

How to Use the CAS Explorer

Important: All data shown in the screenshots is just for illustration. Valid records will be available in the live CAS Explorer.

Contents

How to Use the CAS Explorer.....	1
1. Introduction	2
2. Tables	4
3. Columns	7
4. Subjects.....	9
5. Tumour types.....	10
6. Locations.....	12
7. Time Periods	13
8. Final remarks	14

1. Introduction

The CAS Explorer enables researchers and analysts to get a better understanding of the Cancer Analysis Service (CAS) database within the National Cancer Registration and Analysis Service (NCRAS) in Public Health England (PHE)

CAS Explorer shows the completeness of the NCRAS database, and is designed to be used by researchers and analysts to check the data quality of the table fields before using the Simulacrum or requesting data through the Office of Data Release (ODR) within PHE.

Currently the CAS Explorer includes statistics for the latest CAS snapshot.

The sections of the site show:

Tables - can be used to find column completeness filtered by year and tumour type.

Columns - indicates which columns appear in multiple tables.

Subjects - shows which tables and columns relate to subject areas, for example, treatments, diagnostics and geographies.

Type of Tumour - shows the columns relating to each tumour site.

Location - gives an outline of completeness by cancer registry.

Time Periods - displays the completeness of all the tables aggregated by year.



CAS Explorer shows the completeness of the [NCRAS](#) database, and is designed to be used by researchers and analysts to check the data quality of the table fields before submitting FOI requests to [ODR](#).

For more general information about cancer statistics please visit the [CancerData](#) website. [[see notes](#)]

Currently the CAS Explorer includes statistics for the latest CAS snapshot.

The sections of the site show:

[Tables](#) - can be used to find column completeness filtered by year and tumour type.

[Columns](#) - indicates which columns appear in multiple tables.

[Subjects](#) - shows which tables and columns relate to subject areas, for example, treatments, diagnostics and geographies.

[Type of Tumour](#) - shows the columns relating to each tumour site.

[Location](#) - gives an outline of completeness by cancer registry.

[Time Periods](#) - displays the completeness of all the tables aggregated by year.

Selecting data

Most sections of the site have filter options at the top of the page to specify the extent of the cohort.

The 'Analysis Cohort' option applies the exclusions generally used by analysts and researchers (link to exclusion criteria). The option to turn this off is intended for use by registry staff.

The 'Year' filter under the Tables section is used to get accurate completeness statistics. For many columns completeness has improved markedly for recent data, which means aggregating over all years gives a misleadingly low figure.

Within 'Tumour types' multiple options can be selected to get completeness data for related cancer sites.

Simulacrum - Under the CAS table pages there is a flag indicating which columns are included in the project. Sort by the **'In Simulacrum'** column heading to group these together.

Caveats

Some columns include codes for 'unknown' values. Currently these will be counted as complete because the record has a non null value. This issue will be resolved in a future version of the project.

Some of the columns are tumour specific. In views where all tumours are aggregated, the completeness value displayed will be too low. Use the tumour type filter to obtain accurate values.

* Note that the CAS Explorer data numbers might differ from the CancerData site due to factors such as the exclusion criteria and the dates at which snapshots are taken.

Figure 1 CAS Explorer landing page

2. Tables

The Tables Index page allows you to check the global completeness and the number of rows of each table as seen in the Figure 2.

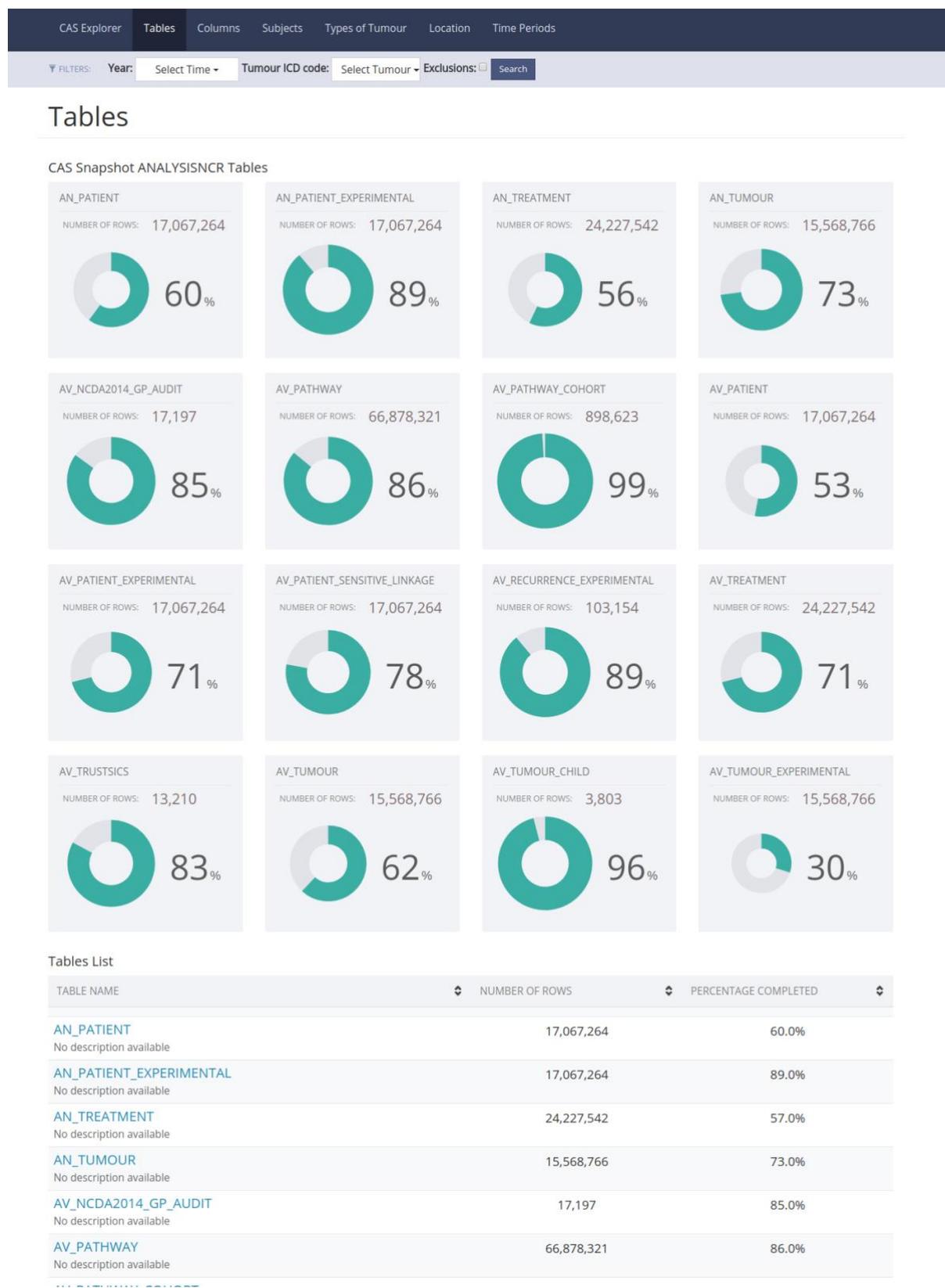


Figure 2 Tables Index Page

The statistics can be filtered by year (Figure 3) and ICD code (Figure 4). There is also a check box on *Exclusions*. This check box removes from the statistics all the records that are not verified under CAS' Standard Operating Procedures.

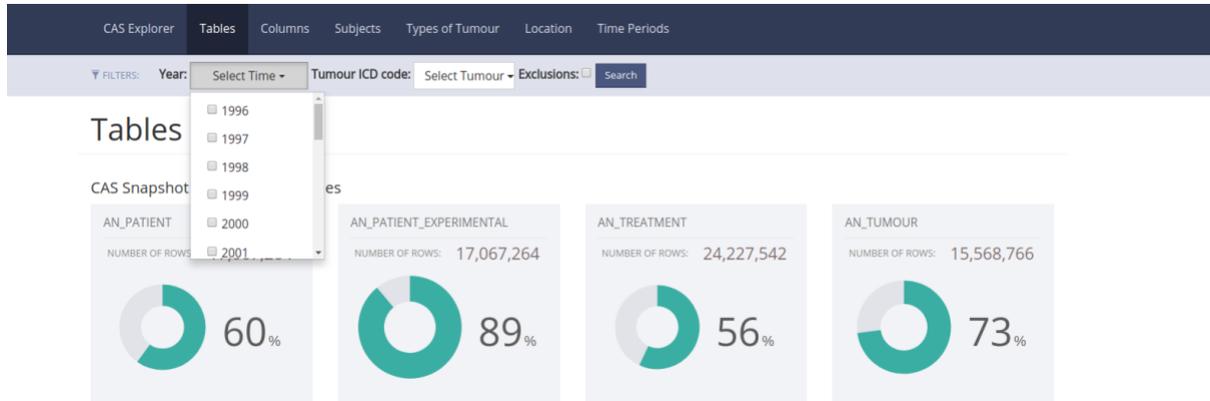


Figure 3 Year dropdown menu

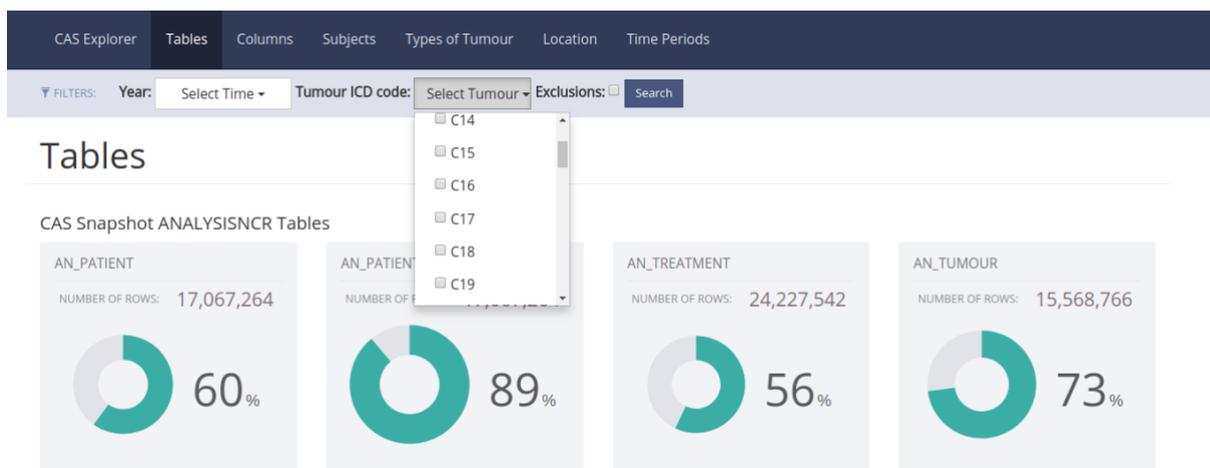


Figure 4 ICD dropdown menu

Each table has its own page with a text description and the number of rows. There is a graph with the global completeness of the table and a list of all columns.

It is possible to check which columns are in Simulacrum, the total number of records, the number of completed records and the percentage of records that are completed (Figure 5).

Tables / AN_PATIENT

AN_PATIENT

Maecenas tempor ligula quis vehicula convallis. Curabitur ullamcorper velit eu quam facilisis, sit amet hendrerit augue mattis. Dummy text.



Related Columns

COLUMN NAME	IN SIMULACRUM	TOTAL RECORDS	COMPLETED RECORDS	PERCENTAGE
PATIENTID	✓	17,067,264.0	17,067,264.0	100.0%
PERSONID		17,067,264.0	17,067,264.0	100.0%
ALIASFLAG		17,067,264.0	17,067,264.0	100.0%
MONTH_DOB		17,067,264.0	16,985,977.0	99.5%
YEAR_DOB		17,067,264.0	16,985,977.0	99.5%
SEX	✓	17,067,264.0	15,603,861.0	91.4%
ETHNICITY	✓	17,067,264.0	10,318,790.0	60.5%
ETHNICITYNAME		17,067,264.0	10,318,790.0	60.5%

Figure 5 AN_Tumour table page

3. Columns

The Columns section provides a list of all columns. It is possible to check if they are in the Simulacrum dataset, which tables include them and the global completeness percentage in all tables.

COLUMN NAME	IN SIMULACRUM	TABLES	PERCENTAGE COMPLETED
ACE27		AV_TUMOUR_EXPERIMENTAL	1.0%
AGE		AV_TUMOUR AV_TREATMENT AV_TUMOUR_CHILD AN_TUMOUR	98.0%
ALIASFLAG		AV_PATIENT AN_PATIENT	100.0%
AVPID		AV_PATHWAY	100.0%
BASISOFDIAGNOSIS		AN_TUMOUR AV_TUMOUR	100.0%
BEHAVIOUR_CODED		AV_TUMOUR AN_TUMOUR	97.0%
BEHAVIOUR_CODED_DESC		AV_TUMOUR AN_TUMOUR	97.0%
BEHAVIOUR_ICD10_O2		AV_TUMOUR AN_TUMOUR	82.0%
BIGTUMOURCOUNT		AV_PATIENT AN_PATIENT	100.0%
BIRTHDATE1		AV_PATIENT	99.0%
BIRTHDATE2		AV_PATIENT	99.0%

Figure 6 Columns Index Page

Every column has its own page with the description, type, status and valid content.

There is a graph showing the global completeness of all tables and a list of tables where that column is. It is possible to read the number of records, the completed records and completeness percentage for each table (Figure 7).

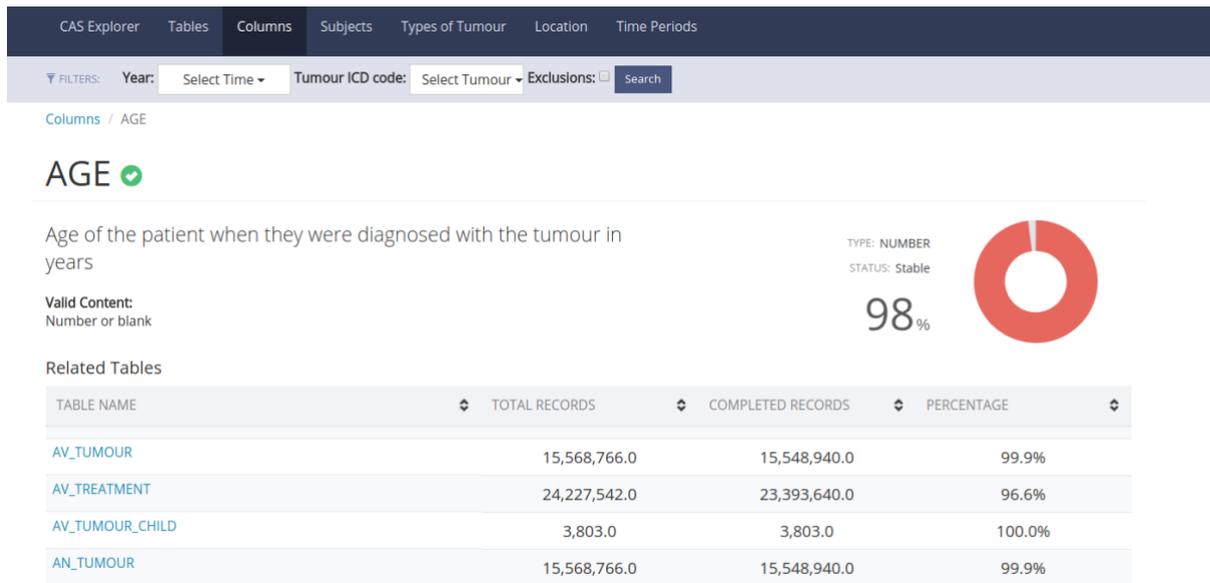


Figure 7 Age Column page

On the column page it is possible to filter according to year, ICD code and Exclusions (Figure 8) as in the tables page.

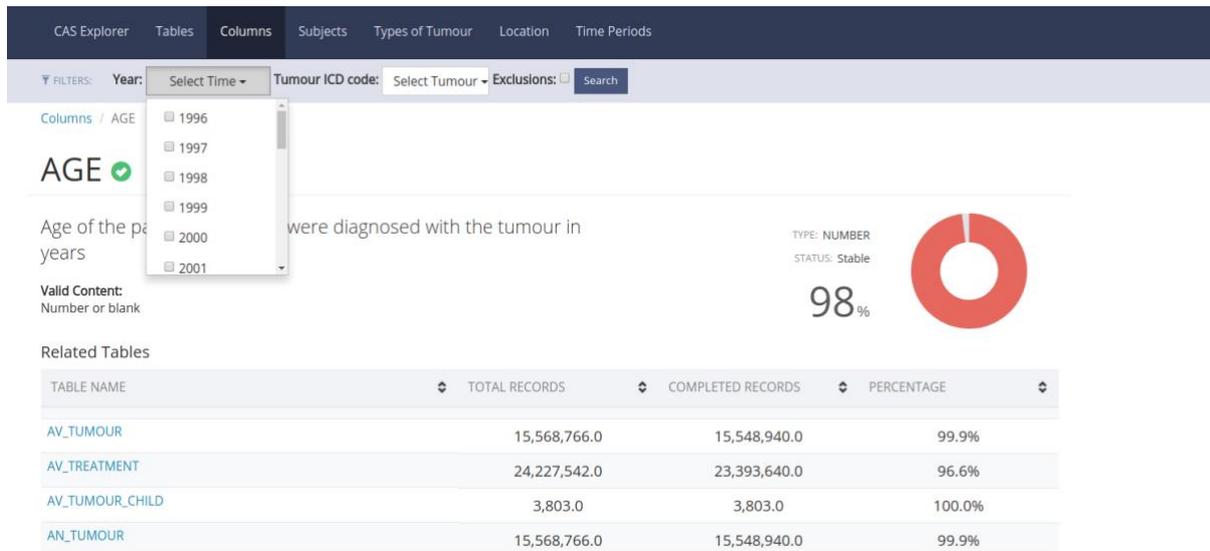


Figure 8 Filters on Age Column page

4. Subjects

In order to understand the data available on CAS the columns are grouped by subjects (e.g. patient personal details, death details, etc.).

The subject page (Figure 9) lists all of them and it is possible to click on them to go to the list of column per page.

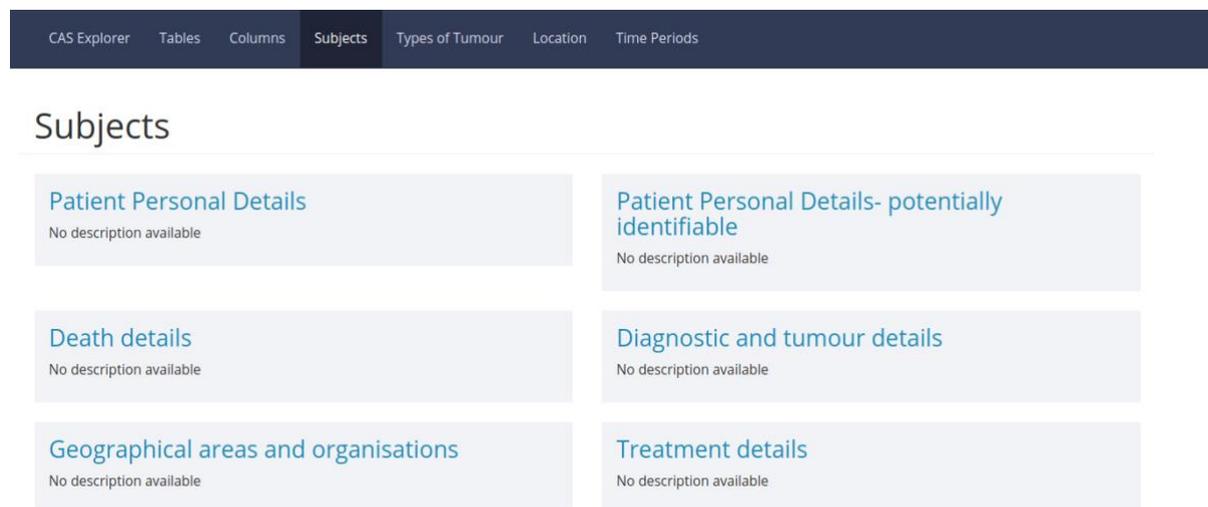


Figure 9 Subjects index page

For each subject it is possible to find the columns and tables which are related to it (Figure 10).

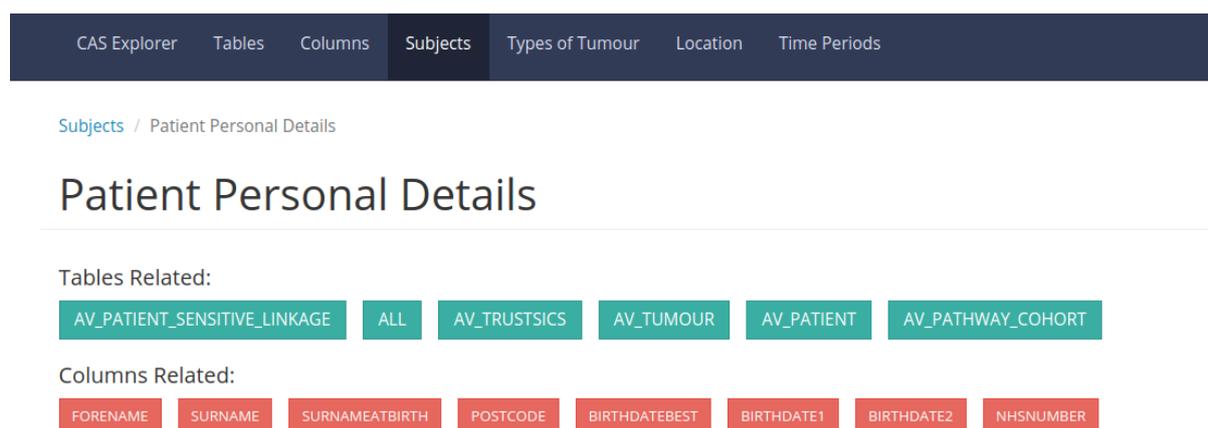


Figure 10 Patient personal details tables and columns

5. Tumour types

It is possible to browse through different tumour types and get completeness statistics (Figure 11).

The screenshot shows the 'Types of Tumour' page in the CAS Explorer interface. The page has a dark blue header with navigation tabs: 'CAS Explorer', 'Tables', 'Columns', 'Subjects', 'Types of Tumour' (selected), 'Location', and 'Time Periods'. Below the header is a search bar with 'Filters', 'Exclusions', and a 'Search' button. The main content area is titled 'Tumour Types' and displays 'Total Completeness: %'. A table lists various tumour types with their corresponding completeness percentages. The 'COMPLETENESS' column header is highlighted with a red border.

COLUMN NAME	COMPLETENESS
KAPOSIS SARCOMA	67%
PLACENTA	67%
SPINAL CORD, CRANIAL NERVES AND OTHER PARTS OF CENTRAL NERVOUS SYSTEM	67%
OTHER AND UNSPECIFIED MALIGNANT NEOPLASMS OF LYMPHOID, HAEMATOPOIETIC AND RELATED TISSUE	67%
OTHER AND ILL-DEFINED DIGESTIVE ORGANS	68%
LIVER AND INTRAHEPATIC BILE DUCTS	68%
TRACHEA	68%
OTHER AND ILL-DEFINED SITES IN THE RESPIRATORY SYSTEM AND INTRATHORACIC ORGANS	68%
MESOTHELIOMA	68%

Figure 11 Tumour types index page

It is possible to click in each tumour type and navigate to its page which includes a description and a list of all columns and pages. The information available for each column includes number of records, completed records and completeness percentage.

CAS Explorer Tables Columns Subjects **Types of Tumour** Location Time Periods

▼ FILTERS Exclusions: Search

Tumour types / Colon

Colon

Tumour description here? Vestibulum dignissim euismod egestas. Curabitur a imperdiet ligula.

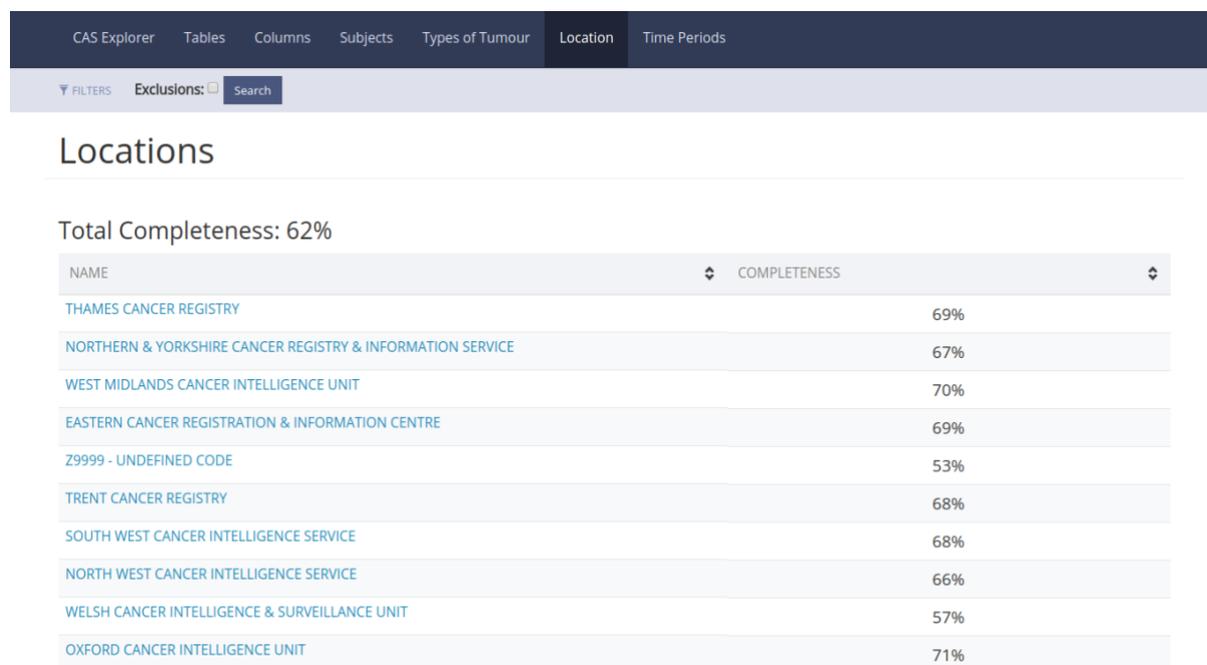
Related Columns:

COLUMN NAME	TOTAL RECORDS	COMPLETED RECORDS	PERCENTAGE
LATERALITY / TABLE: AV_TUMOUR Tumour laterality identifies the side of the body for a tumour relating to paired organs	15,568,766	15,568,766	100.0%
BASISOFDIAGNOSIS / TABLE: AV_TUMOUR Basis of diagnosis of the tumour, according to all data received by the registry	15,568,766	15,568,766	100.0%
BIRTHDATEBEST / TABLE: AV_TUMOUR Date of birth of the patient	15,568,766	15,548,940	99.9%
AGE / TABLE: AV_TUMOUR Age of the patient when they were diagnosed with the tumour in years	15,568,766	15,548,940	99.9%
FIVEYEARAGEBAND / TABLE: AV_TUMOUR Five-year band of age at diagnosis	15,568,766	15,568,766	100.0%
SEX / TABLE: AV_TUMOUR Sex of the patient when the tumour was diagnosed	15,568,766	15,568,676	100.0%

Figure 12 Colon tumour page

6. Locations

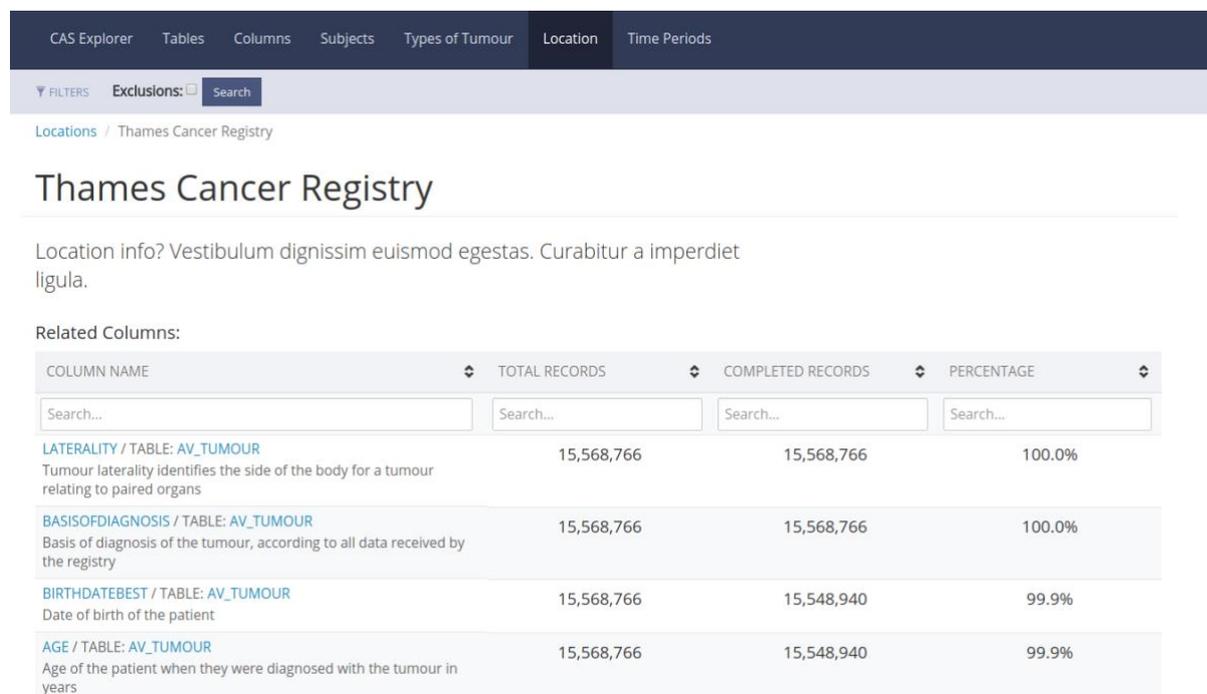
As with tumour types it is possible to find the global completeness statistics per a given location. Currently CAS Explorer has data according to Cancer Registries (Figure 13).



NAME	COMPLETENESS
THAMES CANCER REGISTRY	69%
NORTHERN & YORKSHIRE CANCER REGISTRY & INFORMATION SERVICE	67%
WEST MIDLANDS CANCER INTELLIGENCE UNIT	70%
EASTERN CANCER REGISTRATION & INFORMATION CENTRE	69%
Z9999 - UNDEFINED CODE	53%
TRENT CANCER REGISTRY	68%
SOUTH WEST CANCER INTELLIGENCE SERVICE	68%
NORTH WEST CANCER INTELLIGENCE SERVICE	66%
WELSH CANCER INTELLIGENCE & SURVEILLANCE UNIT	57%
OXFORD CANCER INTELLIGENCE UNIT	71%

Figure 13 Locations index page

By clicking in each location it is possible to find the total records, completed records and completeness percentage per column (Figure 14).



Location info? Vestibulum dignissim euismod egestas. Curabitur a imperdiet ligula.

Related Columns:

COLUMN NAME	TOTAL RECORDS	COMPLETED RECORDS	PERCENTAGE
LATERALITY / TABLE: AV_TUMOUR Tumour laterality identifies the side of the body for a tumour relating to paired organs	15,568,766	15,568,766	100.0%
BASISOFDIAGNOSIS / TABLE: AV_TUMOUR Basis of diagnosis of the tumour, according to all data received by the registry	15,568,766	15,568,766	100.0%
BIRTHDATEBEST / TABLE: AV_TUMOUR Date of birth of the patient	15,568,766	15,548,940	99.9%
AGE / TABLE: AV_TUMOUR Age of the patient when they were diagnosed with the tumour in years	15,568,766	15,548,940	99.9%

Figure 14 Thames cancer registry page

7. Time Periods

Currently CAS Explorer allows the users to browse data per year starting in 1970. In the main page (Figure 15) it is possible to find the global completeness for each year.

PERIOD NAME	COMPLETENESS
2017	Not Available
2016	Not Available
2015	73%
2014	73%
2013	73%
2012	73%
2011	72%
2010	71%
2009	70%
2008	70%

Figure 15 Time periods index page

When clicking on an year CAS Explorer will display its page with the same statistics per column as in Tumour Types and Locations (Figure 16).

Periods / 2013

2013

Description here?

Related Columns:

COLUMN NAME	TOTAL RECORDS	COMPLETED RECORDS	PERCENTAGE
LATERALITY / TABLE: AV_TUMOUR Tumour laterality identifies the side of the body for a tumour relating to paired organs	15,568,766	15,568,766	100.0%
BASISOFDIAGNOSIS / TABLE: AV_TUMOUR Basis of diagnosis of the tumour, according to all data received by the registry	15,568,766	15,568,766	100.0%
BIRTHDATEBEST / TABLE: AV_TUMOUR Date of birth of the patient	15,568,766	15,548,940	99.9%
AGE / TABLE: AV_TUMOUR Age of the patient when they were diagnosed with the tumour in years	15,568,766	15,548,940	99.9%

Figure 16 2013 PAGE

8. Final remarks

CAS Explorer has available all the features mentioned above. Currently the statistics are generic and do not cover columns specific to a tumour.

Filters are only available for the Tumour tables but should be expanded in the future.

CAS Explorer is in continuous development and in the future will have new features.