



CanReg5 Webinar 5: Data Analysis



CanReg 5

Morten Ervik

International Agency for Research on Cancer,
Lyon, France

Lyon, France, 22 October 2012

Outline

Frequencies by year

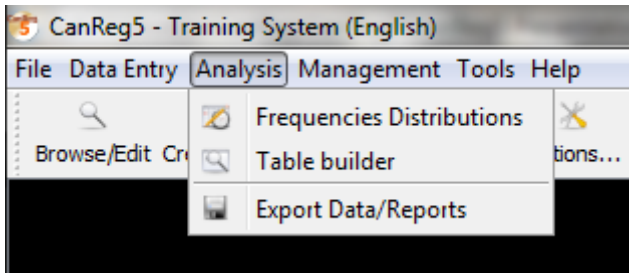
Population Dataset Editor

Table Builder

Export/Report

Live Demo

Summary



Outline

Frequencies by year

Population Dataset Editor

Table Builder

Export/Report

Live Demo

Summary

International Agency for Research on Cancer



Frequencies by year

- ▶ A tool to count the cases in the database.
- ▶ Filter the cases using simple statements
 - ▶ Boolean logic
 - ▶ Subset of the SQL language
 - ▶ Filter wizard to help build filters
- ▶ Stratify the results:
 - ▶ by any variable
 - ▶ by as many variables you need
- ▶ Results can be written to files of standard format for post processing in general spreadsheets (Excel, Calc, Numbers), or other tools.

Example: Cases by year

Frequencies by Year

Table: Tumour+Pati...

Range: Birth date (Patient) Start: End:

AND

Filter: Use Filter Filter Wizard

Refresh Table

Select variables:

- Address Data
- Age Data
- Basis diagnosis Data
- Behaviour Data
- Birth date Data
- Cause of Death Data
- Check status Data
- Civil Status Data
- Date last contact Data
- Date of Death Data
- Extent Data
- Family Name Data
- First names Data
- ICCC code Data

Select all variables

Save table Print Table Pop-out Table

YEAR	CASES
1999	1313
2000	2541
2001	2762
2002	3285
2003	3349
2004	3737
2005	4029
2006	4279
2090	1

Example: Stomach cancers in 2000 by morphology

Frequencies by Year

Table: Tumour+Pat...

Range: Incidence date (Tumour) Start: 2000 End: 2001

AND

Filter: Use Filter top like '16%'

Select variables:

- MSP Tot Data
- Maiden name Data
- MiddleName Data
- Morphology Data
- Multiple Primary Data
- ObsoleteFlagPatientTable Data
- ObsoleteFlagTumourTable Data
- Occupation Data
- PatientCheckStatus Data
- PatientIDTumourTable Data
- PatientRecordID Data
- PatientRecordIDTumourTable Data
- PatientRecordStatus Data
- PatientUpdateDate Data
- PatientI.IndicatorBy... Data

Select all variables

YEAR	MOR	CASES
2000	8140	38
2000	8211	20
2000	8490	19
2000	8145	13
2000	8010	9
2000	8000	7
2000	9590	5
2000	8260	4
2000	8481	4
2000	9680	2
2000	8020	1
2000	8144	1
2000	8480	1
2000	8560	1
2000	9591	1

Example: Sources in one year

Frequencies by Year

Table: Source+Tum...

Range: Incidence date (Tumour) Start: 19990101 End: 19999999

AND

Filter: Use Filter

Select variables:

- Address Data
- Age Data
- Basis diagnosis Data
- Behaviour Data
- Case n° Data
- Check status Data
- Date of Source Data
- Extent Data
- ICCC code Data
- ICD-10 Data
- Incidence date Data
- Laterality Data
- MP Seq Data
- MP Tot Data

Select all variables

YEAR	SOURCE	CASES
1999	016	216
1999	074	179
1999	188	157
1999	057	141
1999	071	118
1999	345	94
1999	261	42
1999	262	38
1999	061	37
1999	349	36
1999	368	31
1999	268	29
1999	366	28
1999	174	24

Example: One source over the years

Frequencies by Year

Table: Source+Tum...

Range: Incidence date (Tumour) Start: 1999 End: 2007

AND

Filter: Use Filter Filter Wizard

source = '016'

Refresh Table

Select variables

- ObsoleteFlagTumourTable Data
- Path lab n° Data
- PatientIDTumourTable Data
- PatientRecordIDTumourTable Data
- Record status Data
- Source Data
- SourceRecordID Data
- Topography Data
- TumourID Data
- TumourIDSourceTable Data
- TumourUnduplicationStatus Data
- TumourUpdatedBy Data
- Unit Data
- Update Date Data

Select all variables

Save table Print Table Pop-out Table

YEAR	CASES
1999	216
2000	987
2001	1183
2002	1349
2003	1399
2004	1564
2005	1790
2006	1696

Outline

Frequencies by year

Population Dataset Editor

Table Builder

Export/Report

Live Demo

Summary

International Agency for Research on Cancer



Population Dataset Editor

- ▶ Lets you add denominators to your database
- ▶ You can store as many as you want:
 - ▶ By year
 - ▶ Sub-populations (relevant cases automatically picked from the database.)
- ▶ Support most age group breakdowns
- ▶ Editor supports copy and paste to and from general spreadsheets like Excel

Enter details

Population Data Set Editor

Details Population Data Set Pyramid

Name: Training, 2005

Filter: Filter Wizard

Source: Estimates

Description:

Age group structure: 0-4, 5-9, 10-14, 15-19, ..., 80+ Other

Date: 20050701

Reference Population: World Standard Population Edit

Delete Toggle lock Cancel Save as New Save

Enter data

Population Data Set Editor

Details Population Data Set Pyramid

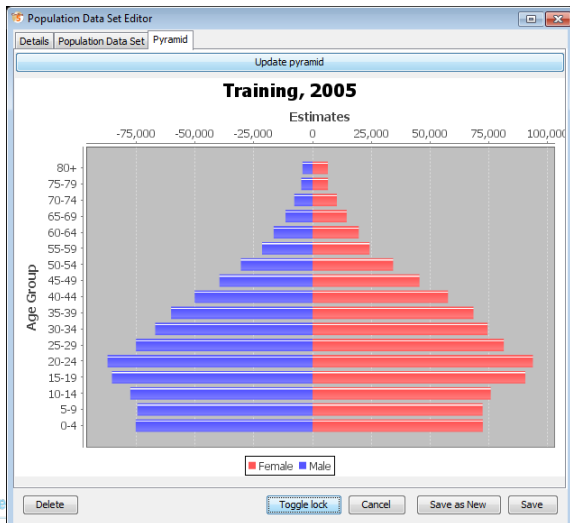
Population Data Set

Age Group	Male	Female
0-4	75217	72421
5-9	74493	72301
10-14	77508	75773
15-19	85325	90473
20-24	87199	93688
25-29	75100	81302
30-34	66898	74392
35-39	60157	68402
40-44	50141	57581
45-49	39614	45484
50-54	30491	34234
55-59	21507	24231
60-64	16566	19625
65-69	11506	14483
70-74	7768	10328
75-79	4786	6543
80+	4269	6523
Total	788545	847784

Delete Toggle lock Cancel Save as New Save

International Agency for Research on Cancer

Preview the population pyramid to visually check for errors



Outline

Frequencies by year

Population Dataset Editor

Table Builder

Export/Report

Live Demo

Summary

International Agency for Research on Cancer



Table builder

- ▶ Generate “ready to print” tables.
- ▶ A wide range of different types of tables, for example:
 - ▶ Incidence tables - rates and number of cases
 - ▶ Age specific incidence curves
 - ▶ Figures of top 10 cancers by sex
 - ▶ Figures of cases by age
 - ▶ Population pyramids
- ▶ A wide range of output formats:
 - ▶ Scalable formats (PDF, PS, SVG)
 - ▶ Image formats (PNG, etc.)
- ▶ Some tables calls transparently R:
 - ▶ Advanced users can write their own R programs that can be called directly from CanReg5.

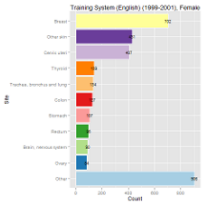
Choose table type

Table builder

Choose table type **Range** Population Data Set **Filter** Make tables

Number of cases in major diagnosis groups in single calendar years of observation
Population Pyramids
Quality indicators
R-test
SEER *prep
Bar Chart - 10 most common cancers by sex
Bar Chart - 10 most common cancers by sex (ASR)
Bar Chart - 10 most common cancers by sex (CUM64)

Bar Chart - 10 most common cancers by sex (ASR)



System	Count
Breast	702
Other skin	437
Cervix uteri	314
Thyroid	284
Trachea, bronchus and lung	274
Colon	271
Stomach	217
Rectum	176
Brain, meningeal system	176
Ovary	117
Other	777

Back Cancel Next

Choose period

Table builder

Choose table type Range Population Data Set Filter Make tables

Start year: 2000

End year: 2002

Mid-year: 2001

Number of years: 3

Back Cancel Next

Choose denominators

Table builder

Choose table type | Range | Population Data Set | Filter | Make tables

Please choose one population dataset per year:

2000: Training, 2000

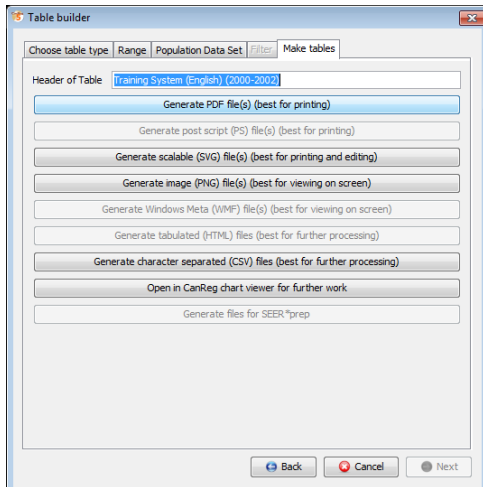
2001: Training, 2001

2002: Training, 2002

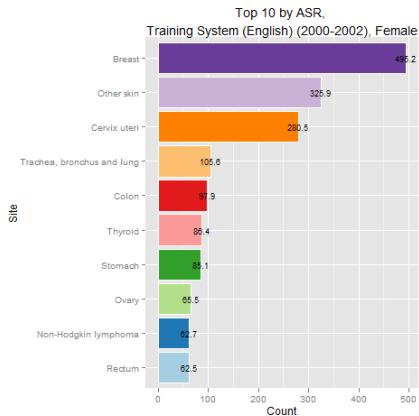
Don't use population dataset(s) - NO RATES WILL BE CALCULATED

Back Cancel Next

Choose file format



Results



Outline

Frequencies by year

Population Dataset Editor

Table Builder

Export/Report

Live Demo

Summary

Export/Report

- ▶ Export data for further work in other tools, or provide to collaborators
 - ▶ individual cases
 - ▶ only selected variables (or all)
 - ▶ choose to export codes and/or labels/descriptions
 - ▶ filter cases as you wish
 - ▶ Character separated files compatible with general spreadsheets, like Excel.

Export/Report

Export

Selection **Export**

Table: Tumour + Pati... Sort by: Incidence date (T... Records: 25296 Show

Range: Birth date (Patient) Start: End:

Filter: Use Filter [Filter Wizard](#)

Export source information

Refresh Table

Select variables

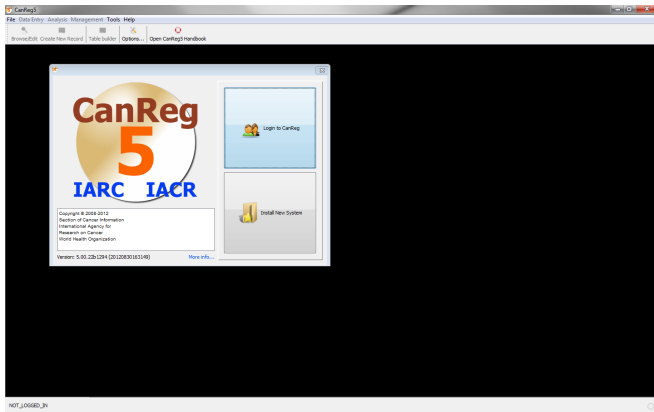
- First names Data Category Description
- ICCC code Data Category Description
- ICD-10 Data Category Description
- Incidence date Data Category Description
- Laterality Data Category Description
- MP Seq Data Category Description
- MP Tot Data Category Description
- Maiden name Data Category Description
- MiddleName Data Category Description
- Morphology Data Category Description
- Multiple Primary Data Category Description
- ObsoleteFlagPatientTable Data Category Description
- ObsoleteFlagTumourTable Data Category Description
- Occupation Data Category Description
- PatientCheckStatus Data Category Description
- PatientIDTumourTable Data Category Description

Select all variables

Preview

INCID	AGE	TOP	MOR	BEH	BAS	I10	SEX
1999...	42	809	8000	3	5	C809	1
1999...	46	539	8070	3	7	C539	2
1999...	63	492	8830	3	7	C492	2
1999...	72	539	8000	3	0	C539	2
1999...	74	199	8211	3	7	C199	2
1999...	76	619	8010	3	2	C619	1
1999...	58	679	8120	3	7	C679	1
1999...	72	619	8140	3	7	C619	1
1999...	36	539	8070	3	7	C539	2
1999...	66	619	8140	3	7	C619	1
1999...	45	539	8077	2	7	D069	2
1999...	62	679	8120	3	7	C679	1
1999...	64	189	8140	3	7	C189	2
1999...	76	349	8042	3	7	C349	1
1999...	92	445	8070	3	7	C445	2
1999...	75	710	8000	3	0	C710	2

Live Demo



International Agency for Research on Cancer



Summary

- ▶ Frequencies by year
 - ▶ stratified by any variable
- ▶ Population Dataset Editor
- ▶ Table builder
 - ▶ for “ready to print” tables.
 - ▶ a multitude of file formats for different uses.
- ▶ Export/report
 - ▶ For further analysis in other programs

Summary

- ▶ Frequencies by year
 - ▶ stratified by any variable
- ▶ Population Dataset Editor
- ▶ Table builder
 - ▶ for “ready to print” tables.
 - ▶ a multitude of file formats for different uses.
- ▶ Export/report
 - ▶ For further analysis in other programs

Summary

- ▶ Frequencies by year
 - ▶ stratified by any variable
- ▶ Population Dataset Editor
- ▶ Table builder
 - ▶ for “ready to print” tables.
 - ▶ a multitude of file formats for different uses.
- ▶ Export/report
 - ▶ For further analysis in other programs

Future webinars

- ▶ Next webinar:
 - ▶ Sometime during the first half of December.
 - ▶ Theme: Customization, etc.
 - ▶ Please send suggestions to ervikm@iarc.fr to propose additional topics.
- ▶ Slides and videos will be put online at the GICR website; <http://gicr.iarc.fr>

For Further Information I



Ervik, Morten

CanReg5 - the handbook
2009-2012

Available online at the IACR web page under software



CanReg5 web page

<http://www.iacr.com.fr/canreg5.htm>



@canreg twitter feed

<http://twitter.com/canreg>