

Introduction

An epidemiologist studies the distribution and determinants of diseases and injuries in human populations. Descriptive epidemiology is the search for the pattern and frequency of health events in a population. For injury prevention, descriptive epidemiology is used to monitor population injury status, to define injury priorities, and to target intervention efforts.

"Epidemiologist in a Box", referred to as "Epi in a Box" throughout this instruction guide, is a tool designed to assist California injury prevention groups in calculating and interpreting rates of injury for specific populations. This tool can be adapted for use by injury prevention groups in other states. It is our hope that this will be a useful tool for those programs that do not have access to an epidemiologist or biostatistician when conducting a needs assessment or program evaluation.

Detailed instructions are included for collecting injury data using the California Department of Health Services (CDHS) Epidemiology and Prevention for Injury Control (EPIC) website, and for collecting population data using the California Department of Finance (DOF) Demographics website. Additionally, basic data terminology, descriptions of other data sources, and interpretations of rates are included in this toolkit.

TO START USING THE TOOL NOW, CLICK "Using Epi in A Box".

Getting Started With Injury Data

Why do we need data?

Injury is the third leading cause of death among all age groups in the United States, and the leading cause of death of children, with more children and adolescents dying of injury than all other causes combined. Not only do children and adults die from injuries, but serious nonfatal injuries can occur, leading to lifelong disability. Most injuries, however, can be prevented through the use of existing technology and environmental changes.

If we are to prevent injuries, we must understand the extent of the problem. Injury data can provide your organization with the information you need to make informed decisions regarding which injuries, populations, and geographic locations should be targeted for intervention.

One way to assess the extent of injury is to calculate baseline rates of both fatal and nonfatal injuries within the target population. The baseline rates of injury can be used to assess need, establish priorities, and give credibility to your intervention efforts. They can also be compared to state and national indicators to identify problems in your target population that need further investigation.

In order to establish and make progress toward your goal, it is important to know at the local level how injury is distributed according to factors such as age, gender, and location so that interventions can be planned effectively. If multiple years of data are available, baseline injury rates should be calculated and compared over time to assess change. Additionally, demographic data can be used to create profiles of the targeted populations to prepare you for your prevention efforts. This data can be shared with local leaders and used to educate the community about risk factors for injury.

Surveillance

Surveillance is one of the first and most basic elements of injury prevention and control. Injury surveillance determines the magnitude of injury morbidity and mortality, the leading causes of injury, and the population groups and behaviors associated with the greatest risk. Not only are surveillance data fundamental to determining priority areas for program planning and prevention, these data are critical to the evaluation of the effectiveness of program activities and in determining problem areas that need additional investigation.

Even if the collection of new data is necessary, the use of secondary data for surveillance can enhance newly collected data. Secondary data refers to existing data collected by city, county, state, and national organizations. A wealth of information on injuries in diverse populations is routinely maintained by various organizations and is often available for public use.

Gathering information

The types of information you need to obtain to determine the extent of injury in your community include: age, gender, intent of injury, and external cause of injury. The intent of an injury is defined as either unintentional, sometimes referred to as accidental, or intentional such as homicide/assault or suicide/self-inflicted. The external cause of injury or E-code is the event that caused the injury, such as fall from height or pedestrian struck by motor vehicle. You may also want to collect information on race/ethnicity, geographic location, injury severity, and key contributing factors.

Frequency Versus Rate

Injury data is generally presented as either a frequency or a rate. A frequency is a count of how many times an event occurred for a particular population. For example, a frequency count would be the number of children between the ages of 1-4 years who died as the result of a motor vehicle accident. The advantage of a frequency is that it answers the question of “how many” events occurred in a population.

Frequency counts can also put the burden of the injury into perspective. For example, a 50 percent increase in the number of drowning deaths for a particular city isn't a big concern if the number of deaths increased from 2 to 3. A 50 percent increase in the number of motor vehicle occupant deaths for this city is more concerning if the number of deaths increased from 90 to 135.

The major problem with frequencies is that you cannot directly compare the number of injuries across populations or subgroups of populations, such as different counties or different age groups, if these groups have different population sizes.

Much of injury assessment involves describing the injuries sustained in a defined community by looking at changes in that population over time or by comparing injuries in that community to injuries in other communities. When making these comparisons, we need to account for the fact that the number of injuries depends in part on the size of the population. To account for growth in a community or to compare communities of

different sizes, we usually develop rates to provide the number of events per population unit.

The frequency with which injuries occur is almost always related to age. For instance, the frequency of falls among 0 to 4 year olds is higher than among 5 to 9 year olds in a given population, but does this mean that they are more common? In order to accurately compare the frequency of injuries between subgroups of a population, you must convert the frequency counts into rates.

Rates are frequencies that have been converted to numbers that share a common denominator¹, usually frequency of occurrence per 100,000 population, thus making all populations equal in size. Rates facilitate the comparison of injury frequency across different groups of people, places, and time periods.

For example, assume your population of children 0 to 4 years old is 180,000 with a frequency of 250 hospitalizations due to falls, and your population of 5 to 9 year olds is 165,000 with a frequency of 240 hospitalizations due to falls. If you compared frequencies directly, you would assume that 0 to 4 year olds were at a slightly increased risk of hospitalization due to falls. However, if you calculate the number of events per a common denominator, such as 100,000, you find that 5 to 9 year olds are actually at a greater risk.

0 to 4 year olds: 250 per 180,000 = events per 100,000 rate = 138.9 per 100,000

5 to 9 year olds: 240 per 165,000 = events per 100,000 rate = 145.5 per 100,000

In order to convert a frequency into a rate, you need to know the size of the population. You should use the estimated population for the year(s) of the data. A rate is calculated by dividing the total number of events in a specified time period by the total number of individuals in the population at risk for these events, multiplied by a constant, usually 100,000.

$$\frac{\text{Number of events in the population per time period}}{\text{Total number of people in the population per time period}} \times 100,000 = \text{Rate}$$

For example, the number of hospitalizations due to falls among 0 to 4 year olds in Your County in 2000 (numerator²) divided by the population of 0 to 4 year olds in Your County in 2000 (denominator) times 100,000 (constant) gives the year 2000 rate of fall related hospitalizations per 100,000 0 to 4 year olds in Your County.

$$\frac{\text{Number of hospitalizations due to falls, 0-4 yrs, 2000}}{\text{Total population, 0-4 yrs, 2000}} \times 100,000 =$$

$$\frac{250}{180,000} \times 100,000 = 138.9 \text{ per } 100,000$$

¹ Denominator: The number that appears on the bottom of a fraction.

² Numerator: The number that appears on the top of a fraction.

With a rate, you can directly compare the rates of injury among population groups to determine which populations are at greater risk of injury.

Small Numbers

Rates based on small numbers of events can fluctuate widely from year to year for reasons other than a true change in the underlying frequency of occurrence of the event. It should be noted that the County of San Diego Emergency Medical Services (EMS) does not calculate rates based on fewer than 5 events. EMS considers rates based on fewer than 5 events to be unstable and to be regarded with caution. A very small numerator means that each event takes on a greater statistical significance, and the rate will have a greater relative standard error. There is no magic “small number”. As with any statistical procedure, when you calculate rates you have to determine the amount of possible error you are comfortable with.

To deal with the problem of a small numerator it is recommended that you combine, or aggregate, three successive years of data to produce a larger number. Keep in mind that when you combine multiple years of injury data to use in rate calculations, you must also combine multiple years of population data.

Consider the following example: Suppose you want to calculate the rate of fatal drownings among 0 to 4 year olds in Your County. You notice that there are only a few deaths each year due to drowning, so you decide to combine multiple years of data. You find that in Your County, there were 2 deaths in 1999, 3 in 2000, and 3 in 2001, for a total of 8 deaths due to drowning among 0 to 4 year olds from 1999 to 2001. The 0 to 4 year old population was 40,000 in 1999, 41,500 in 2000, and 41,900 in 2001, for a total population of 123,400 from 1999 to 2001. The rate would be calculated by dividing 8 by 123,400 and multiplying by 100,000.

$$\frac{(\text{events in year 1} + \text{events in year 2} + \text{events in year 3})}{(\text{population in year 1} + \text{population in year 2} + \text{population in year 3})} \times 100,000 =$$

$$\frac{\text{Number of drowning deaths, 1999-2001, 0 to 4 year olds}}{\text{Total population, 1999-2001, 0 to 4 year olds}} \times 100,000 = 6.5 \text{ per } 100,000$$

The rate of drowning deaths among 0 to 4 year olds in Your County from 1999 to 2001 was 6.5 per 100,000.

Calculating Changes in Rates

Once you have calculated a rate, and identified it as an indicator in need of improvement, you may want to determine what your new rate would be if you were to achieve a 10% improvement, or decrease in rate.

Suppose you determined that the rate of hospitalization due to choking among 0 to 4 year olds in 2003 in Your County was 20.4 per 100,000. Based on the results of choking prevention programs conducted by other organizations, you know that a reasonable goal would be a 10% reduction in the rate of hospitalizations due to choking. To calculate your goal rate, you would multiply $(1 - 0.10) \times 20.4 = (0.90) \times 20.4 = 18.4$ per 100,000.

You can also calculate the change in injury rate from one year to the next. For example, based on your surveillance, you find that among 0 to 4 year olds, the rate of hospitalization due to choking was 21.5 per 100,000 in 2003, and 20.4 per 100,000 in 2004. You can see that there was a decrease in rate, but by how much? To calculate the percent change in rate, use the following formula:

$$\frac{\text{Rate of injury (2004)} - \text{Rate of injury (2003)}}{\text{Rate of injury (2003)}} \times 100 = \frac{20.4 - 21.5}{20.4} \times 100 = -5.4\%$$

From 2003 to 2004, there was a 5.4% decrease in the rate of hospitalization due to choking among 0 to 4 year olds in Your County.

Guidelines for Developing and Comparing Rates

Number of events (numerator)

- When developing a rate, you should count only the events that occur among the population at risk.
- When using data sources where one person may be counted more than once, make note of whether you are counting events or people.
- Tables containing rates should provide precise definitions of the number of events so that people using the information can calculate comparable rates. For example, specify the ICD-9 E-codes or ICD-10 WXYZ codes used to define the injury category.
- If you have fewer than 5 events, combine 3 successive years of data.
- When reporting rates, the data source for the number of events should be specified.

Population at risk (denominator)

- Everyone in the population at risk must be eligible to be counted in the numerator if they have the event of interest. For example, when looking at falls from playground equipment among 5 to 9 year old females, we would only include 5 to 9 year old females in the population at risk, because 5 to 9 year old males would not be included in the number of events.
- If you have to combine numerator data due to small numbers, you must combine the same years of population data for the denominator.
- When reporting rates, the data source for the population at risk should be specified.

Comparing rates

Only compare rates when the numerator and denominator (i.e., events and population) are defined consistently over time and place. Look for:

- Consistency in the definition of the event.
- Consistency in the coding scheme (i.e., injuries coded using ICD-9 E-codes may not be comparable to injuries coded using ICD-10 WXYZ codes).
- Consistency in the population data source.

Data Considerations

Before using any data set to develop rates, several steps must be taken to assure the validity and appropriateness of the measure. Important information to consider includes:

- Data availability.
 - o In addition to on-line resources, what data sources are available in your community?
 - o Is the data available at the individual level or only at the aggregate (group) level?
 - o How current is this data?
 - o Is the data reported and available in a timely manner?
 - o Is there a charge for this data? How much?
 - o How long will it take to obtain this data?
 - o Is the data available in an electronic form? In what format?
- Representativeness.
 - o How is the data collected?
 - o How are each of the variables defined?
 - o Is the data representative of your target population?
 - o Is the data set population based (representing all of the events in the county), or do they represent only those associated with a particular facility or agency? Note: population rates can only be calculated if the data are population-based.
- Consistency.
 - o Is the data source available in the same format over time?
 - o Have the collection methodology, variable definitions, or survey questions, changed over time?
- Completeness and Accuracy.
 - o Is there a high proportion of missing or inaccurate records or variables?
 - o Are the variables forced fields? In other words, is the data recorder required to complete the information?
- Trustworthiness.
 - o Do you trust the collection methodology?
 - o Do you trust the data resource, i.e. the collection agency?

Data Sources

Injury Data

“Epi in a Box” was designed for use with California injury data provided by EPIC. However, this tool will also work with data from other sources and locations. The following list of data sources, compiled by the California Center for Childhood Injury Prevention, can be used to find injury data.

Online Resources

- *EPICenter – California Department of Health Services, EPIC Branch*
<http://www.applications.dhs.ca.gov/epicdata/>
 - o Provides access to state and county-specific injury data.
 - o Morbidity and mortality data is available via custom-build tables.

- o Under the section titled “About Our Data”, click on “EPICenter help area” for a description of the source of EPIC’s fatal and nonfatal data.
- *California Department of Health Services, Vital Statistics Query System*
<http://www.applications.dhs.ca.gov/vsq/default.asp>
 - o Provides access to state and county-specific vital records data.
 - o Mortality data is available by cause of death.
- *WISQARS - Centers for Disease Control and Prevention*
<http://www.cdc.gov/ncipc/wisqars/>
 - o Provides access to state and national injury data.
 - o Mortality data is available via custom reports.
- *WONDER – Centers for Disease Control and Prevention*
<http://wonder.cdc.gov/>
 - o Provides access to county, state, and national data.
 - o Mortality data is available via custom reports for a wide range of public health data.
- *Fatality Analysis Reporting System (FARS) – National Highway Safety Traffic Administration*
<http://www-fars.nhtsa.dot.gov/>
 - o National mortality data is available via on-line queries.
 - o State-level data is available by contacting FARS.
- *Federal Bureau of Investigations (FBI)*
<http://www.fbi.gov/publications.htm>
 - o National, state, and some county-level crime data available via on-line reports.
 - o This is a voluntary reporting system, so it may underestimate the actual incidence of homicides.
- *California Health Interview Survey (CHIS)*
<http://www.chis.ucla.edu/main/default.asp>
 - o Provides California public health information.
 - o State and county-level data are available.
- *Behavioral Risk Factor Surveillance System (BRFSS)*
<http://www.cdc.gov/brfss/>
 - o Provides California public health risk information.
 - o State level data are available.

Local Resources

- Medical Examiner or Coroner's Reports
- Hospital Emergency Departments
- Trauma Departments
- Prehospital Ambulance Response Data (EMS)
- Schools
- Local Police Department

Additional Resources

- *Statewide Integrated Traffic Report System (SWITRS) – California Highway Patrol*
<http://www.chp.ca.gov/>
 - o All reported fatal injury crashes occurring on California roadways are reported.
 - o County-specific data is available via phone, e-mail, fax, and mail requests.
- *California Office of Statewide Health Planning and Development (OSHPD)*
<http://www.oshpd.state.ca.us/index.htm>
 - o Hospital discharge and emergency department data files are reported.
 - o County-specific data is available by request, usually for a fee.
- *California Poison Control System (CPCS)*
<http://www.calpoison.org/>
 - o Maintains computerized records of all calls.
 - o Will generate statistics by request.
- *California Health Kids Survey (CHKS)*
http://www.wested.org/pub/docs/chks_home.html
 - o A comprehensive youth health and risk behavior data collection service available to all California local education agencies.
 - o Provides local, state, and national comparisons.

Population Data

You may have access to population estimates for your county. If not, the following resources can be used to find this data. We recommend using Department of Finance data for California counties.

- *United States Census Bureau*
<http://www.census.gov/popest/counties/asrh/CC-EST2004-alldata.html>
 - o Provides data files for all states, by county.
 - o Age, gender, and race/ethnicity detail is available.
- *California Department of Finance Demographics*
www.dof.ca.gov/HTML/DEMOGRAP/data.htm.
 - o Provides data files for all counties within the State of California.
 - o Age, gender, and race/ethnicity detail is available.

Using *Epi in a Box*

Epi in a Box is a spreadsheet tool designed to help you organize your injury data, calculate and compare rates, make comparisons to population characteristics, and determine appropriate target injury rates. This tool was created based on categories of unintentional injuries, assaults/homicides, and self-inflicted injuries/suicides as defined by the Epidemiology and Prevention for Injury Control (EPIC) branch of the California Department of Health Services. However, *Epi in a Box* can be used with any injury category, from any data source.

Please see the document titled “Cause of Injury Coding” for a breakdown of ICD-9 E-codes and ICD-10 WXYZ codes used by EPIC to define each injury category. The EPICenter Help Area link, located on EPIC’s homepage, will direct you to a description of their fatal and nonfatal injury data sources.

Getting Started

Before you begin using *Epi in a Box*, you should be familiar with the spreadsheet tool, and a few functions of Microsoft Excel.

1. Included on the *Epi in a Box* toolkit, in addition to this guide, are the following documents. Click on the document name to open the item (requires internet access).
 - a. [Epi in a Box_5YearAgeGroups.xls](#)
 - Rates and percentages are calculated based on five-year age groups (0-4, 5-9, etc..)
 - b. [Epi in a Box_EPICAgeGroups.xls](#)
 - Rates and percentages are calculated based on EPIC age groups (<1, 1-4, 5-12, 13-15, 16-20, 21-44, 45-64, 65+)
 - c. [Epi in a Box_IndividualAgeGroups.xls](#)
 - Rates and percentages are calculated based on single years of age, or user-defined age groups.
 - d. [Cause of Injury Coding.pdf](#)
 - Definitions of EPIC injury categories by ICD-9 E-codes and ICD-10 WXYZ codes.
2. Before you begin using a spreadsheet, save a copy in a convenient location on your computer. **You should save a new copy of the spreadsheet on your computer from the internet each time you begin a new project.**
3. Within each of the three spreadsheet documents, referred to individually as a workbook, are seven tabs, or worksheets. The worksheets are described in detail in the next section. The titles are:
 - a. Population
 - b. Unintentional Injury Summary
 - c. Self-Inflicted_Suicide Summary
 - d. Assault_Homicide Summary
 - e. (Blank) Injury Summary
 - f. Individual Age_Sex Comparison
 - g. Rate Change Calculations
4. Each worksheet is protected, meaning you can only enter information in the appropriate cells. Cells that are open for you to enter data are highlighted in **blue**. Cells that contain information that is calculated based on the data you enter are highlighted in **green**. Green cells cannot be changed.
5. All **rates** are calculated as the number of **injuries per 100,000** in the population.
6. To **navigate** to each worksheet, click on the arrows on the bottom left hand corner of your screen, or click on the tab itself.

26	85+	13,860	26,342	40,202
27	All Ages	1,464,555	1,443,950	2,908,505
28				
29				
30				
31				
32				
33				
34				

Total Population by 5-Year Age
(Enter Location & Year)

Population Unintentional Injury Summary Self-Inflicted Suicide Summary

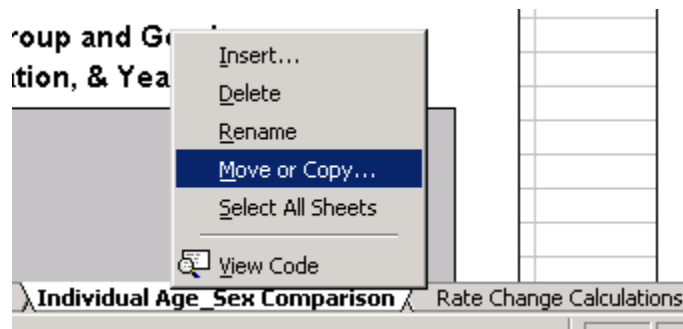
Ready

Start | H:\HOLLY\... | Inbox - Mi... | Epi toolkit ...

- To **change the name of a worksheet or tab**, double click on the title. When it is highlighted in black, you can type in the new name.



- To **change the title of a graph**, left click on the title. A box will appear. Use your cursor to highlight the phrase you want changed. You can now add appropriate text.
- To **copy a worksheet**, right click on the appropriate tab. Left click on "Move or Copy..." Check the box next to "Create a Copy". Indicate where you would like to insert the worksheet by choosing a location under "Before Sheet:".



- Each worksheet is formatted for easy printing. To **print** a single worksheet, go to File, Print. Under "Print what", click on "Active sheet". To print the entire workbook, click on "Entire workbook".
- To **copy a table** from a worksheet into another document, click on a corner cell, hold, and drag over the entire table until it is highlighted. Right click on the highlighted table (make sure to include the title!), and a menu will appear. Click on "Copy", and paste into your document.
- To **copy a graph** from a worksheet, simply right click on the graph, click on "Copy", and paste into your document.

13. To quickly calculate the **sum** of a group of numbers, highlight the list. The sum appears in the bottom right-hand corner of your screen.

0.2	554	710.2	1,300	1,390.4
0.1	552	710.1	1,104	1,420.3
5.8	885	1,255.8	1,770	2,511.6
1.6	321	631.6	642	1,263.2
6.0	123	306.0	246	611.9
9.2	9,867	339.2	19,734	678.5

Range Calculations /

Sum=1,881 NUM

ICent... | Word... | Epi in a ... | 12:10 PM

14. To **remove the protection** from a worksheet, go to Tools, Protection, and click on "Unprotect Sheet". **CAUTION:** Protection should **ONLY** be removed if you are attempting to *delete* rows of data, as described in #15 below. **If this is done incorrectly, you run the risk of damaging formulas, thus producing inaccurate data!!** To turn the protection back on, go to Tools, Protection, and click on "Protect Sheet". Uncheck the "Objects" box, and click "OK".
15. Each of the tables and graphs in this worksheet include all age groups. If you want the tables and graphs to only represent select age groups, such as 0-19 year olds, you can **delete unused rows of data**. To do this, you must first remove protection from the worksheet. Highlight the rows of data that you want to delete by clicking and dragging on the row numbers found on the left-hand side of the screen. Right click on the highlighted data, and select "Delete". Be sure to turn the protection back on immediately, to reduce your risk of damaging formulas within the tables.

Collecting Your County's Population Data

Hopefully you have access to your county's population data by age and gender for the year you are investigating. If not, here's a procedure for finding this data:

1. Open the California Department of Finance Demographics website:
www.dof.ca.gov/HTML/DEMOGRAP/data.htm.
2. Click on "Race/Ethnic Population with Age and Sex Detail, 2000-2050."
3. Click on "Go to Data Files."
4. Scroll down until you find the name of the county that you want data for.
5. RIGHT click on the file name, and select, "Save Target As..."
6. Save the file somewhere convenient.

7. Open the file using MS excel or another spreadsheet program that you're comfortable with. If you are using MS excel:
 - a. Open a blank workbook in excel.
 - b. Click on "File" and "Open."
 - c. Under "Files of Type", select either "text files" or "all files."
 - d. Open the data file from where you saved it.
 - e. A box labeled, "Text Import Wizard, Step 1 of 3" will open.
 - f. Under "Original data type," choose "Delimited."
 - g. Click "Next."
 - h. Under "Delimiters," click on "Comma."
 - i. Click "Next."
 - j. You shouldn't need to do anything with the next screen, so click "Finish." Excel will import all of your data into a spreadsheet.
8. Click on "Data"... "Filter"... "Autofilter." An arrow should appear next to each of your field names.
9. Click on the arrow by the "Year" column, and select the year that you're interested in.
10. Select this array of data by highlighting all columns for that year. Copy this data, and paste to another worksheet.
11. Next to the column titled, "Multirace", type in "Total". Use the sum function to find the sum of all races for each single year of age.
12. You may want to group this data into five-year (or other) age ranges. To do this, set up another table with the desired age groupings; then use the sum function to fill in this data.

Collecting Your Local Injury Data Using EPIC

To access your county's injury data from the California Department of Health Services:

1. Go to the California Department of Health Services EPIC website:
<http://www.applications.dhs.ca.gov/epicdata/default.htm>.
2. Under the "Custom Data Tables" column, choose either *Fatal data* or *Nonfatal data*.
3. Choose your year(s) of data. For example, if you want 2003 data, choose "From 2003 through 2003".
4. Choose your cause of injury. You can choose either a specific cause, such as "Unintentional – Fall" or a grouping of injuries, such as "All unintentional injuries". If you choose a grouping of injuries, you can later stratify by specific cause.
5. Choose your county.

6. Choose your age range. If you want to stratify by single year of age or by age group, choose “all ages.” If you are looking for all ages within a specific age group, choose “From ____ through ____ years old.” For example, if you are looking only at injuries among children 5 to 14 years old, choose “From 5 through 14 years old.”
7. Scroll down to **additional detail**. Here, you can choose to stratify your results by a variety of variables. For example, if you chose “all unintentional injuries” in 2003 for “all ages”, and you want to stratify by specific cause and 5-year age group, you would select the following: First level of detail = Cause of Injury. Second level of detail = Age (5-Year Groups: 0-4, 5-9, etc.).
8. Click on “Display Table”.
9. Right click on the resulting table and copy into a blank excel file, or print the table for later use.

Structure of *Epi in a Box*

The seven worksheets found in each of the three workbooks are identical, except for the age groupings. For the purpose of this description, the workbook titled: *Epi in a Box_5YearAgeGroups.xls* will be used.

Before you begin entering any injury or population data, it is important to document your data source. At the top of each worksheet, you will find a blue box requesting information you should collect about your source. It is also a good idea to change the names on the tab of each worksheet to reflect the data you are collecting. For example, you might want to change “Population” to “Population 2000”, or “Individual Age_Sex Comparison” to a more accurate description of your data, such as “Burn_Hot Object/Substance 2000.”

Several graphs are created by *Epi in a Box*, which will be discussed in more detail in the sections to follow. You should always change the titles of the graphs to reflect the data they are representing. Prompts are given on each graph as a reminder of what information should be included.

Population Worksheet

The *Population* worksheet contains critical data used in rate calculations by each of the other worksheets. Population data should be entered for each age group, by gender. Remember, blue cells are for data entry, green cells are calculated values.

Totals are calculated automatically for each age group. For example, if you enter 42,000 in the 0-4 age group under “Male”, and 43,000 in the 0-4 age group under “Female”, the total population for the 0-4 age group will be calculated, and 85,000 will appear under “Total”.

Totals for all ages, by gender, are also calculated. If you enter population data for each age group, 0-4 through 85+, the “All Ages” total will represent the entire population. If you enter population data for only the age groups 0-4 through 15-19, the “All Ages” total will represent the population aged 0 to 19 years.

It is important that the age groups you enter in your population worksheet match the age groups you enter later in your injury summary worksheets. If you are only evaluating injuries among 0-19 year olds, you should only enter population data for 0-19 year olds. Note: If you do not enter population data for a specific age group, rates will not be calculated for that age group in later worksheets. The symbol, #DIV/0! will appear, indicating that population data is missing.

Two graphs are displayed in the population worksheet. "Total Population by Gender" is a pie graph representing the distribution of the population by gender. "Total Population by 5-Year Age Group and Gender" represents the distribution of the population by age group and gender. Change the titles of the graphs to include geographic location, age groups, and year of data.

Note: The population worksheet is used for calculating rates in each of the other worksheets. You should never make a copy of the population worksheet. If you want to look at more than one year of data, make a copy of the entire workbook, and enter new population data to match the year of injury data you are working with.

Unintentional Injury Summary Worksheet

The *Unintentional Injury Summary* worksheet lists all categories of unintentional injury as defined by EPIC. This worksheet is meant to give you the big picture by calculating rates of fatal and nonfatal injury, as well as the total, for each injury category.

After collecting your injury data, enter the number of fatal and the number of nonfatal injuries for each category, by age group. Gender is not included in this summary table. Rates of both fatal and nonfatal injuries are automatically calculated per 100,000 population. The "Total" column represents the sum of fatal and nonfatal injuries and the corresponding rate.

When the number of injuries is less than 5, a star (*) will appear in the rate column. Rates calculated based on fewer than 5 events are statistically unreliable, and should not be used. For further discussion, see the description in "Small Numbers".

You do not have to enter data for all age groups, or for all injury categories. You could also choose to enter only fatal or nonfatal data. In this case, the "Total" column will be identical to the fatal or nonfatal data column. Enter only the data that you deem important to your efforts. Keep in mind, however, it is better to start with too much information than too little.

The summary worksheet(s) can be used as a surveillance tool. Use a new *workbook* for each new year of injury data that you want to investigate. Remember, the year of your population data must match the year of your injury data.

Self-Inflicted/Suicide Summary Worksheet

The *Self-Inflicted/Suicide Summary* worksheet is identical to the *Unintentional Injury Summary*, but instead lists all categories of self-inflicted injuries (nonfatal) and completed suicides (fatal). See the description above for instructions on how to use this worksheet.

Assault/Homicide Summary Worksheet

The *Assault/Homicide Summary* worksheet is identical to the *Unintentional Injury Summary*, but instead lists all categories of assault injuries (nonfatal) and homicides (fatal). See the description above for instructions on how to use this worksheet.

(Blank) Injury Summary Worksheet

The *(Blank) Injury Summary* worksheet is identical to the three worksheets above, but does not list any categories of injury. This worksheet can be used for injury data collected from a source other than EPIC. Enter the injury type in the blue cell labeled “Enter Injury Category”.

Six blank categories are included in this worksheet. If you need additional categories, you can make additional copies of the worksheet. See instructions under the “Getting Started” section of this manual.

Individual Age/Sex Comparisons Worksheet

This worksheet is designed to give you more detailed information on a specific injury category. For example, after completing the *Unintentional Injury Summary Worksheet*, you noticed that nonfatal rates of Burns, Hot Object/ Substance were especially high for 0-4 year olds when compared to other age groups. You now want to investigate this category by gender. Re-enter the nonfatal data for Burns, Hot Object/Substance, this time by gender, in the table titled “Rates of Injury by Age Group and Gender”. Rates per 100,000 are automatically calculated for each age group, by gender.

Below this table is a graph, titled “Injury Rates by 5-Year Age Group and Gender. This graph displays rates of injury for males (blue bars) and females (red bars), as well as the overall rate of injury (yellow line).

The table titled “Percentage of Injuries Compared to Percentage of Population: Within Each Age Group, by Gender”, compares the distribution of injuries to the distribution of the population. For example, according to the table below, 62% of the injuries sustained by 0-4 year olds are to males, and 38% are to females. In the 0-4 year old population, 50.2% are male, and 49.8% are female. Because the percentage of injury among males is higher than the percentage of the male population, you can conclude that for this data, males were more likely to sustain these injuries than females were. At the bottom of the table, the total, in this example, refers to all children ages 0 to 19 years.

Percentage of Injuries Compared to Percentage of Population Within Each Age Group, by Gender				
	Male		Female	
	Injuries	Population	Injuries	Population
0 to 4	62.0%	50.2%	38.0%	49.8%
5 to 9	53.5%	51.2%	46.5%	48.8%
10 to 14	57.5%	51.5%	42.5%	48.5%
15 to 19	78.9%	53.2%	21.1%	46.8%
20 to 24	-	-	-	-
25 to 29	-	-	-	-
30 to 34	-	-	-	-
35 to 39	-	-	-	-
40 to 44	-	-	-	-
45 to 49	-	-	-	-
50 to 54	-	-	-	-
55 to 59	-	-	-	-
60 to 64	-	-	-	-
65 to 69	-	-	-	-
70 to 74	-	-	-	-
75 to 79	-	-	-	-
80 to 84	-	-	-	-
85+	-	-	-	-
Total	59.9%	51.6%	40.1%	48.4%

Below this table are two pie graphs. "Distribution of Injuries by Gender" displays the distribution of injuries, by gender, for all ages entered into your table. For this example, the graph represents burn injuries sustained by 0-19 year olds. "Distribution of Population by Gender" displays the distribution of the population, by gender, for all ages entered into the population worksheet. If you want this graph to represent only those ages for which you are calculating injury data, delete the unnecessary numbers from the population worksheet.

The table titled "Percentage of the Total Injuries Compared to Percentage of Total Population by Age Group" calculates the percentage of injuries sustained by each age group. For example, according to the table below, 65% of all burn injuries in the 0-19 year population occur among 0-4 year olds, 23.4% occur among 5-9 year olds, and so on. When you compare this to the 0-19 population, 23.8% of the population are 0-4 years old, and 24.7% are 5-9 years old. This tells you that burn injury occurred disproportionately among 0-4 year olds.

Percentage of Total Injuries Compared to Percentage of Total Population by Age Group		
	Total	
	Injuries	Population
0 to 4	65.0%	23.8%
5 to 9	23.4%	24.7%
10 to 14	9.5%	25.8%
15 to 19	2.1%	25.7%
20 to 24	0.0%	0.0%
25 to 29	0.0%	0.0%
30 to 34	0.0%	0.0%
35 to 39	0.0%	0.0%
40 to 44	0.0%	0.0%
45 to 49	0.0%	0.0%
50 to 54	0.0%	0.0%
55 to 59	0.0%	0.0%
60 to 64	0.0%	0.0%
65 to 69	0.0%	0.0%
70 to 74	0.0%	0.0%
75 to 79	0.0%	0.0%
80 to 84	0.0%	0.0%
85+	0.0%	0.0%
Total	100.0%	100.0%

Below this table is a graph titled “Injury Distribution Compared to Population Distribution by Age Group”. The red bars represent the percentage of individuals with the injury within each age group, and the yellow line represents the percentage of the population within each age group. When you see a red bar that is higher than the yellow line, you can conclude that the percentage of injury in that age group is disproportionately high compared to the percentage of that age group in the population.

Remember to change the titles of all graphs to include important information such as the injury type and intent, location, age groups, and year of data.

Rate Change Calculations Worksheet

The *Rate Change Calculations* worksheet can be used after you’ve chosen an indicator that needs improvement. This will help you to establish a reasonable target rate of injury based on your knowledge of previous intervention efforts.

Two tables are presented; the first provides information based on a X% decrease in rate, the second provides information based on a X% increase in rate. Enter the calculated rate of injury, and the corresponding population. A single box titled “Number of Injuries” gives you the number of injuries sustained in the population based on your rate. This should match the number of injuries you entered previously in the summary and comparison worksheets.

In the first table, under the heading, “New rate per 100,000 given a X% decrease”, you will find the expected rate of injury if you were to achieve a X% decrease. Under “Number of injuries expected given a X% decrease”, you will find the number of injuries

to be expected in this population if you were to achieve a X% decrease. Finally, under “Number of injuries prevented given a X% decrease”, you will find the number of injuries you can expect to prevent in this population if you were to achieve a X% decrease.

The second table presents the same information, assuming an increase in rate. Certainly you wouldn’t hope to achieve an increase in the rate of a particular injury, but suppose you’ve noticed a steady 8% increase in the rate of deaths due to bicycle crashes. If you want to predict the rate of bicycle deaths for the following year, assuming the same 8% increase, you would use this table.

Another use of the *Rate Calculations Worksheet* is to determine the percent change in rate. Suppose you found that the rate of drowning deaths among 0-4 year olds was 5.4 per 100,000 in 1999 and 4.8 per 100,000 in 2000. Enter 5.4 in the box labeled “Enter Rate per 100,000”. Scroll through until you find 4.8 under “New rate per 100,000 given a X% decrease.” The corresponding percentages are 11% and 12%. You can state that from 1999 to 2000, the rate of death due to drowning among 0-4 year olds decreased by 11-12%.

Note: The percent increase or decrease in a rate can still be calculated if you do not enter population data. You will not, however, have information on the expected number of injuries. If population data is not entered, the symbol #VALUE! will appear.

Race/Ethnicity

Race/ethnicity is an important demographic factor when considering rates of injury. Not all data sources will offer injury data stratified by race/ethnicity, but when they do, you might consider calculating rates for each of the major subgroups of race/ethnicity. The four major subgroups of race/ethnicity include White, Black, Hispanic, and Asian/Other.

If you want to view rates of injury by race/ethnicity using Epi in a Box, you will need to copy the *entire workbook* four times. Save each workbook according to the racial/ethnic category that it represents.

You will first need to gather and enter your population data by race/ethnicity, age and gender. In other words, in the Hispanic workbook, you will need to enter the Hispanic male population and the Hispanic female population for each age group. Do the same for the white workbook, the black workbook, and the Asian/other workbook.

Once you have entered your population data, you can enter your injury data by race/ethnicity in the summary worksheets and/or the individual age/sex comparison worksheets. The resulting tables and graphs will represent the appropriate racial/ethnic categories. Make sure to change their titles to reflect this.

Missing Data

Most of the data sources described here will have complete information by age and gender. If you notice only one or two injuries with missing age or gender information, it is generally okay to exclude them from your rate calculations. When reporting your rates in a table or a graph, include a note in the footnote, such as “Rates exclude 3 cases with missing age.”

However, if you notice that a significant proportion of your injuries or fatalities is missing age or gender information, you should still consider them when calculating the total rate. You cannot create a detailed rate for a “Missing” category because you do not know what the denominator is. You can include the number of injuries or fatalities missing an age in a total rate for all ages.

For example, using the *Individual Age Groups Workbook*, enter your age groups and a missing category. The missing category will not have population information, and will therefore not calculate rates. However, incidents with missing age information will be factored into the total rate.

Bicycle Related Hospitalization Rates (per 100,000) by Age Group
Your County, Ages 0-19 Years, 2002

	Fatal		Nonfatal		Total	
	Number	Rate	Number	Rate	Number	Rate
0 to 4	3	*	45	34.0	48	18.8
5 to 9	5	2.1	54	22.2	59	12.3
10 to 14	9	2.1	67	15.8	76	8.9
15 to 19	7	1.6	78	18.4	85	9.8
Missing	1	#DIV/0!	5	#DIV/0!	6	#DIV/0!
Total	25	2.0	249	20.4	274	11.2

Source: County of Yours, Division of Epidemiology, 2002

*Rates not calculated on fewer than 5 cases.

Total includes cases with missing age information.

Data Sourcing

When presenting injury data in a table or a graph, it is crucial that you document the exact source of your data so that it can be replicated by someone else should the need arise. Documentation occurs in both the title and in the footnote.

The title of your table or graph should always include the intent of injury (if applicable), the injury category, the age group(s), gender (if applicable), year(s) of data, and geographic location or community. The footnote should include the injury data source, the population data source, the year(s) of data, and who compiled the data if different than the data source. If there are any special notes about the table, such as the meaning of a starred (*) cell, or if the units of the rate are not specified on the table or graph, you should include this information in the footnote as well.

Frequency and Rate (per 100,000) of Hospitalization for Scald Injuries
By Gender, Your County, 0 to 19 Years, 2001

	Male		Female		Total	
	Injuries	Rate	Injuries	Rate	Injuries	Rate
0 to 4	165	86.3	132	80.4	297	83.6
5 to 9	46	22.8	37	19.5	83	21.5
10 to 14	18	6.8	16	6.4	34	6.7
15 to 19	9	3.2	3	*	12	2.4
All Ages	238	31.5	188	34.8	426	33.0

Source: CDHS, EPIC. Compiled by County of Your, Division of Epidemiology, 9/2005 Population

Data: CDOF

*Rates not calculated on fewer than 5 incidents

Interpreting *Epi in a Box*

Before all of this information overwhelms you, remember what your primary purpose is for gathering this data and calculating rates. *Epi in a Box* can be used for a variety of purposes and in multiple ways. Is the data for a needs assessment? Program evaluation? Regardless of your purpose, this tool will help you to answer the following questions.

1. *How do rates for different injuries compare to each other?*

Begin by using the summary worksheets to look at the rates of injury occurring between categories. By scanning across the categories of injury, you'll be able to identify the leading causes of fatal and nonfatal injury overall and by age group. Do any stand out as being especially high or especially low based on state or national indicators?

2. *How are injuries distributed throughout your population?*

Remember that injuries are closely linked to age. Look at each injury category, and compare the rates by age group. You should expect to see different patterns for different age groups, but are injuries occurring disproportionately at a certain age? Use the comparison worksheet to determine how the proportion of injuries in each age group and gender compare to the proportion of the population.

3. *How have the rates of injury changed over time?*

Compare the summary or comparison worksheets you have prepared for multiple years of data. Have the rates for a particular injury changed over time? By how much? Use the rate calculation worksheet to determine the percent change from one year to the next.

4. *How will the rate of injury change if I perform an effective intervention?*

What does an effective intervention mean to you? How many injuries do you hope to prevent? Given your baseline injury rate, determine the expected rates of injury given a percent change by using the rate calculation worksheet.

Conclusion

Injury data is a crucial component to the prevention of injuries. *Epi in a Box* will give you the data you need to determine your population's major injury problems and which populations are at risk. This tool will also provide tables and graphic displays of injury distribution for use in reporting the injury problem to the media and policy makers, for funding intervention efforts, and for conducting needs assessments and program evaluations.

Using your injury data as the foundation, you should be able to combine your existing knowledge regarding effective injury prevention strategies with empirical evidence of need in your community to identify priority injuries for prevention.

For more information, contact:

Alan Smith or Leslie Ray
County of San Diego
Division of Emergency Medical Services
6255 Mission Gorge Road
San Diego, CA 92120
(619) 285-6429
alan.smith@sdcounty.ca.gov or leslie.ray@sdcounty.ca.gov

Epidemiologist in a Box EVALUATION

Thank you for taking the time to complete this survey. Your input will help us to improve future versions of Epidemiologist in a Box. Please e-mail this form with your responses to alan.smith@sdcounty.ca.gov, or mail to the address above.

For each of the statements below, please indicate your agreement or disagreement by placing an X in the appropriate column.

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1. I found the text easy to read.					
2. The spreadsheet was easy to understand.					
3. <i>Epi in a Box</i> was a useful tool.					
4. I will use <i>Epi in a Box</i> on future projects.					
5. I would recommend <i>Epi in a Box</i> to a colleague.					

6. What additional information would you like to see included in this tool?

7. What would make *Epi in a Box* easier to use?

8. Any other comments?

9. May we contact you? Please provide your contact information below (optional).