98a6-910c041ee2b3",
12         "notes": "SDRF in broad.mit.edu_READ.Genome_Wide_SNP_6.mage-tab.1.1003.0 flagged aliquot to be
                   excluded for analysis based on file
                   'SCENA_p_TCGAb29and30_SNP_N_GenomeWideSNP_6_C04_569122.ismpolish.data.txt'.",
13         "creator": "DCC",
14         "submitter_id": "1099",
15         "case_id": "e7503a51-6647-4cc2-80dd-645d0df4db43",
16         "entity_submitter_id": "TCGA-AG-A008-10A-01D-A003-01"
17       },
18       {
19         "category": "Item flagged DNU",
20         "status": "Approved",
21         "entity_id": "d1f35d46-c6c9-4cff-ad95-e86d88b38b51",
22         "classification": "CenterNotification",
23         "entity_type": "aliquot",
24         "created_datetime": 1414794925,
25         "annotation_id": "c6a9e076-bb56-5dd9-89e7-c340594fa8f7",
26         "notes": "SDRF in broad.mit.edu_COAD.Genome_Wide_SNP_6.mage-tab.1.2010.0 flagged aliquot to be
                   excluded for analysis based on file
                   'SNORT_p_TCGA_b89_SNP_N_GenomeWideSNP_6_E05_777376.birdseed.data.txt'.",
27         "creator": "DCC",
28         "submitter_id": "23507",
29         "case_id": "57b0f89f-1b75-453e-922c-01cd4d44ca49",
30         "entity_submitter_id": "TCGA-CK-5914-10A-01D-1649-01"
31       }
32     ],
33     "pagination": {
34       "count": 2,
35       "sort": "",
36       "from": 1,
37       "pages": 12296,
38       "total": 24592,
39       "page": 1,
40       "size": 2
41     }
42   },
43   "warnings": {}
44 }
```

Filters Examples

This section contains additional examples for using the `filters` parameter.

Example: Basic syntax

The following is an example of `filters` syntax, including the JSON object passed to the `filters` parameter, the corresponding API query, and the JSON object returned by the API. The example finds projects where the primary site is Blood.

```
1  {
2    "op": "and",
3    "content": [
4      {
5        "op": "in",
6        "content": {
7          "field": "primary_site",
8          "value": [
9            "Blood"
10           ]
11        }
12      }
13    ]
14  }
```

```
1  curl
        'https://gdc-api.nci.nih.gov/projects?filters=%7b%0d%0a++%22op%22%3a+%22and%22%2c%0d%0a++%22content%22%3a+%5
```

```
1  {
2    "data": {
3      "hits": [
4        {
5          "state": "legacy",
6          "project_id": "TCGA-LAML",
7          "primary_site": "Blood",
8          "disease_type": "Acute Myeloid Leukemia",
9          "name": "Acute Myeloid Leukemia"
10         },
11         {
12           "dbgap_accession_number": "phs000465",
13           "disease_type": "Acute Myeloid Leukemia",
14           "state": "legacy",
15           "primary_site": "Blood",
16           "project_id": "TARGET-AML",
17           "name": "Acute Myeloid Leukemia"
18         },
19         {
20           "dbgap_accession_number": "phs000464",
21           "disease_type": "Acute Lymphoblastic Leukemia",
22           "state": "legacy",
23           "primary_site": "Blood",
24           "project_id": "TARGET-ALL-P2",
25           "name": "Acute Lymphoblastic Leukemia - Phase II"
26         },
27         {
28           "dbgap_accession_number": "phs000515",
29           "disease_type": "Acute Myeloid Leukemia Induction Failure",
30           "state": "legacy",
31           "primary_site": "Blood",
32           "project_id": "TARGET-AML-IF",
33           "name": "Acute Myeloid Leukemia Induction Failure"
34         },
35         {
36           "state": "legacy",
```

```
37        "project_id": "TCGA-LCML",
38        "primary_site": "Blood",
39        "disease_type": "Chronic Myelogenous Leukemia",
40        "name": "Chronic Myelogenous Leukemia"
41      },
42      {
43        "dbgap_accession_number": "phs000463",
44        "disease_type": "Acute Lymphoblastic Leukemia",
45        "state": "legacy",
46        "primary_site": "Blood",
47        "project_id": "TARGET-ALL-P1",
48        "name": "Acute Lymphoblastic Leukemia - Phase I"
49      }
50    ],
51    "pagination": {
52      "count": 6,
53      "sort": "",
54      "from": 1,
55      "page": 1,
56      "total": 6,
57      "pages": 1,
58      "size": 10
59    }
60  },
61  "warnings": {}
62 }
```

Example: Filter cases keeping only 'male'

This is an example of a value-based filter:

```
1 {
2     "op" : "=" ,
3     "content" : {
4         "field" : "cases.clinical.gender" ,
5         "value" : [ "male" ]
6     }
7 }
```

```
1 curl
      'https://gdc-api.nci.nih.gov/cases?filters=%7b%0d%0a+++%22op%22+%3a+%22%3d%22+%2c%0d%0a+++%22content%22+%3a+
```

Example: Filter using a range

This is an example of filtering for age at diagnosis. The request is for cases where the age at diagnosis is between 40 and 70 years. *Note:* `age_at_diagnosis` is expressed in days.

```
1 {
2     "op": "and",
3     "content": [
4         {
5             "op": ">=",
6             "content": {
7                 "field": "cases.clinical.age_at_diagnosis",
8                 "value": [
9                     14600
10                 ]
```

```
11                   }
12           },
13           {
14               "op": "<=",
15               "content": {
16                   "field": "cases.clinical.age_at_diagnosis",
17                   "value": [
18                       25550
19                   ]
20               }
21           }
22       ]
23 }
```

```
1 curl
      'https://gdc-api.nci.nih.gov/cases?filters=%7B%22op%22:%22and%22,%22content%22:%5B%7B%22op%22:%22%3E%3D%22,%
```

Example: Multiple fields

Filter projects for primary_site being Kidney or Brain and program.name being TCGA

```
1 {
2       "op" : "and" ,
3       "content" : [{
4               "op" : "in" ,
5               "content" : {
6                   "field" : "primary_site" ,
7                   "value" : [
8                       "Kidney" ,
9                       "Brain"
10                  ]
11              }
12          }, {
13              "op" : "in" ,
14              "content" : {
15                  "field" : "program.name" ,
16                  "value" : [
17                      "TCGA"
18                  ]
19              }
20          }]
21 }
```

```
1 curl
      'https://gdc-api.nci.nih.gov/projects?filters=%7B%22op%22%3A%22and%22%2C%22content%22%3A%5B%7B%22op%22%3A%22
```

Chapter 8

Appendix A: Available Fields

Appendix A: Available Fields

The GDC API's search and retrieval endpoints provide access to fields that correspond to properties defined in the GDC Data Dictionary. This appendix contains a list of fields available at each endpoint, and a list of field groups accessible via the expand parameter.

Field Listing by Endpoint

Project Fields

Field Name
dbgap_accession_number
disease_type
name
primary_site
project_id
released
state
program.dbgap_accession_number
program.name
program.program_id
summary.case_count
summary.file_count
summary.file_size
summary.data_categories.case_count
summary.data_categories.data_category
summary.data_categories.file_count
summary.experimental_strategies.case_count
summary.experimental_strategies.experimental_strategy
summary.experimental_strategies.file_count

Case Fields

Field Name |
--- |
aliquot_ids |
analyte_ids |
case_id |
created_datetime |
days_to_index |
portion_ids |
sample_ids |
slide_ids |
state |
submitter_aliquot_ids |
submitter_analyte_ids |
submitter_id |
submitter_portion_ids |
submitter_sample_ids |
submitter_slide_ids |
updated_datetime |
annotations.annotation_id |
annotations.case_id |
annotations.case_submitter_id |
annotations.category |
annotations.classification |
annotations.created_datetime |
annotations.creator |
annotations.entity_id |
annotations.entity_submitter_id |
annotations.entity_type |
annotations.legacy_created_datetime |
annotations.legacy_updated_datetime |
annotations.notes |
annotations.state |
annotations.status |
annotations.submitter_id |
annotations.updated_datetime |
demographic.created_datetime |
demographic.demographic_id |
demographic.ethnicity |
demographic.gender |
demographic.race |
demographic.state |
demographic.submitter_id |
demographic.updated_datetime |
demographic.year_of_birth |
demographic.year_of_death |
diagnoses.age_at_diagnosis |
diagnoses.classification_of_tumor |
diagnoses.created_datetime |
diagnoses.days_to_birth |
diagnoses.days_to_death |
diagnoses.days_to_last_follow_up |
diagnoses.days_to_last_known_disease_status |
diagnoses.days_to_recurrence |
diagnoses.diagnosis_id |
diagnoses.last_known_disease_status |
diagnoses.morphology |

diagnoses.primary_diagnosis |
diagnoses.prior_malignancy |
diagnoses.progression_or_recurrence |
diagnoses.site_of_resection_or_biopsy |
diagnoses.state |
diagnoses.submitter_id |
diagnoses.tissue_or_organ_of_origin |
diagnoses.tumor_grade |
diagnoses.tumor_stage |
diagnoses.updated_datetime |
diagnoses.vital_status |
diagnoses.treatments.created_datetime |
diagnoses.treatments.days_to_treatment |
diagnoses.treatments.state |
diagnoses.treatments.submitter_id |
diagnoses.treatments.therapeutic_agents |
diagnoses.treatments.treatment_id |
diagnoses.treatments.treatment_intent_type |
diagnoses.treatments.treatment_or_therapy |
diagnoses.treatments.updated_datetime |
exposures.alcohol_history |
exposures.alcohol_intensity |
exposures.bmi |
exposures.cigarettes_per_day |
exposures.created_datetime |
exposures.exposure_id |
exposures.height |
exposures.state |
exposures.submitter_id |
exposures.updated_datetime |
exposures.weight |
exposures.years_smoked |
family_histories.created_datetime |
family_histories.family_history_id |
family_histories.relationship_age_at_diagnosis |
family_histories.relationship_gender |
family_histories.relationship_primary_diagnosis |
family_histories.relationship_type |
family_histories.relative_with_cancer_history |
family_histories.state |
family_histories.submitter_id |
family_histories.updated_datetime |
files.access |
files.acl |
files.created_datetime |
files.data_category |
files.data_format |
files.data_type |
files.error_type |
files.experimental_strategy |
files.file_id |
files.file_name |
files.file_size |
files.file_state |
files.md5sum |
files.origin |
files.platform |
files.revision |

files.state |
files.state_comment |
files.submitter_id |
files.tags |
files.type |
files.updated_datetime |
files.analysis.analysis_id |
files.analysis.analysis_type |
files.analysis.created_datetime |
files.analysis.state |
files.analysis.submitter_id |
files.analysis.updated_datetime |
files.analysis.workflow_end_datetime |
files.analysis.workflow_link |
files.analysis.workflow_start_datetime |
files.analysis.workflow_type |
files.analysis.workflow_version |
files.analysis.input_files.access |
files.analysis.input_files.created_datetime |
files.analysis.input_files.data_category |
files.analysis.input_files.data_format |
files.analysis.input_files.data_type |
files.analysis.input_files.error_type |
files.analysis.input_files.experimental_strategy |
files.analysis.input_files.file_id |
files.analysis.input_files.file_name |
files.analysis.input_files.file_size |
files.analysis.input_files.file_state |
files.analysis.input_files.md5sum |
files.analysis.input_files.platform |
files.analysis.input_files.revision |
files.analysis.input_files.state |
files.analysis.input_files.state_comment |
files.analysis.input_files.submitter_id |
files.analysis.input_files.updated_datetime |
files.analysis.metadata.read_groups.adapter_name |
files.analysis.metadata.read_groups.adapter_sequence |
files.analysis.metadata.read_groups.base_caller_name |
files.analysis.metadata.read_groups.base_caller_version |
files.analysis.metadata.read_groups.created_datetime |
files.analysis.metadata.read_groups.experiment_name |
files.analysis.metadata.read_groups.flow_cell_barcode |
files.analysis.metadata.read_groups.includes_spike_ins |
files.analysis.metadata.read_groups.instrument_model |
files.analysis.metadata.read_groups.is_paired_end |
files.analysis.metadata.read_groups.library_name |
files.analysis.metadata.read_groups.library_preparation_kit_catalog_number |
files.analysis.metadata.read_groups.library_preparation_kit_name |
files.analysis.metadata.read_groups.library_preparation_kit_vendor |
files.analysis.metadata.read_groups.library_preparation_kit_version |
files.analysis.metadata.read_groups.library_selection |
files.analysis.metadata.read_groups.library_strand |
files.analysis.metadata.read_groups.library_strategy |
files.analysis.metadata.read_groups.platform |
files.analysis.metadata.read_groups.read_group_id |
files.analysis.metadata.read_groups.read_group_name |
files.analysis.metadata.read_groups.read_length |
files.analysis.metadata.read_groups.RIN |

files.analysis.metadata.read__groups.sequencing__center |
files.analysis.metadata.read__groups.sequencing__date |
files.analysis.metadata.read__groups.size__selection__range |
files.analysis.metadata.read__groups.spike__ins__concentration |
files.analysis.metadata.read__groups.spike__ins__fasta |
files.analysis.metadata.read__groups.state |
files.analysis.metadata.read__groups.submitter__id |
files.analysis.metadata.read__groups.target__capture__kit__catalog__number |
files.analysis.metadata.read__groups.target__capture__kit__name |
files.analysis.metadata.read__groups.target__capture__kit__target__region |
files.analysis.metadata.read__groups.target__capture__kit__vendor |
files.analysis.metadata.read__groups.target__capture__kit__version |
files.analysis.metadata.read__groups.to__trim__adapter__sequence |
files.analysis.metadata.read__groups.updated__datetime |
files.analysis.metadata.read__groups.read__group__qcs.adapter__content |
files.analysis.metadata.read__groups.read__group__qcs.basic__statistics |
files.analysis.metadata.read__groups.read__group__qcs.created__datetime |
files.analysis.metadata.read__groups.read__group__qcs.encoding |
files.analysis.metadata.read__groups.read__group__qcs.fastq__name |
files.analysis.metadata.read__groups.read__group__qcs.kmer__content |
files.analysis.metadata.read__groups.read__group__qcs.overrepresented__sequences |
files.analysis.metadata.read__groups.read__group__qcs.per__base__n__content |
files.analysis.metadata.read__groups.read__group__qcs.per__base__sequence__content |
files.analysis.metadata.read__groups.read__group__qcs.per__base__sequence__quality |
files.analysis.metadata.read__groups.read__group__qcs.per__sequence__gc__content |
files.analysis.metadata.read__groups.read__group__qcs.per__sequence__quality__score |
files.analysis.metadata.read__groups.read__group__qcs.per__tile__sequence__quality |
files.analysis.metadata.read__groups.read__group__qcs.percent__gc__content |
files.analysis.metadata.read__groups.read__group__qcs.read__group__qc__id |
files.analysis.metadata.read__groups.read__group__qcs.sequence__duplication__levels |
files.analysis.metadata.read__groups.read__group__qcs.sequence__length__distribution |
files.analysis.metadata.read__groups.read__group__qcs.state |
files.analysis.metadata.read__groups.read__group__qcs.submitter__id |
files.analysis.metadata.read__groups.read__group__qcs.total__sequences |
files.analysis.metadata.read__groups.read__group__qcs.updated__datetime |
files.analysis.metadata.read__groups.read__group__qcs.workflow__end__datetime |
files.analysis.metadata.read__groups.read__group__qcs.workflow__link |
files.analysis.metadata.read__groups.read__group__qcs.workflow__start__datetime |
files.analysis.metadata.read__groups.read__group__qcs.workflow__type |
files.analysis.metadata.read__groups.read__group__qcs.workflow__version |
files.archive.archive__id |
files.archive.created__datetime |
files.archive.data__category |
files.archive.data__format |
files.archive.data__type |
files.archive.error__type |
files.archive.file__name |
files.archive.file__size |
files.archive.file__state |
files.archive.md5sum |
files.archive.revision |
files.archive.state |
files.archive.state__comment |
files.archive.submitter__id |
files.archive.updated__datetime |
files.cases.aliquot__ids |
files.cases.analyte__ids |
files.cases.case__id |

files.cases.created_datetime |
files.cases.days_to_index |
files.cases.portion_ids |
files.cases.sample_ids |
files.cases.slide_ids |
files.cases.state |
files.cases.submitter_aliquot_ids |
files.cases.submitter_analyte_ids |
files.cases.submitter_id |
files.cases.submitter_portion_ids |
files.cases.submitter_sample_ids |
files.cases.submitter_slide_ids |
files.cases.updated_datetime |
files.cases.annotations.annotation_id |
files.cases.annotations.case_id |
files.cases.annotations.case_submitter_id |
files.cases.annotations.category |
files.cases.annotations.classification |
files.cases.annotations.created_datetime |
files.cases.annotations.creator |
files.cases.annotations.entity_id |
files.cases.annotations.entity_submitter_id |
files.cases.annotations.entity_type |
files.cases.annotations.legacy_created_datetime |
files.cases.annotations.legacy_updated_datetime |
files.cases.annotations.notes |
files.cases.annotations.state |
files.cases.annotations.status |
files.cases.annotations.submitter_id |
files.cases.annotations.updated_datetime |
files.cases.demographic.created_datetime |
files.cases.demographic.demographic_id |
files.cases.demographic.ethnicity |
files.cases.demographic.gender |
files.cases.demographic.race |
files.cases.demographic.state |
files.cases.demographic.submitter_id |
files.cases.demographic.updated_datetime |
files.cases.demographic.year_of_birth |
files.cases.demographic.year_of_death |
files.cases.diagnoses.age_at_diagnosis |
files.cases.diagnoses.classification_of_tumor |
files.cases.diagnoses.created_datetime |
files.cases.diagnoses.days_to_birth |
files.cases.diagnoses.days_to_death |
files.cases.diagnoses.days_to_last_follow_up |
files.cases.diagnoses.days_to_last_known_disease_status |
files.cases.diagnoses.days_to_recurrence |
files.cases.diagnoses.diagnosis_id |
files.cases.diagnoses.last_known_disease_status |
files.cases.diagnoses.morphology |
files.cases.diagnoses.primary_diagnosis |
files.cases.diagnoses.prior_malignancy |
files.cases.diagnoses.progression_or_recurrence |
files.cases.diagnoses.site_of_resection_or_biopsy |
files.cases.diagnoses.state |
files.cases.diagnoses.submitter_id |
files.cases.diagnoses.tissue_or_organ_of_origin |

files.cases.diagnoses.tumor__grade |
files.cases.diagnoses.tumor__stage |
files.cases.diagnoses.updated__datetime |
files.cases.diagnoses.vital__status |
files.cases.diagnoses.treatments.created__datetime |
files.cases.diagnoses.treatments.days__to__treatment |
files.cases.diagnoses.treatments.state |
files.cases.diagnoses.treatments.submitter__id |
files.cases.diagnoses.treatments.therapeutic__agents |
files.cases.diagnoses.treatments.treatment__id |
files.cases.diagnoses.treatments.treatment__intent__type |
files.cases.diagnoses.treatments.treatment__or__therapy |
files.cases.diagnoses.treatments.updated__datetime |
files.cases.exposures.alcohol__history |
files.cases.exposures.alcohol__intensity |
files.cases.exposures.bmi |
files.cases.exposures.cigarettes__per__day |
files.cases.exposures.created__datetime |
files.cases.exposures.exposure__id |
files.cases.exposures.height |
files.cases.exposures.state |
files.cases.exposures.submitter__id |
files.cases.exposures.updated__datetime |
files.cases.exposures.weight |
files.cases.exposures.years__smoked |
files.cases.family__histories.created__datetime |
files.cases.family__histories.family__history__id |
files.cases.family__histories.relationship__age__at__diagnosis |
files.cases.family__histories.relationship__gender |
files.cases.family__histories.relationship__primary__diagnosis |
files.cases.family__histories.relationship__type |
files.cases.family__histories.relative__with__cancer__history |
files.cases.family__histories.state |
files.cases.family__histories.submitter__id |
files.cases.family__histories.updated__datetime |
files.cases.files.created__datetime |
files.cases.files.error__type |
files.cases.files.file__id |
files.cases.files.file__name |
files.cases.files.file__size |
files.cases.files.file__state |
files.cases.files.md5sum |
files.cases.files.state |
files.cases.files.state__comment |
files.cases.files.submitter__id |
files.cases.files.updated__datetime |
files.cases.project.dbgap__accession__number |
files.cases.project.disease__type |
files.cases.project.name |
files.cases.project.primary__site |
files.cases.project.project__id |
files.cases.project.released |
files.cases.project.state |
files.cases.project.program.dbgap__accession__number |
files.cases.project.program.name |
files.cases.project.program.program__id |
files.cases.samples.composition |
files.cases.samples.created__datetime |

files.cases.samples.current_weight |
files.cases.samples.days_to_collection |
files.cases.samples.days_to_sample_procurement |
files.cases.samples.freezing_method |
files.cases.samples.initial_weight |
files.cases.samples.intermediate_dimension |
files.cases.samples.is_ffpe |
files.cases.samples.longest_dimension |
files.cases.samples.oct_embedded |
files.cases.samples.pathology_report_uuid |
files.cases.samples.preservation_method |
files.cases.samples.sample_id |
files.cases.samples.sample_type |
files.cases.samples.sample_type_id |
files.cases.samples.shortest_dimension |
files.cases.samples.state |
files.cases.samples.submitter_id |
files.cases.samples.time_between_clamping_and_freezing |
files.cases.samples.time_between_excision_and_freezing |
files.cases.samples.tissue_type |
files.cases.samples.tumor_code |
files.cases.samples.tumor_code_id |
files.cases.samples.tumor_descriptor |
files.cases.samples.updated_datetime |
files.cases.samples.annotations.annotation_id |
files.cases.samples.annotations.case_id |
files.cases.samples.annotations.case_submitter_id |
files.cases.samples.annotations.category |
files.cases.samples.annotations.classification |
files.cases.samples.annotations.created_datetime |
files.cases.samples.annotations.creator |
files.cases.samples.annotations.entity_id |
files.cases.samples.annotations.entity_submitter_id |
files.cases.samples.annotations.entity_type |
files.cases.samples.annotations.legacy_created_datetime |
files.cases.samples.annotations.legacy_updated_datetime |
files.cases.samples.annotations.notes |
files.cases.samples.annotations.state |
files.cases.samples.annotations.status |
files.cases.samples.annotations.submitter_id |
files.cases.samples.annotations.updated_datetime |
files.cases.samples.portions.created_datetime |
files.cases.samples.portions.creation_datetime |
files.cases.samples.portions.is_ffpe |
files.cases.samples.portions.portion_id |
files.cases.samples.portions.portion_number |
files.cases.samples.portions.state |
files.cases.samples.portions.submitter_id |
files.cases.samples.portions.updated_datetime |
files.cases.samples.portions.weight |
files.cases.samples.portions.analytes.a260_a280_ratio |
files.cases.samples.portions.analytes.amount |
files.cases.samples.portions.analytes.analyte_id |
files.cases.samples.portions.analytes.analyte_type |
files.cases.samples.portions.analytes.analyte_type_id |
files.cases.samples.portions.analytes.concentration |
files.cases.samples.portions.analytes.created_datetime |
files.cases.samples.portions.analytes.spectrophotometer_method |

files.cases.samples.portions.analytes.state |
files.cases.samples.portions.analytes.submitter_id |
files.cases.samples.portions.analytes.updated_datetime |
files.cases.samples.portions.analytes.well_number |
files.cases.samples.portions.analytes.aliquots.aliquot_id |
files.cases.samples.portions.analytes.aliquots.amount |
files.cases.samples.portions.analytes.aliquots.analyte_type |
files.cases.samples.portions.analytes.aliquots.analyte_type_id |
files.cases.samples.portions.analytes.aliquots.concentration |
files.cases.samples.portions.analytes.aliquots.created_datetime |
files.cases.samples.portions.analytes.aliquots.source_center |
files.cases.samples.portions.analytes.aliquots.state |
files.cases.samples.portions.analytes.aliquots.submitter_id |
files.cases.samples.portions.analytes.aliquots.updated_datetime |
files.cases.samples.portions.analytes.aliquots.annotations.annotation_id |
files.cases.samples.portions.analytes.aliquots.annotations.case_id |
files.cases.samples.portions.analytes.aliquots.annotations.case_submitter_id |
files.cases.samples.portions.analytes.aliquots.annotations.category |
files.cases.samples.portions.analytes.aliquots.annotations.classification |
files.cases.samples.portions.analytes.aliquots.annotations.created_datetime |
files.cases.samples.portions.analytes.aliquots.annotations.creator |
files.cases.samples.portions.analytes.aliquots.annotations.entity_id |
files.cases.samples.portions.analytes.aliquots.annotations.entity_submitter_id |
files.cases.samples.portions.analytes.aliquots.annotations.entity_type |
files.cases.samples.portions.analytes.aliquots.annotations.legacy_created_datetime |
files.cases.samples.portions.analytes.aliquots.annotations.legacy_updated_datetime |
files.cases.samples.portions.analytes.aliquots.annotations.notes |
files.cases.samples.portions.analytes.aliquots.annotations.state |
files.cases.samples.portions.analytes.aliquots.annotations.status |
files.cases.samples.portions.analytes.aliquots.annotations.submitter_id |
files.cases.samples.portions.analytes.aliquots.annotations.updated_datetime |
files.cases.samples.portions.analytes.aliquots.center.center_id |
files.cases.samples.portions.analytes.aliquots.center.center_type |
files.cases.samples.portions.analytes.aliquots.center.code |
files.cases.samples.portions.analytes.aliquots.center.name |
files.cases.samples.portions.analytes.aliquots.center.namespace |
files.cases.samples.portions.analytes.aliquots.center.short_name |
files.cases.samples.portions.analytes.annotations.annotation_id |
files.cases.samples.portions.analytes.annotations.case_id |
files.cases.samples.portions.analytes.annotations.case_submitter_id |
files.cases.samples.portions.analytes.annotations.category |
files.cases.samples.portions.analytes.annotations.classification |
files.cases.samples.portions.analytes.annotations.created_datetime |
files.cases.samples.portions.analytes.annotations.creator |
files.cases.samples.portions.analytes.annotations.entity_id |
files.cases.samples.portions.analytes.annotations.entity_submitter_id |
files.cases.samples.portions.analytes.annotations.entity_type |
files.cases.samples.portions.analytes.annotations.legacy_created_datetime |
files.cases.samples.portions.analytes.annotations.legacy_updated_datetime |
files.cases.samples.portions.analytes.annotations.notes |
files.cases.samples.portions.analytes.annotations.state |
files.cases.samples.portions.analytes.annotations.status |
files.cases.samples.portions.analytes.annotations.submitter_id |
files.cases.samples.portions.analytes.annotations.updated_datetime |
files.cases.samples.portions.annotations.annotation_id |
files.cases.samples.portions.annotations.case_id |
files.cases.samples.portions.annotations.case_submitter_id |
files.cases.samples.portions.annotations.category |

files.cases.samples.portions.annotations.classification |
files.cases.samples.portions.annotations.created_datetime |
files.cases.samples.portions.annotations.creator |
files.cases.samples.portions.annotations.entity_id |
files.cases.samples.portions.annotations.entity_submitter_id |
files.cases.samples.portions.annotations.entity_type |
files.cases.samples.portions.annotations.legacy_created_datetime |
files.cases.samples.portions.annotations.legacy_updated_datetime |
files.cases.samples.portions.annotations.notes |
files.cases.samples.portions.annotations.state |
files.cases.samples.portions.annotations.status |
files.cases.samples.portions.annotations.submitter_id |
files.cases.samples.portions.annotations.updated_datetime |
files.cases.samples.portions.center.center_id |
files.cases.samples.portions.center.center_type |
files.cases.samples.portions.center.code |
files.cases.samples.portions.center.name |
files.cases.samples.portions.center.namespace |
files.cases.samples.portions.center.short_name |
files.cases.samples.portions.slides.created_datetime |
files.cases.samples.portions.slides.number_proliferating_cells |
files.cases.samples.portions.slides.percent_eosinophil_infiltration |
files.cases.samples.portions.slides.percent_granulocyte_infiltration |
files.cases.samples.portions.slides.percent_inflam_infiltration |
files.cases.samples.portions.slides.percent_lymphocyte_infiltration |
files.cases.samples.portions.slides.percent_monocyte_infiltration |
files.cases.samples.portions.slides.percent_necrosis |
files.cases.samples.portions.slides.percent_neutrophil_infiltration |
files.cases.samples.portions.slides.percent_normal_cells |
files.cases.samples.portions.slides.percent_stromal_cells |
files.cases.samples.portions.slides.percent_tumor_cells |
files.cases.samples.portions.slides.percent_tumor_nuclei |
files.cases.samples.portions.slides.section_location |
files.cases.samples.portions.slides.slide_id |
files.cases.samples.portions.slides.state |
files.cases.samples.portions.slides.submitter_id |
files.cases.samples.portions.slides.updated_datetime |
files.cases.samples.portions.slides.annotations.annotation_id |
files.cases.samples.portions.slides.annotations.case_id |
files.cases.samples.portions.slides.annotations.case_submitter_id |
files.cases.samples.portions.slides.annotations.category |
files.cases.samples.portions.slides.annotations.classification |
files.cases.samples.portions.slides.annotations.created_datetime |
files.cases.samples.portions.slides.annotations.creator |
files.cases.samples.portions.slides.annotations.entity_id |
files.cases.samples.portions.slides.annotations.entity_submitter_id |
files.cases.samples.portions.slides.annotations.entity_type |
files.cases.samples.portions.slides.annotations.legacy_created_datetime |
files.cases.samples.portions.slides.annotations.legacy_updated_datetime |
files.cases.samples.portions.slides.annotations.notes |
files.cases.samples.portions.slides.annotations.state |
files.cases.samples.portions.slides.annotations.status |
files.cases.samples.portions.slides.annotations.submitter_id |
files.cases.samples.portions.slides.annotations.updated_datetime |
files.cases.summary.file_count |
files.cases.summary.file_size |
files.cases.summary.data_categories.data_category |
files.cases.summary.data_categories.file_count |

files.cases.summary.experimental_strategies.experimental_strategy |
files.cases.summary.experimental_strategies.file_count |
files.cases.tissue_source_site.bcr_id |
files.cases.tissue_source_site.code |
files.cases.tissue_source_site.name |
files.cases.tissue_source_site.project |
files.cases.tissue_source_site.tissue_source_site_id |
files.center.center_id |
files.center.center_type |
files.center.code |
files.center.name |
files.center.namespace |
files.center.short_name |
files.downstream_analyses.analysis_id |
files.downstream_analyses.analysis_type |
files.downstream_analyses.created_datetime |
files.downstream_analyses.state |
files.downstream_analyses.submitter_id |
files.downstream_analyses.updated_datetime |
files.downstream_analyses.workflow_end_datetime |
files.downstream_analyses.workflow_link |
files.downstream_analyses.workflow_start_datetime |
files.downstream_analyses.workflow_type |
files.downstream_analyses.workflow_version |
files.downstream_analyses.output_files.access |
files.downstream_analyses.output_files.created_datetime |
files.downstream_analyses.output_files.data_category |
files.downstream_analyses.output_files.data_format |
files.downstream_analyses.output_files.data_type |
files.downstream_analyses.output_files.error_type |
files.downstream_analyses.output_files.experimental_strategy |
files.downstream_analyses.output_files.file_id |
files.downstream_analyses.output_files.file_name |
files.downstream_analyses.output_files.file_size |
files.downstream_analyses.output_files.file_state |
files.downstream_analyses.output_files.md5sum |
files.downstream_analyses.output_files.platform |
files.downstream_analyses.output_files.revision |
files.downstream_analyses.output_files.state |
files.downstream_analyses.output_files.state_comment |
files.downstream_analyses.output_files.submitter_id |
files.downstream_analyses.output_files.updated_datetime |
files.index_files.access |
files.index_files.created_datetime |
files.index_files.data_category |
files.index_files.data_format |
files.index_files.data_type |
files.index_files.error_type |
files.index_files.experimental_strategy |
files.index_files.file_id |
files.index_files.file_name |
files.index_files.file_size |
files.index_files.file_state |
files.index_files.md5sum |
files.index_files.platform |
files.index_files.revision |
files.index_files.state |
files.index_files.state_comment |

files.index__files.submitter__id |
files.index__files.updated__datetime |
files.metadata__files.access |
files.metadata__files.created__datetime |
files.metadata__files.data__category |
files.metadata__files.data__format |
files.metadata__files.data__type |
files.metadata__files.error__type |
files.metadata__files.file__id |
files.metadata__files.file__name |
files.metadata__files.file__size |
files.metadata__files.file__state |
files.metadata__files.md5sum |
files.metadata__files.state |
files.metadata__files.state__comment |
files.metadata__files.submitter__id |
files.metadata__files.type |
files.metadata__files.updated__datetime |
project.dbgap__accession__number |
project.disease__type |
project.name |
project.primary__site |
project.project__id |
project.released |
project.state |
project.program.dbgap__accession__number |
project.program.name |
project.program.program__id |
samples.composition |
samples.created__datetime |
samples.current__weight |
samples.days__to__collection |
samples.days__to__sample__procurement |
samples.freezing__method |
samples.initial__weight |
samples.intermediate__dimension |
samples.is__ffpe |
samples.longest__dimension |
samples.oct__embedded |
samples.pathology__report__uuid |
samples.preservation__method |
samples.sample__id |
samples.sample__type |
samples.sample__type__id |
samples.shortest__dimension |
samples.state |
samples.submitter__id |
samples.time__between__clamping__and__freezing |
samples.time__between__excision__and__freezing |
samples.tissue__type |
samples.tumor__code |
samples.tumor__code__id |
samples.tumor__descriptor |
samples.updated__datetime |
samples.annotations.annotation__id |
samples.annotations.case__id |
samples.annotations.case__submitter__id |
samples.annotations.category |

samples.annotations.classification |
samples.annotations.created_datetime |
samples.annotations.creator |
samples.annotations.entity_id |
samples.annotations.entity_submitter_id |
samples.annotations.entity_type |
samples.annotations.legacy_created_datetime |
samples.annotations.legacy_updated_datetime |
samples.annotations.notes |
samples.annotations.state |
samples.annotations.status |
samples.annotations.submitter_id |
samples.annotations.updated_datetime |
samples.portions.created_datetime |
samples.portions.creation_datetime |
samples.portions.is_ffpe |
samples.portions.portion_id |
samples.portions.portion_number |
samples.portions.state |
samples.portions.submitter_id |
samples.portions.updated_datetime |
samples.portions.weight |
samples.portions.analytes.a260_a280_ratio |
samples.portions.analytes.amount |
samples.portions.analytes.analyte_id |
samples.portions.analytes.analyte_type |
samples.portions.analytes.analyte_type_id |
samples.portions.analytes.concentration |
samples.portions.analytes.created_datetime |
samples.portions.analytes.spectrophotometer_method |
samples.portions.analytes.state |
samples.portions.analytes.submitter_id |
samples.portions.analytes.updated_datetime |
samples.portions.analytes.well_number |
samples.portions.analytes.aliquots.aliquot_id |
samples.portions.analytes.aliquots.amount |
samples.portions.analytes.aliquots.analyte_type |
samples.portions.analytes.aliquots.analyte_type_id |
samples.portions.analytes.aliquots.concentration |
samples.portions.analytes.aliquots.created_datetime |
samples.portions.analytes.aliquots.source_center |
samples.portions.analytes.aliquots.state |
samples.portions.analytes.aliquots.submitter_id |
samples.portions.analytes.aliquots.updated_datetime |
samples.portions.analytes.aliquots.annotations.annotation_id |
samples.portions.analytes.aliquots.annotations.case_id |
samples.portions.analytes.aliquots.annotations.case_submitter_id |
samples.portions.analytes.aliquots.annotations.category |
samples.portions.analytes.aliquots.annotations.classification |
samples.portions.analytes.aliquots.annotations.created_datetime |
samples.portions.analytes.aliquots.annotations.creator |
samples.portions.analytes.aliquots.annotations.entity_id |
samples.portions.analytes.aliquots.annotations.entity_submitter_id |
samples.portions.analytes.aliquots.annotations.entity_type |
samples.portions.analytes.aliquots.annotations.legacy_created_datetime |
samples.portions.analytes.aliquots.annotations.legacy_updated_datetime |
samples.portions.analytes.aliquots.annotations.notes |
samples.portions.analytes.aliquots.annotations.state |

samples.portions.analytes.aliquots.annotations.status |
samples.portions.analytes.aliquots.annotations.submitter__id |
samples.portions.analytes.aliquots.annotations.updated__datetime |
samples.portions.analytes.aliquots.center.center__id |
samples.portions.analytes.aliquots.center.center__type |
samples.portions.analytes.aliquots.center.code |
samples.portions.analytes.aliquots.center.name |
samples.portions.analytes.aliquots.center.namespace |
samples.portions.analytes.aliquots.center.short__name |
samples.portions.analytes.annotations.annotation__id |
samples.portions.analytes.annotations.case__id |
samples.portions.analytes.annotations.case__submitter__id |
samples.portions.analytes.annotations.category |
samples.portions.analytes.annotations.classification |
samples.portions.analytes.annotations.created__datetime |
samples.portions.analytes.annotations.creator |
samples.portions.analytes.annotations.entity__id |
samples.portions.analytes.annotations.entity__submitter__id |
samples.portions.analytes.annotations.entity__type |
samples.portions.analytes.annotations.legacy__created__datetime |
samples.portions.analytes.annotations.legacy__updated__datetime |
samples.portions.analytes.annotations.notes |
samples.portions.analytes.annotations.state |
samples.portions.analytes.annotations.status |
samples.portions.analytes.annotations.submitter__id |
samples.portions.analytes.annotations.updated__datetime |
samples.portions.annotations.annotation__id |
samples.portions.annotations.case__id |
samples.portions.annotations.case__submitter__id |
samples.portions.annotations.category |
samples.portions.annotations.classification |
samples.portions.annotations.created__datetime |
samples.portions.annotations.creator |
samples.portions.annotations.entity__id |
samples.portions.annotations.entity__submitter__id |
samples.portions.annotations.entity__type |
samples.portions.annotations.legacy__created__datetime |
samples.portions.annotations.legacy__updated__datetime |
samples.portions.annotations.notes |
samples.portions.annotations.state |
samples.portions.annotations.status |
samples.portions.annotations.submitter__id |
samples.portions.annotations.updated__datetime |
samples.portions.center.center__id |
samples.portions.center.center__type |
samples.portions.center.code |
samples.portions.center.name |
samples.portions.center.namespace |
samples.portions.center.short__name |
samples.portions.slides.created__datetime |
samples.portions.slides.number__proliferating__cells |
samples.portions.slides.percent__eosinophil__infiltration |
samples.portions.slides.percent__granulocyte__infiltration |
samples.portions.slides.percent__inflam__infiltration |
samples.portions.slides.percent__lymphocyte__infiltration |
samples.portions.slides.percent__monocyte__infiltration |
samples.portions.slides.percent__necrosis |
samples.portions.slides.percent__neutrophil__infiltration |

samples.portions.slides.percent__normal__cells |
samples.portions.slides.percent__stromal__cells |
samples.portions.slides.percent__tumor__cells |
samples.portions.slides.percent__tumor__nuclei |
samples.portions.slides.section__location |
samples.portions.slides.slide__id |
samples.portions.slides.state |
samples.portions.slides.submitter__id |
samples.portions.slides.updated__datetime |
samples.portions.slides.annotations.annotation__id |
samples.portions.slides.annotations.case__id |
samples.portions.slides.annotations.case__submitter__id |
samples.portions.slides.annotations.category |
samples.portions.slides.annotations.classification |
samples.portions.slides.annotations.created__datetime |
samples.portions.slides.annotations.creator |
samples.portions.slides.annotations.entity__id |
samples.portions.slides.annotations.entity__submitter__id |
samples.portions.slides.annotations.entity__type |
samples.portions.slides.annotations.legacy__created__datetime |
samples.portions.slides.annotations.legacy__updated__datetime |
samples.portions.slides.annotations.notes |
samples.portions.slides.annotations.state |
samples.portions.slides.annotations.status |
samples.portions.slides.annotations.submitter__id |
samples.portions.slides.annotations.updated__datetime |
summary.file__count |
summary.file__size |
summary.data__categories.data__category |
summary.data__categories.file__count |
summary.experimental__strategies.experimental__strategy |
summary.experimental__strategies.file__count |
tissue__source__site.bcr__id |
tissue__source__site.code |
tissue__source__site.name |
tissue__source__site.project |
tissue__source__site.tissue__source__site__id |

File Fields

Field Name
access
acl
created__datetime
data__category
data__format
data__type
error__type
experimental__strategy
file__id
file__name
file__size

Field Name

file_state

md5sum

origin

platform

revision

state

state_comment

submitter_id

tags

type

updated_datetime

analysis.analysis_id

analysis.analysis_type

analysis.created_datetime

analysis.state

analysis.submitter_id

analysis.updated_datetime

analysis.workflow_end_datetime

analysis.workflow_link

analysis.workflow_start_datetime

analysis.workflow_type

analysis.workflow_version

analysis.input_files.access

analysis.input_files.created_datetime

analysis.input_files.data_category

analysis.input_files.data_format

analysis.input_files.data_type

analysis.input_files.error_type

analysis.input_files.experimental_strategy

analysis.input_files.file_id

analysis.input_files.file_name

analysis.input_files.file_size

analysis.input_files.file_state

analysis.input_files.md5sum

analysis.input_files.platform

analysis.input_files.revision

analysis.input_files.state

analysis.input_files.state_comment

analysis.input_files.submitter_id

Field Name
analysis.input_files.updated_datetime
analysis.metadata.read_groups.adapter_name
analysis.metadata.read_groups.adapter_sequence
analysis.metadata.read_groups.base_caller_name
analysis.metadata.read_groups.base_caller_version
analysis.metadata.read_groups.created_datetime
analysis.metadata.read_groups.experiment_name
analysis.metadata.read_groups.flow_cell_barcode
analysis.metadata.read_groups.includes_spike_ins
analysis.metadata.read_groups.instrument_model
analysis.metadata.read_groups.is_paired_end
analysis.metadata.read_groups.library_name
analysis.metadata.read_groups.library_preparation_kit_catalog_number
analysis.metadata.read_groups.library_preparation_kit_name
analysis.metadata.read_groups.library_preparation_kit_vendor
analysis.metadata.read_groups.library_preparation_kit_version
analysis.metadata.read_groups.library_selection
analysis.metadata.read_groups.library_strand
analysis.metadata.read_groups.library_strategy
analysis.metadata.read_groups.platform
analysis.metadata.read_groups.read_group_id
analysis.metadata.read_groups.read_group_name
analysis.metadata.read_groups.read_length
analysis.metadata.read_groups.RIN
analysis.metadata.read_groups.sequencing_center
analysis.metadata.read_groups.sequencing_date
analysis.metadata.read_groups.size_selection_range
analysis.metadata.read_groups.spike_ins_concentration
analysis.metadata.read_groups.spike_ins_fasta
analysis.metadata.read_groups.state
analysis.metadata.read_groups.submitter_id
analysis.metadata.read_groups.target_capture_kit_catalog_number
analysis.metadata.read_groups.target_capture_kit_name
analysis.metadata.read_groups.target_capture_kit_target_region
analysis.metadata.read_groups.target_capture_kit_vendor
analysis.metadata.read_groups.target_capture_kit_version
analysis.metadata.read_groups.to_trim_adapter_sequence
analysis.metadata.read_groups.updated_datetime
analysis.metadata.read_groups.read_group_qcs.adapter_content

analysis.metadata.read__groups.read__group__qcs.basic__statistics

analysis.metadata.read__groups.read__group__qcs.created__datetime

analysis.metadata.read__groups.read__group__qcs.encoding

analysis.metadata.read__groups.read__group__qcs.fastq__name

analysis.metadata.read__groups.read__group__qcs.kmer__content

analysis.metadata.read__groups.read__group__qcs.overrepresented__sequences

analysis.metadata.read__groups.read__group__qcs.per__base__n__content

analysis.metadata.read__groups.read__group__qcs.per__base_sequence__content

analysis.metadata.read__groups.read__group__qcs.per__base_sequence__quality

analysis.metadata.read__groups.read__group__qcs.per__sequence__gc__content

analysis.metadata.read__groups.read__group__qcs.per__sequence__quality__score

analysis.metadata.read__groups.read__group__qcs.per__tile__sequence__quality

analysis.metadata.read__groups.read__group__qcs.percent__gc__content

analysis.metadata.read__groups.read__group__qcs.read__group__qc__id

analysis.metadata.read__groups.read__group__qcs.sequence__duplication__levels

analysis.metadata.read__groups.read__group__qcs.sequence__length__distribution

analysis.metadata.read__groups.read__group__qcs.state

analysis.metadata.read__groups.read__group__qcs.submitter__id

analysis.metadata.read__groups.read__group__qcs.total__sequences

analysis.metadata.read__groups.read__group__qcs.updated__datetime

analysis.metadata.read__groups.read__group__qcs.workflow__end__datetime

analysis.metadata.read__groups.read__group__qcs.workflow__link

analysis.metadata.read__groups.read__group__qcs.workflow__start__datetime

analysis.metadata.read__groups.read__group__qcs.workflow__type

analysis.metadata.read__groups.read__group__qcs.workflow__version

annotations.annotation__id

annotations.case__id

annotations.case__submitter__id

annotations.category

annotations.classification

annotations.created__datetime

annotations.creator

annotations.entity__id

annotations.entity__submitter__id

annotations.entity__type

annotations.legacy__created__datetime

annotations.legacy__updated__datetime

annotations.notes

annotations.state

Field Name
annotations.status
annotations.submitter_id
annotations.updated_datetime
archive.archive_id
archive.created_datetime
archive.data_category
archive.data_format
archive.data_type
archive.error_type
archive.file_name
archive.file_size
archive.file_state
archive.md5sum
archive.revision
archive.state
archive.state_comment
archive.submitter_id
archive.updated_datetime
associated_entities.case_id
associated_entities.entity_id
associated_entities.entity_submitter_id
associated_entities.entity_type
cases.aliquot_ids
cases.analyte_ids
cases.case_id
cases.created_datetime
cases.days_to_index
cases.portion_ids
cases.sample_ids
cases.slide_ids
cases.state
cases.submitter_aliquot_ids
cases.submitter_analyte_ids
cases.submitter_id
cases.submitter_portion_ids
cases.submitter_sample_ids
cases.submitter_slide_ids
cases.updated_datetime
cases.annotations.annotation_id

Field Name

cases.annotations.case_id

cases.annotations.case_submitter_id

cases.annotations.category

cases.annotations.classification

cases.annotations.created_datetime

cases.annotations.creator

cases.annotations.entity_id

cases.annotations.entity_submitter_id

cases.annotations.entity_type

cases.annotations.legacy_created_datetime

cases.annotations.legacy_updated_datetime

cases.annotations.notes

cases.annotations.state

cases.annotations.status

cases.annotations.submitter_id

cases.annotations.updated_datetime

cases.demographic.created_datetime

cases.demographic.demographic_id

cases.demographic.ethnicity

cases.demographic.gender

cases.demographic.race

cases.demographic.state

cases.demographic.submitter_id

cases.demographic.updated_datetime

cases.demographic.year_of_birth

cases.demographic.year_of_death

cases.diagnoses.age_at_diagnosis

cases.diagnoses.classification_of_tumor

cases.diagnoses.created_datetime

cases.diagnoses.days_to_birth

cases.diagnoses.days_to_death

cases.diagnoses.days_to_last_follow_up

cases.diagnoses.days_to_last_known_disease_status

cases.diagnoses.days_to_recurrence

cases.diagnoses.diagnosis_id

cases.diagnoses.last_known_disease_status

cases.diagnoses.morphology

cases.diagnoses.primary_diagnosis

cases.diagnoses.prior_malignancy

Field Name
cases.diagnoses.progression_or_recurrence
cases.diagnoses.site_of_resection_or_biopsy
cases.diagnoses.state
cases.diagnoses.submitter_id
cases.diagnoses.tissue_or_organ_of_origin
cases.diagnoses.tumor_grade
cases.diagnoses.tumor_stage
cases.diagnoses.updated_datetime
cases.diagnoses.vital_status
cases.diagnoses.treatments.created_datetime
cases.diagnoses.treatments.days_to_treatment
cases.diagnoses.treatments.state
cases.diagnoses.treatments.submitter_id
cases.diagnoses.treatments.therapeutic_agents
cases.diagnoses.treatments.treatment_id
cases.diagnoses.treatments.treatment_intent_type
cases.diagnoses.treatments.treatment_or_therapy
cases.diagnoses.treatments.updated_datetime
cases.exposures.alcohol_history
cases.exposures.alcohol_intensity
cases.exposures.bmi
cases.exposures.cigarettes_per_day
cases.exposures.created_datetime
cases.exposures.exposure_id
cases.exposures.height
cases.exposures.state
cases.exposures.submitter_id
cases.exposures.updated_datetime
cases.exposures.weight
cases.exposures.years_smoked
cases.family_histories.created_datetime
cases.family_histories.family_history_id
cases.family_histories.relationship_age_at_diagnosis
cases.family_histories.relationship_gender
cases.family_histories.relationship_primary_diagnosis
cases.family_histories.relationship_type
cases.family_histories.relative_with_cancer_history
cases.family_histories.state
cases.family_histories.submitter_id

Field Name
cases.family_histories.updated_datetime
cases.files.created_datetime
cases.files.error_type
cases.files.file_id
cases.files.file_name
cases.files.file_size
cases.files.file_state
cases.files.md5sum
cases.files.state
cases.files.state_comment
cases.files.submitter_id
cases.files.updated_datetime
cases.project.dbgap_accession_number
cases.project.disease_type
cases.project.name
cases.project.primary_site
cases.project.project_id
cases.project.released
cases.project.state
cases.project.program.dbgap_accession_number
cases.project.program.name
cases.project.program.program_id
cases.samples.composition
cases.samples.created_datetime
cases.samples.current_weight
cases.samples.days_to_collection
cases.samples.days_to_sample_procurement
cases.samples.freezing_method
cases.samples.initial_weight
cases.samples.intermediate_dimension
cases.samples.is_ffpe
cases.samples.longest_dimension
cases.samples.oct_embedded
cases.samples.pathology_report_uuid
cases.samples.preservation_method
cases.samples.sample_id
cases.samples.sample_type
cases.samples.sample_type_id
cases.samples.shortest_dimension

Field Name

cases.samples.state

cases.samples.submitter_id

cases.samples.time_between_clamping_and_freezing

cases.samples.time_between_excision_and_freezing

cases.samples.tissue_type

cases.samples.tumor_code

cases.samples.tumor_code_id

cases.samples.tumor_descriptor

cases.samples.updated_datetime

cases.samples.annotations.annotation_id

cases.samples.annotations.case_id

cases.samples.annotations.case_submitter_id

cases.samples.annotations.category

cases.samples.annotations.classification

cases.samples.annotations.created_datetime

cases.samples.annotations.creator

cases.samples.annotations.entity_id

cases.samples.annotations.entity_submitter_id

cases.samples.annotations.entity_type

cases.samples.annotations.legacy_created_datetime

cases.samples.annotations.legacy_updated_datetime

cases.samples.annotations.notes

cases.samples.annotations.state

cases.samples.annotations.status

cases.samples.annotations.submitter_id

cases.samples.annotations.updated_datetime

cases.samples.portions.created_datetime

cases.samples.portions.creation_datetime

cases.samples.portions.is_ffpe

cases.samples.portions.portion_id

cases.samples.portions.portion_number

cases.samples.portions.state

cases.samples.portions.submitter_id

cases.samples.portions.updated_datetime

cases.samples.portions.weight

cases.samples.portions.analytes.a260_a280_ratio

cases.samples.portions.analytes.amount

cases.samples.portions.analytes.analyte_id

cases.samples.portions.analytes.analyte_type

Field Name
cases.samples.portions.analytes.analyte_type_id
cases.samples.portions.analytes.concentration
cases.samples.portions.analytes.created_datetime
cases.samples.portions.analytes.spectrophotometer_method
cases.samples.portions.analytes.state
cases.samples.portions.analytes.submitter_id
cases.samples.portions.analytes.updated_datetime
cases.samples.portions.analytes.well_number
cases.samples.portions.analytes.aliquots.aliquot_id
cases.samples.portions.analytes.aliquots.amount
cases.samples.portions.analytes.aliquots.analyte_type
cases.samples.portions.analytes.aliquots.analyte_type_id
cases.samples.portions.analytes.aliquots.concentration
cases.samples.portions.analytes.aliquots.created_datetime
cases.samples.portions.analytes.aliquots.source_center
cases.samples.portions.analytes.aliquots.state
cases.samples.portions.analytes.aliquots.submitter_id
cases.samples.portions.analytes.aliquots.updated_datetime
cases.samples.portions.analytes.aliquots.annotations.annotation_id
cases.samples.portions.analytes.aliquots.annotations.case_id
cases.samples.portions.analytes.aliquots.annotations.case_submitter_id
cases.samples.portions.analytes.aliquots.annotations.category
cases.samples.portions.analytes.aliquots.annotations.classification
cases.samples.portions.analytes.aliquots.annotations.created_datetime
cases.samples.portions.analytes.aliquots.annotations.creator
cases.samples.portions.analytes.aliquots.annotations.entity_id
cases.samples.portions.analytes.aliquots.annotations.entity_submitter_id
cases.samples.portions.analytes.aliquots.annotations.entity_type
cases.samples.portions.analytes.aliquots.annotations.legacy_created_datetime
cases.samples.portions.analytes.aliquots.annotations.legacy_updated_datetime
cases.samples.portions.analytes.aliquots.annotations.notes
cases.samples.portions.analytes.aliquots.annotations.state
cases.samples.portions.analytes.aliquots.annotations.status
cases.samples.portions.analytes.aliquots.annotations.submitter_id
cases.samples.portions.analytes.aliquots.annotations.updated_datetime
cases.samples.portions.analytes.aliquots.center.center_id
cases.samples.portions.analytes.aliquots.center.center_type
cases.samples.portions.analytes.aliquots.center.code
cases.samples.portions.analytes.aliquots.center.name

Field Name
cases.samples.portions.analytes.aliquots.center.namespace
cases.samples.portions.analytes.aliquots.center.short__name
cases.samples.portions.analytes.annotations.annotation__id
cases.samples.portions.analytes.annotations.case__id
cases.samples.portions.analytes.annotations.case__submitter__id
cases.samples.portions.analytes.annotations.category
cases.samples.portions.analytes.annotations.classification
cases.samples.portions.analytes.annotations.created__datetime
cases.samples.portions.analytes.annotations.creator
cases.samples.portions.analytes.annotations.entity__id
cases.samples.portions.analytes.annotations.entity__submitter__id
cases.samples.portions.analytes.annotations.entity__type
cases.samples.portions.analytes.annotations.legacy__created__datetime
cases.samples.portions.analytes.annotations.legacy__updated__datetime
cases.samples.portions.analytes.annotations.notes
cases.samples.portions.analytes.annotations.state
cases.samples.portions.analytes.annotations.status
cases.samples.portions.analytes.annotations.submitter__id
cases.samples.portions.analytes.annotations.updated__datetime
cases.samples.portions.annotations.annotation__id
cases.samples.portions.annotations.case__id
cases.samples.portions.annotations.case__submitter__id
cases.samples.portions.annotations.category
cases.samples.portions.annotations.classification
cases.samples.portions.annotations.created__datetime
cases.samples.portions.annotations.creator
cases.samples.portions.annotations.entity__id
cases.samples.portions.annotations.entity__submitter__id
cases.samples.portions.annotations.entity__type
cases.samples.portions.annotations.legacy__created__datetime
cases.samples.portions.annotations.legacy__updated__datetime
cases.samples.portions.annotations.notes
cases.samples.portions.annotations.state
cases.samples.portions.annotations.status
cases.samples.portions.annotations.submitter__id
cases.samples.portions.annotations.updated__datetime
cases.samples.portions.center.center__id
cases.samples.portions.center.center__type
cases.samples.portions.center.code

Field Name
cases.samples.portions.center.name
cases.samples.portions.center.namespace
cases.samples.portions.center.short_name
cases.samples.portions.slides.created_datetime
cases.samples.portions.slides.number_proliferating_cells
cases.samples.portions.slides.percent_eosinophil_infiltration
cases.samples.portions.slides.percent_granulocyte_infiltration
cases.samples.portions.slides.percent_inflam_infiltration
cases.samples.portions.slides.percent_lymphocyte_infiltration
cases.samples.portions.slides.percent_monocyte_infiltration
cases.samples.portions.slides.percent_necrosis
cases.samples.portions.slides.percent_neutrophil_infiltration
cases.samples.portions.slides.percent_normal_cells
cases.samples.portions.slides.percent_stromal_cells
cases.samples.portions.slides.percent_tumor_cells
cases.samples.portions.slides.percent_tumor_nuclei
cases.samples.portions.slides.section_location
cases.samples.portions.slides.slide_id
cases.samples.portions.slides.state
cases.samples.portions.slides.submitter_id
cases.samples.portions.slides.updated_datetime
cases.samples.portions.slides.annotations.annotation_id
cases.samples.portions.slides.annotations.case_id
cases.samples.portions.slides.annotations.case_submitter_id
cases.samples.portions.slides.annotations.category
cases.samples.portions.slides.annotations.classification
cases.samples.portions.slides.annotations.created_datetime
cases.samples.portions.slides.annotations.creator
cases.samples.portions.slides.annotations.entity_id
cases.samples.portions.slides.annotations.entity_submitter_id
cases.samples.portions.slides.annotations.entity_type
cases.samples.portions.slides.annotations.legacy_created_datetime
cases.samples.portions.slides.annotations.legacy_updated_datetime
cases.samples.portions.slides.annotations.notes
cases.samples.portions.slides.annotations.state
cases.samples.portions.slides.annotations.status
cases.samples.portions.slides.annotations.submitter_id
cases.samples.portions.slides.annotations.updated_datetime
cases.summary.file_count

Field Name
cases.summary.file_size
cases.summary.data_categories.data_category
cases.summary.data_categories.file_count
cases.summary.experimental_strategies.experimental_strategy
cases.summary.experimental_strategies.file_count
cases.tissue_source_site.bcr_id
cases.tissue_source_site.code
cases.tissue_source_site.name
cases.tissue_source_site.project
cases.tissue_source_site.tissue_source_site_id
center.center_id
center.center_type
center.code
center.name
center.namespace
center.short_name
downstream_analyses.analysis_id
downstream_analyses.analysis_type
downstream_analyses.created_datetime
downstream_analyses.state
downstream_analyses.submitter_id
downstream_analyses.updated_datetime
downstream_analyses.workflow_end_datetime
downstream_analyses.workflow_link
downstream_analyses.workflow_start_datetime
downstream_analyses.workflow_type
downstream_analyses.workflow_version
downstream_analyses.output_files.access
downstream_analyses.output_files.created_datetime
downstream_analyses.output_files.data_category
downstream_analyses.output_files.data_format
downstream_analyses.output_files.data_type
downstream_analyses.output_files.error_type
downstream_analyses.output_files.experimental_strategy
downstream_analyses.output_files.file_id
downstream_analyses.output_files.file_name
downstream_analyses.output_files.file_size
downstream_analyses.output_files.file_state
downstream_analyses.output_files.md5sum

Field Name
downstream_analyses.output_files.platform
downstream_analyses.output_files.revision
downstream_analyses.output_files.state
downstream_analyses.output_files.state_comment
downstream_analyses.output_files.submitter_id
downstream_analyses.output_files.updated_datetime
index_files.access
index_files.created_datetime
index_files.data_category
index_files.data_format
index_files.data_type
index_files.error_type
index_files.experimental_strategy
index_files.file_id
index_files.file_name
index_files.file_size
index_files.file_state
index_files.md5sum
index_files.platform
index_files.revision
index_files.state
index_files.state_comment
index_files.submitter_id
index_files.updated_datetime
metadata_files.access
metadata_files.created_datetime
metadata_files.data_category
metadata_files.data_format
metadata_files.data_type
metadata_files.error_type
metadata_files.file_id
metadata_files.file_name
metadata_files.file_size
metadata_files.file_state
metadata_files.md5sum
metadata_files.state
metadata_files.state_comment
metadata_files.submitter_id
metadata_files.type

Field Name
metadata_files.updated_datetime

Annotation Fields

Field Name |
— |
annotation_id |
case_id |
case_submitter_id |
category |
classification |
created_datetime |
entity_id |
entity_submitter_id |
entity_type |
legacy_created_datetime |
legacy_updated_datetime |
notes |
state |
status |
submitter_id |
updated_datetime |
project.code |
project.dbgap_accession_number |
project.disease_type |
project.name |
project.primary_site |
project.program.dbgap_accession_number |
project.program.name |
project.program.program_id |
project.project_id |
project.released |
project.state |

Field Group Listing by Endpoint

Projects Field Groups

Field Group Name
program
summary
summary.data_categories
summary.experimental_strategies

Cases Field Groups

Field Group Name
annotations
demographic
diagnoses
diagnoses.treatments
exposures
family_histories
files
files.analysis
files.analysis.input_files
files.analysis.metadata
files.analysis.metadata.read_groups
files.analysis.metadata.read_groups.read_group_qcs
files.archive
files.cases
files.cases.annotations
files.cases.demographic
files.cases.diagnoses
files.cases.diagnoses.treatments
files.cases.exposures
files.cases.family_histories
files.cases.files
files.cases.project
files.cases.project.program
files.cases.samples
files.cases.samples.annotations
files.cases.samples.portions
files.cases.samples.portions.analytes
files.cases.samples.portions.analytes.aliquots
files.cases.samples.portions.analytes.aliquots.annotations
files.cases.samples.portions.analytes.aliquots.center
files.cases.samples.portions.analytes.annotations
files.cases.samples.portions.annotations
files.cases.samples.portions.center
files.cases.samples.portions.slides
files.cases.samples.portions.slides.annotations
files.cases.summary
files.cases.summary.data_categories
files.cases.summary.experimental_strategies
files.cases.tissue_source_site

Field Group Name
files.center
files.downstream_analyses
files.downstream_analyses.output_files
files.index_files
files.metadata_files
project
project.program
samples
samples.annotations
samples.portions
samples.portions.analytes
samples.portions.analytes.aliquots
samples.portions.analytes.aliquots.annotations
samples.portions.analytes.aliquots.center
samples.portions.analytes.annotations
samples.portions.annotations
samples.portions.center
samples.portions.slides
samples.portions.slides.annotations
summary
summary.data_categories
summary.experimental_strategies
tissue_source_site

Files Field Groups

Field Group Name
analysis
analysis.input_files
analysis.metadata
analysis.metadata.read_groups
analysis.metadata.read_groups.read_group_qcs
annotations
archive
associated_entities
cases
cases.annotations
cases.demographic
cases.diagnoses

Field Group Name
cases.diagnoses.treatments
cases.exposures
cases.family_histories
cases.files
cases.project
cases.project.program
cases.samples
cases.samples.annotations
cases.samples.portions
cases.samples.portions.analytes
cases.samples.portions.analytes.aliquots
cases.samples.portions.analytes.aliquots.annotations
cases.samples.portions.analytes.aliquots.center
cases.samples.portions.analytes.annotations
cases.samples.portions.annotations
cases.samples.portions.center
cases.samples.portions.slides
cases.samples.portions.slides.annotations
cases.summary
cases.summary.data_categories
cases.summary.experimental_strategies
cases.tissue_source_site
center
downstream_analyses
downstream_analyses.output_files
index_files
metadata_files

Annotations Field Groups

Field Group Name
project
project.program

Chapter 9

Appendix B: Key Terms

Appendix B: Key Terms

The following table provides definitions and explanations for terms and acronyms relevant to the content presented within this document.

Term	Definition
ACL	Access Control List
API	Application Programming Interface
BAM	Bidirectional Associative Memory
BCR	Biospecimen Core Resource
CSV	Comma Seperated Values
DCC	Data Coordinating Center
eRA	Electronic Research Administration
GDC	Genomic Data Commons
HTTP	Hypertext Transfer Protocol
IDE	Integrated Development Environment
JSON	JavaScript Object Notation
MAGE-TAB	Microarray Gene Expression - Tabular format
NCI	National Cancer Institute
NIH	National Institutes of Health
Op	Operations
REST	Representational State Transfer
TARGET	Therapeutically Applicable Research to Generate Effective Treatments
TCGA	The Cancer Genome Atlas
TSV	Tab Seperated Values
UUID	Universally Unique Identifier
URL	Universal Resource Locator
XML	Extensible Markup Language

Chapter 10

Appendix C: Format of Submission Queries and Responses

Appendix C: Format of Submission Requests and Responses

Format of Submission Request

The general format of JSON objects submitted to the GDC API is as follows:

```
1  {
2      "type": string,
3      "id": string,
4      "submitter_id": string,
5      "<properties>": any type,
6      "<relationship_name>": [
7          {
8              "id": string,
9              "submitter_id": string
10         },
11         ...
12     ]
13 }
```

The request must specify either an `id` or a `submitter_id`.

`id`: A string specifying the `id` of the node that the user is creating or updating. This is the persistent GDC UUID4 for the node. If it is preferred to refer to the node using a custom id, users can do so with the `submitter_id` field (described below).

`submitter_id`: A string specifying the custom id of the object the user is creating or updating. This is not the official GDC ID for the node.

`<properties>`: These key-value pairs will be applied to properties on the referenced node.

`<relationship_name>`: A JSON object that specifies a relationship (link) between the node and other nodes. Links are typically established using the `submitter_id` or `id` of the neighboring node.

Format of API Response to a Submission Request

The following fields are included in all API responses to submission requests.

```
1  {
2    "cases_related_to_created_entities_count": int,
3    "cases_related_to_updated_entities_count": int,
```

138

```
 4    "code": int,
 5    "created_entity_count": int,
 6    "entities": [entities],
 7    "entity_error_count": int,
 8    "message": string,
 9    "success": boolean,
10    "transaction_id": string,
11    "transactional_error_count": int,
12    "transactional_errors": [transactional_errors],
13    "updated_entity_count": int
14 }
```

cases_related_to_created_entities_count: Number of cases related to the created entities.

cases_related_to_updated_entities_count: Number of cases related to the updated entities.

code: The HTTP status code of the response message.

created_entity_count: Number of entities created.

entities: A list of entities of the form:

```
 1 {
 2    "action": string,
 3    "errors": [entity_errors],
 4    "id": string,
 5    "related_cases": [object],
 6    "type": string,
 7    "unique_keys": [unique_keys],
 8    "valid": boolean,
 9    "warnings": [object]
10 }
```

entity_errors: A list of errors that occurred while parsing, validating, or performing a CRUD operation on a specific entity. Entity errors are of the form:

```
1 {
2     "keys": [string],
3     "message": string
4 }
```

unique_keys: Properties, or combinations of properties, that can be used to uniquely identify the node in the GDC. Unique_keys are of the form:

```
1 {
2     "project_id": string,
3     "submitter_id": string
4 }
```

entity_error_count: Number of entities that were not successful.

message: A human-readable message describing the transaction.

success: A boolean value stating whether the transaction was successful. If the value is False, then no changes will be made to the database.

transaction_id: A string specifying the transaction id.

transactional_error_count: Number of transactional errors that occurred.

transactional_errors: A list of transactional errors that have occurred. These errors are errors that are not specific to an individual entity. Transactional errors are of the form:

```
1 {
2     "message": string
3 }
```

`updated_entity_count`: Number of existing entities updated by the transaction.

Error Types

`EntityNotFoundError` A referenced entity was not found among existing entities and entities specified in the transaction.

`MissingPropertyError` A required property was not provided.

`ValidationError` A provided property did not pass a validation test.

Status Messages

API responses will contain a status for each entity specified in the request:

`success`: The desired transaction was sucessful and the entity's state was modified in the GDC.

`valid`: The desired transaction was not sucessful, but the trasaction was not aborted because of this entity.

`error`: The desired transaction was not sucessful, and the transaction was aborted because of this entity. This entity did not pass validation or an internal error occured when attempting to complete the transaction. The error state will be accompanied by a list of errors recorded about the entity (see label-error-messages).

Note: GDC API requests are transactional. An error with processing a node specified in the transaction will abort the transaction and will result in no changes being applied for any node involved in the transaction.

Chapter 11

Release Notes

API Release Notes

v1.5.0

- **GDC Product**: Application Programming Interface (API)
- **Release Date**: October 31, 2016

New Features and Changes

- API responds with an error when the request specifies an unsupported combination of `filters` and `facets`.
- In TSV submissions, trailing and leading whitespace, including non-ASCII whitespace characters, are stripped from property names and values.
- For released projects, any updates to previously submitted entities (i.e. `"state": "submitted"`) will be included in the following GDC data release.
- Performance improvements for manifest generation.

Bugs Fixed Since Last Release

- Uploading certain unsupported metadata files caused the associated submission transactions to remain stuck in pending state.

Known Issues and Workarounds

- Fields are not counted as missing if parent field is also missing. This may occur with queries of nested fields in the Data Portal Advanced Search or an API query using a filter. This behavior could impact results reported using search parameters of "IS MISSING" or "NOT MISSING".
- Certain very large API requests will time out. It is recommended to break up very large requests into a series of smaller requests.

v1.4.0

- **GDC Product**: Application Programming Interface (API)
- **Release Date**: September 23, 2016

New Features and Changes

- Submission transaction log includes additional information to assist in tracking.
- Submission project state transitions are disabled temporarily while project release features are being improved.
- GDC data dictionary changes:
  - The `submittable` property was added to all entity types in the GDC data model. It indicates whether the entity type can be submitted by users.
  - Category of Read Group entities in the GDC Data Model has changed from `data_bundle` to `biospecimen`.
  - Analyte entities support an expanded set of `analyte_type` values.

Bugs Fixed Since Last Release

- None to report

Known Issues and Workarounds

- API search & retrieval queries that do not include a `sort` parameter may return results in different order each time they are executed. This is a particular problem for paginated responses (i.e. responses to queries for which the number of results is greater than the `size` parameter).
  - **Workaround:** Include a `sort` parameter in API search & retrieval queries.
- Fields are not counted as missing if parent field is also missing. This may occur with queries of nested fields in the Data Portal Advanced Search or an API query using a filter. This behavior could impact results reported using search parameters of "IS MISSING" or "NOT MISSING".
- Certain very large API requests will time out. It is recommended to break up very large requests into a series of smaller requests.

v1.3.1

- **GDC Product**: Application Programming Interface (API)
- **Release Date**: September 7, 2016

New Features and Changes

- Successful `_dry_run` submission transactions can be committed to the GDC data model without having to re-upload metadata. The transactions can also be closed to prevent them from being committed in the future.
- Submission transactions can be submitted in asynchronous mode. In this mode, the GDC API will issue an immediate acknowledgement of the transaction, along with the `transaction_id`. The status of the transaction can be verified by the user at a later time by specifying the `transaction_id`. Users submitting large transactions may find this mode helpful.
- New submission transaction properties can be queried with GraphQL
- GDC Data Dictionary changes:
  - Clinical Supplement entities can have `data_format` set to OMF.
  - Biospecimen Supplement entities can have `data_format` set to SSF or PPS.
  - Read group `instrument_model` can be set to "Illumina HiSeq 4000".
  - Category of Slide entities in the GDC Data Model has changed from `data_bundle` to `biospecimen`.

Bugs Fixed Since Last Release

- Incorrect BMI calculation in the import of BCR XML files.

Known Issues and Workarounds

- API search & retrieval queries that do not include a `sort` parameter may return results in different order each time they are executed. This is a particular problem for paginated responses (i.e. responses to queries for which the number of results is greater than the `size` parameter).
  - **Workaround:** Include a `sort` parameter in API search & retrieval queries.
- Fields are not counted as missing if parent field is also missing. This may occur with queries of nested fields in the Data Portal Advanced Search or an API query using a filter. This behavior could impact results reported using search parameters of "IS MISSING" or "NOT MISSING".
- Certain very large API requests will time out. It is recommended to break up very large requests into a series of smaller requests.

v1.2.0

- **GDC Product**: Application Programming Interface (API)
- **Release Date**: August 9, 2016

New Features and Changes

- Tarballs generated by the `data` endpoint in response to multi-file data download requests now include a folder structure that puts each file in a folder whose name is the file's UUID.
- UUIDs in clinical XML files are no longer treated in a case-sensitive way by the `submission` endpoint.
- Improved performance of `submission` endpoint for transactions that include many cases.
- Speed improvements for the `submission` endpoint.
- BCR XML is no longer validated against its XSD at submission.

Bugs Fixed Since Last Release

- Fixed handling of `POST` requests to address problems with cart functionality in older versions of Firefox
- Files of category `related_files` can now be downloaded from the `data` endpoint.
- Allowed submission by projects in certain dbGaP registration states that were previously blocked.

Known Issues and Workarounds

- API search & retrieval queries that do not include a `sort` parameter may return results in different order each time they are executed. This is a particular problem for paginated responses (i.e. responses to queries for which the number of results is greater than the `size` parameter).
  - **Workaround:** Include a `sort` parameter in API search & retrieval queries.
- Fields are not counted as missing if parent field is also missing. This may occur with queries of nested fields in the Data Portal Advanced Search or an API query using a filter. This behavior could impact results reported using search parameters of "IS MISSING" or "NOT MISSING".
- Certain very large API requests will time out. It is recommended to break up very large requests into a series of smaller requests.

v1.1.0

- **GDC Product**: Application Programming Interface (API)
- **Release Date**: May 25, 2016

New Features and Changes

- BAM index files (.bai) are now automatically downloaded with parent BAM.

Bugs Fixed Since Last Release

- None to report

Bugs Fixed Since Last Release

- Sorting by file `submitter_id` no longer causes an internal server error
- BAM index files are now included with harmonized BAM files
- Certain very long API requests will time out. It is recommended to break up into a series of smaller requests.

Known Issues and Workarounds

- Fields are not counted as missing if parent field is also missing. This may occur with queries of nested fields in the Data Portal Advanced Search or an API query using a filter. This behavior could impact results reported using search parameters of "IS MISSING" or "NOT MISSING".
- Certain very large API requests will time out. It is recommended to break up very large requests into a series of smaller requests.

v1.0.1

- **GDC Product**: Application Programming Interface (API)
- **Release Date**: May 16, 2016

New Features and Changes

- HTTP interface that uses JSON as the primary data exchange format
- Programmatic access to functionality provided by GDC Data and Submission portals, via `projects`, `cases`, `files`, `annotations`, `data`, `slicing`, `status`, and `submission` endpoints
- Programmatic access to GDC Legacy Archive via `legacy` endpoint
- Token-based authentication for secure access to controlled data and to submission functionality
- RESTful search that supports simple and complex queries via `filters`, `fields`, and `facets` parameters, and `project`, `files`, `cases`, and `annotations` endpoints.
- Search results can be sorted using `sort` parameter, paginated using `size` and `from` parameters, and output in JSON, TSV, and XML using `format` and `pretty` parameters.
- `_mapping` endpoint enables user discovery of fields available for data search and retrieval operations
- Support for downloading of individual files and of archives containing multiple files
- Generation of download and upload manifests for use with the GDC Data Transfer Tool
- BAM slicing functionality for downloading part(s) of a BAM file specified using chromosomal coordinates or HGNC gene names
- Transactional submission system that links individual data elements according to a graph-based GDC Data Model
- Two data entity identifiers: UUIDs, which are consistent across GDC, and Submitter IDs, for compatibility with submitters' tracking systems

Bugs Fixed Since Last Release

- None to report

Known Issues and Workarounds

- Use of non-ascii characters in token passed to Data Transfer Tool will produce incorrect error message "Internal server error: Auth service temporarily unavailable".
- Use of a decimal in an integer search field produces unexpected error.
- Certain very large API requests will time out. It is recommended to break up very large requests into a series of smaller requests.