CanReg5 Webinar 5: Data Analysis

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
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

<table>
<thead>
<tr>
<th>YEAR</th>
<th>CASES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>1313</td>
</tr>
<tr>
<td>2000</td>
<td>2541</td>
</tr>
<tr>
<td>2001</td>
<td>2762</td>
</tr>
<tr>
<td>2002</td>
<td>3285</td>
</tr>
<tr>
<td>2003</td>
<td>3349</td>
</tr>
<tr>
<td>2004</td>
<td>3737</td>
</tr>
<tr>
<td>2005</td>
<td>4029</td>
</tr>
<tr>
<td>2006</td>
<td>4279</td>
</tr>
<tr>
<td>2007</td>
<td>1</td>
</tr>
</tbody>
</table>
Example: Stomach cancers in 2000 by morphology
Example: Sources in one year
Example: One source over the years
Outline

Frequencies by year

Population Dataset Editor

Table Builder

Export/Report

Live Demo

Summary
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: 
Source: Estimates
Description: 

Age group structure: 0-4, 5-9, 10-14, 15-19, ..., 80+
Date: 20050701

Delete | Toggle lock | Cancel | Save as New | Save
Enter data

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>75217</td>
<td>72421</td>
</tr>
<tr>
<td>5-9</td>
<td>7443</td>
<td>72301</td>
</tr>
<tr>
<td>10-14</td>
<td>77508</td>
<td>75773</td>
</tr>
<tr>
<td>15-19</td>
<td>85325</td>
<td>90473</td>
</tr>
<tr>
<td>20-24</td>
<td>87199</td>
<td>93688</td>
</tr>
<tr>
<td>25-29</td>
<td>75100</td>
<td>81302</td>
</tr>
<tr>
<td>30-34</td>
<td>66898</td>
<td>74392</td>
</tr>
<tr>
<td>35-39</td>
<td>60157</td>
<td>68402</td>
</tr>
<tr>
<td>40-44</td>
<td>50141</td>
<td>57581</td>
</tr>
<tr>
<td>45-49</td>
<td>39614</td>
<td>45484</td>
</tr>
<tr>
<td>50-54</td>
<td>30491</td>
<td>34234</td>
</tr>
<tr>
<td>55-59</td>
<td>21507</td>
<td>24231</td>
</tr>
<tr>
<td>60-64</td>
<td>16866</td>
<td>19625</td>
</tr>
<tr>
<td>65-69</td>
<td>11506</td>
<td>14463</td>
</tr>
<tr>
<td>70-74</td>
<td>7768</td>
<td>10328</td>
</tr>
<tr>
<td>75-79</td>
<td>4786</td>
<td>6543</td>
</tr>
<tr>
<td>80+</td>
<td>4269</td>
<td>6523</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>788545</strong></td>
<td><strong>847784</strong></td>
</tr>
</tbody>
</table>
Preview the population pyramid to visually check for errors.
Outline

Frequencies by year

Population Dataset Editor

Table Builder

Export/Report

Live Demo

Summary
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 period

- **Start year:** 2000
- **End year:** 2002
- **Mid-year:** 2001
- **Number of years:** 3
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

<table>
<thead>
<tr>
<th>Choose table type</th>
<th>Range</th>
<th>Population Data Set</th>
<th>Filter</th>
<th>Make tables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Header of Table</td>
<td>Training System (English) (2000-2002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generate PDF file(s) (best for printing)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generate post script (PS) file(s) (best for printing)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generate scalable (SVG) file(s) (best for printing and editing)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generate image (PNG) file(s) (best for viewing on screen)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generate Windows Meta (WMF) file(s) (best for viewing on screen)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generate tabulated (HTML) files (best for further processing)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generate character separated (CSV) files (best for further processing)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open in CanReg chart viewer for further work</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generate files for SEER*prep</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Back  Cancel  Next
Outline

Frequencies by year

Population Dataset Editor

Table Builder

Export/Report

Live Demo

Summary
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
Outline

Frequencies by year

Population Dataset Editor

Table Builder

Export/Report

Live Demo

Summary
Live Demo
Outline

Frequencies by year

Population Dataset Editor

Table Builder

Export/Report

Live Demo

Summary
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
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