Death Risk RatingsTutorial

These tutorials provide users with a basic guideline on how to use the **DeathRiskRankings.org** website. The interface is designed to be intuitive and might allow you to jump right in; however, spending a few minutes walking through the tutorials will help you to use the system more effectively and to understand the results better.

It is important to remember that though the interface on this website allows for quick exploration and easy comparison the of causes of death; it does not provide information as to ***why*** these differences exist. To answer those questions, medical researchers, health psychologists, and accident specialists will have to be consulted.

Tutorial 1: Getting Started

Tutorial 2: What Is a MicroMort?

Tutorial 3: Searching Using One variable

Tutorial 4: Searching Using Two variables

Tutorial 5: Customizing Search

Tutorial 6: Changing the Output Displays

Tutorial 7: Simpson's Paradox and Interpreting the Results

Tutorial 1: Getting Started

The website home page is: www.DeathRiskRankings.org

The home page provides basic information on the site and links to background information on data sources and interpretation of query results. The concept of a *MicroMort* – a one-in-a-million chance of dying – is critical to understanding the outputs. The second tutorial explains the MicroMort concept in detail.

Figure 1-1: Homepage

There are eight tabs across the top of each page of the website:

- **Home:** The website home page shown in Figure 1-1
- **Death Comparisons:** The page where users query death statistics from the underlying database. This is where you do your searches and where the output is displayed. Figure 1-2 provides an example.
- **Underlying Data**: Information about the source data provided by the system.
- **Calculations**: An explanation of calculations source and methods.
- **Tutorial**: Basic guidelines for using the system and discussion of some fundamental issues, including interpretation of the data. There are seven tutorials
    1. Getting started
    2. What is a MicroMort?
    3. Searching using one variable
    4. Searching using two variables
    5. Customizing searches
    6. Manipulating output information
    7. Simpson's Paradox

- **About This Project**: A description of the project and its genesis..
- **Contact Us**: Contact information if you have questions or would like to provide constructive feedback.
- **Other Calculators**: Links to sites that have life-expectancy calculators

Figure 1-2: Death Comparisons

Tutorial 2: What Is a MicroMort?

In short, a “MicroMort” is a one-in-a-million chance of dying.

One way to think about a MicroMort is to consider a thought experiment where your mortality fate plays out as a game of chance. Suppose that you had a giant urn filled with one million balls (1,000,000). Some of these balls are purple and some are green. Every year, you reach into the urn and randomly draw a ball. If it is green, you will live one more year, but if it is purple, you will die within the year. We call this thought experiment the *purple balls of death*, and emphasize that each purple ball represents one chance in a million of dying in the next year – a MicroMort.

DeathRiskRankings.org will tell you how many purple balls someone like you has in their urn. But before using this website, what do you think your risk of dying in the next year is? The table below will help you relate the number of purple balls in your urn to other ways to think about the risk.

Number of Purple Balls (MicroMorts)	Risk of Dying	Percent per Year	Number of Coin Flips All Coming Up Heads
100	1 in 10,000	0.01%	~13
1,000	1 in 1,000	0.1%	~10
10,000	1 in 100	1%	~7
100,000	1 in 10	10%	~3
500,000	1 in 2	50%	1

Figure 2-1: Purple Balls, MicroMorts, Risk of Dying, Percent Chance of Dying, Coin-flip Equivalent

People do not typically think about risk in this fashion and doing this mental exercise may be difficult, fun, or scary.

You must remember that the MicroMort estimates on this site are based on actual death certificates from the past seven years, and therefore, the website calculates death risks for the average person in each population group (or cohort). The question is, for your, age, lifestyle, and health choices, how many purple balls are in your urn? Depending on your lifestyle, your risk could be dramatically higher (if you are an obese, drunken skydiver) or lower (if you are a

vegetarian aerobics instructor) than the values shown, but the results provide a good starting point to think about the risks that you face.

Each purple ball can be tied to some cause of death. Some are related to accidents, some to illness. This website provides MicroMort estimates for up to 66 different causes (the number of causes shown depends on the display that you have selected). You can check to see what the number one cause of death is for your demographic group and see the distribution of the underlying causes of purple balls in your urn (or the urns of any cohorts that you choose to search).

In addition, this website allows comparisons across multiple variables, including age, gender, race, and geographic regions. The geographic comparisons can be done at the individual state level in the US or the country level in Europe or by regions in the US and Europe. For example, how do the death risks compare across age groups for males in California and in France?

The MicroMort concept was first articulated by Stanford University Professor Ron Howard in 1968 and has been used when discussing regulatory risk trade-offs. For more on the concept of a MicroMort and for examples of how it has been used to focus policy debates, the following references will be helpful:

- Professor Ron Howard's slides on MicroMorts: http://stanford-online.stanford.edu/sdrmda61w/session10b/slides/sld001.htm
- Professor Ron Howard's 1989 article on medical decision making and MicroMorts: http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=4034472
- Professor David Spiegelhalter's article in the Times of London on MicroMorts http://www.timesonline.co.uk/tol/comment/columnists/guest_contributors/article5696688.ece
- Professor Ron Howard's article in April 1984 issue of Management Science http://mansci.journal.informs.org/cgi/content/abstract/30/4/407

Tutorial 3: Searching Using One Variable

A search using one input criteria is the most straightforward to do and the easiest to interpret. Starting on the "Death Comparison" webpage, you select one of seven options in "Step 1: Primary Comparison" (See Figure 3-1). For example, to compare the risk of dying for males and females, select "Gender." To compare the risks for different age groups, select "Age."

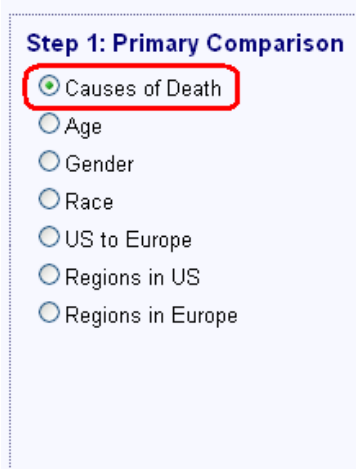


Figure 3-1: Choosing primary category

A basic search of immediate interest might be to find out your own risk of dying in the next year and from what cause. DeathRiskRankings.org cannot give you your exact profile, but it can provide the average risk for someone of your age, gender, race (in the US), and region. The search procedure is straightforward.

- **Step 1: Primary Comparison:** Select *Causes of Death* (Figure 3-1)
- **Step 2: Secondary Comparison:** Select *No Secondary Comparison* (more on the other options in Step 2 in the next tutorial).
- **Step 3: Search Criteria:** Select specific demographics about the person of interest (yourself). (Figure 3-2).
  - **Age:** 21 *(Note: individual ages are found after the age categories on the pull-down menu)*
  - **Gender:** Male
  - **US or Europe:** United States
  - **Region:** Pennsylvania
  - **Race:** White *(Note: race information is only available for the US comparisons)*
  - **Cause of Death:** Not selectable because it's selected in Step 1
  - **Simple Language:** Use the default, that is the less technical medical terms

- **Forecast Period:** Dying within 1 year

- **Step 4: Submit:** Click on the *Submit* button to complete the search (Figure 3-3)

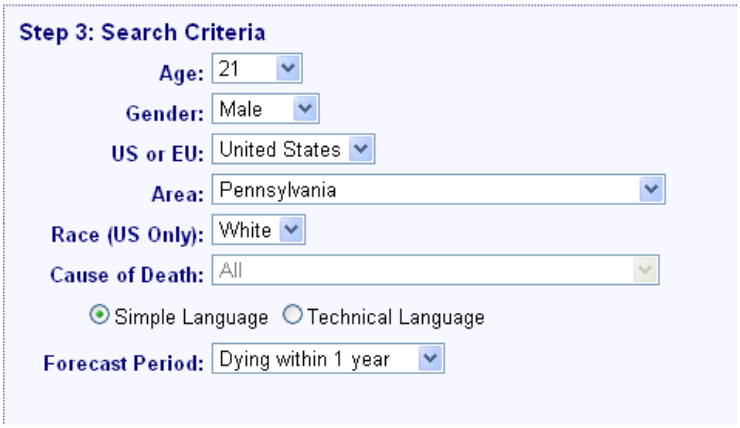


Figure 3-2: Search details

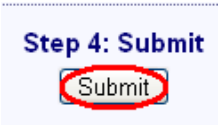


Figure 3-3: Completing the search

The search results will appear in the same window below the input options. A summary of the search will appear to make sure that the output displayed matches the search that you wanted and did not make a mistake (Figure 3-4). The *Search Parameters* table boxed in red, shows the specifics of your search. Using the buttons in the *Display Options* box, you can change the metric with which your results are displayed. These options will be discussed in Tutorial 6. For this first example, we use the default settings.

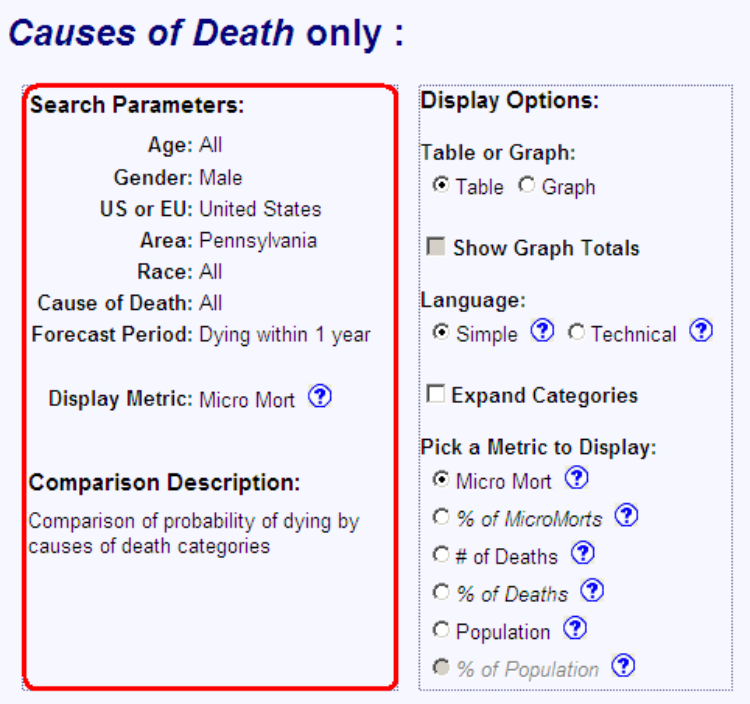


Figure 3-4: Output information

Figure 3-5 shows the output table for this search. The default output metric is the MicroMort, or one-in-a-million chance of dying. For more details on MicroMorts see Tutorial 2. A simple guiding principle is that MicroMorts are bad and that having a lot is not good.

The output table shows that for a 21-year old, white, male living in Pennsylvania, the chance of dying next year is 1,320 MicroMorts. So, if there were 1 million white 21-year olds in Pennsylvania, 1,320 would die next year. This is approximately a 1-in-750 chance of dying (that is 1,000,000/1,320 = 757). In percentage terms, there is a 0.13% chance of dying next year (1,320/1,000,000).

The table also shows details as to the causes of this chance of dying. The default display lists up to 20 summary categories of causes in alphabetical order and their associated MicroMorts. The sum of all the causes will equal the total. To sort based on MicroMorts, simply click on the *MicroMort* header in the table. Multiple clicks will alternate the sort between increasing or decreasing in value. Clicking on the *Cause of Death* header will sort the table alphabetically by the causes. By visually scanning down the list, it is easy to pick out the top causes: Accidents (with 757 MicroMorts), Suicides (with 213 MicroMorts), and Homicide (with 81 MicroMorts).

Cause of Death	MicroMort
Total	**1,320**
Accidents	757
Accidents - Other	7
Blood diseases	4
Cancers	57
Circulatory system diseases	44
Congenital defects	12
Digestive system diseases	7
Endocrine/metabolic diseases	13
Homicide	81
Ill-defined symptoms/causes	39
Infectious and parasitic diseases	10
Mental and behavioral disorders	9
Musculoskeletal diseases	2
Nervous system diseases	21
Non-cancerous growths	5
Respiratory diseases	17
Suicide	213
Undetermined - Homicide Suicide or Poisoning?	18
Urinary tract diseases	4

Figure 3-5: MicroMort output display

It is interesting to note that these three top causes (totaling 1,051 MicroMorts or 80% of the total of 1,320) have nothing to do with disease. Cancers are the fourth leading cause with 57 MicroMorts (comprising approximately 4% of total mortality risk); however, 57 MicroMorts translates to about a 1 in 17,500 chance of dying in the coming year. This is just as likely as flipping 16 fair coins and having all of them come up heads.

This single variable search allows the user to explore death risk by one dimension. However, in many cases, you want to know how risks compare between different groups (for example, males versus females, or US versus Europeans). To do this, you need to run a two-variable search, which is described in Tutorial 4.

Tutorial 4: Searching Using Two Variables

Searching with two variables allows a user to compare the risks of dying across multiple groups. With this site, many questions can be answered. For example: For which causes do men have a higher death risk than females? Or, alternatively, how does accident risk across different age groups compare between the US and Europe?

To complete a two-variable search, you need to make selections in both **Step 1: Primary Comparison** and **Step 2: Secondary Comparison**. The option selected in Step 1 will determine what items will be in each row of the output table. For example, if you select *Age* in Step 1, then each row will be a different age category. The option selected in Step 2 will determine what items will be in each column. For example, if you select *Gender* in Step 2, then there will be two columns, one for males and one for females. Step 3 still allows you to make the search specific for a particular group or cause. Figure 4-1 illustrates the general layout for these four steps. ..

Figure 4-1: Two-variable search

Walking through an example will help explain the different options. Suppose you wanted to compare the one-year US suicide risk for males and females across different age categories.

To do this (see Figures 4-2 and 4-3),

- **Step 1: Primary Comparison:** Select *Age*
- **Step 2: Secondary Comparison:** Select *Gender*
- **Step 3: Search Criteria:** Select specific demographics of the population of interest
  - **Age:** Not selectable because it's selected in Step 1
  - **Gender:** Not selectable because it's selected in Step 2

- **US or EU:** United States
- **Region:** All
- **Race:** All
- **Cause of Death:** Suicide
- **Simple Language:** Use the default *Simple Language*
- **Forecast Period:** Dying within 1 year is only available because *Age* is selected in Step 1.

- **Step 4: Submit:** Click on the *Submit* button to complete the search

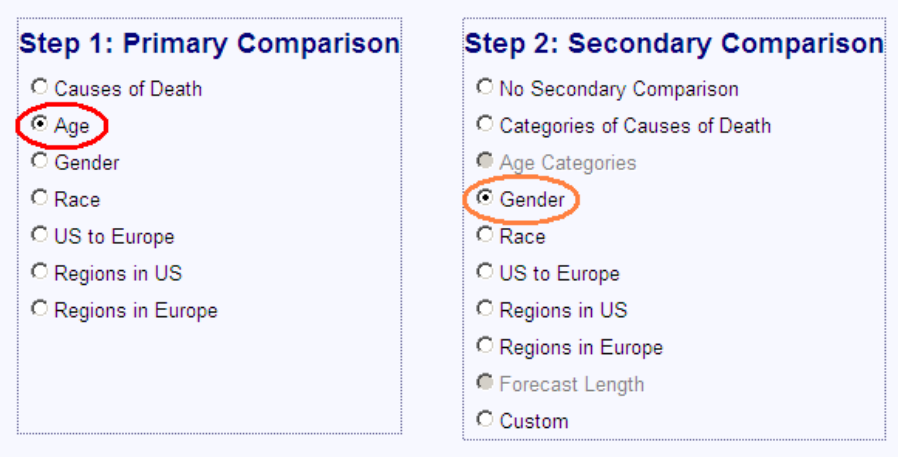


Figure 4-2: Setting categories

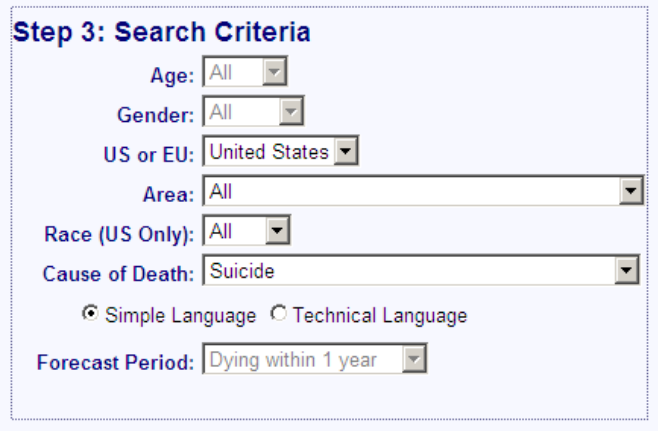
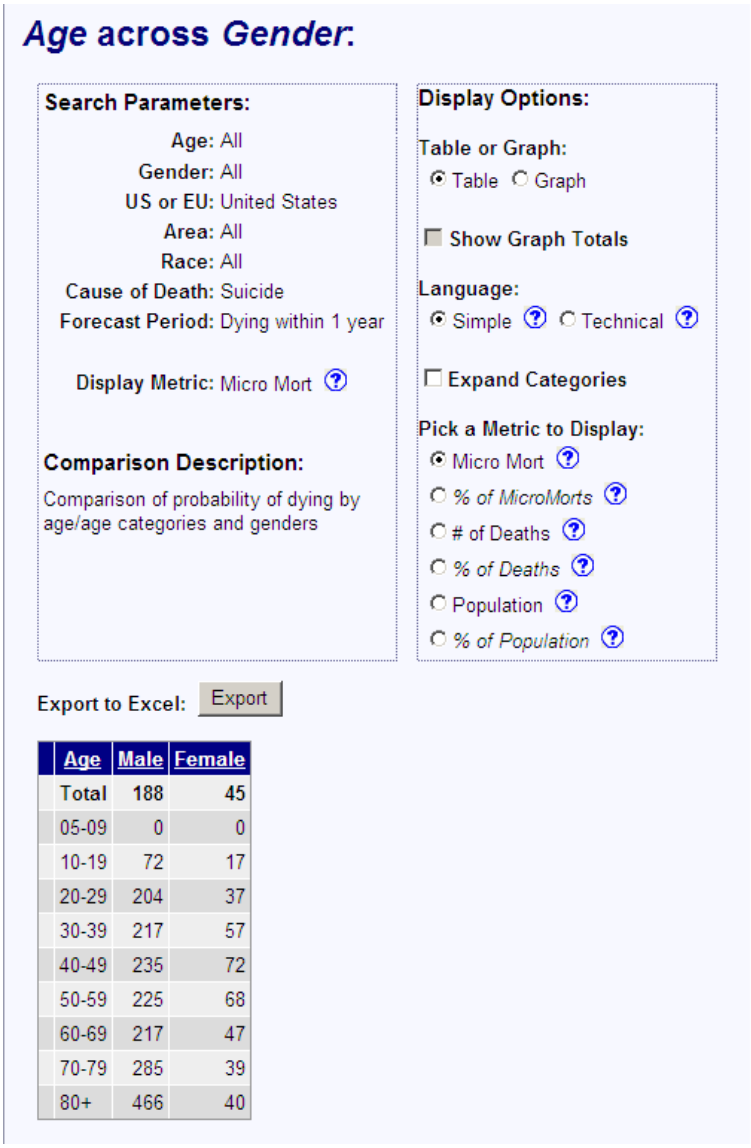


Figure4-3: Specifications

The results from this search, including the summary of the search criteria, are shown in Figure 4-4. Double check to make sure that you have completed the search that you wanted to do. Figure 4-4 uses the default display settings. We explain further display options in Tutorial 5.



Age across Gender:

Search Parameters:

Age: All
Gender: All
US or EU: United States
Area: All
Race: All
Cause of Death: Suicide
Forecast Period: Dying within 1 year

Display Metric: Micro Mort

Comparison Description:

Comparison of probability of dying by age/age categories and genders

Display Options:

Table or Graph:
(•) Table ( ) Graph

[ ] Show Graph Totals

Language:
(•) Simple ( ) Technical

[ ] Expand Categories

Pick a Metric to Display:
(•) Micro Mort
( ) % of MicroMorts
( ) # of Deaths
( ) % of Deaths
( ) Population
( ) % of Population

Export to Excel: Export

Age	Male	Female
Total	188	45
05-09	0	0
10-19	72	17
20-29	204	37
30-39	217	57
40-49	235	72
50-59	225	68
60-69	217	47
70-79	285	39
80+	466	40

Figure 4-4: Resulting chart form age and gender comparison for suicide risk in the US

The output table reveals a number of interesting results. The first column shows the age categories in ten-year blocks, except the first (5-9 years old) and last (80 and older). The second and third columns show the risk (in MicroMorts) for men and women, respectively. The output reveals that male suicide risks are greater than females for every age category (except the youngest where there is no risk). The suicide risk for women increases to age 40-49, and then decreases as women get older. For men, however, the suicide generally continue to increase, with the 80+ category having the highest risk, with males being more than 11 times more likely to commit suicide than women.

To compare two specific groups, a custom search is needed. Tutorial 5 shows how this is done.

Tutorial 5: Customized Searches

The **Custom** option in **Step 2 Secondary Comparison** provides users with an opportunity to compare two specific populations (for example, compare the risk and causes of deaths for 21-year old males in Pennsylvania and France).

To do this (see Figures 5-1),

- **Step 1: Primary Comparison:** Select *Causes of Death*
- **Step 2: Secondary Comparison:** Select *Custom*

As soon as *Custom* is selected a new input box appears (*Comparison Group 2*).

- **Step 3: Search Criteria:** Select specific demographics about the person of interest (yourself).
    - **Age:** 21
    - **Gender:** Male
    - **US or EU:** United States
    - **Region:** Pennsylvania
    - **Race:** White
    - **Cause of Death:** Not selectable because it's selected in Step 1
    - **Simple Language:** Use the default *Simple Language*
    - **Forecast Period:** Dying within 1 year is only available with *Custom* searches

Repeat the step for **Comparison Group 2**, but this time for a 21-year old in France.

    - **Age:** 21
    - **Gender:** Male
    - **US or EU:** Europe
    - **Region:** France
    - **Race:** Not selectable because race data is not available for Europe
    - **Cause of Death:** Not selectable because it's selected in Step 1
    - **Forecast Period:** Dying within 1 year

- **Step 4: Submit:** Click on the *Submit* button to complete the search

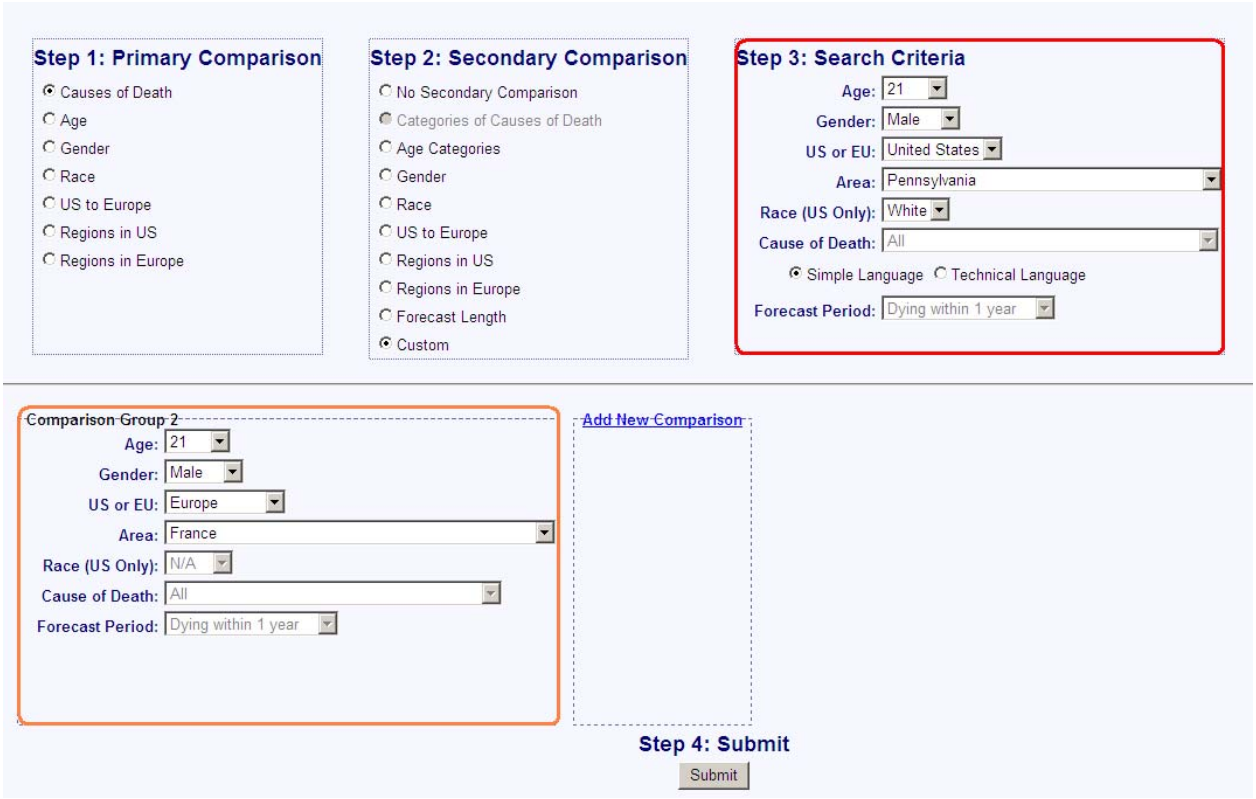


Figure 5-1: Criteria for individual search

To add other groups to the comparison (for example, a 21-year old German male), the user can click on the *Add New Comparison* link as seen in Figure 5-2. This will new input boxes as shown in Figure 5-3.

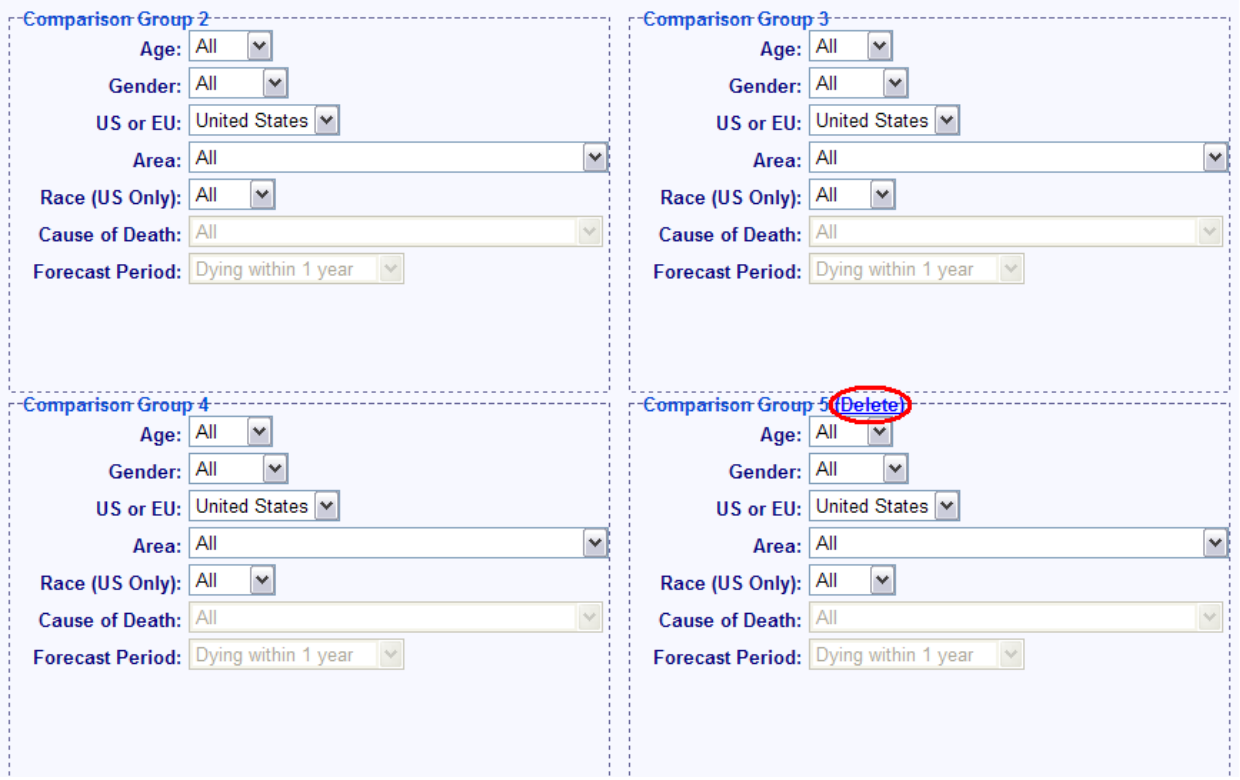


Figure5-2: Expanding search

Because two groups are being compared, there are two entries (separated by commas) for each parameter in the Search Parameter box.

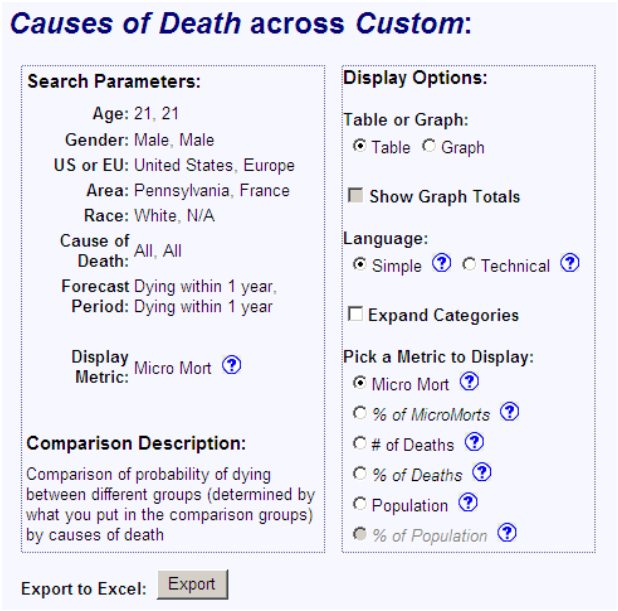


Figure 5-3: Details about the search

Cause of Death	Group 1	Group 2
Total	**1,320**	**929**
Accidents	757	488
Accidents - Other	7	1
Blood diseases	4	3
Cancers	57	56
Circulatory system diseases	44	25
Congenital defects	12	12
Digestive system diseases	7	4
Endocrine/metabolic diseases	13	9
Homicide	81	15
Ill-defined symptoms/causes	39	84
Infectious and parasitic diseases	10	7
Mental and behavioral disorders	9	12
Musculoskeletal diseases	2	1
Nervous system diseases	21	32
Non-cancerous growths	5	5
Respiratory diseases	17	10
Skin diseases	0	0
Suicide	213	149
Undetermined - Homicide Suicide or Poisoning?	18	16
Urinary tract diseases	4	0

Figure 5-4: Customization outputs

Output from this search is shown in Figure 5-4. Group 1 is the 21-year old male in Pennsylvania (the first set of search criteria) and Group 2 is corresponding person in France. The results can be sorted by clicking on Group 1 or Group 2.

The output reveals that the Frenchman has a lower risk of dying (929 MicroMorts compared to 1,320 MicroMorts for the Pennsylvanian, approximately 30% lower risk). Most of this difference can be attributed to the risk of death from **a**ccidents, **s**uicides, and **h**omicides (**ASH**). The total of these differences is 405 MicroMorts. ASH causes account for 80% (1,058/1,320) of the Pennsylvanian's risk but only 70% (653/929) of the Frenchman's risk.

Tutorial 6: Changing the Output Display

So far in all the tutorials, the outputs have been shown in MicroMorts (chance in a million of dying). However, the web site allows for the display of other output metrics. This tutorial explains the available options. Throughout this tutorial, a search comparing causes of death between US and European 50-59 year old males will be used to demonstrate the options. The details of the search parameters for this search are in Figure 6-1.

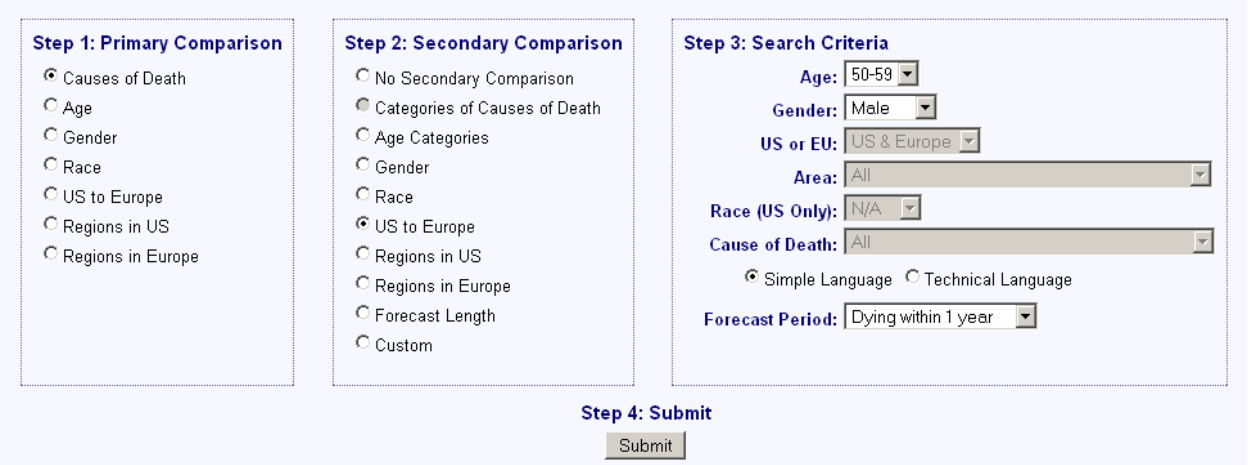


Figure 6-1: Details of the input search parameters

After the search has been completed, the *Display Option* box, which is next to the *Search Parameter* summary, controls the output metrics and display characteristics (see Figure 6-2).

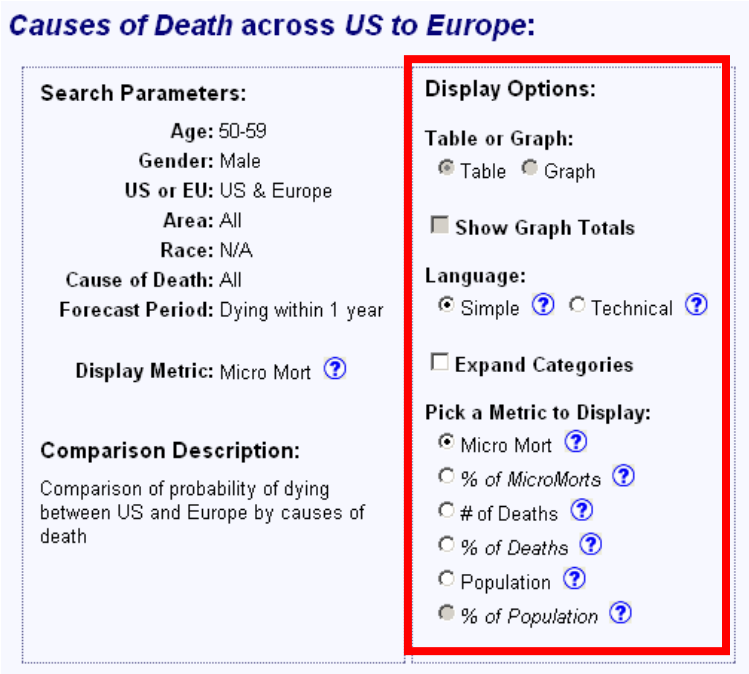


Figure 6-2: Display options

The output table with the default MicroMorts metric is shown in Figure 6-3. The output shows that for 50-59 year old males, the risk of dying in the next year is higher for Europeans.

Cause of Death	United States	Europe
Total	**7,810**	**8,379**
Accidents	479	463
Accidents - Other	25	9
Blood diseases	27	12
Cancers	2,302	3,017
Circulatory system diseases	2,564	2,554
Congenital defects	21	16
Digestive system diseases	504	702
Endocrine/metabolic diseases	343	146
Homicide	57	24
Ill-defined symptoms/causes	99	299
Infectious and parasitic diseases	338	119
Mental and behavioral disorders	130	142
Musculoskeletal diseases	27	14
Nervous system diseases	134	131
Non-cancerous growths	31	44
Perinatal problems	0	0
Respiratory diseases	369	309
Skin diseases	8	4
Suicide	225	272
Undetermined - Homicide Suicide or Poisoning?	25	47
Urinary tract diseases	102	55

Figure 6-3: Output table in MicroMorts

Just by scanning the output table, it may be hard to determine which causes lead to this difference. However, by single clicking on the *United States* (or *Europe*) title, the table can be sorted from high to low or low to high for that column (see Figure 6.4). The sort shows that there is essentially no difference in the circulatory system risk, but there are large differences in cancer risks. Follow-on queries could investigate whether this difference between US and Europeans holds for women, other age categories, sub-regions/states/countries, and specific types of cancers.

The following **Display Options** are available:

- **Table or Graph:** Some of the output tables allow for graphical displays. The graphs use default settings that may not be the most useful. For more sophisticated graphs, we recommend that you export the output data to Excel to facilitate custom tailoring of your graphics. Export options are discussed later in this tutorial.

- **Language: Simple or Technical**: This function enables the user to display certain parts

the output to professional terms found in the ICD 10 (see Figure 6.5). The results of selecting technical terminology can be seen in Figure 6-6.

	Cause of Death	United States	Europe
	Total	**7,810**	**8,379**
	Circulatory system diseases	2,564	2,554
	Cancers	2,302	3,017
	Digestive system diseases	504	702
	Accidents	479	463
	Respiratory diseases	369	309
	Endocrine/metabolic diseases	343	146
	Infectious and parasitic diseases	338	119
	Suicide	225	272
	Nervous system diseases	134	131
	Mental and behavioral disorders	130	142
	Urinary tract diseases	102	55
	Ill-defined symptoms/causes	99	299
	Homicide	57	24
	Non-cancerous growths	31	44
	Blood diseases	27	12
	Musculoskeletal diseases	27	14
	Accidents - Other	25	9
	Undetermined - Homicide Suicide or Poisoning?	25	47
	Congenital defects	21	16
	Skin diseases	8	4
	Perinatal problems	0	0

Figure 6-4: Output table in MicroMorts sorted by causes in the US

- **Expanding categories:** This allows the *Primary Comparison* categories (shown on the left side of the table) to be shown in greater detail. For example, the 21 "Cause of Death" categories can be expanded to 66 categories (see Figure 6-7). The 10-year age categories can be expanded to individual ages, regions in the US can be expanded to individual states, and regions in Europe can be expanded to individual countries. Step 2: *Secondary Comparison* categories cannot be expanded.

- **Pick a metric to display**: Using this option, the default output metric, MicroMorts can be changed to percentages (see Figure 6-8), number of deaths (see Figure 6-9 and 10), and population (see Figure 6-11).

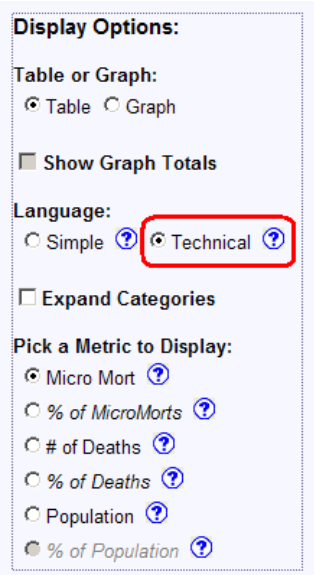


Figure 6-5: Output table in MicroMorts

Cause of Death	United States	Europe
Total	**7,810**	**8,379**
Accidents	479	463
Certain conditions originating in the perinatal period	0	0
Congenital malformations and chromosomal abnormalities	21	16
Diseases of the Blood	27	12
Diseases of the circulatory system	2,564	2,554
Diseases of the digestive system	504	702
Diseases of the genitourinary system	102	55
Diseases of the musculoskeletal system/connective tissue	27	14
Diseases of the nervous system and the sense organs	134	131
Diseases of the respiratory system	369	309
Diseases of the skin and subcutaneous tissue	8	4
Endocrine, nutritional and metabolic diseases	343	146
Events of undetermined intent	25	47
Homicide, assault	57	24
Infectious and parasitic diseases	338	119
Malignant neoplasms	2,302	3,017
Mental and behavioural disorders	130	142
Other external causes of injury and poisoning	25	9
Other Neoplasms	31	44
Suicide and intentional self-harm	225	272
Symptoms, signs, abnormal findings, ill-defined causes	99	299

Figure 6-6: Technical language

	Cause of Death	United States	Europe
	Total	**7,810**	**8,380**
	Accidents: Falls	48	86
	Accidents: Other	100	163
	Accidents: Others	25	9
	Accidents: Poisoning	106	45
	Accidents: Transportation	225	169
	AIDS	136	19
	Alcohol abuse	97	123
	Arthritis	4	2
	Asthma	13	16
	Blood disease	27	12
	Cancer: Bladder	39	65
	Cancer: Breast	4	5
	Cancer: Colon	171	165
	Cancer: Esophagus	109	146
	Cancer: Kidney	80	82
	Cancer: Lip/mouth	72	212

...

	Pneumonia	82	101
	Respiratory system: Other	273	191
	Skin disease	8	4
	Sudden infant death syndrome (SIDS)	0	1
	Suicide	225	272
	Tuberculosis	5	38
	Ulcers	14	35
	Undetermined: Homicide, suicide, or poisoning?	25	47
	Unknown/unspecified	80	254
	Urinary system: Other	10	6

Figure 6-7: Expanding categories

	Cause of Death	United States	Europe
	Total	**100.0**	**100.0**
	Accidents	6.1	5.5
	Accidents - Other	0.3	0.1
	Blood diseases	0.3	0.1
	Cancers	29.5	36.0
	Circulatory system diseases	32.8	30.5
	Congenital defects	0.3	0.2
	Digestive system diseases	6.5	8.4
	Endocrine/metabolic diseases	4.4	1.7
	Homicide	0.7	0.3
	Ill-defined symptoms/causes	1.3	3.6
	Infectious and parasitic diseases	4.3	1.4
	Mental and behavioral disorders	1.7	1.7
	Musculoskeletal diseases	0.3	0.2
	Nervous system diseases	1.7	1.6
	Non-cancerous growths	0.4	0.5
	Perinatal problems	0.0	0.0
	Respiratory diseases	4.7	3.7
	Skin diseases	0.1	0.0
	Suicide	2.9	3.2
	Undetermined - Homicide Suicide or Poisoning?	0.3	0.6
	Urinary tract diseases	1.3	0.7

Figure 6-8: Percent of MicroMorts assigned to each cause

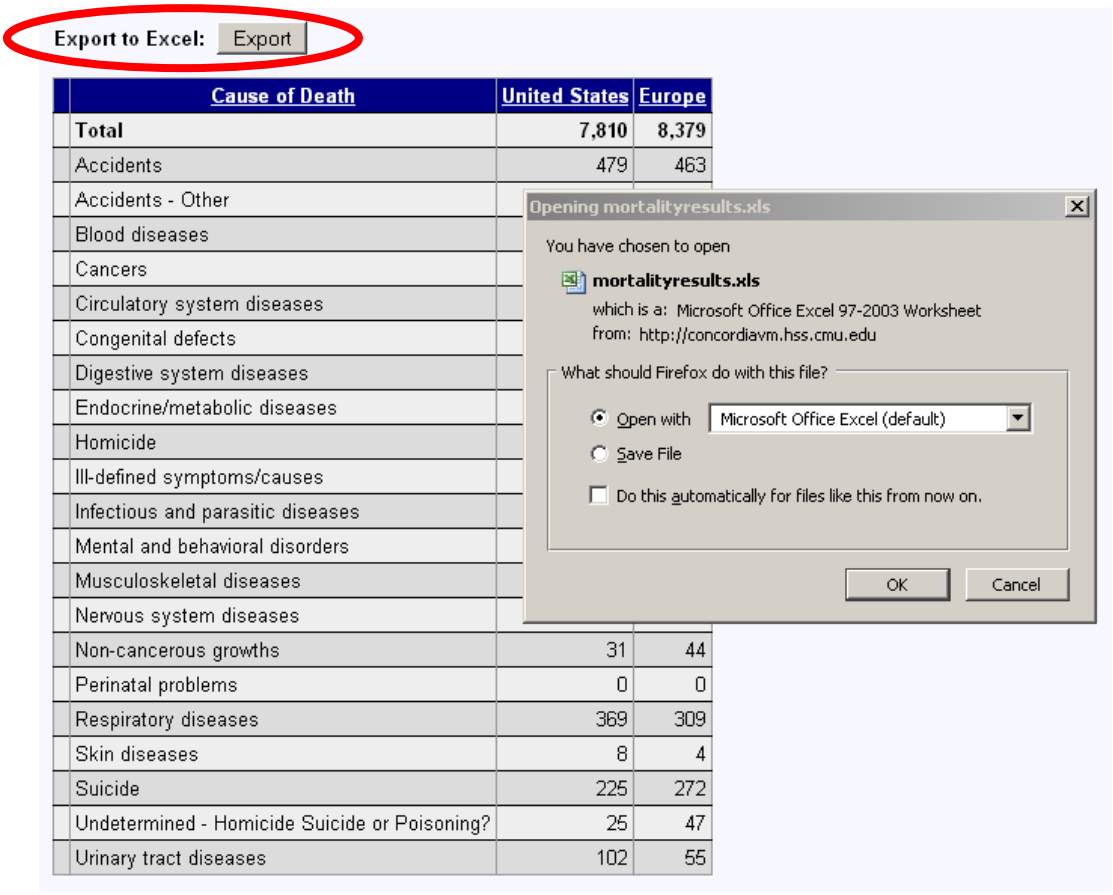
	Cause of Death	United States	Europe
	Total	**124,729**	**253,417**
	Accidents	7,645	14,006
	Accidents - Other	399	272
	Blood diseases	433	361
	Cancers	36,772	91,261
	Circulatory system diseases	40,946	77,252
	Congenital defects	335	477
	Digestive system diseases	8,047	21,237
	Endocrine/metabolic diseases	5,479	4,421
	Homicide	904	712
	Ill-defined symptoms/causes	1,574	9,041
	Infectious and parasitic diseases	5,397	3,588
	Mental and behavioral disorders	2,084	4,295
	Musculoskeletal diseases	423	418
	Nervous system diseases	2,142	3,972
	Non-cancerous growths	500	1,321
	Perinatal problems	1	2
	Respiratory diseases	5,899	9,359
	Skin diseases	131	106
	Suicide	3,590	8,222
	Undetermined - Homicide Suicide or Poisoning?	405	1,431
	Urinary tract diseases	1,623	1,663

Figure 6-9: Output in Expected # of Deaths in a Year

	Cause of Death	United States	Europe
	Total	15,971,881	30,246,368

Figure 6-10: Sorting with different units of measure

- **Export to Excel**: All output tables can be exported to Microsoft Excel. This allows for additional comparisons to be made, search results to be archived, and graphs to be produced. Just above and to the right of the output table is the *Export* button. Selecting it will lead to a pop-up window (see Figure 6-11).



Cause of Death	United States	Europe
Total	7,810	8,379
Accidents	479	463
Accidents - Other		
Blood diseases		
Cancers		
Circulatory system diseases		
Congenital defects		
Digestive system diseases		
Endocrine/metabolic diseases		
Homicide		
Ill-defined symptoms/causes		
Infectious and parasitic diseases		
Mental and behavioral disorders		
Musculoskeletal diseases		
Nervous system diseases		
Non-cancerous growths	31	44
Perinatal problems	0	0
Respiratory diseases	369	309
Skin diseases	8	4
Suicide	225	272
Undetermined - Homicide Suicide or Poisoning?	25	47
Urinary tract diseases	102	55

Figure 6-11: Exporting information to excel

Tutorial 7: Simpson's Paradox

When viewing output from death comparison searches, you must be extremely careful in how you interpret the results. A paradox often taught in introductory statistic classes is common.

For example, a simple comparison between **white and black females in the US for all causes across the age categories** shows the following results (in MicroMorts):

Age	White	Black
05-09	127	190
10-19	271	307
20-29	456	816
30-39	874	1,848
40-49	1,905	3,978
50-59	4,342	8,092
60-69	11,380	17,842
70-79	30,192	37,219
80+	98,116	101,836

Figure 7-1: MicroMort comparison between black and white US females by age categories

For each age category, black females have much higher death risk then white females. In fact for the "30-39" and "40-49" age categories, the black female risk of dying in the next year is over twice that of the white females. This is an important fact that should generate a focused policy discussion as to the causes of the disparity.

Suppose that instead of looking at each age category, you looked at the same search, but did for all ages at once. You would get the following results:

Age	White	Black
Total	9,537	7,612

Figure 7-2: MicroMort comparison between black and white US females for all ages

How is it possible that **overall** black females have much lower death risk? ***In every age category, whites are less then blacks, but over all blacks are less then whites*!!** This is a classic paradox.

This paradox was first identified in 1951 by Edward Simpson and two of his colleagues Karl

Pearson and Udny Yule and came to be known as *Simpson's Paradox*. It occurs because blacks and whites have different age distributions. On average, white females are older than black females, and older females have much higher risk than younger females. So, in fact when you compare the overall risk, you are comparing old white women with young black women, and the black population has less risk.

Here is the age distribution for the two groups:

Age	White	Black
05-09	6.9%	9.1%
10-19	14.1%	18.1%
20-29	13.7%	16.3%
30-39	15.3%	16.6%
40-49	16.1%	15.6%
50-59	12.7%	10.7%
60-69	8.9%	6.5%
70-79	6.9%	4.4%
80+	5.4%	2.8%

Figure 7-2: Population distribution comparison between black and white US females

For every age category "30-39" and younger, the black population has a higher percent than the white, and for every age category "40-49" and older, the white population has a higher percent than the black. This results in the average white woman being older and having higher risk than the average black woman.

A policy discussion that only looked at the overall average risk would reach the completely wrong conclusion.

Simpson's paradox is not limited to race differences among US women. It is found throughout the death comparison data. For example, when comparing US men and women, overall there appears to be no difference, but if broken down by age categories, men have much higher risks for every age group. This is because, on average, the female population is older then the male population. US and European comparisons are also susceptible to the paradox. The European population is on average older than the US.

In order to not fall for the paradox, **it is critical that any comparison between two populations must be done using age categories.**