

VitalPro User's Guide

Texas ICD-10 Underlying Cause Death



Expert Health Data Programming, Inc.

VitalPro User's Guide

Texas ICD-10 Underlying Cause Death

Expert Health Data Programming, Inc.

**VitalPro User's Guide
Texas ICD-10 Underlying Cause Death**

Copyright 1998-2006 by Expert Health Data Programming, Inc. All rights Reserved

Published 2006 by Expert Health Data Programming, Inc., Renton, WA USA

This user's guide has been carefully prepared, but is not guaranteed against errors or omissions. Please report any errors, omissions, or suggestions concerning the user's guide, VitalPro, or Vitalnet, using the contact system at the www.ehdp.com web site. The web site also has current information on VitalPro software and other Vitalnet systems.

Birtha, Epidemic, Epigram, MedTrend, MultiCod, Oncogram, PopTrend, PregData, Vitalnet, VitalWeb, VitalPro, and VitalWeb Wizard are trademarks of Expert Health Data Programming, Inc. Other product and company names mentioned may be the trademarks of their owners.

Table of Contents

1: Introduction	2
Advantages and Benefits	3
Data Within Epigram	4
Access Methods	5
DSHS Network Access	5
Confidentiality Policy	6
2: Understanding Results	7
Typical Epigram Table	8
Table #1 to Interpret	9
Table #1, Interpreted	10
Table #2 to Interpret	11
Table #2, Interpreted	12
3: Using the Epigram Interface	13
Epigram Main Menu	14
Epigram Submenus	15
Geographic Submenu	16
Select Another Submenu	17
Years Submenu	18
Produce a Table	19
Save Results to Log File	20
Formats for Exporting	21
Explore the Data	22
Continue with Analyses	23
Outlines of Submenus	24
Submenus for Designing Table Layout	25
Submenus for Modifying Data Variables, Other Settings	26
More Practice Analyses	27
4: Epigram Parameters	28
Statistic [Outcome]	29
Row Variable, Row Sorting, Column Variable	30
Multi-Tables	31
Geo Selection [Area Sets]	32
Causes of Death [ICD]	32
Other Variables	33
Race Groups	33
Age Groups	33
Sex	33
Years	33
Other Settings	34
Confidence Intervals	34
Decimal Places	34
Cell Suppression	35
Trend Analysis	36
Glossary	37

1: Introduction

Vitalnet is a comprehensive, integrated system for analyzing health data. Texas Epigram is the Vitalnet module for analyzing Texas mortality data. The software is designed to be easily used without a user's guide. However, many users learn better from a written text, and all users will benefit from an overview of what the software can do before using it.

This user's guide describes Epigram Professional Version (VitalPro), which runs directly on a PC or LAN. Web Versions of Vitalnet (VitalWeb) are described at the www.ehdp.com web site. Either Vitalnet system will greatly ease and speed your work.

Notes for Adobe Acrobat pdf version: The screen snapshots are best viewed at a 125% magnification, and will be somewhat fuzzy at other magnifications. The snapshots will be sharp when printed.

This user's guide incorporates a tutorial. Carrying out the tutorial will help you quickly become an expert at using Vitalnet. All procedures you are expected to carry out as part of the tutorial are highlighted by a different text appearance, as shown in this example:

Sample Tutorial Step - Press 'A' to add Texas as one area set.

Here's how this user's guide is organized:

Chapter 1: Introduction - Describes general characteristics of Texas Epigram, access methods, and confidentiality requirements.

Chapter 2: Understanding Results - Explains the organization and content of Epigram tables (the results you get from using Epigram). Shows and explains typical tables.

Chapter 3: Using the Epigram Interface - Describes how to use Epigram menus to select parameters, produce tables, and save output, with examples. Guides you step-by-step to produce your first table. Outlines available menus.

Chapter 4: Texas Epigram Parameters - Describes allowed selections for age groups, area sets (counties and regions), ICD-10 sets (causes of death), race groups, years, and other options.

Glossary - Defines terms related to analyzing mortality data and using Epigram.

Citation for Epigram - Expert Health Data Programming, Inc., Texas VitalPro User's Guide: Data Warehouse Software for Analyzing Texas ICD-10 Underlying Cause Death. Renton, Washington. 1998-2006. Browse www.ehdp.com for more information about the software or to contact EHDP.

Acknowledgements - We gratefully acknowledge the cooperation and assistance of staff from the Texas Department of State Health Services and other users.

Advantages and Benefits

Epigram makes it easy to analyze Texas mortality data.

- **Fast** - You get results in seconds or minutes. Depending on the analysis, alternate methods could easily require hours to weeks to complete.
- **Flexible** - A wide variety of tables may be produced. Set rows and columns however you want. Standard parameters may be selected and combined as needed.
- **Efficient** - You can make a whole series of tables with one keystroke (multi-tables). Bar graphs allow for quick scanning for trends.
- **Easy to use** - Operations are menu-driven. Scrolling windows are used to select items from lists. You don't need to know any special codes such as FIPS codes.
- **Well-documented** - Although Epigram has been designed to be self-explanatory, it also includes extensive on-line help. Each menu has its own help screen. Also, you may select from a list of on-line help topics, providing advice on all topics related to Epigram. The on-line help complements the information included in this user's guide. Useful on-line reference materials are also provided, such as tables of standard populations used for age-adjustment. Finally, all output is fully documented.
- **Integrates with other software** - Tables may be saved to a log file in ASCII format for subsequent editing and printing with any editor. You may also save Epigram tables in CSV format or as a dBASE III file, for easy importing into data analysis, spreadsheet, mapping, graphing or other presentation software. Or, tables may be saved as HTML for display in a web browser.

Data Within Epigram

Epigram links and analyzes geographic, population, mortality, and ICD-10 data.

- **Geographic information** - Epigram includes a database of Texas counties and regions, linked to the population and mortality data sets.
- **Population data** - The Center for Health Statistics at DSHS provides 1980-1999 Texas population data. The Texas State Data Center at Texas A&M provides Texas population data for 2000 and beyond. The Texas State Data Center periodically revises its estimates for years between censuses, but the changes are typically small. The population variables included within Epigram are age, county of residence, race, sex, and year. Population figures are used for calculating mortality rates. A separate Vitalnet interface, PopTrend, analyzes population data for demographic trends and to obtain denominators for analyzing other data sets.
- **Mortality data** - The Bureau of Vital Statistics at the Texas Department of State Health Services provides all Texas mortality data. For an estimate on when the next data file will be loaded, contact the Bureau. Mortality variables within Epigram include age of deceased, cause of death, county of residence, race, sex, and year.

Epigram analyzes by place of residence of the deceased. For example, if a Dallas resident died in a motor vehicle collision in Houston, Epigram would classify the death under Dallas. Standard mortality reports usually use place of residence (the other system is "place of occurrence").

Epigram analyzes by underlying cause (the cause which initiated the sequence of events leading to death). For example, if a death certificate lists rheumatoid arthritis, myocardial infarction (MI), and cardiac arrhythmia secondary to MI, the underlying cause is MI. Standard mortality reports usually use underlying cause.

- **Data linking** - Four data sets (geographic data, population data, mortality data, and ICD data) are linked with each other, so the Epigram data warehouse is greater than the sum of its parts.

Access Methods

Both local and remote access are available for Texas Epigram.

- **Local access (PC's and local area networks)** - If you are a Texas Department of State Health Services employee, you will typically use Texas Epigram from the DSHS local area networks. Texas Epigram may also be installed to a stand-alone IBM-compatible PC or laptop running any version of Windows.

Execution speed - Epigram is fast. Analyses that might otherwise take hours to weeks to set up and run are done in seconds or minutes. For those who access Texas Epigram on a PC network, the program execution speed depends on the type of PC you are using and the characteristics of the network you are working on. The program has been tested to work well on all IBM-compatible PC's, and simply runs faster on faster PC's.

DSHS Network Access

DSHS employees may access Texas Epigram from the DSHS networks.

Windows - The following assumes that you are a Texas Department of State Health Services employee, and are using a PC attached to a DSHS network. Using any version of Windows, click on the Epigram icon to start the program. The icon looks like a doctor's bag. If you do not have an icon, request that your network manager set things up so that Epigram can be run from your network.

When you are done - After you finish using Texas Epigram, you will still be logged on to your own server, so you do not need to reboot your PC. If you have a problem accessing Texas Epigram from a DSHS network, ask your local network manager for help first.

Confidentiality Policy

Texas Epigram users must comply with confidentiality requirements.

Confidentiality policy - Your use of Texas Epigram indicates your agreement to the following conditions: You will not try to use Texas Epigram results nor let anyone else use Texas Epigram results to learn the identity of a reported death, or for any purpose other than statistical analysis. If you discover the identity of a reported death, you will advise the Director of the Bureau of Vital Statistics at the Texas Department of State Health Services of the incident, will safeguard or delete the information that would identify the individual, will make no use of the knowledge, and will inform no one else of the discovered identity.

Cell suppression - If you so desire, Texas Epigram can suppress cells that have fewer than a user-defined number of deaths. Cell suppression can increase the confidentiality of written reports in some cases. Cell suppression [is described more fully later](#).

2: Understanding Results

Review - Texas Epigram is a powerful tool for analyzing Texas mortality data. The program compares just about anything with anything, and makes just about any kind of output table you want. Chapter 1 explained how Epigram makes analyzing mortality data easy, listed the linked data sets used by Texas Epigram, described the access methods, and spelled out the confidentiality policy.

What's in this chapter - Before jumping in and using the program (Chapter 3), it is advised to get a good understanding of the results of the program. This chapter explains the organization and content of Epigram tables (the results you get from using Epigram). This chapter explains the four sections of a table:

- **Header** - Basic analysis parameters.
- **Data section** - Numerical results.
- **Bar graphs** - Graphical results.
- **Footnotes** - Other analysis parameters.

In addition, this chapter shows examples of actual Texas Epigram tables, to give you an idea of what is possible. You are given an opportunity to practice interpreting sample tables.

Typical Epigram Table

Each table has four parts: header, data section, bar graphs, footnotes.

```
Deaths
Tabulated by Age and Sex
Years: 1999
Place of Residence: Galveston, Harris
ICD J40-J47: Chronic Lower Respiratory Diseases

Age          Male    Female   Total
-----
Birth-19      7       1        8
20-39         2       4        6
40-59        34      33       67
60-99+      359     399     758
-----
Total         402     437     839

Horizontal Bar Graphs (X = 26.6 Deaths, x = 13.3):

Age          Male          Female
-----
Birth-19     x
20-39
40-59       Xx              X
60-99+     XXXXXXXXXXXXXXXx XXXXXXXXXXXXXXXX
-----

Analysis Footnotes:
Unique ID, for Keeping Track of Analyses: 428HBCSW
Output Produced: Sat Apr 28 17:02:53 2001, by Texas VitalNet
Deaths Classified by: ICD-10, Underlying Cause, County of Residence
Mortality Data Source: TDH Bureau of Vital Statistics
```

The **header** documents the analysis. The table analyzes 1999 deaths for Galveston and Harris counties. Chronic lower respiratory disease (ICD J40-J47) is analyzed.

The **data section** contains the results, organized into columns and rows. In this example, there is one column for each sex and one row for each of four age groups. Verify there were 33 deaths in women age 40-59, 402 deaths in males, and a total of 839 deaths. Each result, such as 33, 402, and 839, is called a "cell".

The **horizontal bar graphs** represent the results in a simple graphical format. In this example, each 'X' (big X) symbol represents 26.6 deaths. The bar graphs are optional. The bar graphs clearly show that the great majority of deaths in this analysis were in the 60-99+ age group, for both sexes.

The **footnote** documents less important aspects of the analysis, such as when the table was produced, and assigns a unique ID to the table for future reference.

Table #1 to Interpret

Here is a sample table (footnotes omitted) for you to practice on. Try interpreting the table by filling in the spaces below. Then, check your answers with those on the next page.

Death Rate (per 100,000)			
Tabulated by Area Set and Sex			
Age: 20-39 Years			
Years: 1999			
Place of Residence: PHR 3, PHR 6, PHR 7, PHR 8			
ICD B20-B24: Human Immunodeficiency Virus [HIV] Disease			
Area Set	Male Rate, Deaths	Female Rate, Deaths	Total Rate, Deaths
PHR 3	13.6 120	2.1 18	7.9 138
PHR 6	16.2 120	6.0 43	11.1 163
PHR 7	10.1 34	4.8 15	7.6 49
PHR 8	12.7 39	2.0 6	7.4 45
Total	13.8 313	3.7 82	8.8 395

Horizontal Bar Graphs (X = 1.1 Deaths / 100,000, x = 0.55):

Area Set	Male Rate	Female Rate	Total Rate
PHR 3	XXXXXXXXXXXXXx	XX	XXXXXXXXXx
PHR 6	XXXXXXXXXXXXXXXx	XXXXXx	XXXXXXXXXXXXx
PHR 7	XXXXXXXXXXx	XXXXx	XXXXXXX
PHR 8	XXXXXXXXXXXXXx	XX	XXXXXXX
Total	XXXXXXXXXXXXXXX	XXXx	XXXXXXXXX

Header Basic result type (statistic): _____

Years analyzed: _____

Geographic areas analyzed: _____

ICD groups analyzed: _____

Age groups analyzed: _____

Data Section Rows variable: _____

Columns variable: _____

How many females died in PHR 6: _____

Death rate for males in PHR 7: _____

Bar Graphs Death rate symbolized by big 'X': _____

Table #1, Interpreted

Here is the same table, and a suggested interpretation. If you had a problem, try reading through the answers again or ask a local data analyst for help.

Area Set	Male Rate, Deaths	Female Rate, Deaths	Total Rate, Deaths
PHR 3	13.6 120	2.1 18	7.9 138
PHR 6	16.2 120	6.0 43	11.1 163
PHR 7	10.1 34	4.8 15	7.6 49
PHR 8	12.7 39	2.0 6	7.4 45
Total	13.8 313	3.7 82	8.8 395

Horizontal Bar Graphs (X = 1.1 Deaths / 100,000, x = 0.55):

Area Set	Male Rate	Female Rate	Total Rate
PHR 3	XXXXXXXXXXXXXx	XX	XXXXXXXXXx
PHR 6	XXXXXXXXXXXXXXXx	XXXXXx	XXXXXXXXXXXXx
PHR 7	XXXXXXXXXXx	XXXXx	XXXXXXX
PHR 8	XXXXXXXXXXXXXx	XX	XXXXXXX
Total	XXXXXXXXXXXXXXX	XXXx	XXXXXXXXX

Header	Basic result type (statistic): Death rate (per 100,000) Years analyzed: 1999 Geographic areas analyzed: Public Health Regions 3, 6-8 ICD groups analyzed: ICD B20-B24 (HIV/AIDS) Age groups analyzed: 20-39
Data Section	Rows used in this example: Row for each area set Columns used in this example: Column for each sex How many females died in PHR 6: 43 Death rate for males in PHR 7: 10.1 per 100,000
Bar Graphs	Death rate symbolized by big 'X': 1.1 per 100,000

Table #2 to Interpret

Here is another table. Fill in the spaces below. Then, check your answers on the next page.

Age-Adjusted Death Rate (per 100,000)						
Age Adjustment Standard Population: 2000 US						
Tabulated by Race-Ethnicity and Sex						
Years: 1999						
Place of Residence: Texas						
ICD C33: Malignant Neoplasm Of Trachea						
ICD C34: Malignant Neoplasm Of Bronchus And Lung						
		Male		Female		Total
Race		Rate,	Deaths	Rate,	Deaths	Rate, Deaths

White	86.8	4,201		46.4	3,055	62.7 7,256
Black	123.9	721		45.0	387	76.2 1,108
Hispanic	43.6	505		14.0	231	26.1 736
Other	41.4	48		16.7	32	27.1 80

Total	81.0	5,475		39.7	3,705	56.4 9,180

Horizontal Bar Graphs (X = 8.3 Deaths / 100,000, x = 4.1):						
		Male		Female		Total
Race		Rate		Rate		Rate

White		XXXXXXXXXXx		XXXXXx		XXXXXXXXXx
Black		XXXXXXXXXXXXXXXXXX		XXXXXx		XXXXXXXXXX
Hispanic		XXXXXXx		Xx		XXX
Other		XXXXX		XX		XXXx

Total		XXXXXXXXXX		XXXXX		XXXXXXXX

Header Basic result type (statistic): _____

Years analyzed: _____

Geographic areas analyzed: _____

ICD groups analyzed: _____

Data Section Rows used in this example: _____

Columns used in this example: _____

How many Black males died: _____

Death rate for Hispanic females: _____

Overall death rate: _____

Bar Graphs Death rate symbolized by big 'X': _____

Table #2, Interpreted

Here is the second table again, and our interpretation of the results.

Age-Adjusted Death Rate (per 100,000)						
Age Adjustment Standard Population: 2000 US						
Tabulated by Race-Ethnicity and Sex						
Years: 1999						
Place of Residence: Texas						
ICD C33: Malignant Neoplasm Of Trachea						
ICD C34: Malignant Neoplasm Of Bronchus And Lung						
Race	Male		Female		Total	
	Rate	Deaths	Rate	Deaths	Rate	Deaths
White	86.8	4,201	46.4	3,055	62.7	7,256
Black	123.9	721	45.0	387	76.2	1,108
Hispanic	43.6	505	14.0	231	26.1	736
Other	41.4	48	16.7	32	27.1	80
Total	81.0	5,475	39.7	3,705	56.4	9,180

Horizontal Bar Graphs (X = 8.3 Deaths / 100,000, x = 4.1):

Race	Male	Female	Total
	Rate	Rate	Rate
White	XXXXXXXXXXx	XXXXXx	XXXXXXXXx
Black	XXXXXXXXXXXXXXXXXX	XXXXXx	XXXXXXXXXX
Hispanic	XXXXXXx	Xx	XXX
Other	XXXXX	XX	XXXx
Total	XXXXXXXXXX	XXXXX	XXXXXXXX

Header	Basic result type (statistic):	Age-adjusted rate
	Years analyzed:	1999
	Geographic areas analyzed:	Texas
	ICD groups analyzed:	ICD C33-C34
Data Section	Rows used in this example:	Row for each race
	Columns used in this example:	Column for each sex
	How many Black males died:	721
	Death rate for Hispanic females:	14.0 per 100,000
	Overall death rate:	56.4 per 100,000
Bar Graphs	Death rate symbolized by big 'X':	8.3 per 100,000

3: Using the Epigram Interface

Review - Chapter 1 explained how Epigram makes it easy to analyze Texas mortality data, listed the linked data sets contained within Texas Epigram, discussed the access methods, and described the confidentiality policy. Chapter 2 explained the layout of an output table, and presented sample tables for discussion and interpretation.

What's in this chapter - Finally, you will get to use the program! This is probably what you have been waiting for! You will learn to navigate the interface, select parameters, and produce a few tables like those you learned about in the previous chapter. The chapter also has a schematic overview of all Texas Epigram menus, for your reference.

Access the program - A tutorial runs through this chapter. All procedures you are expected to do as part of the tutorial are highlighted in a different type style, as shown below. At this point, do the following:

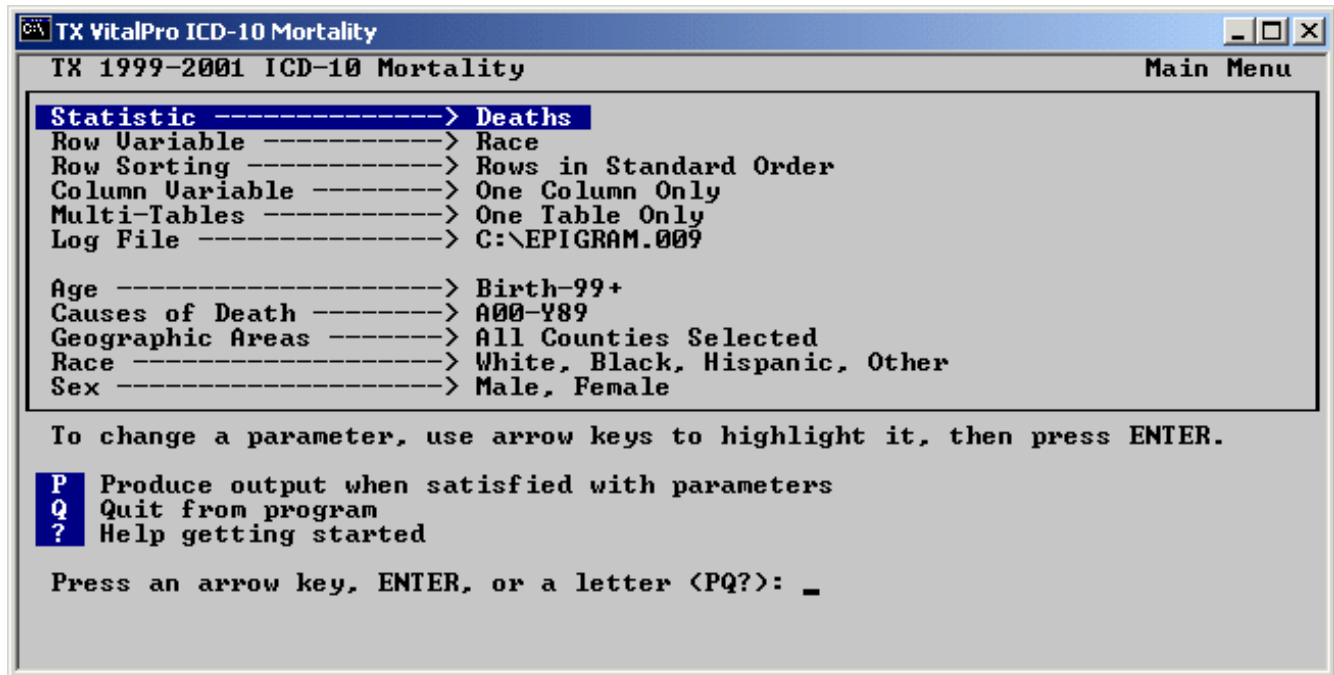
Access and start Texas Epigram by clicking on the icon.

The icon looks like a doctor's bag. If you do not have an icon, request that your network manager set things up so that Epigram can be run from your network. If needed, get assistance from another Texas Epigram user. You will start at the Main Menu (shown on next page).

Epigram Main Menu

You will constantly return to the Main Menu.

After Epigram starts, you are presented with the Main Menu, similar to the following:



```
TX VitalPro ICD-10 Mortality
TX 1999-2001 ICD-10 Mortality Main Menu
Statistic -----> Deaths
Row Variable -----> Race
Row Sorting -----> Rows in Standard Order
Column Variable -----> One Column Only
Multi-Tables -----> One Table Only
Log File -----> C:\EPIGRAM.009

Age -----> Birth-99+
Causes of Death -----> A00-Y89
Geographic Areas -----> All Counties Selected
Race -----> White, Black, Hispanic, Other
Sex -----> Male, Female

To change a parameter, use arrow keys to highlight it, then press ENTER.

P Produce output when satisfied with parameters
Q Quit from program
? Help getting started

Press an arrow key, ENTER, or a letter <PQ?>: _
```

The Main Menu is the "command center" - You move to submenus to modify parameters, but return to the Main Menu to make a table. The Main Menu lists all parameters currently selected, providing an overview.

Highlighting a parameter - One of the parameters, such as "Outcome", is highlighted with a "light bar" that you move by pressing an arrow key. The parameter list will scroll when you reach the bottom. Note that you do not use a mouse to run the program. Do the following:

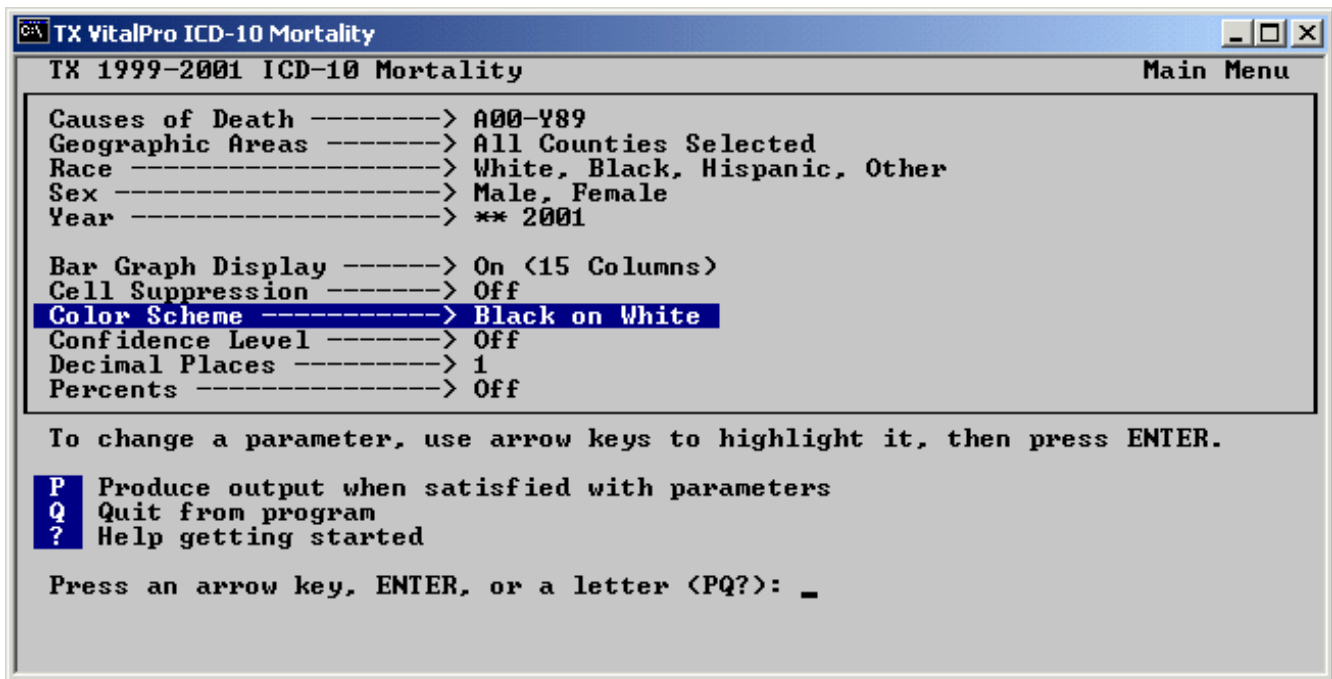
Press down arrow key to move light bar down. Keep pressing to see the parameter list scroll down.

The prompt is worth reading - The Main Menu (and every other Epigram menu) has a prompt at the bottom with guidance on what to do next. On the Main Menu, you can press an arrow key, the ENTER key, one of two letters (P or Q), or '?' for help.

Changing window appearance if running directly on PC - The font size for the window running VitalPro may be changed from the Windows toolbar, to suit your screen. The size "8 x 12", as shown, is a reasonable size. The window running VitalPro may be toggled to full screen (and back) by pressing ALT-ENTER (hold down the ALT key and press the ENTER key).

Epigram Submenus

Access a submenu by highlighting an item and pressing ENTER.



Selecting parameter to change - To change one of the parameters, highlight the parameter (using the arrow keys) and press ENTER.

Quick changes - Some parameters have a very simple submenu. Try this:

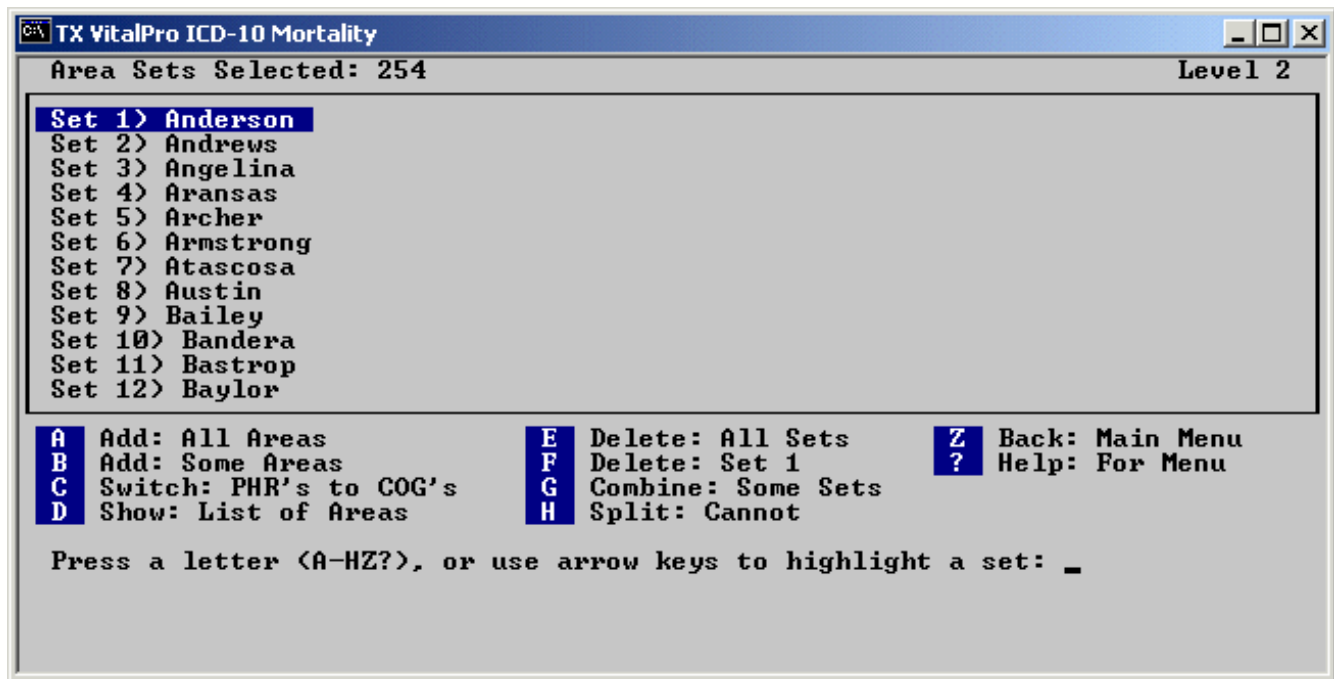
Highlight "Color Scheme" parameter by using the arrow keys. Press ENTER key to access simple submenu. Select different color scheme. Press ESCAPE key or 'Z' to return to main menu.

More complex changes - For more complex parameters, a more complicated submenu will appear and help you change the parameter. For example, to change the geographic selection, you would highlight the parameter "Geographic Areas", and press ENTER. Do the following:

Use arrows keys to highlight "Geographic Areas". Then, press ENTER to access the submenu.

Geographic Submenu

Submenus guide you step-by-step in selecting parameters.



Submenu with scrolling window - This is a typical Epigram submenu. It has a scrolling window with a list of items (currently list of selected Counties). One or more of the items may be highlighted by using the arrow keys.

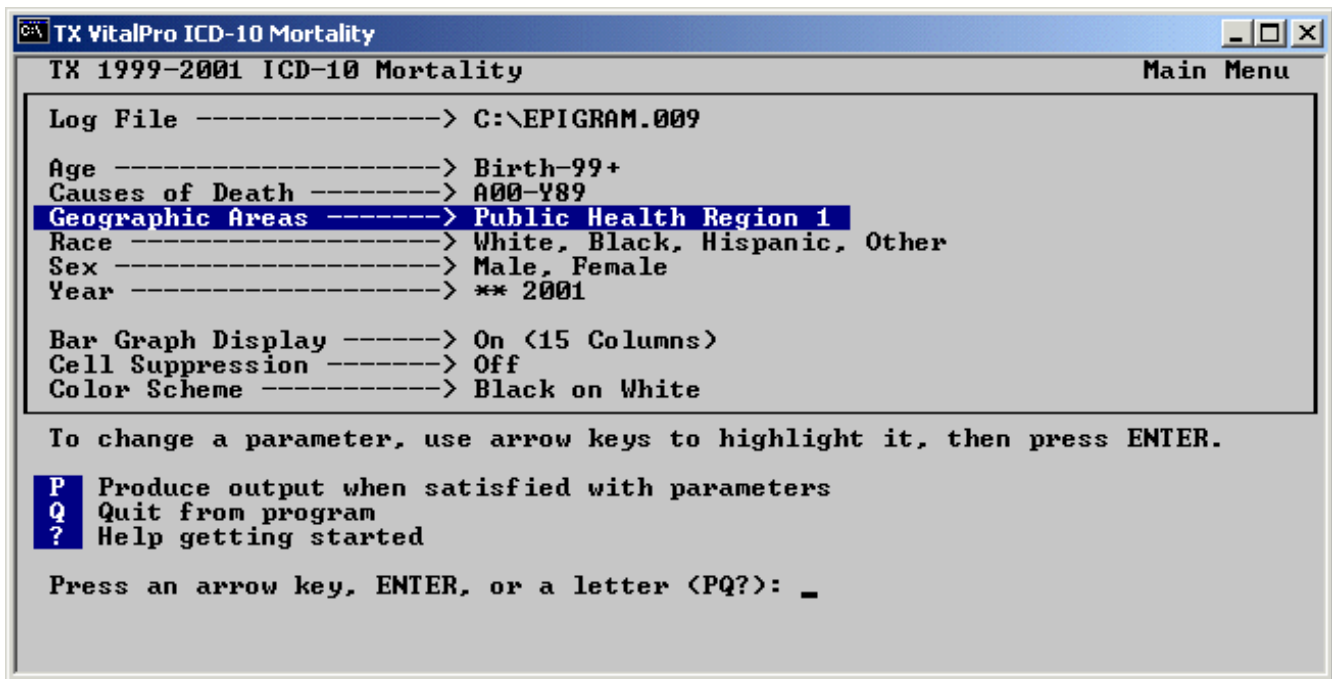
Letters are action items - This submenu has a list of actions, such as " E Delete all sets". Pressing the letter (such as 'E') carries out the action.

Adding an area - Carry out the following steps to add an area:

1. Press 'E' to delete all areas sets. You may be asked to confirm.
2. Then, press 'B' to add some areas.
3. Highlight "Public Health Region 1", and press ENTER to add.
4. Then, press ESCAPE key to return to previous menu.
5. Note that an item is added to the scrolling window.
6. Finally, press ESCAPE key to return to the Main Menu.

Select Another Submenu

Access another submenu by highlighting an item and pressing ENTER.

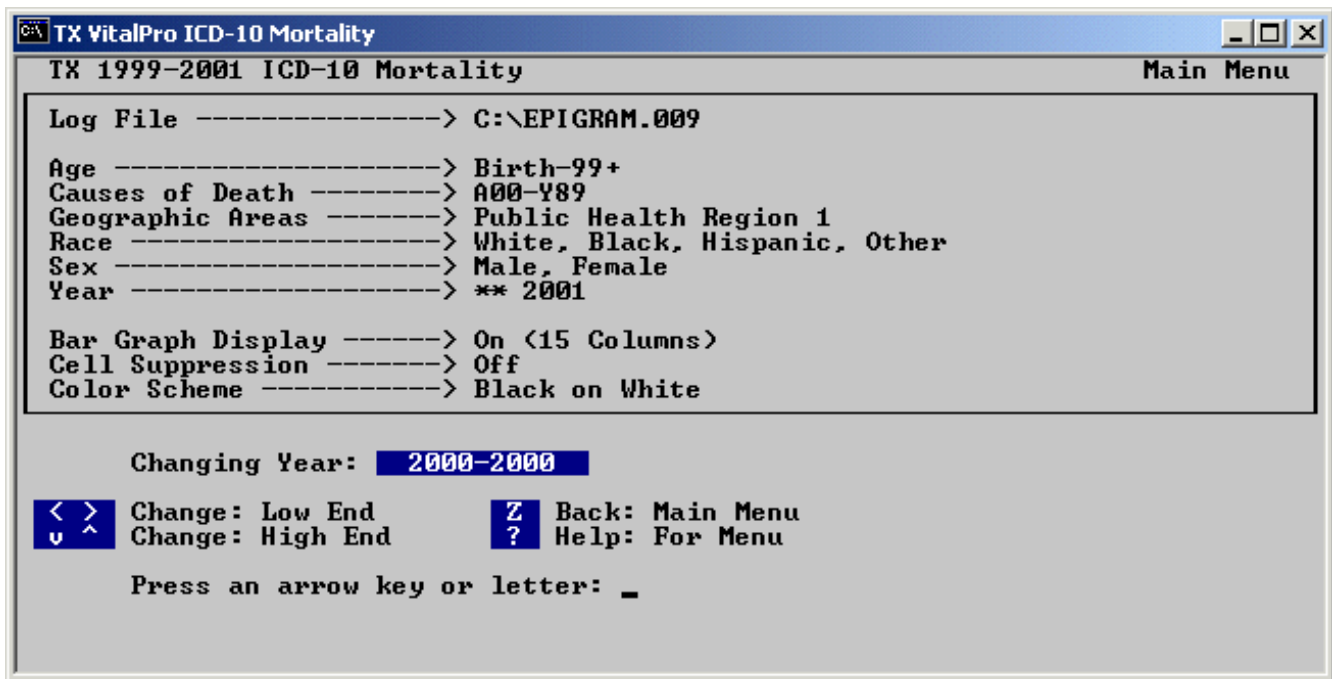


Make sure you're back - You should now be back at the Main Menu. Note that "Geographic Areas" has been modified (it is now set to Public Health Region 1). If you are not back at the Main Menu, get assistance from a coworker familiar with using Texas Epigram. Next, do the following steps, to access the submenu for modifying years:

1. Use the arrow keys to highlight "Year".
2. Then, press ENTER to move to a submenu for modifying the years.

Years Submenu

Submenus help you quickly and easily change parameters.



Submenu for a single range - Your screen should look similar to that shown above. This is the type of submenu used to modify a single range, such as a range of years. Next, do the following to modify the range and return to the Main Menu:

1. Press LEFT or RIGHT arrow key to change the low end.
2. Press UP or DOWN arrow key to change the high end.
3. Press HOME or END key to select the maximum range.
4. Keep modifying range until you have selected a single year, such as 1999.
5. When satisfied, press 'Z' (or ESCAPE key) to return to the Main Menu.

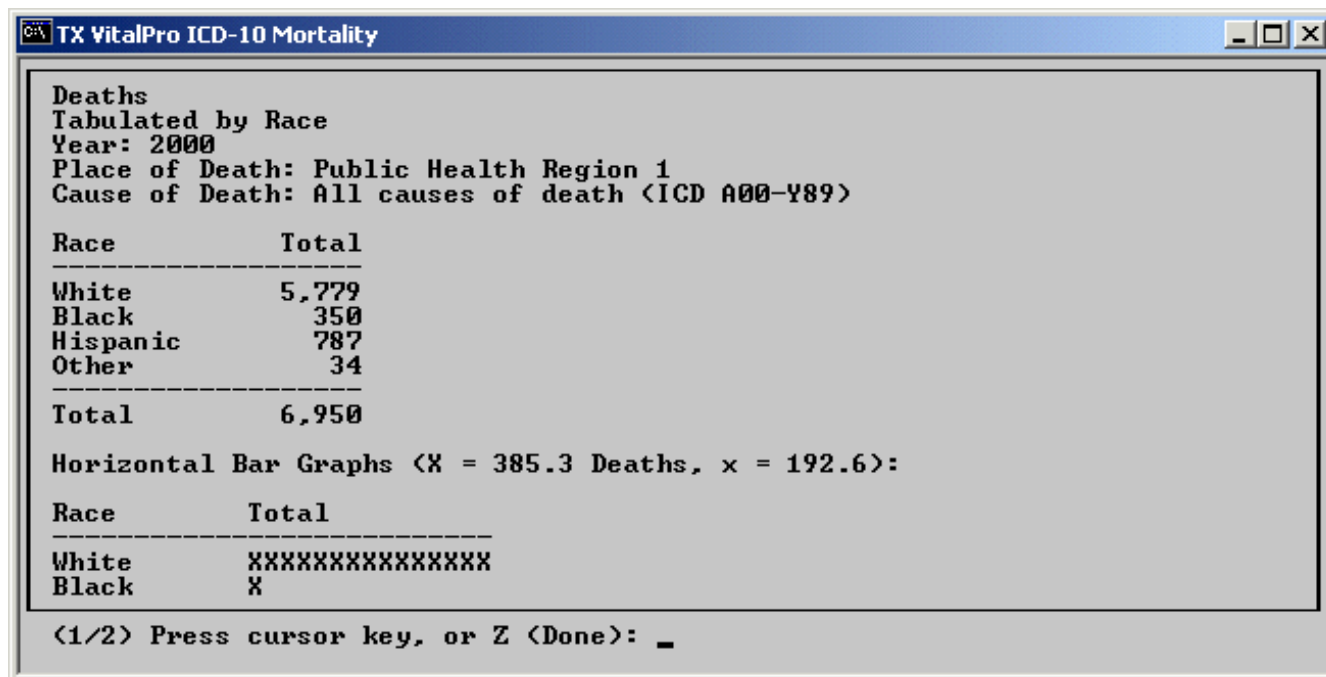
Produce a Table

Press 'P' at Main Menu when you are ready to make a table.

Check that you at the Main Menu - You should now be back at the Main Menu. You have used the submenus to select a few parameters, and are now ready to produce your first table!

Produce a table - From the Main Menu, you can press 'P' to produce output. Go ahead and do it:

Press the letter 'P' to produce a table similar to the following:



```
TX VitalPro ICD-10 Mortality
Deaths
Tabulated by Race
Year: 2000
Place of Death: Public Health Region 1
Cause of Death: All causes of death <ICD A00-Y89>

Race          Total
-----
White         5,779
Black         350
Hispanic      787
Other         34
-----
Total        6,950

Horizontal Bar Graphs <X = 385.3 Deaths, x = 192.6>:

Race          Total
-----
White         XXXXXXXXXXXXXXXXXXXX
Black         X

<1/2> Press cursor key, or Z <Done>: _
```

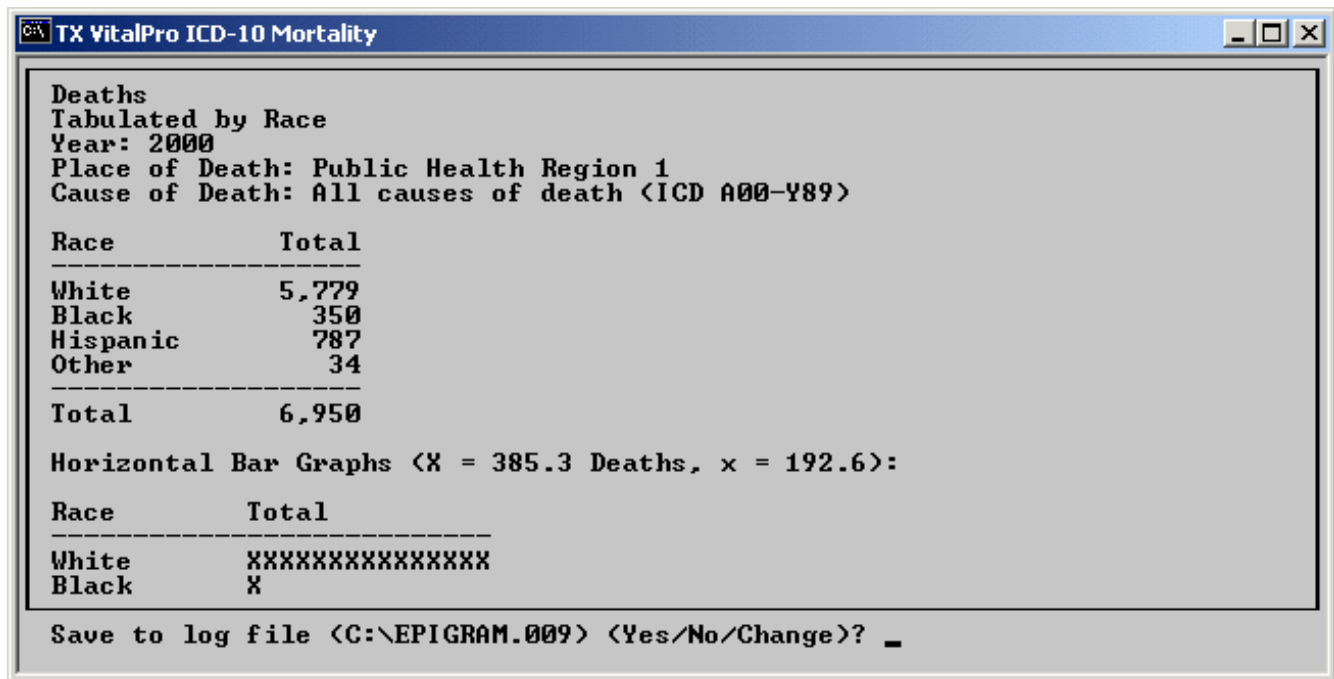
Time required - After you press 'P' from the Main Menu, Epigram calculates the results. The amount of time required depends on which parameters are selected and what kind of computer you are using, but is typically seconds. When Epigram finishes calculating the results, the output table will appear in a new window, as shown above.

Examining the table - You may browse the table with the arrow keys and other cursor keys (PgUp, PgDn, HOME, END). When you are finished examining the table, press the ESCAPE key (or 'Z'). Epigram will ask if you want to save the table to your log file (see next page). Depending on your access method and which menu options are selected, you may also be asked if you want to print the table or to save to an [alternate format](#) (CSV, dBASE III, HTML). Try this:

Press down arrow key and other cursor keys to examine the results. Press ESCAPE or 'Z' when you are finished viewing the results. For now, press 'N' when asked to save or print the table.

Save Results to Log File

Each table may be saved to your log file.



Purpose of the log file - Epigram always has a log file ready for saving the results exactly as you see them on the screen. After you finish viewing a table, Epigram always asks if you want to save it to your log file. If you press 'Y' the table will be appended to the log file (added to the end) as ASCII text. At any time, you may close the current log file and open a new one with a new name.

Log file location (local PC or LAN access) - If you are using Texas Epigram on a PC local area network or a stand-alone PC, the log file will be saved directly on a network disk or local disk. You can save to any directory where you have write permission. Epigram always tells you the location, such as "C:\EPIGRAM.001".

Word processing hints - After you finish using Epigram, you may edit and print the log files using any word processor. Use a non-proportional font such as Courier to keep the columns aligned. If needed, adjust the page orientation, margins or font size so the text does not wrap to the next line.

Directly printing the results - Local users (stand-alone PC, LAN PC, possibly UNIX) have the additional option of immediately printing an output table. After you view the output, Epigram may ask if you want to print the results. If you never want to print, you may disable printing from the Main Menu. NOTE: If you are connected to a network printer, your local network manager may need to enable printing from command windows.

Formats for Exporting

Results may be saved to a CSV, DBF, or HTML file.

Saving to a second format - After you finish viewing the results, Epigram may ask if you want to save to a comma-separated-value (csv), HTML (htm), or dBASE III (dbf) file. If you want to save to one of these file formats, set the "Second Format" option on the Main Menu.

CSV file (Comma-separated-value) (best for spreadsheets) - CSV output has a comma between each output item, and each text item is surrounded by "double quotes". Bar graphs are not included in CSV output since CSV output is usually imported into presentation software capable of more sophisticated graphics. CSV format is ideal for importing into spreadsheet software.

```
"Deaths"
"Tabulated by Race-Ethnicity and Sex"
"Years: 1999"
"Place of Residence: PHR 1"
"Causes of Death (ICD-10): A00-Y99: All Causes Of Death"

"Race", "Male", "Female", "Total"
"-----"
"White", 2797, 3063, 5860
"Black", 206, 174, 380
"Hispanic", 508, 335, 843
"Other", 15, 10, 25
"-----"
"Total", 3526, 3582, 7108

"Analysis Footnotes:"
"Unique ID, for Keeping Track of Analyses: 428GNXCA"
"Output Produced: Sat Apr 28 17:09:40 2001, by Texas VitalNet"
"Deaths Classified by: ICD-10, Underlying Cause, County of Residence"
"Mortality Data Source: TDH Bureau of Vital Statistics"
```

Example of CSV Output

DBF file (useful for many other applications) - Results may also be saved to a dBASE III database file (dbf extension). Database files are excellent for importing into almost any data analysis, graphics, spreadsheet, mapping or other presentation software. Field names are automatically imported along with the data. Suppressed cells are represented by the number "-1".

DBF file limitations - 1) Header and footer information listing analysis parameters is not included in database files. To help out, you may want to make the name of the dBASE file the same as the table ID, such as "405MFAQB.DBF" so you can refer to the table later. 2) No more than 128 output columns may be saved to a dBASE III file. This rarely presents a problem, because an unlimited number of rows is allowed.

Explore the Data

Epigram allows rapid exploratory data analysis.

The analysis cycle - Exploratory data analysis means you systematically refine your analyses. Once you have made a table, look it over. If it meets your needs, save it to your log file. Possibly save the results to a CSV or dBASE III file. Next, look over the parameters on the Main Menu, highlight the parameter you want to change, and press ENTER. A submenu will lead you through the modification process. After setting all parameters as desired, return to the Main Menu and produce another table by pressing 'P'. Try it:

At the Main Menu, modify a parameter or two. Then press 'P' to produce and view another output table. Try some of the [sample analyses](#) shown later.

Hints for avoiding misinterpretation -

- Double-check table header and footer to verify it was the analysis you intended.
- Double-check the ICD codes to make sure they are the diseases and injuries of interest.
- Find the right balance between:
 - Aggregation (fewer rows and columns, easier to present, more stable rates).
 - Stratification (more rows and columns, more information, shows differences).
- Acknowledge limitations in mortality data, including:
 - Misdiagnosis of the cause of death.
 - Changes in coding practices by certifiers.
- Acknowledge limitations in population data, including:
 - Possible census miscounts.
 - Difficulty projecting estimates to years between censuses.
- Use confidence intervals to help decide if differences are statistically significant.
- If uncertainties remain, contact local data experts for advice.

Continue with Analyses

Epigram lets you quickly carry out a wide variety of analyses.

It's easy - Once you get the hang of it, using Epigram is a snap. If you have gotten this far, you should be able to carry out about any analysis you desire. Explore the menus - that way, you'll know what is available. Don't be afraid to experiment and try out different options. Refer to the help files and to other sections of this user's guide.

On-line help - Although Epigram has been designed to be as self-explanatory as possible, it also includes extensive on-line help. To access on-line help from any menu, press '?'. A help screen will appear, with information related to the current menu. After viewing the help screen, a scrolling list of help topics may be viewed, for your selection. Try it:

At the Main Menu, press '?' and explore the help system.

Quitting Epigram - The 'Z' or ESCAPE key always returns to the previous menu. Later, when you are done, you may press the 'Q' key from the Main Menu to quit Epigram.

What other menus and options are available? - If you have done the examples in this chapter, you should have a good understanding of how to use Epigram menus to select parameters. However, you have just scratched the surface of the capabilities and power of Epigram. For your reference, the following pages list all of the Texas Epigram menus. The purpose of the list of menus is to help you become aware of what is available so that you can take fullest advantage of the software. Do this:

Referring to the outline on the next pages, explore the menus to learn what is available.

Outlines of Submenus

Using Texas Epigram, explore the menus listed below.

Main Menu allows you to:

1. First, design table layout:

- [Statistic Menu](#)
- [Row Variable Menu](#)
- [Row Sorting Menu](#)
- [Column Variable Menu](#)
- [Multi-Tables Menu](#)

2. Next, modify data variables:

- [Age Groups Menu](#)
- [Area Sets Menu](#)
- [Cause of Death Menu](#)
- [Race Groups Menu](#)
- [Sex Menu](#)
- [Year Ranges Menu](#)
- [Other Settings from Main Menu](#)

3. When ready, carry out an action:

- Produce Table (s)
- Exit Epigram

Submenus for Designing Table Layout

Using Texas Epigram, explore menus for designing table layout.

Statistic Menu sets type of result:

- Age-adjusted death rates
- Crude death rates
- Death counts
- Standardized mortality ratios
- Years of potential life lost (YPLL)
- YPLL rates

Row / Column / Multi-Table Menus.

Row / Column / Table for each selected:

- Only one row / column / table
- Age group
- Area set
- ICD set (cause of death)
- Race group
- Sex
- Year

Row Sorting Menu options:

- Rows unsorted
- Rows sorted by total data, high to low
- Rows sorted by total data, low to high

Submenus for Modifying Data Variables, Other Settings

Using Texas Epigram, explore menus for modifying variables.

Race Menu / Sex Menu:

- Select categories as sets
- Delete one / all sets
- Combine sets into new set
- Split previously combined set

Year Ranges Menu:

- Use a standard set, such as 2-year
- Extend upper / lower end of range
- Split range into individual years
- Delete a year range

Age Groups Menu:

- Use standard set, such as 10-year
- Extend upper / lower end of group
- Split group into separate groups
- Delete an age group

Cause of Death (ICD) Menu:

- Add causes from simple menu (includes most common causes)
- Add from detailed menus (includes complete ICD system)
- Add an ICD range (includes complete ICD system)
- Show list of ICD codes, for display and printing
- Delete one / all ICD sets
- Combine ICD sets into new set
- Split ICD set into component ICD groups

Area Sets Menu:

- Add all areas
- Add areas / area sets (submenu)
- Change region type
- Show and print list of areas
- Delete one / all area sets
- Combine area sets into new set
- Split area set into component areas

Other settings from Main Menu:

- Change log file for saving output
- Enable display of bar graphs
- Enable dBASE, CSV, HTML output
- Enable row / column percents
- Set color scheme
- Set confidence level, or turn off
- Set level of detail for trends
- Set number of decimal places

More Practice Analyses

For additional practice, and to gain more understanding of how Texas Epigram can speed and simplify data analysis, carry out the following sample analyses. Columns are by sex for each practice table, unless otherwise noted. Do the following:

1. For each practice analysis, select the parameters as shown below.
2. Then, press 'P' from the Main Menu to produce output.

1. Current leading causes of death for Texas:

ICD sets - NCHS rankable causes	Area sets - Texas	Statistic - Deaths
Row sort - Sorted high to low	Row for each - ICD set	Races - All
Table for each - One only	Years - Most recent year	

2. 1999 HIV (ICD B20-B24) age-specific death rates for one County:

Statistic - Death rate	ICD sets - B20-B24	Years - 1999
Row for each - Age group	Area sets - Travis	Races - All
Table for each - One only	Ages - 10-year groups	

3. 1999 lung cancer (ICD C33-C34) death rates ranked by area:

Statistic - Age-adjusted rate	Row for each - Area set	ICD sets - C33-C34
Adjustment standard - 2000 US	Table for each - One only	Races - All
Row sort - Sorted high to low	Area sets - All counties	Years - 1999

4. Texas diabetes (ICD E10-E14) time trends, one table for each race:

Statistic - Age-adjusted rate	Row for each - Year	ICD sets - E10-E14
Adjustment standard - 2000 US	Table for each - Race	Races - All
Area sets - Texas	Years - All years	

4: Epigram Parameters

Review - In the previous chapters, you have learned what Epigram is and how to use it. Chapter 1 explained that Epigram analyzes Texas mortality data, listed the linked data sets, discussed access methods, and outlined confidentiality requirements. Chapter 2 described the layout and interpretation of an Epigram output table. Chapter 3 showed how to use the interface to select parameters and produce output tables.

What's in this chapter - This chapter lists and explains the parameters and special options you may modify using the Texas Epigram interface. Epigram allows great flexibility in selecting and modifying parameters. The program allows you to mix and match parameters just about any way needed. Also, Epigram prevents you from selecting incompatible parameters.

Here are the parameters, options and concepts covered in this chapter:

- Age groups
- Causes of death (ICD sets)
- Cell suppression
- Confidence intervals
- Decimal places
- Geographic selection (area sets)
- Multi-tables
- Race
- Sex
- Statistic (outcome)
- Table columns
- Table rows
- Table row sort settings
- Trend analysis
- Years

Statistic [Outcome]

The statistic is the basic type of number in a table.

What is a statistic? - Every table has a statistic. The statistic (or outcome) is the basic type of data generated in an output table. Each statistic is defined in the glossary, and on-line help files give calculation methods.

Reliability Indicator		Statistic
Number of Deaths	Confidence Interval	
--	Y	Number of deaths
Y	Y	Death rate per 100,000
Y	Y	Mean age of death
Y	Y	Age-adjusted death rate per 100,000
Y	Y	Standardized mortality ratio (SMR)
Y	Y	Years of potential life lost (YPLL)
Y	Y	YPLL rate per 100,000

Reliability indicator - As shown in the table above, a reliability indicator (either number of deaths or confidence interval) is displayed along with most statistics. The **number of deaths** is normally displayed along with each rate and YPLL. Optionally, you may request that **confidence intervals** also be displayed as a reliability indicator.

Row Variable, Row Sorting, Column Variable

Table rows and columns may be set however you want.

Rows and columns - Rows are horizontal lines in a data table. Columns go up and down.

Every table has a row variable and a column variable. Rows and columns may be set to any of the following:

- Only one row / column (not set to a variable)
- One row / column per selected age group
- One row / column per selected area set
- One row / column per selected sex
- One row / column per selected ICD set
- One row / column per selected race set
- One row / column per selected year set

Sorted rows - Any table may be sorted as follows:

- Rows in standard order (not sorted by data). For example:
 - Area sets in alphabetical order, such as Bexar before El Paso.
 - ICD sets in alpha-numerical order, such as C34 before K74.
- Rows sorted from high to low, by total data
- Rows sorted from low to high, by total data

Multi-Tables

You may automatically produce a series of tables.

```
Deaths
Tabulated by Race-Ethnicity and Sex
Years: 1999
Place of Residence: Anderson
ICD A00-Y99: All Causes Of Death

Race          Male    Female  Total
-----
White         226    202    428
Black         98     47    145
Hispanic      21     5     26
Other         2      0     2
-----
Total         347    254    601

*** Tables omitted to save space ***

Deaths
Tabulated by Race-Ethnicity and Sex
Years: 1999
Place of Residence: Zavala
ICD A00-Y99: All Causes Of Death

Race          Male    Female  Total
-----
White          7      5     12
Black          0      0     0
Hispanic       54     33    87
Other          0      0     0
-----
Total         61     38    99
```

Purpose of multi-tables - Suppose you want to make a separate table for each county. It would be tedious to select the first county, make a table, select the second county, make a table, etc. Multi-tables automates the production of such a series of tables into one operation. The multi-table setting may be one of the following:

- Only one table (the default)
- One table per selected age group
- One table per selected area set
- One table per selected ICD set
- One table per selected race set
- One table per selected year set

Geo Selection [Area Sets]

Every table has a geographic specification.

Geographical Areas - Texas Epigram analyzes data to the county level. Texas has 254 counties, organized into regions (a group of several counties), as follows:

- 11 Public Health Regions (PHR's) (Effective 3/1/93)
- 24 Councils of Government (COG's)

You may select any combination of areas - Epigram makes it easy to compare results between different groupings ("sets") of geographic areas (Counties and County groupings), or limit analysis to specific geographic areas. One or more geographic areas may be combined into an area set, and compared with other sets.

Selection is quick and easy - All geographic operations are grouped together on one submenu, to allow quick and easy selection of any combination of areas desired. [A previous interface snapshot](#) shows the main geographic menu.

Causes of Death [ICD]

Mortality tables always include causes of death.

ICD codes - Epigram classifies causes of death with ICD codes. ICD stands for "International Classification of Diseases", the standard system for classifying causes of illness and death. Each disease or condition has a separate 3-digit ICD code (character followed by two digits) (for example, C53 for cervical cancer), or 4-digit ICD code (for example, G71.0 for muscular dystrophy). Data from 1980 to 1998 use the 9th revision of the ICD system (ICD-9). As of 1999, mortality data use ICD-10.

ICD groups and sets - An ICD group is one or more consecutive ICD codes (for example, C00-D48 for neoplasms). An ICD set is one or more ICD groups combined. For example, ICD C50 for breast cancer and ICD C53 for cervical cancer may be combined into an ICD set. At the broadest level, the ICD-10 system has 19 major disease categories. At the most detailed level, there are about 7,600 four-digit codes.

Selecting ICD groups - Epigram allows you to select and combine ICD groups to meet your analysis needs. You may select ICD sets by entering the range (for example, enter C16 for stomach cancer, or B20-B24 for HIV/AIDS). Or, you may select from ICD selection menus, organized to allow you to narrow your focus without having to know the specific code. Or, you may select one of several standard ICD lists. For example, for easily calculating leading causes of death, select the NCHS list of 50 rankable causes.

Texas Epigram automatically formats leading cause reports, to save you the trouble of converting ICD codes such as "ICD E10-E14" and set numbers such as "Set #18" to descriptive terms such as "Diabetes" and "Heart Disease".

Other Variables

Race Groups

You may select any combination of race groups to analyze. Races may be combined into "sets" as needed. A separate submenu allows easy selection. Keep in mind that differences in health status between races may be due to socio-economic differences.

Each death is classified as White, Black, Hispanic, or Other, as follows: 1) If race is reported as Black, the category is "Black". 2) If race is reported as Chinese, Japanese, Hawaiian, Filipino, Asian Indian, Korean, Samoan, Vietnamese, or Guamanian, the category is "Other". 3) Of the remainder, those said to be of Hispanic origin are counted as "Hispanic". 4) Of the remainder, if race is reported as Other, North American Indian, Central or South American Indian, or not classifiable, the category is "Other". 5) All remaining are classified as "White".

Age Groups

Epigram lets you combine and analyze age groups in just about any way needed. You may select any contiguous combination of age groups to analyze, such as 22-34. Or, you may select a set of ranges, such as birth-19, 20-39, 40-64, 65-99+.

The program will let you know which age groups are available for use. Standard age groupings, such as 5-year, 10-year, and 20-year age groups are easily selected. Age groups may be combined in just about any way desired.

Up to age 21, 1-year ages are used (for example, 3 or 11-17). For 22 and over, an age group may use any combination of 5-year groups (for example, 40-44 or 30-49). The highest group is currently 75-99+ (75 and over).

Sex

You may select Male, Female, or both sexes combined.

Years

You may select any continuous range of one or more years to analyze, such as 1980-1983. In addition, you may select a set of year ranges for comparing, such as 1990-1991, 1992-1993, 1994-1995. The program will let you know which years are available for use. Standard groupings, such as 1-year, 2-year, and 3-year groups are easily selected. Year groups may be combined in just about any way desired.

Multi-year death rates are calculated by adding all of the deaths over the time period, and dividing by the sum of the populations over that same period.

Other Settings

Confidence Intervals

Statistical reliability - Results that are based on a smaller number of events are less reliable than those based on a larger number. Confidence intervals allow you to estimate the statistical reliability of your results.

Confidence interval definition - A confidence interval (also known as confidence limits) is the range of values within which the true value of a variable is thought to occur, with a specified confidence level (95%, 90%, 80%, etc.). A higher confidence level (for example, 99%) is more stringent and results in a smaller interval than a lower confidence level (for example, 80%). Use the 95% confidence level unless you have a reason to do otherwise.

Methods for confidence intervals - Epigram uses the Poisson distribution to calculate most confidence intervals, using the method described in "Scientific Tables", Diem and Lentner (ed), Geigy, 1970 (equations 802a and 802b on page 189). Confidence intervals for standardized mortality ratios are calculated using the method of Rothman and Boice described in "Research Methods in Occupational Epidemiology", Checkoway, Pearce, and Crawford-Brown, Oxford University Press, 1989, (equations 5.7 and 5.8 on page 127). Confidence intervals for trend analysis are calculating using Student's t-distribution.

Decimal Places

Epigram also allows specification of the number of decimal places in your results. For example, the number 64.29 has two decimal places. The number 64 has zero decimal places.

Cell Suppression

You may suppress cells with low numbers of deaths.

How cell suppression works - Cells with low numbers of deaths may be blanked out with an asterisk (*). You set the level at which results will be suppressed. Bar graphs are not included if any cells are suppressed, because the graphs would not be accurate.

Suppression of row / column totals - A row total will be **suppressed** if there is exactly one suppressed cell in the row, or if the row total is low. A row total will be **displayed** if there is more than one suppressed cell in the row, unless the row total is low. The same rules apply to columns.

Use of cell suppression - Cell suppression may be used for certain data tables to be released to the public where there is a concern that low numbers should not be published due to potential confidentiality issues. Concerns about reliability of small numbers are usually better addressed by using confidence intervals, or by including the number of deaths as a reliability indicator.

Cell suppression in dBASE III files - A suppressed result is indicated as "-1".

```
Deaths
Tabulated by Age and Sex
Years: 1999
Place of Residence: Travis
ICD B20-B24: Human Immunodeficiency Virus [HIV] Disease

Age          Male   Female  Total
-----
Birth-14     *      *       *
15-24        *      *       *
25-34         5      *       *
35-44        28     4       32
45-54        16     *       *
55-99+       3      *       *
-----
Total        52     8       60

Analysis Footnotes:
Unique ID, for Keeping Track of Analyses: 428ZRYWE
Output Produced: Sat Apr 28 17:13:57 2001, by Texas VitalNet
Deaths Classified by: ICD-10, Underlying Cause, County of Residence
Suppress Cell [*]: If 2 or Fewer Deaths In Cell
Suppress Row/Column Total [*]: If Exactly One Suppressed Cell in Row/Column
Mortality Data Source: TDH Bureau of Vital Statistics
```

Table with Suppressed Cells

Trend Analysis

Epigram automatically carries out trend analyses.

Purpose of trend analysis - When analyzing data with a year for each row, you usually want to know: Is there a trend up or down? Is the rate increasing or decreasing? Statistical analysis is helpful in quantifying the answer.

Least-squares line shows the trend - Epigram automatically carries out a "least-squares" analysis for a time series. This finds the best straight line to describe the data for each data column. The line is defined by the value for the first year ("Y-intercept"), and the amount of change per year ("slope").

Confidence intervals show if significant trend - Epigram calculates a confidence interval for the slope. If the confidence interval doesn't include 0, there is a significant trend.

Simplified presentation - Epigram allows the user to substitute a simplified presentation (not shown) that only says whether the trend is significantly up or down for each column.

```

Death Rate (per 100,000)
Age: Birth-11m Years
Place of Residence: Texas
ICD 001-999: All Causes Of Death

      Year      Male      Female      Total
      Rate      Rate      Rate
-----
1990      XXXXXXXXXXXXXXXX XXXXXXXXXXXXXx  XXXXXXXXXXXXXXXx
1991      XXXXXXXXXXXXXXXx  XXXXXXXXXXXXXx  XXXXXXXXXXXXXXX
1992      XXXXXXXXXXXXXXXX  XXXXXXXXXXXXXXX  XXXXXXXXXXXXXXX
1993      XXXXXXXXXXXXXXXx  XXXXXXXXXXXXXx  XXXXXXXXXXXXXXXx
1994      XXXXXXXXXXXXXXXx  XXXXXXXXXXXXXx  XXXXXXXXXXXXXXX
1995      XXXXXXXXXXXXXXX  XXXXXXXXXXXXX  XXXXXXXXXXXXXXX
-----
Total      XXXXXXXXXXXXXXXX  XXXXXXXXXXXXXx  XXXXXXXXXXXXXXXx

Detailed Least-Squares Analysis, to Detect Linear Trend:

Column      1990 Y-intercept      Slope      95% CI (4 df, t=2.7764)
-----
Male          941.1          -41.76      -59.37 to -24.16 **
Female        742.9          -29.64      -44.32 to -14.96 **
Total         844.1          -35.81      -50.58 to -21.05 **
-----

Y-intercept and slope may be used to draw least-squares line.
If confidence interval (CI) does not include 0, trend is significant [**].
    
```

Significant Downward Trends for Death Rate Under One

Glossary

Age-adjusted death rate - Deaths per 100,000 population, adjusted to a standard population (such as US 1940 or US 2000), by the direct method. Age-adjusted rates are often better for making comparisons than unadjusted rates, because they adjust for differences in age distribution between populations. An age-adjusted rate is a summary measure. Besides calculating overall age-adjusted rates, it is also recommended to compare age-specific rates.

Age-adjustment standard - A standard population for calculating an age-adjusted death rate. The 1940 and 2000 US Census population are the most common standards.

Area set - One or more areas combined.

ASCII file - A text file, with only alphabetical, numerical, and punctuation characters, like you would see in normal text. Vitalnet can produce output in ASCII format.

Bar graphs - A section of a Vitalnet table. Gives a useful graphical representation of the data. May be omitted from the output table.

Cause of death - Any condition which leads to or contributes to death and is classifiable according to the International Classification of Diseases (ICD) system.

CDC - Centers for Disease Control and Prevention. US federal health agency.

Cell - A space for a single numerical result in a table, at a row-column intersection.

Cell suppression - An asterisk "*" is placed in cells with fewer deaths than a limit set by the user. Row / column totals with exactly one suppressed cell in the row / column are also suppressed. If more than one cell in the row / column is suppressed, the row / column total is displayed.

Columns - Vertical groupings of data in a Vitalnet table, such as a column for each race group.

Confidence interval (confidence limits) - A range of values within which the true value of a variable is thought to lie, with a specified level of confidence. For a result of 23.5, a confidence interval might be (23.1-23.9). The smaller the interval, the more reliable the result. If the 95% confidence intervals do not overlap, there is a statistically significant difference. Vitalnet uses several methods to calculate confidence intervals. The output table documents which method was used.

Confidence level - The likelihood that the true value of a variable is within a confidence interval. For example, for confidence intervals at the 95% level, we are statistically 95% certain that the actual value of the variable is within the interval.

Councils of Government (COG's) - Groupings of Texas counties. There are 24 COG's.

CSV format - Also called comma-separated-value format. A type of computer output that is readily imported into other software, especially spreadsheet software. Each output item is separated by a comma from surrounding items, and each output text item is surrounded by "double quotes". Vitalnet produces CSV output. A comma-separated-value file has "csv" extension.

Data mining - Finding unexpected relationships in a data set. Similar to exploratory data analysis. Vitalnet is excellent at data mining. Of course, keep in mind that the more you look, the more unusual events you will find, just by chance.

Data warehouse - Software system, such as Vitalnet, making large complex databases readily available for querying and analysis.

dBASE III format - A widely used file format derived from the database software of the same name. Files in this format may be readily imported into almost any data analysis, graphing, mapping, or other presentation software. Has dbf extension. Suppressed cells are represented as the number "-1". Vitalnet produces dBASE III output.

Death rate - Deaths per 100,000 population. May be used to compare the burden of disease between different groups. Also called crude death rate.

Denominator - The number on the bottom of a fraction. Population data are often referred to as "denominator data", as they are used as denominators to calculate population-based rates.

Export - Produce output that can be read into other computer programs. Vitalnet produces ASCII text, comma-separated-value (csv extension), HTML (htm extension), and dBASE III (dbf extension) files for export.

Filter variable - A variable solely used to filter which records are included in the output. For example, for a single table with race rows and sex columns, age is a filter variable.

Footer - Last part of a Vitalnet table. Lists other details of the analysis, such as the date and time produced, and the data sources. Each table is assigned a unique ID, listed in the footer, to assist you in keeping track of and organizing analyses.

Header - First part of a Vitalnet table. Lists basic parameters you selected to define the table, such as years analyzed.

ICD - International Classification of Diseases. A widely used system of classifying diseases and injuries. Each disease or set of diseases has an ICD code or ICD group assigned to it. Vitalnet uses the ICD-9 system, which has been in use since 1979, and the ICD-10 system, which started with 1999 mortality data.

ICD-10 code - A single ICD number representing a single disease or injury (for example, ICD E10 for insulin-dependent diabetes).

ICD-10 group - A range of continuous ICD-10 codes (for example, ICD E10-E14 for diabetes mellitus).

ICD-10 set - One or more ICD-10 groups or codes combined, for example ICD C50 (breast cancer) and ICD C53 (cervical cancer).

Import - Read information into a computer program. ASCII text, CSV files, HTML, and dBASE III files from Vitalnet are easily imported into word processing, spreadsheet, data analysis, mapping, graphing, and other presentation software programs.

Leading causes for ICD-10 - The ten causes of death with the highest number of deaths, out of a standard National Center for Health Statistics list of 50 rankable causes. Vitalnet makes it easy to select and rank the 50 cause list.

Least-squares - A standard method for fitting the best straight line to a set of points. Produces a Y-intercept and a slope defining the least-squares line.

Mean age of death - If the ages of death were 50, 51, and 58, the mean age of death is $(50 + 51 + 58) / 3 = 53$

Multiple age groups - One age group for each row or column of a table. Example: Birth-19, 20-39, 40-59, 60-99+.

Multiple cause mortality data - Data which include all causes of death listed on the death certificate. Contrast with underlying cause mortality data.

Multiple causes of death - All diseases or injuries which led directly to death, or all circumstances of the accident or violence which produced the fatal injury.

NCHS - National Center for Health Statistics. US health statistics agency. Part of the CDC.

PHR - Public Health Region. Groupings of Texas counties. There are 11 PHRs.

Place of occurrence mortality data - Data compiled by the location where the death occurred, without regard to the place of residence of the deceased. Vitalnet does not currently analyze for place of occurrence.

Place of residence mortality data - Data compiled by the usual place of residence of the deceased, without regard to the location where the death occurred. Vitalnet analyzes mortality data by place of residence.

Population - The number of people living in an area.

Primary Statistic - The basic type of numerical result displayed in a table, chart, or map. For example, death rate.

Public Health Regions (PHR) - Groupings of Texas counties. There are 11 Public Health Regions.

Rows - Horizontal lines in a Vitalnet table, such as a row for each race group.

Row sort settings - Vitalnet rows may be sorted in ascending or descending order.

Set - A combination of one or more things. For example, several areas may be combined into an area set.

Single age group - Only one age group (30-49, for example) is selected. A single age group is used for tables that do not have age columns or age rows.

Standardized mortality ratio (SMR) - The ratio of the expected number of deaths in a population to the observed number of deaths. The expected number of deaths is derived from applying a standard set of rates (usually state or national rates) to the population. SMRs are useful for assessing whether the mortality in a population is higher than expected.

Statistic - See "Primary Statistic".

Table - A set of results produced by Vitalnet. A table has several parts:

1. Header - lists basic analysis settings
2. Data section - numerical results
3. Horizontal bar graphs - graphical representation of the data
4. Footer - lists other analysis settings

Underlying cause - The disease or injury that initiated the train of events leading directly to death, or the circumstances of the accident or violence which produced the fatal injury. A single underlying cause is assigned to each death.

Underlying cause mortality data - Data which include only the underlying cause of death listed on the death certificate. Contrast with multiple cause mortality data.

UNIX - A widely used computer operating system. Vitalnet can run on UNIX, either remotely, locally, or by client-server.

Unknown Values - Unknowns are automatically inserted into a Vitalnet table. For example, a separate row (or column) for unknown race. The rate is assigned as zero for an unknown category, since there is no population denominator to use. Some fields, such as sex for certain data sets, are never unknown, so unknowns are left off the table. When, such as for age-adjusted rates, the unknown variable (age) is different from the rows or columns, the number of unknowns for age is shown below the table.

Windows - A widely used set of PC operating systems, including Windows 3.1, 95, 98, NT, 2000, and XP. Vitalnet will run under any version of Windows.

World Wide Web (WWW) - A widely used part of the internet that may be easily accessed with a web browser. Vitalnet runs on the WWW.

Years of potential life lost (YPLL) - Sum of the years of life lost by persons who suffered early deaths. Early death is usually defined as death occurring before the age of 65 (the YPLL age limit). For example, death at age 40 (40.5) results in 24.5 YPLL to age 65. YPLL is a widely used measure of premature mortality.

YPLL age limit - The age used for calculating YPLL. The most common age limit is 65, but other age limits may be selected from within Vitalnet.

YPLL rate - YPLL per 100,000 population in the appropriate age category. For example, the YPLL rate up to age 65 is calculated as follows: (YPLL up to age 65) / (population for age group 0-64). YPLL rate is not commonly used.