

## Appendix 8: Guide to Data, Data Cleaning and Measurement

In this section we'll discuss how to ensure that your data are collected and stored correctly in the clinical software and how to use Pen CAT to support work outlined in this toolkit.

If you have not already done so, please contact your pathology providers and ensure that they are sending results to you electronically and in the correct format (HL7). Some pathology providers will have preferences set for your Health Service and it's essential that these are set correctly. If you are receiving pathology results electronically in the right format your software will process these results into the collect location and no human intervention is needed.

For Pen CAT to extract, report and visualise cancer screening data, data must be recorded in the clinical information software properly. Pen CAT will only recognise cancer screening data when it is entered in the correct format and in the expected location. The location(s) and format(s) vary slightly across clinical information software (and sometimes between software versions). In some cases, you may need to enter data manually, before doing this please check they meet Pen CAT requirements (location and terminology). Pen CAT has a detailed user guide available online that shows how to manually record cancer screening data for clinical information software (in the mapping sections). The following pages provide links to the inclusions and test names that Pen CAT will extract for bowel, breast and cervical cancer screening for the common clinical software programs.

The Pen CAT user guide can be access via the following link:  
<http://help.pencs.com.au/display/CG>

### Clinical software recording of cancer screening data

#### Best Practice

##### Bowel Cancer Screening

Pathology Data Mapping

<http://help.pencs.com.au/display/ADM/General+Data+Category+Mappings+BP>

General Data Category Mappings

<http://help.pencs.com.au/display/ADM/General+Data+Category+Mappings+BP>

Exclusions

<http://help.pencs.com.au/display/ADM/Conditions+Data+Category+Mappings+BP>

##### Breast Cancer Screening

Pathology Data Mapping

<http://help.pencs.com.au/display/ADM/General+Data+Category+Mappings+BP>

General Data Category Mappings

<http://help.pencs.com.au/display/ADM/General+Data+Category+Mappings+BP>

Exclusions

<http://help.pencs.com.au/display/ADM/Conditions+Data+Category+Mappings+BP>

##### Cervical Cancer Screening

Pathology Data Mapping

<http://help.pencs.com.au/display/ADM/Pathology+Data+Mapping+BP>

General Data Category Mappings

<http://help.pencs.com.au/display/ADM/General+Data+Category+Mappings+BP>

Exclusions

<http://help.pencs.com.au/display/ADM/General+Data+Category+Mappings+BP>

<http://help.pencs.com.au/display/ADM/Conditions+Data+Category+Mappings+BP>

## Medical Director

### Bowel Cancer Screening

Pathology Data Mapping

<http://help.pencs.com.au/display/ADM/General+Data+Category+Mappings+MD3>

General Data Category Mappings

<http://help.pencs.com.au/display/ADM/General+Data+Category+Mappings+MD3>

Exclusions

<http://help.pencs.com.au/display/ADM/Conditions+Data+Category+Mapping+MD3>

### Breast Cancer Screening

Pathology Data Mapping

<http://help.pencs.com.au/display/ADM/General+Data+Category+Mappings+MD3>

General Data Category Mappings

<http://help.pencs.com.au/display/ADM/General+Data+Category+Mappings+MD3>

### Cervical Cancer Screening

Pathology Data Mapping

<http://help.pencs.com.au/display/ADM/Pathology+Data+Mappings+All+Systems>

General Data Category Mappings

<http://help.pencs.com.au/display/ADM/General+Data+Category+Mappings+MD3>

Exclusions

<http://help.pencs.com.au/display/ADM/Conditions+Data+Category+Mapping+MD3>

---

## MedTech

### Bowel Cancer Screening

Pathology Data Mapping

<http://help.pencs.com.au/display/ADM/Mapping+Medtech+Appendix+-+Details>

General Data Category Mappings

<http://help.pencs.com.au/display/ADM/Mapping+Medtech+Appendix+-+Details>

### Breast Cancer Screening

Pathology Data Mapping

<http://help.pencs.com.au/display/ADM/Medtech+Data+Mapping?preview=/1477037/5636107/CAT4%20Data%20Mapping%20Medtech.pdf>

General Data Category Mappings

<http://help.pencs.com.au/display/ADM/Mapping+Medtech+Appendix+-+Details>

Exclusions

None

### Cervical Cancer Screening

Pathology Data Mapping

<http://help.pencs.com.au/display/ADM/Medtech+Data+Mapping?preview=/1477037/5636107/CAT4%20Data%20Mapping%20Medtech.pdf>

General Data Category Mappings

<http://help.pencs.com.au/display/ADM/Mapping+Medtech+Appendix+-+Details>

## ZedMed

### Bowel Cancer Screening

#### Pathology Data Mapping

<http://help.pencs.com.au/display/ADM/General+Data+Mapping+Zedmed>

#### General Data Category Mappings

<http://help.pencs.com.au/display/ADM/General+Data+Mapping+Zedmed>

#### Exclusions

<http://help.pencs.com.au/display/ADM/Conditions+Data+Category+Mappings+ZEDMED>

### Breast Cancer Screening

#### Pathology Data Mapping

<http://help.pencs.com.au/display/ADM/General+Data+Mapping+Zedmed>

#### General Data Category Mappings

<http://help.pencs.com.au/display/ADM/General+Data+Mapping+Zedmed>

#### Exclusions

<http://help.pencs.com.au/display/ADM/Conditions+Data+Category+Mappings+ZEDMED>

### Cervical Cancer Screening

#### Pathology Data Mapping

<http://help.pencs.com.au/display/ADM/Pathology+Data+Mapping+Zedmed>

#### General Data Category Mappings

<http://help.pencs.com.au/display/ADM/General+Data+Mapping+Zedmed>

#### Exclusions

<http://help.pencs.com.au/display/ADM/Conditions+Data+Category+Mappings+ZEDMED>



## Pen CAT Cancer Screening Reports

Pen CAT provides bowel, breast and cervical screening reports and these can be found under the "Screening" tab.

In the first instance, please set and save each report for your Health Service Population as discussed in the "Measuring Your Progress Through Data". Make sure you "Clear Filters" in the top right and then set filters as appropriate, such as post codes you may want to include or exclude. Once you have set the appropriate filter(s), click on "Recalculate" in the top left and then save your search by clicking on "Saved Filters" On the top menu bar and then "Save New Filter". Enter a name for your search and click "OK".

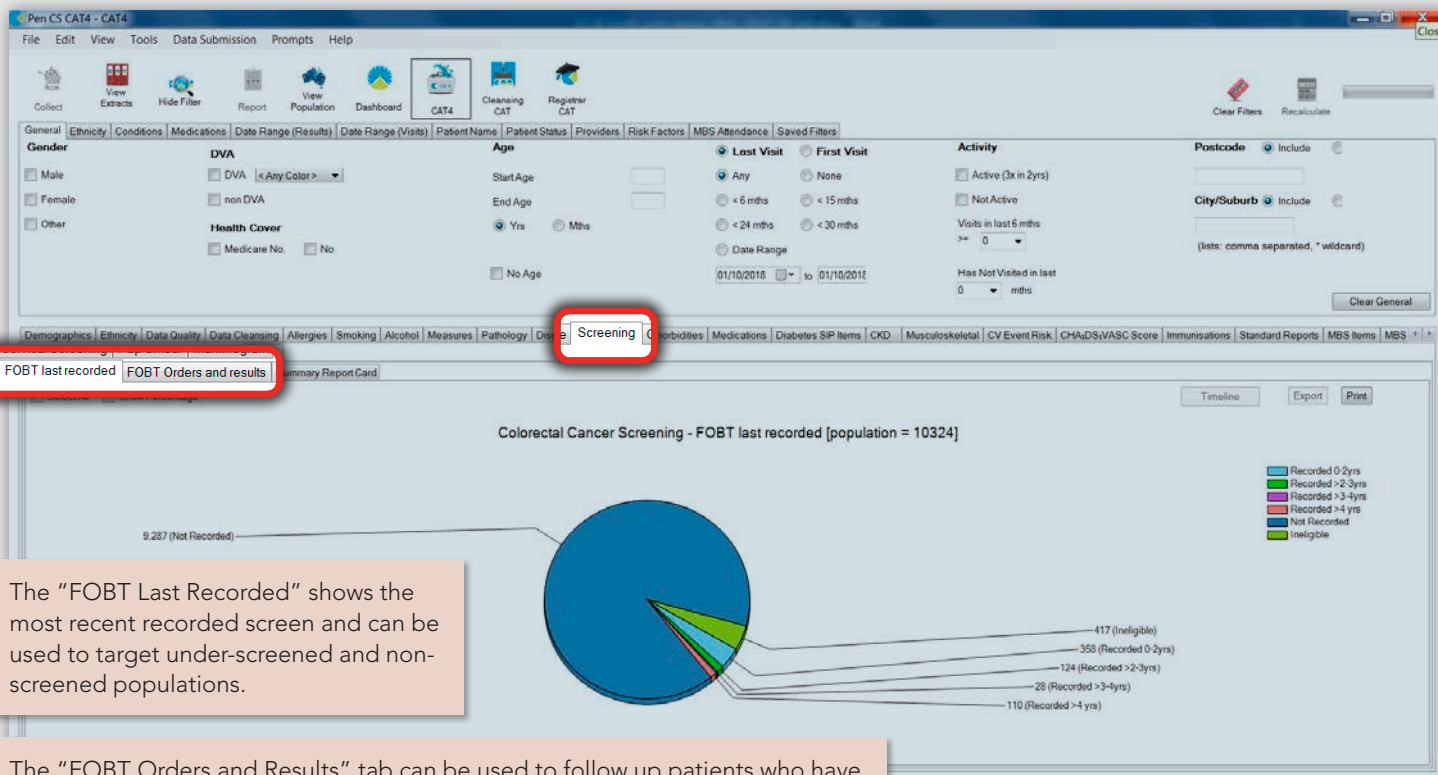
Once you've saved a search for bowel, breast and cervical cancer screening, you will be able to use these searches over time to monitor your progress in each area.

Data can be further filtered, if required, to target specific patient groups to support measurement you may want to use in MFI and/or PDSA work.

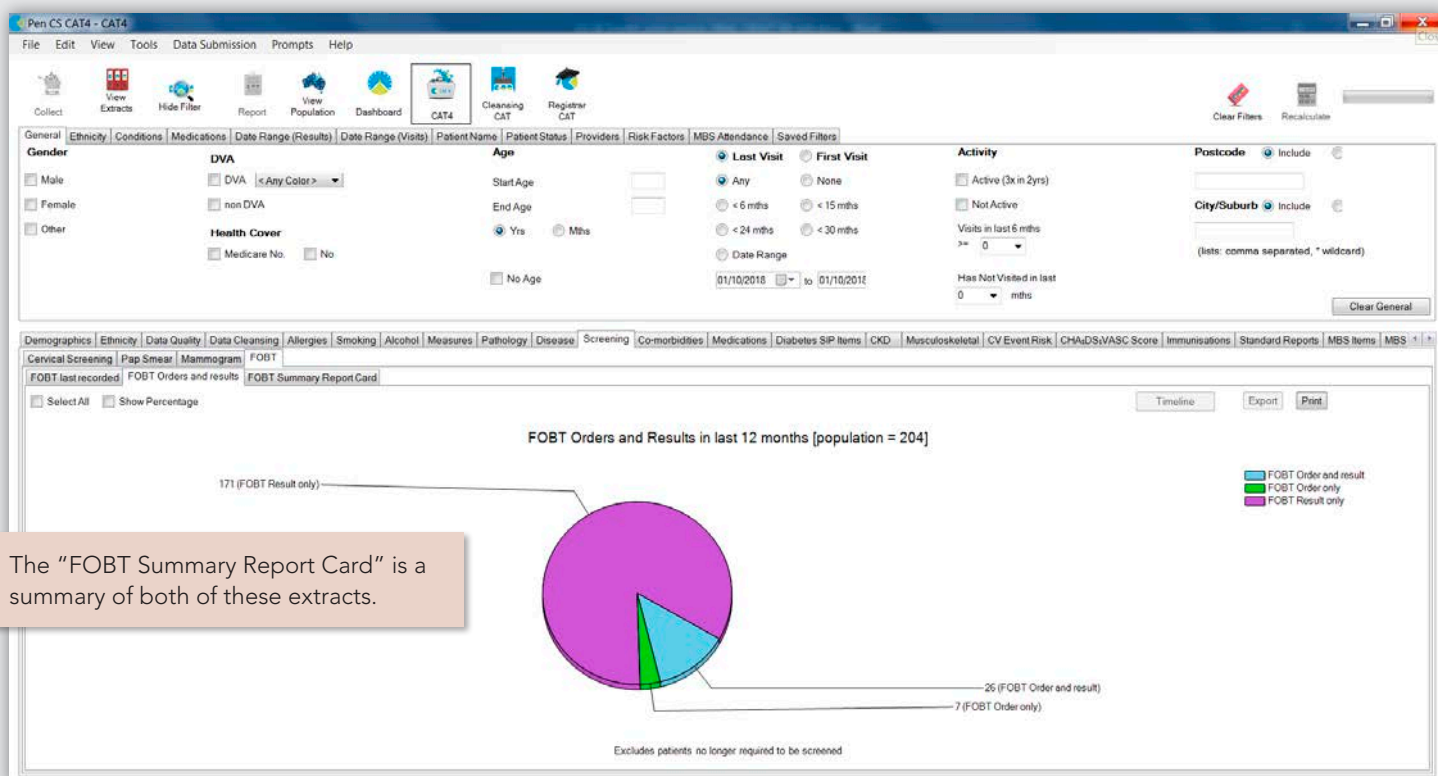
The general data for each of the screening types are shown next.

## Bowel Cancer

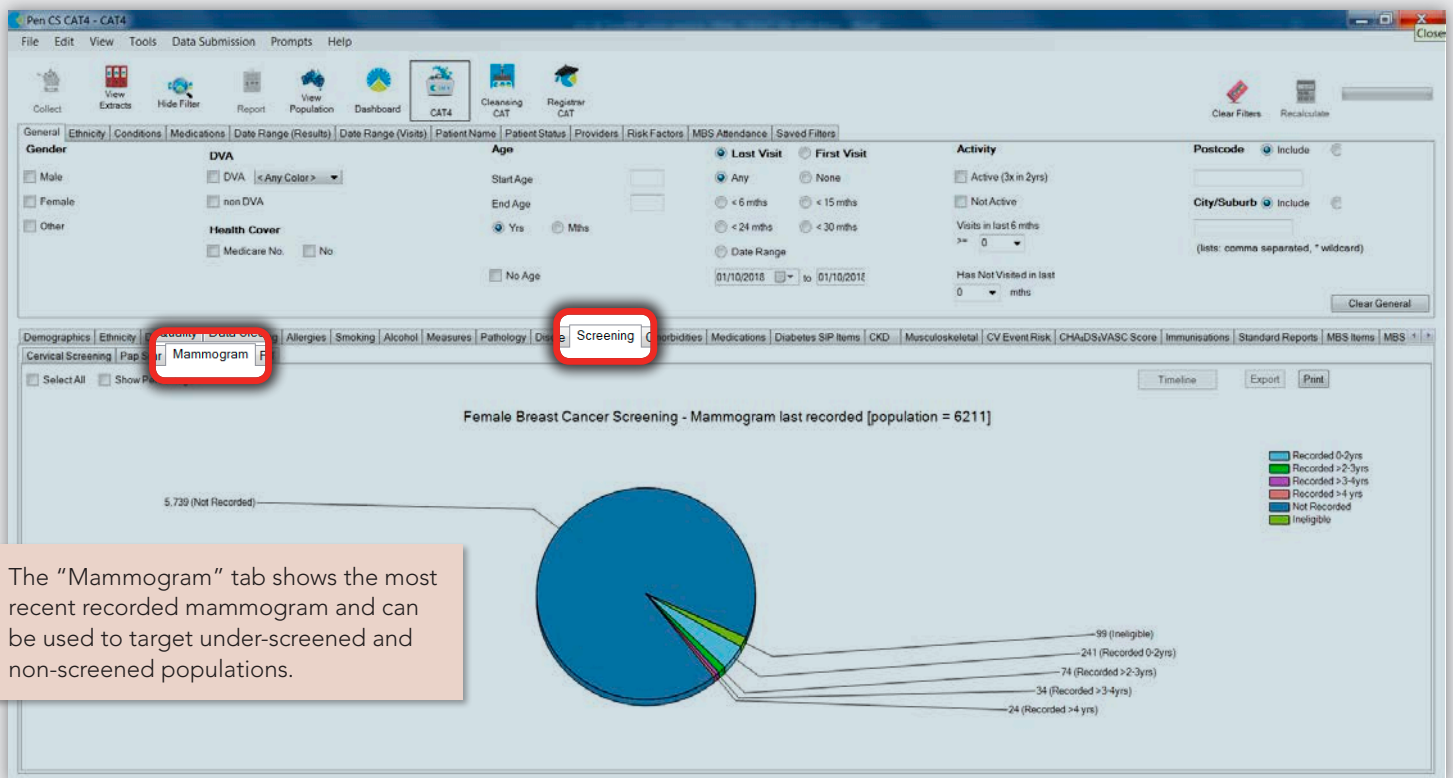
Three different extracts are available for bowel screening.



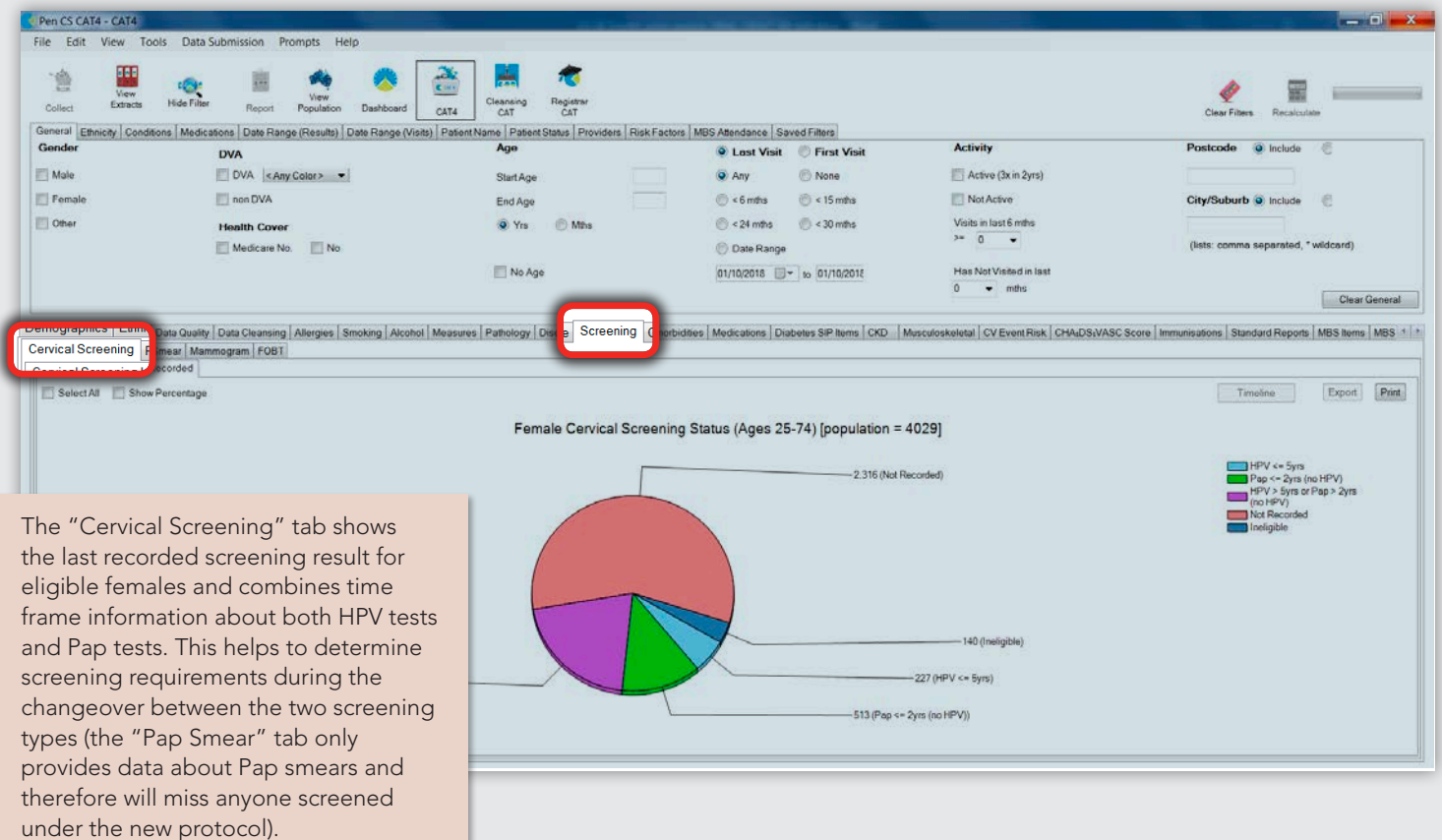
The "FOBT Orders and Results" tab can be used to follow up patients who have an ordered FOBT tests but have no results. The number of results is likely to be much greater than the number of tests ordered due to the various ways FOBT tests can be undertaken (for example, GP-ordered or government-issued tests).



## Breast Cancer



## Cervical Cancer



## Pen CAT Cancer Screening Searches

Pen CAT provides detailed instructions on how to identify patients who are eligible for cancer screening but have no results recorded. Please use the links below to access these instructions.

### Bowel Cancer

<http://help.pencs.com.au/display/CR/Find+patients+who+do+not+have+an+FOBT+recorded>

### Breast Cancer

<http://help.pencs.com.au/display/CR/Find+patients+who+have+not+had+a+mammogram+recorded>

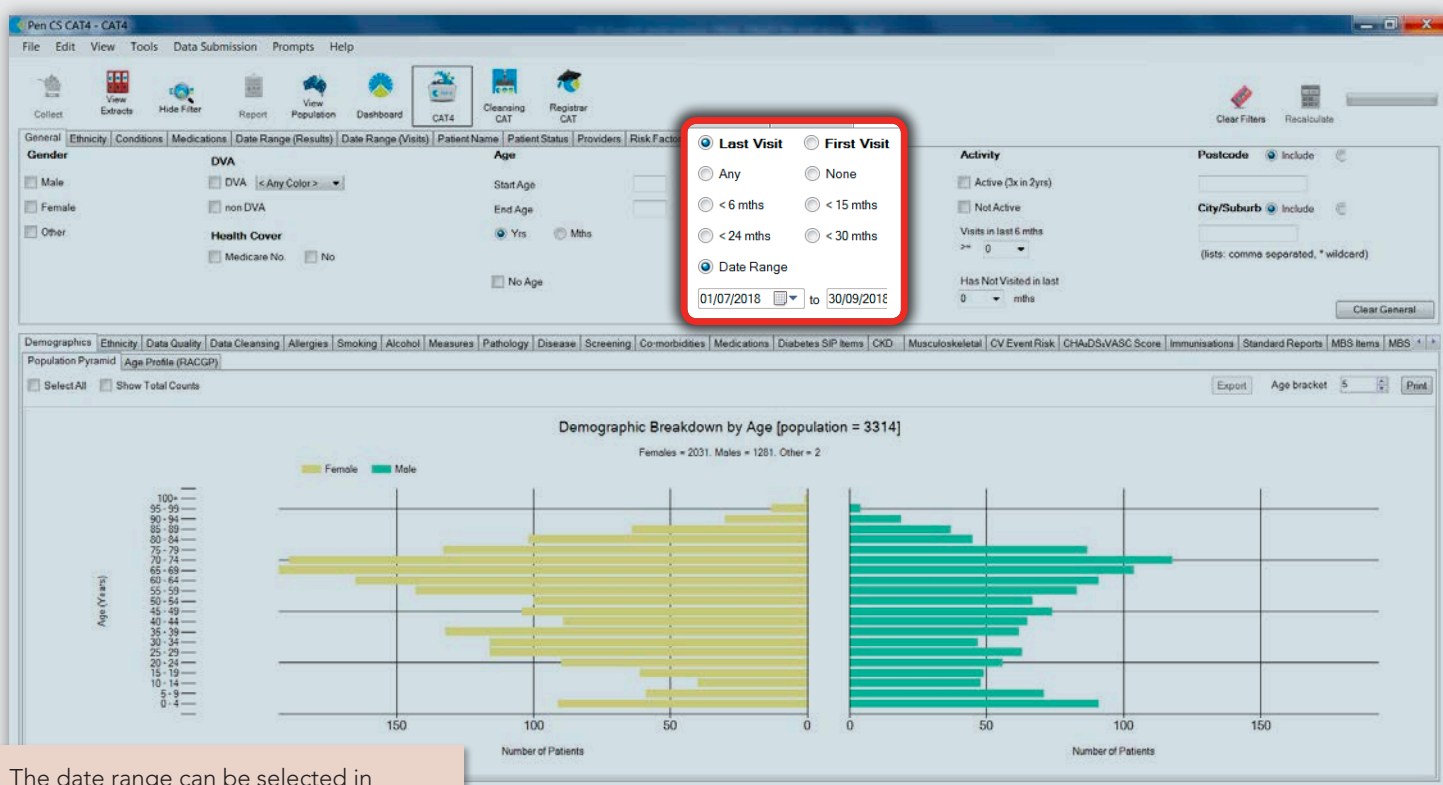
### Cervical Cancer

<http://help.pencs.com.au/display/CR/Find+patients+eligible+for+cervical+screening>

## Data Filtering

Pen CAT reports can be filtered by the timing of the patient's visit (first or most recent visit). Filtering can also target high-risk or under-screened populations based on variables such as ethnicity or age.

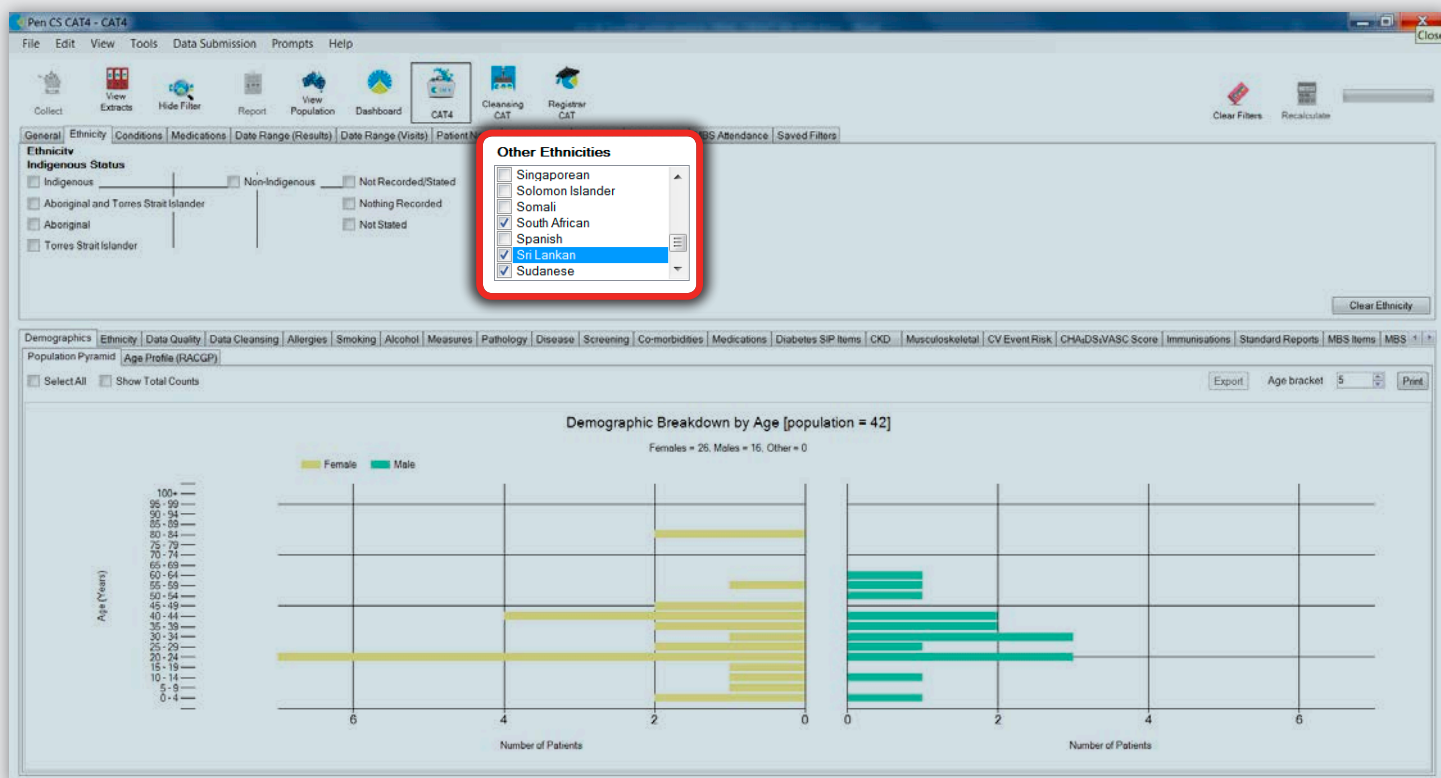
### Date Range



The date range can be selected in several ways, either via last visit or first visit. Default "Last Visit" options are either <6 months or <12 months, or a specific date range. The date range can also be selected when filtering for the first visit.



## Ethnicity



Apart from the standardised four ethnicity options (Non-Indigenous, Aboriginal, Torres Strait Islander, Aboriginal and Torres Strait Islander), other ethnicity options will be available depending on the clinical information software used. Multiple selections are possible, if needed.

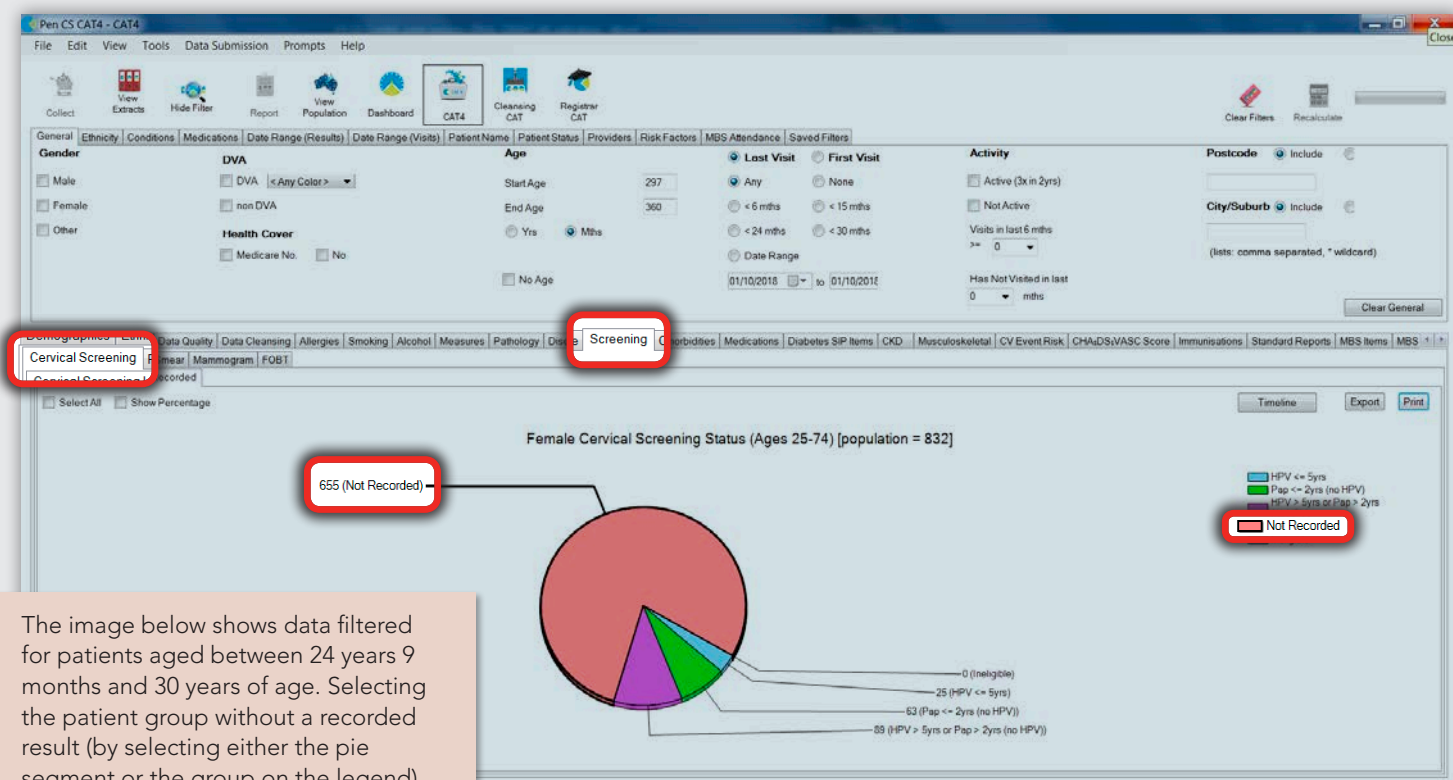
The ethnicity filtering is limited by the list of ethnicities used in your clinical information software system and the data quality. There may also be mismatching of data when converting from one clinical information software system to another.

Where available, you will be able to use this filter to identify vulnerable populations.



## Age

Filtering via age can also be used to target specific populations. Entering the age in months rather than years can help when allowing for a lead time prior to the patient reaching the target age (for example, using 297 months allows an alert to be sent to a patient for an HPV test 3 months before their 25th birthday).



The image below shows data filtered for patients aged between 24 years 9 months and 30 years of age. Selecting the patient group without a recorded result (by selecting either the pie segment or the group on the legend) allows for these patients to be targeted.

## Data Cleaning

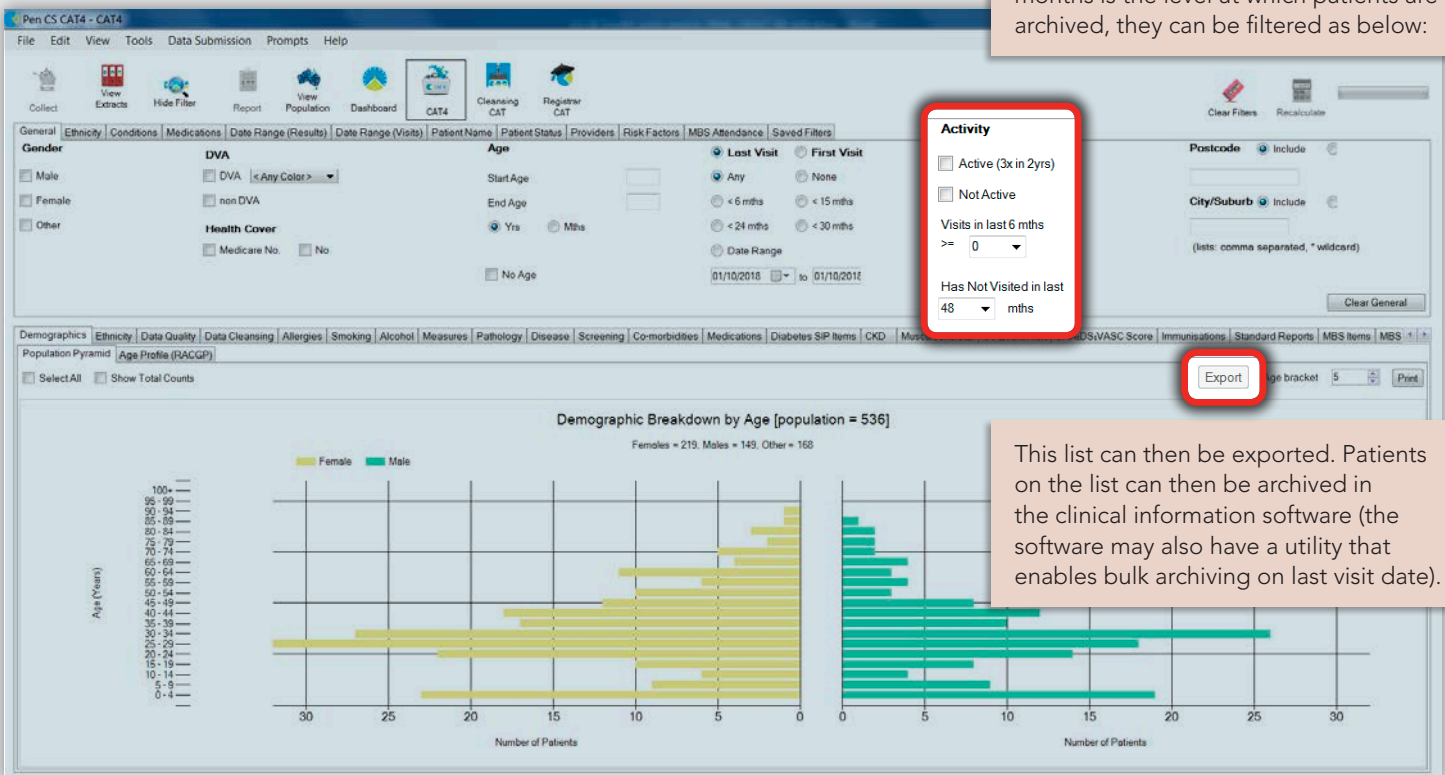
Prior to undertaking any data cleansing procedures, it is important to have processes and procedures in place to prevent the same data issues from reoccurring. All staff need to be made aware of any changes to procedures, and regular monitoring of the data can help address issues early.

Time can be saved during data cleansing when processes are completed in an efficient order. This reduces the time spent on cleaning, or trying to clean, records that are then archived. For this reason, archiving inactive records should be the first activity undertaken.

## Archiving

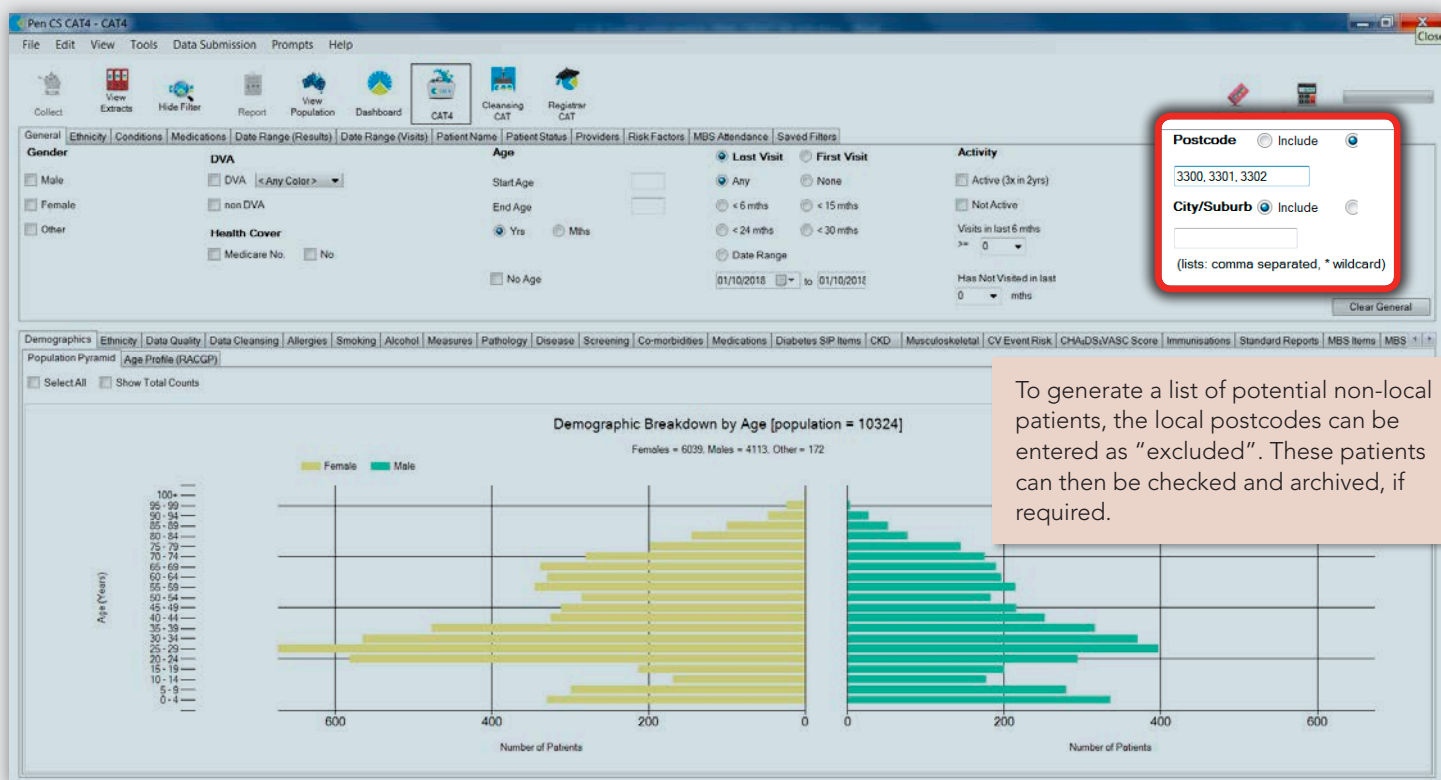
The approach used to archive will be decided by your Health Service and is generally based on the time lapsed since the patient has last visited the Health Service. You can then use Pen CAT to find patients that need to be archived.

For example, if no visit in the past 48 months is the level at which patients are archived, they can be filtered as below:



This list can then be exported. Patients on the list can then be archived in the clinical information software (the software may also have a utility that enables bulk archiving on last visit date).

Depending on need, patients may also be excluded and archived due to postcode. This is most likely in areas with high levels of tourism, for example coastal towns. Patients may only have a single visit and, therefore, are more likely to have missing information. These patients can also skew rates of conditions and management and should be removed to improve data quality.



Archiving can create problems with the creation of duplicate patients if staff are not aware to check archived patients when a patient isn't initially found in the system. This is one of the most common causes of ongoing data issues. Potential duplicate patients can be identified using Pen CAT under the "Data Quality" tab. As these lists are generated using an algorithm, many may not be duplicates (for example, twins will have multiple duplicate demographics and therefore will appear on the list), however they are a very useful place to start when undertaking initial data cleaning activities.

Pen CS CAT4 - CAT4

File Edit View Tools Data Submission Prompts Help

Collect View Extracts Hide Filter Report View Population Dashboard CAT4 Cleaning CAT Register CAT

General Ethnicity Conditions Medications Date Range (Results) Date Range (Visits) Patient Name Patient Status Providers Risk Factors MBS Attendance Saved Filters

Gender: Male, Female, Other

DVA: DVA, <Any Color>, non DVA

Health Cover: Medicare No, No

Age: Start Age, End Age, Yrs, Mths, No Age

Last Visit: Any, < 6 mths, < 15 mths, < 24 mths, < 30 mths, Date Range

First Visit: None, < 6 mths, < 15 mths, < 24 mths, < 30 mths, Date Range

Activity: Active (3x in 2yrs), Not Active, Visits in last 6 mths, Has Not Visited in last 0 mths

Postcode: Include, 3300, 3301, 3302

City/Suburb: Include

(lists: comma separated, \* wildcard)

Clear General

Demographics Ethnicity Data Quality Data Cleaning Allergies Smoking Alcohol Measures Pathology Disease Screening Co-morbidities Medications Diabetes SIP Items CKD Musculoskeletal CV Event Risk CHA2DS2-VASc Score Immunisations Standard Reports MBS Items MBS

Report Date: 01/10/2018 3:04 AM

Practice Name: Deidentified Practice

Duplicate Number Patient Report

Match on ANY of: Medicare number, HCC number or DVA number

Number Matches [patient count = 10324]

Surname	First Name	Sex	D.O.B.	Age	Address	City	Postcode	Medicare	HCC No	DVA No	ID
Surname	First Name	M	01/08/2012	6	12 Jogger St	Suburb Town	5992	123412341234			9886
Surname	First Name	F	01/10/1983	35	12 John St	Suburb Town	5708	123412341234			1032
Surname	First Name	M	01/10/1951	67	12 Jogger St	Suburb Town	5930	123412341234			7711
Surname	First Name	M	01/10/1971	47	12 Jogger St	Suburb Town	4292	123412341234			5231
Surname	First Name	F	01/10/1974	44	12 John St	Suburb Town	3434	123412341234			3238
Surname	First Name	F	01/10/1977	41	12 John St	Suburb Town	4220	123412341234			9964
Surname	First Name	F	01/10/1987	31	12 John St	Suburb Town	3350	123412341234			3479
Surname	First Name	M	01/10/1994	24	12 Jogger St	Suburb Town	5476	123412341234			9960

The "Duplicate Name Patient Report" tab checks for duplication of all of the surname, first name initial, gender and date of birth.

**Duplicate Name Patient Report**

Report Date: 01/10/2018 3:04 AM  
Practice Name: Deidentified Practice

Match on ALL of: surname, first name initial, gender and DoB

Surname	First Name	Sex	D.O.B.	Age	Address	City	Postcode	Medicare	HCC No	DVA No	ID
Surname	Firstname	M	01/08/2012	6	12 Jogger St	Suburb Town	5992	123412341234			9886
Surname	Firstname	F	01/10/1983	35	12 John St	Suburb Town	5708	123412341234			1032
Surname	Firstname	M	01/10/1951	67	12 Jogger St	Suburb Town	5930	123412341234			7711
Surname	Firstname	M	01/10/1971	47	12 Jogger St	Suburb Town	4292	123412341234			5231
Surname	Firstname	F	01/10/1974	44	12 John St	Suburb Town	3434	123412341234			3238
Surname	Firstname	F	01/10/1977	41	12 John St	Suburb Town	4220	123412341234			9964
Surname	Firstname	F	01/10/1987	31	12 John St	Suburb Town	3350	123412341234			3479

## Missing Demographics

As cancer screening is dependent on age, and, for cervical and breast cancer, gender, it is important to complete this missing demographic data. This can be done either via the "Data Cleansing" tab or separately via gender and age.

**Data Cleansing**

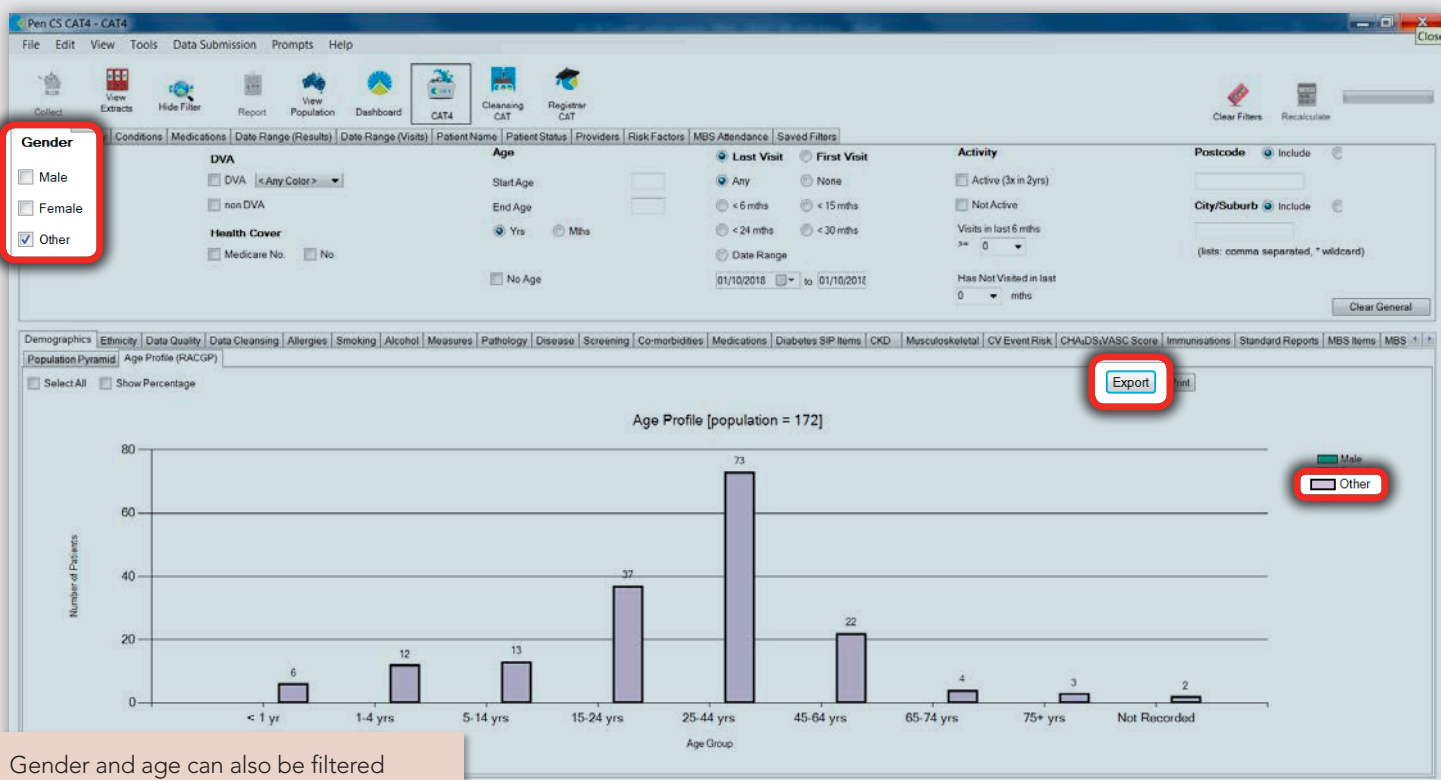
Missing Demographics

9 [count = 172]

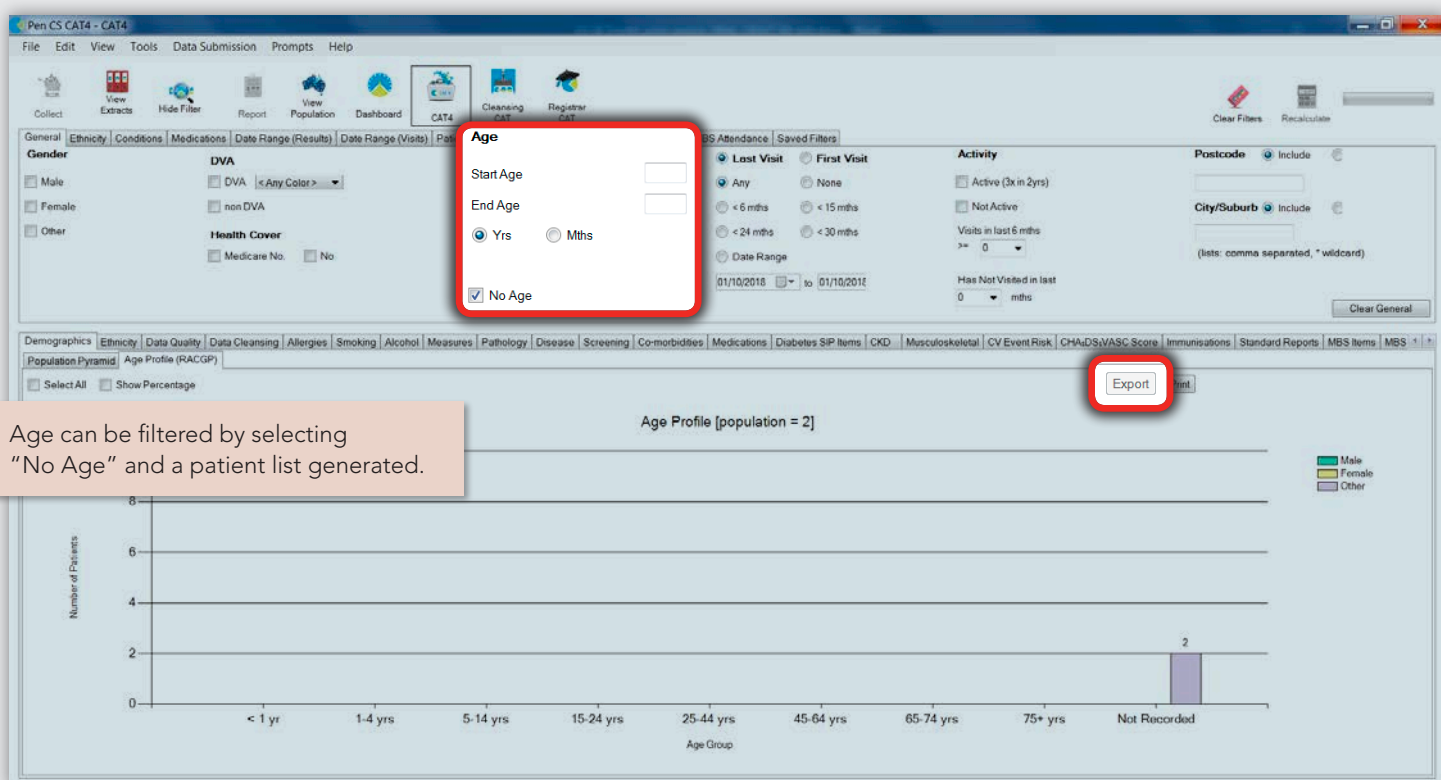
Surname	First name	Date of Birth	Sex	Address	Suburb	Postcode	Home Phone	Work Phone	Mobile Phone	Assigned Provider
Surname	Firstname_10_	01/10/1983		12 Jogger St	Suburb Town	5932	07 50505050	07 50509999	0444444444	
Surname	Firstname_10_	01/06/2015		12 Jogger St	Suburb Town	4164	07 50505050	07 50509999	0444444444	
Surname	Firstname_10_	01/10/1972		12 Jogger St	Suburb Town	3834	07 50505050	07 50509999	0444444444	
Surname	Firstname_10_	01/10/1985		12 Jogger St	Suburb Town	5275	07 50505050	07 50509999	0444444444	
Surname	Firstname_10_	01/10/1996		12 Jogger St	Suburb Town	2595	07 50505050	07 50509999	0444444444	
Surname	Firstname_10_	01/10/1993		12 Jogger St	Suburb Town	4964	07 50505050	07 50509999	0444444444	
Surname	Firstname_10_	01/10/1989		12 Jogger St	Suburb Town	5199	07 50505050	07 50509999	0444444444	
Surname	Firstname_10_	01/10/1984		12 Jogger St	Suburb Town	4583	07 50505050	07 50509999	0444444444	
Surname	Firstname_10_	01/10/1997		12 Jogger St	Suburb Town	2576	07 50505050	07 50509999	0444444444	
Surname	Firstname_10_	01/10/1992		12 Jogger St	Suburb Town	4117	07 50505050	07 50509999	0444444444	
Surname	Firstname_124	01/03/2014		12 Jogger St	Suburb Town	3338	07 50505050	07 50509999	0444444444	
Surname	Firstname_13_	01/10/1999		12 Jogger St	Suburb Town	2254	07 50505050	07 50509999	0444444444	
Surname	Firstname_13_	01/10/1984		12 Jogger St	Suburb Town	2390	07 50505050	07 50509999	0444444444	
Surname	Firstname_13_	01/10/1999		12 Jogger St	Suburb Town	3812	07 50505050	07 50509999	0444444444	
Surname	Firstname_13_	01/10/1992		12 Jogger St	Suburb Town	5602	07 50505050	07 50509999	0444444444	
Surname	Firstname_14_	01/10/1986		12 Jogger St	Suburb Town	3301	07 50505050	07 50509999	0444444444	

The "Data Cleansing" tab will list all patients with missing demographics, not just gender and/or date of birth. Any missing data will be represented by a red box. If Pen CAT has been set-up to do so, double-clicking on a patient in the list will open that patient within the clinical software allowing for immediate updating of patient information.





Gender and age can also be filtered separately. Patients not recorded as either male or female can be filter by selecting the gender as "other". Depending on the clinical software used, this may also flag patients who have a non-male/female gender recorded.



Age can be filtered by selecting "No Age" and a patient list generated.



## Diagnoses in free text form

Diagnoses that are not entered correctly, particularly due to free text entries, greatly reduce the functionality of your clinical information software and the ability for Pen CAT to report and visualise data accurately. Although there are methods for mapping these entries to specific diagnoses within the clinical information software, it is time-consuming and may be less accurate when mapped. This is one area where time spent cleaning the data can be wasted if a Health Service wide procedure is not implemented and recording diagnosis in free text is allowed to continue, as data will need frequent cleansing.

Pen CAT has recipes available for bulk cleaning of free text diagnoses for both Best Practice and Medical Director. Other clinical software may also have systems for bulk cleaning - check with your software provider.

### Best Practice

<http://help.pencs.com.au/display/CR/Bulk+clean+up+of+free+text+diagnosis+-+BP+users>

### Medical Director

<http://help.pencs.com.au/display/CR/Bulk+clean+up+of+free+text+diagnosis+-+MD3+users>

### Zedmed

Diagnoses mapping from free text to ICPC coding is also available in Zedmed. It can be found in: Clinical Records > Utilities-Clinical > ICPC Problem Mapper

## Tips for Clean Data

- Keep staff informed and involved
- Decide on changes to processes with input from those who perform the task (there may be unseen issues that can be picked up before a procedure is changed)
- Check your patient registration sheet and ensure it captures all the required information and matches the clinical software options
- Always check archived patients prior to creating a new patient
- Use data management tools, such as TopBar, for alerts for missing data at the time of patient presentation
- Decide on the coding to be used for diagnoses, screening, etc.
- Develop cheat sheets and make them easily accessible (for example, data entry/checking archived patients at the front desk, or coding in the consultation rooms)
- Set up regular archiving and data monitoring processes (check for duplicates, free text diagnoses, etc.)
- If recurring problems are noted, confirm processes with staff (does a procedure need to be changed?)