

# EPI INFO

version 5

extracto del manual

CURSO-TALLER SOBRE  
USO DE LA EPIDEMIOLOGIA EN LA  
PLANIFICACION Y EVALUACION DE  
PROGRAMAS DE SALUD

7 - 11 OCTUBRE 1991

UNFPA CEPAL CELADE

OPS-OMS

Programa de Postgrado en Población  
y Desarrollo

Curso-Taller

USO DE LA EPIDEMIOLOGIA EN LA PLANIFICACION  
Y EVALUACION DE PROGRAMAS DE SALUD.

MATERIALES DE TRABAJO

Santiago de Chile, 7-11 de Octubre, 1991.

UNFPA CELADE CEPAL OPS-OMS

## CHAPTER 2

### WHAT IS EPI INFO, VERSION 5?

Epi Info is a series of microcomputer programs for handling epidemiologic data in questionnaire format and also for organizing study designs and results into text that may form part of written reports. A questionnaire can be set up and processed in a few minutes, but Epi Info can also form the basis for a powerful disease surveillance system database with many files and record types. It includes the features most used by epidemiologists in statistical programs such as SAS or SPSS and database programs like dBASE combined in a single system that may be freely copied and given to friends and colleagues.

There are three levels of facilities in Epi Info for processing questionnaire or other structured data. Using Epi Info on the simplest level, you can computerize a questionnaire or investigation form in a few minutes by:

- Running the main menu.
- Typing the questionnaire or form with EPED, the word processor, or another word-processing program.
- Entering data in the questionnaire using the ENTER program.
- Analyzing the data using the ANALYSIS program to produce lists, frequencies, cross tabulations, means, graphs, and accompanying statistics.

As your knowledge of the program grows, you will use additional features to shape data entry and analysis to your needs. You will probably want to:

- Insert error checking, skip patterns, and automatic coding in the questionnaire using the CHECK program.
- Select records, create new variables, recode data, manipulate dates, and carry out conditional operations with IF statements during ANALYSIS.

- Incorporate these operations into program files so that they can be performed repeatedly, or by other persons unfamiliar with programming.
- Import and export files from other systems like SAS, SPSS, dBASE, and Lotus 1-2-3.
- Change the names of variables in the data file using the CHECK program.

The third level of features is important if you are setting up a permanent database system, a large study or want to customize Epi Info's operations to suit special needs. For such purposes, you can:

- Program the data entry process to include mathematical operations, logical checks, color changes, pop-up boxes, and custom routines written in other languages.
- Specify the format of reports from ANALYSIS to produce customized tables.
- Enter data into more than one file during the same session, moving automatically among several questionnaires within ENTER.
- Link several different types of files together in ANALYSIS so that questions can be answered that require data from more than one file.
- Compare duplicate files entered by different operators to detect data-entry errors.

EPED, the editor included with Version 5, contains a programming system for text called EPIAID that parallels Epi Info's facilities for structured data. Programs are included to guide you through the creation of questionnaires and the design of an epidemic investigation. The text produced can be used as part of an investigation report. EPIAID programs can be written to assist in letter writing or production of reports for other purposes.

Version 5 of Epi Info contains a host of new features, such as handling of several related files at one time, easy revision of data file formats, browsing and updating of files during analysis, ability to analyze dBASE files directly, missing value indicators, programming languages for data entry and for text editing, ability to call Pascal or assembly language programs from the data-entry program, fixed decimal points, European dates, and many other facilities. Files created by previous versions of Epi Info can be used in Version 5. With a few revisions, programs written for Version 3 can be run in Version 5.

The programs included in Epi Info are:

- EPI -- The main menu that allows access to other programs.
- EPED -- A word processor for creating questionnaires or for general use. A special facility called EPIAID in EPED runs programs that assist in text preparation and in the design of epidemiologic investigations.
- ENTER -- Produces a data file automatically from a questionnaire created in EPED or another word processor. Manages entry of data from a questionnaire into a disk file. Allows revision of file format even after records have been entered.
- ANALYSIS -- Produces lists, frequencies, cross tabulations, and a variety of other results from Epi Info or dBASE files. Appropriate epidemiologic statistics, such as odds ratios, relative risk, confidence limits, the Fisher exact test, and chi square tests accompany the tables. Stratified analyses with the Mantel-Haenszel technique, one-way ANOVA and Kruskal-Wallis tests, linear regression, and analysis of matched case-control studies are also offered. Records may be selected or sorted using defined variables, "if" statements, and mathematical and logical operations. Graphing, complex report formatting, generation of new data sets for testing, and a programming language are included in ANALYSIS. Data items spread over several files can be linked and analyzed as a single unit.
- CHECK -- Sets up ranges, legal values, automatic coding, and skip patterns for the ENTER program, if desired. Mathematical and logical operations between fields, complex jump patterns, access to several files in the same entry process, and use of custom programs provided by the user are also supported by CHECK.
- STATCALC -- Calculates statistics from table values entered from the keyboard. STATCALC performs calculations for single and stratified 2-by-2 tables, sample size, and single and stratified trend analysis.
- CONVERT -- Converts data files from Epi Info into 12 other formats for a variety of database and statistics programs.
- IMPORT -- Brings in files from other systems so that they can be used in Epi Info.
- MERGE -- Merges files produced from questionnaires with the same or different formats. This allows combining data files entered on different computers, and updating of previous records using data from new entries.

## CHAPTER 7

### CREATING QUESTIONNAIRES USING THE EPED EDITOR AND EPIAID

#### Tutorial

Run EPED from the main menu. Press <F3> to see the EPIAID menu, and choose "Make Epi Info Questionnaire" by moving the cursor bar with the down arrow key and pressing <Enter> to make the choice.

Follow the instructions in the tutorial carefully and make a questionnaire. When you are asked for a file name, use the first eight letters of your first name followed by a period and .QES, as in MARY.QES

#### OVERVIEW

Before beginning to enter data, you must provide Epi Info with a questionnaire. The ENTER program described in Chapter 8 uses the questionnaire to construct a data file, called a .REC file, automatically. To do this, ENTER reads a questionnaire that has been produced on a word processor in ordinary text format. It constructs variable names from the text in your questionnaire and derives data types and lengths of variables from the blanks you have inserted in the questionnaire. After creating the data file, ENTER presents the questionnaire on the computer screen and allows entry of data within a few seconds.

The word processor, EPED, is provided for convenience in constructing questionnaires. EPED can also be used for correspondence, reports, and other purposes, as described in the preceding chapter. This chapter tells how to make a questionnaire in EPED. If you have another favorite word processor that can produce a simple text (ASCII) file or a Wordstar file, the instructions that follow will tell how to produce the questionnaire, but you must rely on other sources for information about the word processor itself.



## FIELD TYPES IN EPI INFO

Text or "underline" fields, indicated by continuous underline characters. Any printable character can be entered in this field. Blank fields (spaces) are interpreted as missing data in ANALYSIS. The length of the variable or field will be the number of underline characters used. The maximum length of a text field is 80 characters.

#  
##.#  
etc.

Numeric fields. Only numbers or spaces will be accepted. If nothing is entered, the result will be a blank, which is interpreted as a missing value and displayed as a period (.) in ANALYSIS. The number of digits is indicated by the number of "#"s. If a decimal point is given, the field will be in "fixed decimal" format, allowing exactly the indicated number of digits to the right of the decimal point. For monetary amounts from 000.00 to 999.99, for example, ###.## would be the appropriate field. A number field can be up to 14 characters long, counting the decimal point as 1 character.

<A>  
<A >

Upper-case fields. These are similar to general-purpose or "underline" fields, but entries will be converted to upper case. The length is indicated by the number of characters between the "less-than" and "greater-than" symbols.

<Y>

A "Yes/No" field. Only Y, N, and SPACE or <Enter> are accepted. Spaces or the <Enter> key are interpreted as missing data in ANALYSIS. Entries are converted to upper case after they are entered.

<mm/dd/yy>  
<mm/dd>  
<dd/mm/yy>  
<dd/mm>

U.S. and European Date fields. Dates will be checked when entered to be sure they are valid. When entering data, you type only the numbers in the date; the program provides the separating slash marks.

<phonenum>

A local phone number, entered as xxx-xxxx where the x's are digits.

<long distance>

A long-distance phone number, entered as (xxx)xxx-xxxx, where the x's are digits.

<today>  
<today/yy>  
<today/yyyy>

Today's-date or date-of-last-change field, which will automatically enter the date when the computer record was last saved. If the record is later edited and saved again, the contents of the field will be the latest date on which the record was saved. The three forms allow the field to contain month and day only, or month, day, and two-digit year, or month, day, and four-digit year.

The date which automatically appears in this field is the "system date" from the computer. Be sure to enter today's date each time when turning the computer on, unless your computer automatically restores the correct date on powering up. Within Epi Info, the date in this field is supplied automatically, and you will not be able to enter anything but a carriage return (<Enter>).

UPDATING FILE 1 FROM MATCHING RECORDS IN FILE 2 -- CHOICE 3

In disease surveillance work, frequently the first report received is incomplete or not entirely correct. In conducting field studies, not all blanks in a questionnaire may be completed at the first interview, and the information may need to be supplemented later. This requires that the initial records be updated.

A convenient way to do this with MERGE is to create a second file of records with the same unique identifier, such as person identification number. Blanks in the second set of records are completed only if the corresponding blank on the first record is to be updated (overwritten). If a field in file 2 has the missing value, the corresponding field in file 1 will NOT be replaced, and whatever was in the field before will remain.

Suppose you wish to update the file CASES.REC from a file called CASESUPP.REC and create a new file called REVCASES.REC.

CASES.REC contains

CASENO	AGE	SEX
1001	33	
1002	15	M
0999	30	F

CASESUPP.REC contains:

CASENO	AGE	SEX
1001	0	F
1002	16	
1005	56	M

Now run MERGE with the appropriate file names entered, Merge Option 3, and CASENO as the MatchField. If you answer "Y" to the question about including records in File 2 that are not present in File 1, the results will be:

REVCASES.REC

CASENO	AGE	SEX
0999	30	F
1001	33	F
1002	16	M
1005	56	M

If there is more than one blank on a line, Epi Info begins the search for a useful field name after the preceding blank. Thus

Name \_\_\_\_\_ Age \_\_\_\_ Sex \_

produces three fields entitled NAME, AGE, and SEX. Now comes the tricky part; suppose that there is no preceding text on the same line, as in

Address \_\_\_\_\_  
\_\_\_\_\_

Epi Info will call these two fields ADDRESS and ADDRESS2. The program also detects duplications, such as the occurrence of another

Address \_\_\_\_\_

many lines removed from the first two. Rather than making a duplicate field name, the program would call this one ADDRESS3. ENTER can handle up to 99 such duplications of a single field name.

If you wish to number your fields, this can be done as

1. Name \_\_\_\_\_

To prevent trouble in programs requiring that a field name start with a letter, Epi Info places an "N" in front of numbers if they are the first item in a variable name. You may prefer to avoid this and use:

Name 1. \_\_\_\_\_

Field names surrounded by curly brackets {} will be used in preference to those normally generated, so that you can control the process. Thus, What is your {name} will result in the field name NAME rather than WHATISYOUR. More than one set of curly brackets can be used to form a field name if they are on the same line, as in {What} is your {name}, giving the field name WHATNAME.

It is also possible to change field names after the data file has been created, as described in the chapter on the CHECK program.

## CHAPTER 8

### ENTERING DATA USING THE ENTER PROGRAM

#### Tutorial

Create a questionnaire as described in the previous chapter. Be sure that it has a name ending in ".QES".

Run ENTER from the main menu.

In response to the request for "Data file:", type the name of your questionnaire, but leave off the ".QES". If your questionnaire was "MARY.QES", then type "MARY".

Press <Enter>. The program asks for the name of the questionnaire file. To see a list of .QES files, press <F1> and then move the bar cursor with the arrow keys and press <Enter> to make a choice. Here, of course, you type the name of your questionnaire file, e.g. "MARY", and then press <Enter>.

The ENTER program will now make a data file (.REC file) from your questionnaire and display the questionnaire on the screen, ready to receive data.

Fill in the blanks in the questionnaire, pressing <Enter> after each until you have completed the form. The message "Write data to disk? (Y/N)" will appear at the bottom of the screen. Answer "Y" or "y". Note that record number 2 is now displayed in the lower right corner of the screen.

Complete several questionnaire records. Try the arrow keys and the <PgUp>, <PgDn>, <Home>, and <End> keys.

After saving the last questionnaire, use the <F7> and <F8> keys to move back and forth through the records and review your work.

Edit one of the entries by placing the cursor in any field and changing the contents. Then press the <End> key and "Y" to save the edited copy.

See the section on "Finding Records That Match Criteria" in this chapter and try a few searches. Note that <Ctrl-F> displays the function keys that are used in finding records.

To return to entry mode and add a new record, press <Ctrl-N> for "New record." Note that you are now in the next available empty record.

To quit the program, press <F10>.

To use ENTER, first create a questionnaire as described in the previous chapter. This file should have the suffix ".QES". If this is done in the EPED program described above, you are ready to run ENTER.

If you used another word processor such as Word Perfect or Multimate, the questionnaire may have to be converted to a plain text or ASCII file, using a utility provided with the word processor. If your word processor adds its own suffix, you can rename the file from the DOS prompt as follows:

```
A>RENAME  OUTBREAK.WPS  OUTBREAK.QES
```

This changes the file name from OUTBREAK.WPS to OUTBREAK.QES.

#### RUNNING ENTER

On the main EPI menu, move the cursor bar to the ENTER program choice and press <Enter> or simply type "N" and press <Enter>. A screen will appear, asking for the name of the data file.

## MAKING A NEW DATA (.REC) FILE

The data file is constructed from your .QES file by the ENTER program. When you are asked for the data file name, give either the name of the file you wish to create or the name of an existing data file and then press <Enter>. If there is no data file by the name given, ENTER asks for the name of the questionnaire file so that a data file can be created. Usually it is convenient to use the same file name for the data and questionnaire files, calling them BOTULISM.REC and BOTULISM.QES, for example. There is no need to type the suffixes .REC and .QES because they are assumed by the program.

When the cursor is in either the "Data file" or "Questionnaire file" blanks, you may press <F1> to see the available files. To see those in another directory or on another disk, type the disk letter and/or directory and then press <F1>. Moving the cursor bar around with the arrow keys and then pressing <Enter> will choose a file from the list.

Questionnaire file names must end in .QES to be recognized. Suppose that you have created a questionnaire called OUTBREAK.QES. When you enter OUTBREAK in the data file blank, you will be asked for the questionnaire name, and you should enter OUTBREAK.

When you have specified both files, ENTER takes a few seconds to read the questionnaire and make a data file and then displays the questionnaire on the screen. You are now ready to enter data.

## ENTERING DATA

Entering data means typing the appropriate responses in the blanks on the screen. The cursor will jump automatically from blank to blank.

At the bottom of the screen the prompt line displays the available commands with the current record number at the right end of the line. Pressing <Ctrl-F> (for "Find") will display other commands used for searching and moving from record to record.

Just above the prompt line, the type of data that will be accepted in the field containing the cursor is displayed. These limits can be changed using the CHECK program described in a later chapter, but for now, they reflect only numeric, general purpose, date, and other field types.

The following commands have special meaning during data entry:

Backspace	Deletes the character to the left of the cursor.
Right Arrow Left Arrow	Moves one character right or left.
Up Arrow Down Arrow	Moves to the previous or next entry field.
<PgUp> <PgDn>	Moves the screen up or down a page.
<Home>	Moves to the first entry field in the questionnaire.
<End>	Moves to the end of the questionnaire.
<Ins>ert	Turns on and turns off the insert mode alternately. When insert mode is off, new characters will overwrite (replace) old characters. When it is on, new characters will be inserted between the old characters.
<Del>ete	Deletes a character at the cursor location.
<F10>	Closes files and exits from the program. Other function keys are described in the following sections and summarized at the end of this chapter.

When you enter data,

- Fields will accept only the type of data shown above the prompt line.
- Pressing <Enter> without any other entry signifies a missing value.
- When a field is full, the cursor moves automatically to the next field unless this feature has been turned off as described at the beginning of Chapter 35.
- Errors on entry are signaled by a "beep" and can be corrected immediately.
- Each blank is called a "field." Each copy of the questionnaire you complete is called a "Record." The records are stored in the file with the name ending in .REC, called either the data file or the .REC file.

After the information has been entered in the last field of a questionnaire, the question "Write data to disk (Y/N)?" appears at the bottom of the screen. Replying "Y" or "y" saves the record and brings up the next available empty record. Note that the record number on the lower right changes. If the reply is "N", the cursor jumps to the first field in the questionnaire and you have the opportunity to further edit the record.

To exit from the program at any time, use the F10 function key. If you haven't saved the current record, ENTER asks if you would like to and then returns to the EPI menu or to DOS.

#### EDITING AN EXISTING DATA FILE

When ENTER creates a data file, it stores the entire questionnaire in the file so that subsequent use of the file does not require the original questionnaire. The file name has the suffix ".REC". Epi Info will find such an existing file as soon as you type the name in the "Data file:" blank, and you will see the next available blank record from that file on the screen.

#### MOVING FROM RECORD TO RECORD

The current record number is always shown in the lower right corner of the screen. To move to the previous record, press the <F7 key>. <F8> will go to the next record in the file, if any.

If you have been "browsing" through the file with <F7> and <F8> and now wish to enter another record, hold down the <Ctrl> (Control) key and press "N" for "New." (You may need to press <Esc> first if a message appears as you encounter either end of the file.) The <Ctrl-N> will clear the screen and move to the next available new record for data entry.

## FINDING RECORDS THAT MATCH CRITERIA

ENTER allows you to search the data file and find records that match your criteria. To find a record, first be sure the current record has been saved, if necessary. Then press <Ctrl-F> for "Find," and type in the items you want to find, followed by <F1> for "Find 1st". You might type SMITH in a "Lastname" field, for example, and then press the <F1> key. The first SMITH in the file will be presented on the screen, if one exists. If you wish to find the next "SMITH" in the file, press <F4>, and the search starts from your current position in the file. If you wish to go backward, hit <F3>, and the search starts at your current position in the file and goes backward.

ENTER will search in this way using any combination of fields that you enter before issuing the search command--a multifield search on any combination of fields. You could have entered "Smith", "John", and "47" for the age. The search would then be confined to records containing 47-year-old John Smiths.

Sometimes it is useful to do a "fuzzy" search, looking for all names beginning with "Peter", for example, if you are not sure how "Petersen" or "Peterson" was spelled. You can use the "wildcard" character to specify the search, entering "Peter\*" in this case. Wildcards are only used in non-numeric fields.

If you have only a few hundred records, searches will be reasonably fast. Epi Info provides an indexing facility so that searches of thousands of records can be done almost instantaneously. Creating indexed fields is described in Chapter 10 on the CHECK program. From the user's point of view there is no difference in the commands for searching; indexing merely makes the search faster for large files.

A command called AUTOSEARCH is also described in Chapter 22. AUTOSEARCH performs a search for duplicate records as soon as you have entered a selected data item in the questionnaire. It might be set up to perform the search after you have entered first and last names, for example. A matching record already in the file would be called to your attention so that additions can be made or exact duplication avoided.

## EDITING RECORDS

To edit a previously entered record, first find and retrieve it as described above. Then change any of the items in the record, entering a new item with the aid of the arrow, delete, and insert keys. Be sure to press <Enter> when leaving a field that has been changed. When you have made all the desired changes, go to the last field in the record and press <Enter>, or merely press the END key. The question "Write data to disk (Y/N)?" will

appear at the bottom of the screen. If you reply "Y", the record as it now appears will replace the old record in the file. If you type "N," the cursor will jump to the first field in the questionnaire. If you move to another record with <F7> or <F8> without saving the edits you have made, they will be discarded and the edited record will revert to its previous form in the file.

Another method of editing and permanently changing records will be described in the chapter on the ANALYSIS program under the UPDATE command. This allows editing in a spreadsheet format that may be more convenient for changing the same field in many records, for example.

#### DELETING OR UNDELETING RECORDS

Pressing the <F6> key will mark the current record as deleted. An asterisk (\*) will appear next to the record number in the lower right corner of the screen. The data items in the record are still visible in the ENTER program, but tabulations done in the ANALYSIS program will skip this record. The WRITE RECFILE command, described later in the chapter on ANALYSIS, will create a copy of the data file without the deleted records and thus provides a permanent way to rid the file of such records, if desired.

A record, once deleted, may be made active again (undeleted) by pressing the <F6> key again. This key thus alternates (toggles) between deleting and undeleting.

#### REVISING THE STRUCTURE OF THE DATA FILE

You can revise a questionnaire (.QES file), using EPED, and then create a new data file in the revised format. This often occurs when you have entered a number of records and suddenly have an inspiration about adding a few questions or wish to correct an error in the questionnaire.

Revising the structure of the data file should be carefully distinguished from simply revising the data, as described above for the editing process. To revise the questionnaire structure, return to EPED and make appropriate changes in your original .QES questionnaire file. Then return to the ENTER program. Enter the name of the .REC file under "Data file:", but do not press <Enter>. Notice the instructions on the screen that say "F6 to revise questionnaire." Press <F6> and you will be asked for the name of the questionnaire file. Type the name of the revised .QES file and then press <Enter>.

ENTER will automatically make a .REC file in the new format, merge the old data into the new file, and then rename the old .REC file to .OLD. The entire process is described on the screen as it is performed.

The process works correctly only if the field names in the .REC file have not been changed. If you have used the CHECK program to change the field names in the .REC file, you will need to edit the questionnaire to produce these same field names, or the previous field names will be constructed from the questionnaire again.

#### PRINTING THE QUESTIONNAIRE AND THE DATA

To print a copy of the questionnaire or to list records from the file, press <F5>. ENTER asks if you want a copy of the questionnaire to be printed. If you do, enter 'Y'; otherwise, enter 'N'. The questionnaire will be printed with all entry blanks indicated by underline characters, suitable for pencil-and-paper entry.

You are then asked for the starting and ending record numbers. Be sure that your printer is ready, and enter the numbers of the first and last records to be printed, or zero for both if you want only the questionnaire. Answer "Y" to Single Sheets(Y/N) only if you want the printing to stop after every page for insertion of a new sheet of paper in the printer. After a short pause, ENTER sends the data to the printer.

Records are summarized in condensed form, for example:

```
Record number 60
AGE          :53      SEX          :F          TIMESUPPER      :1930          ILL              :Y
ONSETDATE    :04/18   ONSETTIME       :2330          BAKEDHAM         :Y
SPINACH      :Y       MASHEDPOTA     :Y          CABBAGESAL      :Y          JELLO           :Y
ROLLS        :N       BROWNBREAD     :Y          MILK             :N          COFFEE          :Y
WATER        :Y       CAKES          :Y          VANILLA         :Y          CHOCOLATE       :Y
FRUITSALAD  :N
```

Another way to list the records is provided in the analysis program, where the LIST \* command will print the data in column-and-row format. Either one may be used for various purposes such as hard-copy backup or for hand checking the data.

#### USING ENTER WHEN A CHECK FILE IS PRESENT

".CHK" files are explained in the chapter on the program called CHECK. They work with the ENTER program to check and edit data items while they are being entered.

As the cursor moves to each field in ENTER, the name of a variable appears at the bottom of the screen with a message giving the type of entry which will be accepted. If no .CHK file is present, this may be "All entries allowed," or "Integers allowed."

If a check file is present with the same name as the data file (e.g., OSWEGO.CHK for OSWEGO.REC), messages about legal values, ranges, and codes may appear for some of the fields. These messages should be self-explanatory; they indicate what type of entry will be accepted in each field. If the data item entered does not meet the criteria, the program will beep and request another entry. If you are absolutely sure that you want to enter something other than the required values, you can escape from a field by pressing the Down-Arrow key after entering the "illegal" value.

If the field you are entering is specified as a "jump" field in the CHECK program, the cursor may skip over parts of the questionnaire when you enter certain values, skipping the symptom questions if you indicate that no illness occurred, for example. This allows use of a single questionnaire for several purposes.

If a field contains codes set up through CHECK, entering a value may cause the corresponding member of a code pair to appear in another field, or even in the same field. Thus, entering 101 may cause "Hepatitis A" to appear, or "A" may result in "Hepatitis A".

In some fields, pressing <F9> will display a window of legal values or codes on the screen. One of these can be chosen by moving the cursor bar with the arrow keys and pressing <Enter>. Typing one or more letters will position the cursor bar on the closest match. Typing "J" will move the cursor to the first code beginning with "J," for example.

For further details on setting up .CHK files, please see the chapter on CHECK. These files are completely optional and offer a convenient way to set up constraints for data entry that will help to eliminate errors at this stage. CHK files can be set up or changed at any time, even though some records have already been entered. In the ENTER program, they only perform checks on data entered while the CHK file is present, not on previously entered records.

## SUMMARY OF SPECIAL FUNCTION KEYS

The special function keys are listed at the bottom of the screen. Those used for finding records become visible when you press <Ctrl-F>.

- <F1> Searches for a record with information matching the information found in all non-blank fields except <IDNUM> and <Today> fields. The search starts from the beginning of the file and goes forward. Partial matches ("SMI\*" is a partial match of "SMITH") can be specified on non-numeric fields by using the asterisk or "wildcard" character.
- <F2> Finds a particular record number. Enter a number, and the record with that number will be brought to the screen.
- <F3> Same as <F1>, except the search starts from the current file position and goes backward. Finds the previous occurrence of a record matching the search criteria.
- <F4> Same as <F1>, except the search starts from the current file position and goes forward. Finds the next occurrence of a record matching the search criteria.
- <F5> Allows you to print the questionnaire or the data stored in specified records. Prompts will appear at the bottom of the screen to guide you through the printing process.
- <F6> Marks the current record as deleted. Such records will be ignored in the ANALYSIS program. If the record has already been deleted, pressing this key will undelete (restore) the record to active status. Deleted records are indicated by an asterisk in the lower right corner of the screen beside the record number.
- <F7> Moves back to the record just before the one on the screen.
- <F8> Moves forward to the next record, if any.
- <F9> Displays codes, if any, that are set up for automatic coding in the current field. Choices may be made from the menu that appears by moving the bar cursor with the arrow keys and then pressing <Enter>. Typing a letter of the alphabet will move the cursor to the first choice beginning with that letter. A choice can also be typed after pressing <Esc> to remove the menu.
- <F10> Leaves the ENTER program and returns to the EPI menu or to DOS.

CHECK FEATURES COVERED ELSEWHERE

ENTER can be programmed to do many interesting and complex operations not discussed in this chapter, including mathematical and logical operations with more than one field, changing the colors on the screen, popping up help windows, and calling Pascal or Assembly language programs that act on the contents of fields during data entry. These functions are described in Chapters 22 and 23.

## CHAPTER 15

### STATCALC: A STATISTICAL CALCULATOR FOR SINGLE AND STRATIFIED TABLES

#### Tutorial

Run STATCALC from the EPI menu. Choose "Tables" by pressing <Enter>.

Enter four numbers, pressing <Enter> after each number. Then press <Enter> again or <F4> to perform calculations. Follow the screen instructions to see exact confidence limits.

Repeat the process, but this time press <F2> for another stratum after entering four numbers and seeing the results. Enter four numbers for stratum 2. Press <Enter> again or <F4> to see the calculations for both strata combined.

Enter 6, 8, or 10 numbers and then press <F4>. Press <F10> to return to the STATCALC menu.

Choose "Sample size." Experiment with each of the three types of sample size by entering numbers and performing calculations.

Choose "Chi Square for Trend." Enter data from the example in this chapter, pressing <F2> after each stratum and <F4> when you are ready to calculate the final results.

## OVERVIEW

STATCALC does statistics for data entered from the keyboard into tables on the screen. Facilities are provided to perform analyses of single and stratified 2-by-2 tables, to calculate sample sizes for studies involving proportions, and to analyze trends with or without stratification.

## RUNNING STATCALC

Run STATCALC from the main EPI menu. A menu will appear, containing the following choices:

Tables (2 x 2, 2 x n)  
Sample size  
Chi square for trend

To make a selection, move the cursor bar to the selection with the arrow keys and press <Enter> or press the letter key that is highlighted, such as T for the first choice. Help screens containing information about the program and the statistical methods are displayed by pressing <F1>. We recommend that you read these help screens carefully.

## ANALYSIS OF SINGLE AND STRATIFIED TABLES (2-by-2 to 2-by-9)

This choice presents a 2-by-2 table on the screen. To do statistics for a single 2-by-2 table, you simply enter four numbers in the table and then press <F4> to perform the calculations. STATCALC calculates the odds ratio and relative risk with confidence limits, three types of chi-square tests, and, if appropriate, Fisher's exact probability calculation. After these have appeared, you may choose to do exact confidence limits. This is most useful if the numbers in the table are small.

For 2-by-2 tables, additional strata may be entered by pressing <F2> after the statistics have appeared. After all strata have been entered, pressing <F4> will present the results of the stratified analysis and the opportunity to do exact confidence limits, if desired.

For single tables with up to 9 cells in the vertical dimension, simply continue entering numbers, followed by an extra carriage return. STATCALC will do a chi square for the 2-by-n table.

The statistics produced in STATCALC are similar to those done from data files by the ANALYSIS program, except that STATCALC does both Cornfield and exact confidence limits for single and for stratified tables. The confidence limits for relative risk in stratified and single tables are those of Greenland and Robins (Biometrics 1985; 41:55-68). Exact confidence limits based on the algorithm of Mehta, Patel, and Gray ( J Am Statistical Assoc 1985;80:969-973) are provided for use with tables in which the Cornfield confidence limits are not sufficiently accurate (sparse contingency tables). Further information on statistical methods is given in Chapter 32.

#### SAMPLE SIZE

If you choose "Sample size," a secondary menu is presented, with the following choices:

- Population survey
- Cohort or cross-sectional
- Unmatched case-control

All calculations are designed for studies in which the results are proportions expressed in percentages. In a population survey or descriptive study, finding the proportion of persons answering yes to a particular question is a typical goal. The program asks you to specify probabilities that your sample will predict the true situation in the population(s) being sampled and the amount of inaccuracy you are willing to tolerate, and then it calculates the sample size based on these assumptions.

These programs assume random sampling of the population, 100% participation, 100% accuracy in obtaining and recording answers, and other conditions that are seldom present in real life. If you expect 90% participation rather than 100%, you will need to multiply the sample size by 100/90 to obtain a more realistic estimate.

Calculated sample sizes are only a guide, and in many studies, the number of available cases or the number of interviews that can be done within the study budget is the final deciding factor. These programs provide a useful way to obtain estimates against which other factors can be weighed. In the long run the assumptions about what result you will obtain in the study will be replaced by actual results, and the assumed confidence levels will be replaced by p values or confidence limits derived from the results. Until such hindsight is available, however, sample size calculations provide a degree of foresight.

Operation of the three calculations proceeds by filling in a confidence level, power, and assumptions about how close an estimate of the actual proportion, relative risk, or odds ratio you would like. For the latter, choose the value furthest from the real population value that you are willing to tolerate. For example, in the Prospective or Cross-Sectional Study, suppose that you choose 10% as the hypothetical proportion of "Yes" answers to a question in the unexposed persons. You may wish to enter 20% as the closest value you would be able to distinguish in the exposed persons, meaning that any value of 20% or over would give a p value in the final study of .05 or less. This corresponds to a relative risk of 2.0 or an odds ratio of 0.5. Any of these three values may be entered initially to obtain the same results.

After assumptions have been entered, pressing <F4> will result in calculations. The result for the exact assumptions entered will be highlighted, and a series of other values based on varying the assumptions will also be presented. The results can be printed or sent to a file as described below.

#### CHI SQUARE FOR TREND

The test for trend is often used for dose-response studies and can also be used to test for trends with age, passage of time, or any ordered variable. The data can be stratified on other variables like age and sex to eliminate confounding from these variables.

The Extended Mantel-Haenszel chi square that is calculated reflects the departure of a linear trend from horizontal (i.e., no trend). If the associated p value is less than .05, there is 95% probability that a trend exists in the underlying population.

Data for analysis of trend by this method must be grouped according to an ordered numerical sequence. The simplest groups are arbitrary ones of 0, 1, 2, 3, etc. These are the "observations" or "scores." Other scores might be the mid-points of groups, such as 0 for 0 glasses of water, 2 for 1 to 3 glasses, 4 for 3 to 5, etc. These are the groups over which the presence of trend will be measured. Strata are determined by the confounding variables; common strata would be age or age-sex groups. The following table may be used as an example.

Frequency of Cigarette Smoking Among Women with of Myocardial  
Infarction (MI) and Controls (C), Stratified by Age

Stratum Group Cig./day	Score	1 25-29		2 30-34		3 35-39		4 40-44		5 45-49	
		MI	C								
None	0	1	131	0	188	3	161	11	169	23	157
1-24	1	1	101	6	152	12	130	21	134	42	97
>24	2	4	51	15	83	22	65	39	68	34	52

From Schlesselman; JJ. Case-Control Studies. Oxford Univ. Press, NY, 1982, p. 205.

In this example you would enter Stratum 1 on the first screen, then press <F2> and enter Stratum 2 on the second screen. After entering all five strata in this way, you would press <F4> to see the results.

PRINTING OR SENDING RESULTS TO A FILE

To send calculations for the current screen to the printer, press <F5>. If you prefer to have output go to a file rather than to the printer, press <F6> instead. The first time you press <F6> a prompt will request a file name. Once the file is open, subsequent <F5> commands will add material to the file rather than sending it to the printer. Pressing <F6> again closes the file. If you give the name of an existing file for the <F6> command, new material will be added to the existing file without harming its original contents.

RELATED TOPICS IN OTHER CHAPTERS

Formulas and references for the statistical procedures are given in Chapter 32.

## CHAPTER 18

### MERGE: MERGING AND UPDATING FILES AND RECORDS

#### Tutorial

##### Combining Data from Files with the Same Structure

Suppose that several people enter data from the same questionnaire on different computers, each with its own .REC file. All the .REC files were made from the same .QES file and have similar structures. The following technique will put all the records in the same file so that they can be analyzed together. To simulate the situation, we will use two copies of OSWEGO and pretend that they contain different records.

Run ANALYSIS from the Epi Menu.

Make a copy of OSWEGO.REC by typing

```
EPI>COPY OSWEGO.REC OSWEGO1.REC
```

Press <F10> to leave ANALYSIS.

Run MERGE from the EPI menu. Give the file names OSWEGO, OSWEGO1, and RESULTS. Choose option number 1, "Concatenate similar files," and then press <F4>. After the files have been concatenated, press <F10>, return to the EPI menu, and use ENTER to examine the records in RESULTS and verify that there are 150 records rather than the original 75 in OSWEGO.

Other functions of MERGE link related records in different files. See this chapter and Chapter 25 on the relational features of Epi Info for more information.

## SETTING UP EPED TO MAKE A QUESTIONNAIRE

Run EPED from the EPI menu. Now press <F2> for the file menu and find "Open File This Window." Move to this option with the arrow keys if necessary and press the Enter key. EPED will ask for the name of the file. Respond with a file name having eight or fewer letters, followed by ".QES". If the questionnaire already exists, it will be presented on the screen for editing. From the <F6> or Set menu, you should also set the right margin to 80 and turn word wrap OFF.

## INDICATING DATA ENTRY LOCATIONS IN THE QUESTIONNAIRE

When a questionnaire is being developed for use in ENTER, a few simple conventions are necessary to tell the program where to create data entry fields or variables and what kind of data to accept at these locations. ENTER will give names to these fields, and they will become the variables which can be manipulated in the ANALYSIS program. The field formats are described below and on the back cover of this manual.

There is a special command in EPED to make it easy to insert fields in your questionnaire. To see it in action, type <Ctrl-Q Q> (hold down the control key, type Q once, and then let go of the control key and type a single Q). It can also be accessed from the Text menu, but <Ctrl-Q Q> is easy to remember. A menu of field types like the ones on the back of the manual will appear on the screen. To insert one in your questionnaire, move the bar cursor to the choice and press <Enter>. For some field types, EPED may ask about the length of the field or the number of digits. After you have provided this information, the field will be inserted in the questionnaire at the current position of the cursor.

If you are using a word processor other than EPED or find it easier, simply type the proper letters for the field into your questionnaire without using the <Ctrl-Q Q> command. The field types are shown below, with information about their meaning.

File 1 is the main .REC file for the merge. File 2 is the .REC file containing records that are to be merged into or on top of those in file 1. NewFile is a .REC file to be created by MERGE that will contain the results of the merge.

#### CONCATENATING SIMILAR FILES TOP TO BOTTOM -- CHOICE 1

Two files that have the same format (came from the same questionnaire) can be concatenated, with one appended to the end of the other, as follows:

File 1	Name of one file
File 2	Name of the other file
New file	Name of a new file to be created
Merge Option	1
Match Fields	Leave blank
Y/N question	Answer "Y"

After you have answered the questions on the screen, press F4. The program will convert the file and then return to the merge screen. Press <F10> to return to the EPI menu or to DOS.

This option is used for combining files made from the same questionnaire but entered at different times or on different computers. This often happens during an outbreak, when several workers enter data to save time. It also occurs when a number of persons on a network submit periodic reports as separate files, and the results must be combined for further analysis.

#### MERGING ALL FIELDS OF DISSIMILAR FILES (SIDE TO SIDE) -- CHOICE 2

This command does the same thing as RELATING the two files in ANALYSIS and writing a new file with WRITE RECFILE. This is discussed in Chapter 25.

Suppose that information has been collected on individuals and on their households in a study. Different questionnaires were used for each purpose, and the result is a file called PERSON.REC, containing one record for each individual, and a file called HOUSE.REC, containing one record for each household. The investigator was careful to see that a field called "HOUSEID" appears on both questionnaires so that they can be linked together.

Now run MERGE as follows:

File 1	HOUSE
File 2	PERSON (The one with multiple records for each one in File 1)
New file	Name of a new file--HHPERS
Merge Option	2
Match Fields	HOUSEID (could be several fields, separated by spaces)
Y/N question	Answer "N"

For each person record that matches a household record number, the new file will contain a record combining the contents of both the PERSON record and the HOUSE record. These new person records can now be processed for any of the household variables. Processing information for households alone, however, can still be done with the HOUSE.REC file, which has not been disturbed by the merging process. Some questions that can be answered are:

How many households are there? The answer is given by any frequency or table using HOUSE.REC. Note that using the merged file would give the wrong answer, since household information has been replicated for each person in the study.

How many persons under the age of 10 are there? Selecting for AGE < 10 in the file PERSON.REC or the new file, HHPERS.REC, followed by a frequency or table, will give the answer.

How many persons under age 10 live in houses having fewer than five rooms? This can be answered only by using the new file, since neither of the original files contains the complete story. Using HHPERS.REC, first select for AGE < 10 and then for ROOMS <= 5, followed by FREQ AGE. Or the answer may be easily extracted from the table that is produced by TABLES AGE ROOMS.

Once you understand the method of combining files and replicating information in this way, you will be able to handle two or more levels of information such as HOUSEHOLDS, INDIVIDUALS, and multiple VISITS to individuals, using Epi Info files and MERGE. If such work is an objective, you should also read Chapter 25 on the relational features of Epi Info.

<IDNUM>  
<IDNUM >

This is a special-purpose field that maintains sequential identification numbers. The first number in the file will be number 1, and each succeeding record will have an IDNUM one higher than the previous record. This field is used to maintain unique identification numbers automatically.

Since IDNUM fields are filled automatically, the cursor skips them during data entry. If you wish to set the value of an <IDNUM> field to a number other than 1 in the first record of a file, use the up arrow key to enter the field and type the desired number before saving the record.

You should not use the underline, number (#) sign, or less-than or greater-than characters in the questionnaire except in the formats shown above, as the program will think you are trying to create an incorrect field. Specifically, < > and unpaired ">" or "<" signs are not allowed. The fields with special words such as <today/yy> must be exactly the length shown to contain the date or telephone number information correctly.

#### HELPING EPI INFO TO CREATE USEFUL VARIABLE NAMES

Unlike most data entry programs, Epi Info creates variable or field names automatically from the original questionnaire. When it encounters an underline or other special character, it looks for text on the same line which can be the basis for a variable name. The first 10 non-punctuation characters become the name. This name is used in the ANALYSIS program to refer to the data entered in this field. If the line is

Name \_\_\_\_\_

the field will be called NAME. If it is

What is your name? \_\_\_\_\_

the field will be called ISYOURNAME, since selected words like "what" are discarded automatically, along with spaces and punctuation. It is wise to put significant words within the first 10 characters of the text preceding the blank, or to mark them with curly brackets ( ) as described below.

If you answered "N" to the last question, then CASENO 1005 will not appear in REVCASES.REC.

This procedure provides a handy method for updating old records that might already have been sent to another location in a surveillance system, for example. Only the update information needs to be sent; the previous records can then be updated by the recipient using MERGE.

Note that if you wish to erase an item in File 1, it is not sufficient to send blank information in File 2. You will need to adopt a convention for missing data other than 0 or blank for this purpose, such as XXX or --- for an underline field and 99 for a numeric field. This information in File 2 will then replace the item in File 1.

#### TRANSFERRING A FILE INTO ANOTHER FORMAT (REVISE) -- CHOICE 4

In conducting an investigation, suppose that after entering a number of records you decide to revise the questionnaire. You therefore edit the original .QES file, adding several questions and perhaps subtracting several others.

The following steps are performed automatically by the ENTER program when you press <F6> to revise a questionnaire. In case you wish to do this within MERGE or understand better what happens, the steps are as follows.

First, create a .REC file from your new questionnaire by using ENTER and giving a new name for the data file. Instead of entering data, however, use <F10> to leave the program and save the (empty) new .REC file.

Now choose MERGE from the menu, and enter the following information:

File 1	Name of the new, empty file
File 2	Name of the other file, containing the data
New file	Name of a new file to be created
Merge Option	4
Match Fields	Leave blank
Y/N question	Answer "Y"

When you press <F4>, MERGE will combine the two files so that the revisions are included. Some kinds of revisions may work better if you reverse the order of the two files, and this might be worth trying if difficulties occur with the order suggested.

## RELATED TOPICS IN OTHER CHAPTERS

Revising file data structure by revising a questionnaire, creating a new data file, and merging existing data into the new data file is done automatically by the ENTER program, as described in Chapter 8. ENTER actually performs this function by calling the MERGE program, but the details are taken care of internally.

MERGE produces files that contain permanent results of merging. It merges two files "relationally" in a batch mode. Almost all of the operations that MERGE performs can be duplicated in a dynamic way by using the relational features of ENTER and ANALYSIS. Chapter 25 describes these in detail. Once files have been RELATED in ANALYSIS, they can be made permanent by using the WRITE RECFILE command, thus accomplishing the same thing as a merge.

**VALIDATE** -- Compares two Epi Info files entered by different operators and reports any differences.

**Help files** -- Much of the material in the reference section of this manual is available in help files accessed by pressing <F1> from within Epi Info programs.

**Sample Programs** -- Two complete sample surveillance systems, files from two epidemic investigations and a system for anthropometric calculations in nutritional surveys, are provided as examples of Epi Info use. There are numerous other small data files, and many programming examples are provided in the manual.

**Tutorials** -- Complete interactive tutorials are provided to introduce the features of EPED and ANALYSIS.

Epi Info, Version 5, requires an IBM-compatible microcomputer running the PC-DOS or MS-DOS operating system (Version 2.0 or higher), 512 K (524,288) bytes of random-access memory (RAM), and at least one floppy disk drive. A graphics adapter (board) is necessary to produce graphs. Adapters using the Hercules or IBM monochrome, EGA, VGA, IBM3270, IBM8514, and AT&T standards are supported. An IBM- or Epson-compatible printer or Hewlett Packard-compatible laser printer can be used to print results. For most effective use of Epi Info with large data sets, a hard-disk drive, 640 K of memory, and a color graphics board and monitor are recommended.

Data files can consist of as many records as DOS and your disk storage can handle (up to four billion). A questionnaire can have up to 500 lines or approximately 20 screens. The number of variables is not limited, except that they must fit in the 500 lines. The maximum length for a text variable is 80 characters. The total length of variables in one file must not exceed 2048 characters. Given the ability to link several related files, this should not be a practical problem, since you can spread data over several files and link them during analysis. The programs in Epi Info require about one megabyte of disk space, but ANALYSIS, the largest program, will file on one 362 K disk.

Versions 1-3 of Epi Info were produced in the Epidemiology Program Office of the Centers for Disease Control (CDC). Version 4 was produced and distributed on an experimental basis by the Surveillance, Forecasting, and Impact Assessment Unit, Global Programme on AIDS, World Health Organization (WHO), Geneva, Switzerland, in collaboration with CDC. Version 5, the current version, is a joint CDC/WHO effort.

The programs are made available by WHO and CDC and are not copyrighted. Making copies for others is permitted and encouraged. Please provide copies of the manual if you distribute the programs.

ILL	Minimum	25%ile	Median	75%ile	Maximum	Mode
+	3.000	17.000	38.500	59.000	77.000	15.000
-	7.000	14.000	35.000	50.000	69.000	11.000

#### ANOVA

(For normally distributed data only)

The p value is equivalent to that for the Student's T Test, since there are only 2 samples.

Variation	SS	df	MS	F statistic	p-value
Between	712.655	1	712.655	1.560	0.213137
Within	33340.732	73	456.722		
Total	34053.387	74			

Bartlett's test for homogeneity of variance

Bartlett's chi square = 0.119 deg freedom = 1 p-value = 0.729786

The variances are homogeneous with 95% confidence.

If samples are also normally distributed, ANOVA results can be used.

Mann-Whitney or Wilcoxon Two-Sample Test (Kruskal-Wallis test for two groups)

Kruskal-Wallis H (equivalent to Chi square) = 1.161  
 Degrees of freedom = 1  
 p value = 0.281226

At first this may be more numbers than you ever wanted, but the overall results can be quickly understood. The means of the two sets of ages are 39 and 33. Now look at the p values under the ANOVA and Kruskal-Wallis tests. Since both are more than 0.05, we can conclude that the difference in mean age between the ILL and NOT ILL groups is not "significant." Bartlett's test helps you decide which one of the two methods to choose, but in this case, both tests lead to the same conclusion. Further details are given in Chapter 32 on Statistics.

Note on missing values: In the MEANS, TABLES, and FREQ procedures, missing values will be ignored if they were entered by pressing <Enter> in the ENTER program and if SET IGNORE is ON (normally it is; see the SET command in Reference Chapter 36). If, however, you have used another code, such as 99, for missing values, be sure to select only the non-missing values before using the means procedure. This can be done by using SELECT AGE <> 99, for example. Be particularly aware of this point if the data have been imported from another system in which missing values may be coded differently.

## TITLES

The TITLE command allows you to specify up to five lines of text that will appear at the top of a table, frequency, chart, or graph. TITLE can be used before the TABLES, FREQ, and graphics commands to produce an appropriate title for the results. To cancel the effect after the command has been executed, TITLE can be used again with blank entries or new text for the next command. The following commands will define a two-line title for subsequent commands:

```
TITLE 1 Number of Cases by Date of Onset
TITLE 2 North Chicago, January 1990
```

To remove this title for later commands use:

```
TITLE 1 (All titles with numbers 1 or greater are removed.)
```

## CHARTS AND GRAPHS

ANALYSIS produces histograms, scatter plots, pie charts, and bar and line graphs directly from data files. Making a graph requires a single command; the following examples will give the idea:

```
HISTOGRAM ONSETDATE
PIE RACE
BAR AGEGROUP
LINE DAY
SCATTER SYSBP BODYMASS
```

If /R is placed after the SCATTER command, a least squares regression line will be drawn through the data points, as in:

```
SCATTER SYSBP BODYMASS /R
```

Sometimes it is important to compare two graphs and to make the scales on the vertical axes the same for this purpose. Adding a slash (forward slash, not backslash) to the command and specifying the maximum value of the Y axis will accomplish this, as in:

```
EPI>Bar ROLLS /Y=0-75
```

(The lower value must be 0.) For scatter plots, the maximum and minimum numbers on the X (horizontal) and Y (vertical) axes can be made more extreme by specifying new numbers:

```
EPI>Scatter age BP /X=0-300 /Y=0-100
```

The minimum cannot be increased nor the maximum decreased from the values chosen automatically.

#### SENDING RESULTS TO THE PRINTER OR TO A FILE

Before using the print commands, be sure that a printer is connected to the computer, turned on, and supplied with paper. Pressing the Printer On key <F5> will send all subsequent output to the printer. It does not print results already on the screen. The best way to do this is to turn the printer on with <F5> and then press the up arrow to recall the previous command. Pressing <Enter> will run the command again, and this time the results will be printed. To turn the printer off, press <F5> again.

Note that when you press <F5>, the command ROUTE PRINTER or ROUTE SCREEN appears. This key provides a shortcut to these commands, but you can also type them on the command line to control printing.

The ROUTE command will also direct the results to a file by the name of your choice.

```
EPI>ROUTE C:RESULTS.TXT
```

will send all results to the file RESULTS.TXT on drive C, until another ROUTE command sends results to the SCREEN, PRINTER, or another file or files are closed by leaving ANALYSIS with <F10>.

When a graph or chart is requested as in the preceding section and printing is on, the graph will be printed; otherwise, it will appear only on the screen. Printing of graphs and charts can be done on Hewlett-Packard Laserjet compatible printers and plotters and on IBM and Epson compatible dot-matrix printers, although other printers will usually work for tables and lists. Graphs may print with other printers but graphics are more subject to incompatibility problems than plain text. See the section of Chapter 36 on the SET command for methods of setting up ANALYSIS for laser printers and plotters.

#### RECALLING ANALYSIS COMMANDS

As commands are entered, ANALYSIS saves each line for possible reuse. When the cursor is at the EPI> prompt, pressing the <up arrow> key will bring back previous commands and allow you to correct errors or change the commands.

Once a line has been recalled, it may be edited. When the <Enter> key is pressed, the command will be executed just as though it had just been typed.

Previous commands are stored in a space called the Command Stack. ANALYSIS normally saves the last 20 commands, but this can be increased by resetting the size of the Command Stack. This is described in the reference chapter for ANALYSIS under the SET command.

#### RELATED TOPICS IN OTHER CHAPTERS

The commands we have described so far will produce results from a questionnaire, but we have not yet discussed how to omit selected records from the ANALYSIS, how to group or recode data items, take action only if certain conditions are met, reorganize and write new files, and many other operations that are essential to complete analysis of a data set. As commands become more complex, it is useful to save them for repeated use in a program file. These and other procedures are discussed in Chapter 11, which should be consulted as soon as you have mastered the elementary operations in ANALYSIS and need to begin serious data analysis.

ANALYSIS also contains features useful in setting up permanent systems such as those used in U.S. state health departments for maintaining communicable disease reports. Two examples of such systems are included on the Epi Info disks and described in Chapters 14 and 26. Commands of interest to those programming such systems are described in Chapters 13, 22, 24, 25, and 33-36.

## CHAPTER 10

### THE CHECK PROGRAM: OPTIONAL ERROR CHECKING, CODING AND SKIP PATTERNS DURING DATA ENTRY

#### Tutorial

Run the CHECK program from the EPI menu.

Enter the file name OSWEGO and respond that you are "Ready" by typing "Y" and pressing <Enter>.

You will see the OSWEGO questionnaire on a screen that looks like the ENTER screen but that has a different series of prompts at the bottom.

Enter "1" in the AGE field. Now press <F1> to make this the minimum age allowed in the field. Note the change in the line just above the function key prompts. Now enter "90" in this field and press <F2> to define the maximum of the range.

Move the cursor to the field labelled SEX. Enter "M" and press <F6> to make this a "legal value." Do the same for "F" and for "U" (Unknown). Press <F4> to make this a "Must Enter" field. The result is that only M, F, and U will be accepted in this field, and a missing value is not allowed, since we have decided to use U for this purpose.

Now let's install a conditional jump or skip pattern so that date and time of onset are skipped if there is no illness. Place the cursor in the ILL field, enter "N", and press <F7>. Following the instructions on the screen move the cursor to BAKED HAM and press <F7> again. The cursor will now jump automatically to BAKED HAM when N is entered for ILL. If Y or blank is entered, the cursor will proceed normally to Date of Onset.

Press <F10> and affirm that you wish to save the results. The EPI menu will appear.

Run the ENTER program with the file OSWEGO. Enter a record to verify that the range, legal values, and jump work properly. Do not save the results, but press <F10> and "N" to return to the menu.

Other features of CHECK that were not demonstrated include automatic coding, repeat fields, file indexing, and changing the names of variables. These are described in this chapter and in Chapters 22 and 23.

## OVERVIEW

Often it is useful to have the computer check for errors during the data entry process, to do automatic coding of entries, and to skip over parts of the questionnaire if certain conditions are met. The CHECK program makes it possible to instruct ENTER to perform such operations automatically. By using CHECK and investing a little extra time in setting up conditions, you can protect your data against many common types of error and also make data entry easier and more automatic. This is particularly helpful if several people or those with minimal computer experience will be entering the data.

CHECK makes a file with a name ending in .CHK. The .CHK file contains instructions for ENTER to restrict the data entered in specified fields. When ENTER is run, it automatically looks for a file with the same name as the .REC file but ending in .CHK. Using CHECK is optional, and the ENTER program will function just as well if you decide not to make a .CHK file.

## RUNNING CHECK

Before running CHECK, a .REC file must already exist for the questionnaire you wish to enhance. Use ENTER to make a .REC file from the .QES file if necessary before running CHECK. Do not enter any data items, but exit immediately from ENTER with <F10> to make an empty .REC file; then run CHECK from the main menu. Enter the name of the .REC file in the data file space in CHECK and answer "Y" in the "Ready?" blank.

CHECK presents the questionnaire on the screen with the following function keys indicated on the bottom line:

F1/F2-Min/Max    F5-Link fields    Legal: F6-Add Shift-F6-Display Ctrl-F6-Delete  
F3-Repeat        F9-Edit field     Jump : F7-Add Shift-F7-Display Ctrl-F7-Delete  
F4-MustEnter    F10-Quit         Codes: F8-Add Shift-F8-Display Ctrl-F8-Delete

Information on the status of functions set up for the field in which the cursor rests will be displayed just above this series of function-key prompts.

The function keys act on the field in which the cursor is currently located. Pressing a function key once sets the condition, and in some cases, pressing it a second time will remove the condition. Thus, if you wish to make CITY a Repeat field, as described below, place the cursor in CITY, and press the <F3> key. To remove the Repeat condition, press <F3> again. The condition currently in effect is shown on the line above the function-key information. Several different conditions may be entered for one field.

#### RANGE CHECKING (MIN/MAX) <F1><F2>

When a minimum and maximum are specified, the value entered in a field must be greater than or equal to a specified lower bound and less than or equal to a given upper bound. Any field type may have a range, although the meaning is somewhat different if the field is non-numeric. In this case, the name must lie within the two bounds when sorted in alphabetic order.

To enter a lower or upper range, enter a value in the appropriate field and then press <F1> or <F2>. Additional values can be entered as "Legal", using the <F6> key so that "1-3, 9" can be the final result.

#### Examples:

Lower range 1, Upper range 3, will accept only values of 1, 2, or 3.  
Lower range Smith, Upper range Thompkins, will accept "Smith",  
"Smithson", and "Thomas", but reject "Thompson".

#### LEGAL VALUES <F6>

Legal values are those that will be accepted as entries in a field. When a list of legal values is specified, an entry must match one of a specified list of values, or it will be rejected. Pressing <Enter> for a missing (blank) value is accepted unless "Must Enter" is also specified as discussed below.

To enter a legal value, type the value in the appropriate field and press <F6>. The value will be added to the list of legal values. To remove an existing legal value, place the value in the field and then press <Ctrl-F6>, i.e., hold down the <Ctrl> key and press <F6>. Pressing <Shift-F6> will display current legal values for the field if there are more than will fit on the line at the bottom of the screen.

#### MUST ENTER <F4>

When Must Enter is specified, missing (blank) values are not allowed. The value entered must be another valid entry.

To specify Must Enter for a field, place the cursor in the proper field and press <F4>. Pressing <F4> again removes the Must-Enter condition.

#### REPEAT <F3>

During entry, the value of the field will be set automatically to the value in the last record entered or displayed. This is useful when a series of questionnaires all contain the same name, for city or state, for example. The few that do not can be changed during entry.

To designate a repeat field, place the cursor in this field and press <F3>. Pressing <F3> again removes the Repeat condition.

#### CONDITIONAL JUMPS (SKIP PATTERNS) <F7>

Frequently questionnaires contain sections that are to be skipped, depending on the answer to a prior question. If no illness occurred, for example, the cursor should jump over the section on symptoms and proceed to the next useful item. It is also useful at times to have the cursor jump to an entirely different part of the questionnaire containing special-purpose questions. If "Rabies" is entered for DISEASE, the cursor could jump to a series of questions specific for animal rabies, and then after the last question, jump back to the question following DISEASE.

To set up conditional jumps, enter a value in the appropriate field, e.g., N for ILL, and press <F7>. You will be asked to indicate the destination field for the jump. Put the cursor in this field, and press <F7> again. In the example given if ILL is entered as N, the ENTER program will automatically jump to the designated field, skipping over questions pertaining to illness. Any other answer, such as Y, will cause the cursor to move on to the next field as usual.

You can specify several values with different conditional jumps for a field if desired. To specify a jump regardless of the entry in the field, press <F7> while the field is blank. Such unconditional jumps are useful for changing the order in which fields are entered on a questionnaire.

To jump to one field on "1" and to another on all other entries, you would create one general jump with a blank field and then a different jump for the "1" value. Entry of values for which a jump is not specified will, of course, cause the cursor to go to the next field as usual.

To display all current jumps for a field, press <Shift-F7>. A jump may be removed with <Ctrl-F7>.

#### LINKING OF FIELDS AND AUTOMATIC CODING <F5> AND <F6>

If automatic coding is specified, an entry will be compared with a table consisting of pairs of codes, words, or phrases, and converted to the value of the other member of the pair. The result will be placed in the same field, or, if another field has been specified with the link command, in the linked field.

Suppose that you link two fields called DISEASE and DISEASECOD and set up codes for DISEASE. When you enter "Hepatitis B" in a field called DISEASE, the corresponding code (say, 10) would be found automatically and placed in a field called DISEASECOD. If you enter 10 in DISEASECOD, "Hepatitis B" will be placed in DISEASE.

Unless you wish the entries to be recoded to the same field, you must set up the Link between fields before setting up the coding. To do this, place the cursor in the first field and press <F5>. Then following the directions on the screen, move the cursor to the field to be linked and press <F5> again. The fields are now linked and you can proceed with setting up the codes.

Suppose again that there fields for DISEASE and DISEASECOD and these have been linked as described above. When you enter a value in DISEASE, like "Hep A", and then press <F8>, the cursor will jump to DISEASECOD, the linked field. You then enter the code, in this case 10, and press <F8> again. The cursor will jump back to DISEASE, and you are ready to enter as many other codes as you wish, using the same method.

Coding restricts the entries allowed, as though the values had been designated LEGAL. The ENTER program treats upper and lower case alike in coded fields so that "Hepatitis B", "HEPATITIS B", and "hepatitis b" are all treated as the same value for DISEASE.

During data entry, if you must enter a value which is not a valid code or legal value, this can be done using the down-arrow key instead of <Enter> to record the value. You can use the coding facility to accept several synonyms and recode them to a single uniform entry. To do this, OMIT the linking step, thereby linking the field to itself. Then enter the synonyms, such as "Hepatitis B", "Hep B", and "B", giving "HEPATITIS B" as the code for each. This will have the effect of coding all three entries to the same final data item.

Codes can be removed by pressing <Ctrl-F8> with the code to be removed and the cursor in one of the two relevant fields. <Shift-F8> will display all codes for a field.

#### THE EDIT FIELD COMMAND <F9>

The .CHK file contains a series of instructions to the ENTER program. The Edit Field function allows you to see and edit the commands that are inserted automatically in the file by CHECK. It confines the view to commands for the current field--the one containing the cursor.

Try entering one or more CHECK commands and then examine the CHECK file by pressing <F9> with the cursor in the same field. In the example given above for the coding function, if there are many disease codes, you may prefer to use the Edit function to type them directly into the .CHK file.

To do this, press <F9> with the cursor in DISEASE. An editing window appears with the portion of the .CHK file for DISEASE displayed. You will note code pairs in the section called "Codes", terminated by the word END. The first would be "HEP A" 10. You may type in other code pairs, taking care to put quotation marks around any codes that contain spaces. To leave the editing window, press <ESC>.

Editing the file directly is more likely to result in errors than working through the automatic coding process, but if you are careful not to disrupt the structure of the .CHK file, it should be more efficient for large code sets like all the counties in a state or a series of diseases. Several hundred codes can be handled this way, if desired.

#### CHANGING VARIABLE NAMES

The variable names that Epi Info constructs automatically from text preceding the entry blank are not always the best representation of the meaning of the variable. These names can be changed in the data file from the Editor in the CHECK program. To change a variable name, place the

cursor in the corresponding field and press <F9>. Then press <F1> for "Change name" as indicated at the bottom of the screen and enter the new name for the field. The variable name will be changed in both the .CHK file and the .REC file.

#### COMMENT LEGAL CODES

Sometimes it is useful to be able to display codes with comments on their meaning, but without automatic recoding. Using the Edit Field function, a LEGAL statement may be changed to COMMENT LEGAL to display descriptions of legal values. Suppose that a LEGAL statement appears in the .CHK file as:

```
SEX
  Legal
    M
    F
    U
  END
END
```

Pressing <F9> from the main CHECK screen when the cursor is in the SEX field will display these statements.

Add the word COMMENT and the comments or labels, as follows:

```
SEX
  Comment Legal
    M Male
    F Female
    U Unknown
  END
END
```

Now the ENTER program will display the codes in a pop-up window if the <F9> or SEE CODES key is pressed during data entry. This is useful if there are a number of codes and you do not wish to have the longer versions actually entered into the data file, as would occur if you set up codes in CHECK.

The indentation to the left of the items in .CHK files is helpful to understanding the file, but is ignored by the ENTER program and is optional if you wish to enter more items.

<Ctrl-N> Clears all the fields in the current record and sets the record position to point to the next New record to be added to the data file. Use this command if you have finished editing records and wish to enter new records, or to leave the search (<Ctrl-F>) mode.

<Ctrl-F> Prepares the screen to Find a record.

#### RELATED TOPICS IN OTHER CHAPTERS

Example Chapters 14, 23, and 26 describe questionnaires on the Epi Info disks that can serve as models for more advanced data entry functions. The CHECK program described in Chapters 10 and 22 can be used to add many features to the data entry process such as range checking, legal values, skip patterns, automatic coding, help windows, crossfield checking, mathematical and logical functions, and even linking in programs written in other languages. The results of such programming can be seen in the example chapters.

## CHAPTER 9

### ANALYSIS: PRODUCING LISTS, FREQUENCIES, TABLES, STATISTICS, AND GRAPHS FROM EPI INFO FILES

#### Tutorial

Run the ANALYSIS program from the main menu.

When the prompt "EPI>" appears, type:

```
RUN TUTOR1
```

Enjoy the tutorial.

To leave the program at any time, press <F10>

#### OVERVIEW

ANALYSIS produces lists, frequencies, tables, statistics, and graphs from Epi Info or dBASE files. Simple commands will cause ANALYSIS to select records using specified criteria, sort or list records, do frequencies or cross tabulations, do logical or mathematical operations on a variable, put the results in a new variable, and direct the results to the screen, a printer, or a disk file.

The commands form a mini programming language, and they may be entered one by one from the keyboard or placed in a program file that is then "run" from ANALYSIS. In this chapter, we will describe only the simpler operations--those that might be used during a field investigation for preliminary analysis of a questionnaire. The writing of programs to "clean" data sets and for permanent disease surveillance systems is described in later chapters.

## GETTING ACQUAINTED

By running ANALYSIS from the EPI menu, you will see a lower window for entering commands and a larger one above where the results of the commands will appear. At the bottom of the screen some key commands are indicated, and information about data files and available memory is displayed at the top.

Whenever the cursor is at the EPI> prompt, commands may be entered from the keyboard. If you make a mistake or want to change part of the line being entered, use the <Backspace> or <Left arrow> key to return to the correct position and then enter the new material. The <Del> and <Ins> keys work as in EPED, deleting characters and turning the Insert mode on and off. The <Home> and <End> keys skip to the beginning or end of the current line.

The function keys at the bottom of the screen allow selecting help topics, commands, and variable names from lists that appear on the screen when the key is pressed. After pressing one of the function keys, you choose an item by moving the highlight bar with the up- and down-arrow keys and pressing <Enter>.

As results appear in the upper window, those that have scrolled out of sight at the top of the screen may be reviewed by pressing the <PgUp> and <PgDn> keys to move up and down one screen at a time. For finer movements, <Ctrl-PgUp> and <Ctrl-PgDn> will move one line at a time.

We suggest that you run the tutorial as instructed above. It covers most of the material in this chapter interactively. Chapters 11, 13, 25, and many other chapters contain further information about ANALYSIS. For particular commands, the reference section in Chapter 36 provides much more detail about additional uses and options than can be given in this introductory chapter. This chapter describes the basic commands used to analyze a questionnaire and produce line lists, frequencies, cross tabulations, means, and graphs.

## THE HELP SYSTEM

To see a list of the commands in ANALYSIS, press <F2>, the Commands key. A list of commands will appear in a window on the screen. Moving the cursor to a command and pressing <Enter> will place that command next to the prompt, although, of course, you may also type commands from the keyboard.

Whenever a command has been entered--READ, for example--pressing <F1> will bring up further information about that command, including the format in

which field names and other information must be supplied. If no command has been entered, <F1> brings up a list of topics; choosing a topic then produces information on that topic.

For many commands, such as READ, typing the command and then <Enter> will bring up a menu of appropriate choices--in this case, the files that can be read. Choosing a file then carries out the command with the appropriate file or variable.

To summarize, help is never more than one keystroke away:

- For general information on a topic, press <F1> without a command on the EPI> line, and then choose a topic from the list.
- For command information, press <F2> for commands.
- For more information about a command, enter the command and then press <F1>.
- For file information, type READ <Enter>, and then choose the file from the menu that appears.
- For a list of variables in the current file, press <F3> for variables and choose from the list. Several variables can be "tagged" with the <+> key or "untagged" with the <-> key before pressing <Enter> to finalize the choice(s).

The help system is intelligent enough to change the order of words in a command; if you choose a variable and then a command from the menus, their order is automatically reversed.

#### THE FIRST STEP--READING A FILE

Analysis must be performed on the records in a file. The file may be either an Epi Info file, produced by entering data with the ENTER program, or a dBASE file from another source. Epi Info files can be produced from many other formats by the IMPORT program (Chapter 17), and the resulting file can be used in ANALYSIS.

The command that tells ANALYSIS what file to use is READ <file name>, and this is usually the first command given in ANALYSIS.

To see a list of available files, type

EPI> READ

and press <Enter>. To see the files in another directory, such as c:\Data\, you would first type READ c:\Data\ before pressing <Enter>. To see all dBASE files, type READ \*.dbf. A directory of files will appear in a window. Move the cursor bar with the arrow keys and choose a file by pressing <Enter>.

Whether you use the file directory or simply type READ and the file name, ANALYSIS will use this file for all subsequent operations until another READ is performed. It has become the "active data set."

ANALYSIS will also read and process files made in the dBASE format. To use a dBASE file rather than an Epi Info file, simply type the file name followed by .DBF. The only operation in ANALYSIS that cannot be performed with a dBASE file is UPDATEing of records, described in a later section.

Allowable READ commands include the following:

```
READ OSWEGO                (OSWEGO.REC is understood)
READ OSWEGO.REC
READ OSWEGO.DBF           (A dBASE file)
READ OSWEGO1 OSWEGO2 OSWEGO3
    (Three files with similar formats that you wish to analyze
    as a single unit)
```

For the rest of this chapter, we will assume that you have chosen the file OSWEGO.REC, supplied with the Epi Info system, by giving the command:

```
EPI>READ OSWEGO
```

OSWEGO.REC contains interview information from a foodborne outbreak often used as a teaching exercise. The file contains 75 questionnaires obtained from attendees at a church supper that was followed by a gastroenteritis outbreak.

#### PRODUCING A LINE LISTING

To produce a listing of the records in the file, type:

```
EPI>LIST
```

The result is:

REC	AGE	SEX	TIMESUPPER	ILL	ONSETDATE	ONSETTIME	BAKEDHAM	SPINACH	MASHEDPOTA
1	11	M	9999	N		0	N	N	N
2	52	F	1500	Y	04/19	30	Y	Y	Y
					---	ETC.	---		



CROSS TABULATIONS

The TABLES command will count the records in which the values fulfill criteria for two fields at the same time. Thus, the command

EPI>TABLES SEX ILL

will result in:

SEX	ILL		Total
	+	-	
F	30	14	44
M	16	15	31
Total	46	29	75

Single Table Analysis

Odds ratio  
Cornfield 95% confidence limits for OR            0.70 < OR < 5.85            2.01

Relative risk of (ILL=+) for (SEX=F)            1.32  
Greenland, Robins 95% conf. limits for RR    0.89 < RR < 1.96  
(Biometrics 1985;41:55-68)

Ignore relative risk if case control study.

	Chi-Squares	P-values
Uncorrected:	2.11	0.14679237
Mantel-Haenszel:	2.08	0.14951418
Yates corrected:	1.46	0.22620244

Note that the interpretation of relative risk (risk ratio) depends on the orientation of the table, and that not all relative risks will be meaningful. The relative risk given is the rate of the first factor on the first line compared with the rate of the first factor on the second line or  $(a/a+b)/(c/c+d)$  as defined in Chapter 32. An interpretation is printed with the relative risk value. If risk factors are on the left side of the table and disease across the top, with presence indicated first in each case, the relative risk represents the risk of disease for persons with the first factor relative to those with the second factor. As indicated in the note, relative risk should be ignored in a case-control study, since it cannot be interpreted meaningfully.

STRATIFIED ANALYSES

Series of stratified tables can be produced by listing more than two variables in the TABLES command. The variables after the first two serve as the basis for dividing the tables into levels or strata, one for each combination of variables after the second. For a quick preview, try

EPI>TABLES VANILLA ILL SEX  
(Tables of VANILLA consumption by ILLness, stratified by SEX)

This will result in separate tables for male and female cases. A Mantel-Haenszel stratified analysis will follow the two tables. The summary of the statistical combination of the two strata appears as follows for VANILLA ILL SEX:

\*\* Summary of 2 Tables With Non-Zero Margins \*\*  
N for sets w/ discordant results = 75

SUMMARY ODDS RATIOS

Crude OR for sets with discordant results	23.45
Mantel-Haenszel Weighted Odds Ratio	37.64
95% confidence limits for M-H OR	7.19 < OR < 197.01
(Robins, Greenland, Breslow AJE 1986; 124:719-23)	

SUMMARY RELATIVE RISKS  
(Ignore if Case-Control Study)

Crude RR	5.57
Summary RR of (ILL=+) for (Vanilla=+)	5.73
95% confidence limits for summary RR	1.99 < RR < 16.51
(Greenland, Robins Biometrics 1985;41:55-68)	

M-H Summary Chi Square	25.28
P value	0.00000050

Unable to do Woolf's Test due to zero cell values  
\*\* End of Stratified Analysis \*\*

The very large odds ratio indicates a strong association between eating vanilla ice cream and the occurrence of illness. The Mantel-Haenszel weighted odds ratio, summarizing the results of stratification by sex, is even larger. Both the confidence limits and the extremely small p value indicate that the Mantel-Haenszel weighted odds ratio is significantly different from 1.0, the value indicating no association.

Both the crude and stratified relative risks are also significantly greater than 1.0. In this outbreak of staphylococcal food poisoning in Oswego, New York, the vanilla ice cream was indeed identified as the cause, through both statistical and microbiologic evidence.

An asterisk or star(\*) may be used in the TABLES command to indicate all variables. Thus, TABLES ILL \* would produce tables of illness status against each variable in the questionnaire. This is a quick way to obtain preliminary analysis of a questionnaire for which comparison of two groups is the key feature.

Woolf's test, when available, indicates differences in the odds ratios between strata, if the p value is small. Such differences suggest that the Mantel-Haenszel weighted odds ratio, while still valid as a summary of all strata, does not necessarily reflect the value for one of the strata.

#### THE MEANS COMMAND

The TABLES command is for data items that are arranged in categories and counted. There is another whole world of statistics for numbers that are continuous, such as height, weight, and age. The MEANS command produces a table that displays continuous or ordinal data and then performs appropriate statistics.

The MEANS command requires two items of information--the variable containing data to be analyzed and the variable that indicates how groups will be distinguished. The command is:

MEANS [Numeric variable to be analyzed] [Variable for Grouping]  
If you prefer not to display the table of values, append "/N" to the command to indicate "No tables."

Using the OSWEGO.REC file, MEANS AGE ILL will compare ages for persons in whom ILL ="Y" with those for whom ILL = "N".

Using the MEANS AGE ILL example, you should obtain the following results:

ILL	Obs	Total	Mean	Variance	Std Dev
+	46	1806	39.261	477.264	21.846
-	29	955	32.931	423.709	20.584
Difference			6.330		

# **LOS SERVICIOS DE SALUD EN LAS AMERICAS. ANALISIS DE INDICADORES BASICOS**

La preparación de esta publicación estuvo a cargo del Programa de Desarrollo de Servicios de Salud de la OPS. Fueron responsables: del contenido del documento el Dr. Carlos Castillo S., asesor; de la estructura el Dr. José María Paganini, Coordinador del Programa, y del diseño de la base de datos y la edición la Lic. Lucía Ruggiero. Se seleccionaron apartes del artículo Características de la Atención Médica en América Latina, por el Dr. José Quíñones. En la elaboración de los cuadros sobre el "Índice de necesidad en salud", se contó con valiosos comentarios y recomendaciones del Dr. Bahjat Oaqlish, del Departamento de Estadística de la Escuela Johns Hopkins de Salud Pública.

Cuademo Técnico No. 14

**ORGANIZACION PANAMERICANA DE LA SALUD**  
Oficina Sanitaria Panamericana, Oficina Regional de la  
**ORGANIZACION MUNDIAL DE LA SALUD**

525 Twenty-third Street, NW  
Washington, DC, 20037, EUA

1988

## 2.6 INDICE DE NECESIDAD EN SALUD

Durante los apartados anteriores, se presentó una descripción general de las tendencias que a partir de 1957 han tenido los cinco indicadores básicos de los servicios de salud de las Américas. Para complementar esta descripción y contribuir a una visión más amplia de las necesidades actuales de salud de los países de las Américas se incluye los resultados obtenidos al elaborar un "Índice de Necesidad en Salud".

El esquema seguido para la construcción del mismo se inicia del reconocimiento de la existencia de grandes diferencias entre los países de la Región de las Américas en el área socioeconómica y en el estado de salud de los diversos grupos sociales que los componen.

La existencia de esta desigualdad entre las diversas regiones y países ha sido explícitamente reconocida en múltiples declaraciones a nivel internacional (Alma-Ata) (70). Así mismo ha sido una de las bases que ha motivado la adopción de la Meta de Salud para Todos en el Año 2000, y la estrategia de atención primaria. Estas metas representan un esfuerzo para disminuir o eliminar las distancias en el campo de salud que median entre países y grupos sociales. Sin entrar a realizar un juicio de valor definitivo, un ejemplo indicativo de las disparidades entre las subregiones y países de las Américas es la gran diferencia que se observa entre los valores de los cinco indicadores descritos anteriormente. Sin embargo, las razones-promedios de los indicadores en la región, pocas veces reflejan los valores de las mismas a nivel sub-regional o de los países a nivel individual. Teniendo en cuenta las consideraciones anteriores y como una contribución al análisis de las diferencias entre los países, se consideró conveniente tener un conjunto de indicadores que se relacionen entre la situación de salud y atención de salud en cada país.

Para la elaboración del índice se seleccionaron 20 indicadores de cinco categorías socioeconómicas y de salud. Estas categorías incluyen 5 indicadores demográficos, 4 indicadores del estado de salud, 4 indicadores de cobertura de salud, 4 indicadores de los recursos de los servicios de salud y 3 indicadores económicos y gasto en salud.

Conviene señalar que estos indicadores representan a su vez, los rubros seleccionados para monitorizar el alcance de las metas mínimas adoptadas en el Plan de Acción para la instrumentación de las estrategias regionales de la OPS (71, 72, 73). El Cuadro 2.66 contiene un listado de estos indicadores y en los Cuadros 1-21 del Anexo I se encuentran los valores correspondientes a los países incluidos.

Definidos los 20 indicadores se procedió a la selección del método estadístico\* que permitiera la integración de estos indicadores en el índice de necesidad en salud y consecuentemente se pudiera calcular el grado de desigualdad y necesidad en salud de los distintos países de la región. Es pertinente destacar que no existe una forma única de medir grados de necesidad en salud, ni tampoco acuerdo en cuanto al tipo específico de indicadores o procedimientos estadísticos que deban utilizarse. El grado de desagregación de la información utilizado en la construcción del indicador fue por país (Anexo I).

- 
- \* En relación con el procedimiento estadístico, se seleccionó el método de componentes principales del análisis factorial. Este método permite sintetizar el comportamiento de los indicadores seleccionados en una nueva variable o índice (73, 74, 75, 76).

Diversos autores sugieren la utilización de este procedimiento cuando el objeto de estudio (en este caso la necesidad en salud) está asociado a un gran número de indicadores, y se desea obtener una sola variable (componente principal) que combine las variables o indicadores originales. En este sentido, el primer componente principal obtenido es la combinación lineal ponderada de las variables incluidas, que representa la mayor cantidad de variabilidad total en los datos. La determinación de los valores del índice de necesidad en salud fueron calculados sin rotación de las matrices factoriales.

Los valores finales obtenidos en el índice son expresión de la suma de los valores estandarizados y ponderados de las variables originales utilizadas. De esta forma, el método sintetizó el comportamiento de los 20 indicadores seleccionados en una nueva variable denominada "Índice de Necesidad en Salud". El paquete estadístico empleado para el cálculo del índice fue el "Statistical Package for Social Sciences" (SPSS-X) versión 2.1 para computadora IBM VM/CMS 4381 (74). La elaboración estadística se realizó en las instalaciones del Centro de Datos y Computación de la Escuela Johns Hopkins de Salud Pública en Baltimore, Maryland.

CUADRO 2.66

LISTADO DE LOS DIVERSOS INDICADORES UTILIZADOS EN  
LA ELABORACION DEL INDICE DE NECESIDADES EN SALUD,  
REGION DE LAS AMERICAS, ALREDEDOR DE 1984

---

A. INDICADORES SOCIODEMOGRAFICOS

1. Tasa de crecimiento anual (X)
2. Porcentaje de población menor de 15 años
3. Porcentaje de población de 60 años y más
4. Tasa de fecundidad (nacimientos vivos por 1,000 mujeres 15-49 años)
5. Porcentaje de población alfabetizada (mayor de 15 años)

B. INDICADORES DEL ESTADO DE SALUD

1. Esperanza de vida al nacer
2. Tasa de mortalidad infantil (por 1,000 nacimientos vivos)
3. Tasa de mortalidad en 1-4 años (por 1,000 niños de 1-4 años)
4. Porcentaje de nacimientos con bajo peso al nacer (menos de 2,500 gramos)

C. INDICADORES DE COBERTURA DE SERVICIOS DE SALUD

Porcentaje de niños menores de 1 año totalmente inmunizados contra:

1. Difteria, tosferina y tétanos (3 dosis)
2. Poliomielitis (3 dosis)
3. Número de egresos hospitalarios por 100 habitantes
4. Número de camas por 1,000 habitantes

D. INDICADORES DE RECURSOS HUMANOS

1. Médicos por 10,000 habitantes
2. Dentistas por 10,000 habitantes
3. Enfermeras por 10,000 habitantes
4. Auxiliares de enfermería por 10,000 habitantes

E. INDICADORES ECONOMICOS Y DEL GASTO EN SALUD

1. PIB - Producto Interno Bruto, en US dólares
2. Gasto total en salud como porcentaje del PIB\*
3. Porcentaje del presupuesto total de salud correspondiente al Gobierno Central

---

\* Incluye servicios personales y no personales de salud

No se incluyeron en el análisis los siguientes países: Uruguay, Puerto Rico, Cuba, Anguila, Bermuda, Martinica, Guyana Francesa y Antillas Neerlandesas, por no contar con la información necesaria. Los 35 países incluidos fueron clasificados en cuatro niveles de desigualdad y necesidad en salud: muy alta, alta, media y baja.

Este índice ofrece una descripción útil de la distribución geopolítica de los niveles de necesidad en salud en las Américas. Debe aclararse también, que este índice de necesidad en salud, representa la posición que cada país tiene en relación al resto de los países, una vez que se toman en cuenta el conjunto de los 20 indicadores seleccionados. De tal manera, este índice no es una medida que pretende establecer una relación absoluta de la necesidad en salud.

Así mismo, como el índice representa un valor promedio de la necesidad de cada uno de los países con respecto al resto de los países, puede ocurrir que no obstante un país determinado alcance un rango de necesidad media o baja, al interior de dicho país, pueden existir y de hecho existen áreas o unidades geopolíticas donde se tiene una gran disparidad y necesidad en salud. Lo opuesto también acontece.

Para la interpretación de los grados de necesidad en salud encontrada en los valores del índice, se tiene una relación opuesta, es decir a menor valor del índice, mayor necesidad, y a mayor valor del índice, menor necesidad (ver Cuadro 2.67). Con el propósito de facilitar la interpretación, se presentan los valores ordenados de menor a mayor, así mismo se le asigna el rango correspondiente a cada país. De la misma manera, se presentan los países agrupados en rangos de acuerdo a su posición en cada uno de los 20 indicadores.

El rango de variación del índice fluctúa de -1.99 en el caso de Haití (mayor grado de necesidad) hasta 2.79 que correspondió a los Estados Unidos (menor necesidad). Este rango de variación de casi 5 puntos permitió dividir los 35 países en 4 estratos: los de necesidad en salud muy alta, con valores de -0.9 y menos; los de necesidad alta, de -0.01 a -0.8; los de necesidad media, de 0.0 a 0.9, y los de necesidad baja, de 1.0 y más. De los 35 países considerados, 6 (el 17%) se ubicaron en el estrato de necesidad en salud muy alta, 12 (el 34%) presentaron un nivel alto de necesidad; 14 (el 40%) correspondieron al nivel medio de necesidad y sólo 3 (8.6%) se situaron en el grupo de baja necesidad.

En el Cuadro 2.67 se presentan los resultados obtenidos de los países que se ubican en el estrato de necesidad muy alta, tres se localizan en Centroamérica: Guatemala, Nicaragua y Honduras, dos en el Caribe Latino: Haití y República Dominicana y uno en el Área Andina: Bolivia. En el otro extremo del índice, es decir con una necesidad baja, destacan Estados Unidos y Canadá en la América del Norte y Barbados en el Caribe. Los países agrupados en el estrato de necesidad alta son: Perú, El Salvador, México, Ecuador, Belice, Paraguay, Brasil, Colombia, Jamaica, Santa Lucía, Venezuela y Guyana. En el caso del estrato de necesidad media, se encontraban sólo 4 países de América Latina: Panamá, Costa Rica Chile y Argentina; el resto corresponden a la subregión del Caribe

de habla inglesa: Islas Vírgenes (UK), Bahamas, Montserrat, Trinidad y Tobago, Antigua, Dominica, San Vicente, Suriname, Grenada y San Cristóbal y Nieves.

Los contrastes observados en los valores del indicador de necesidad en salud confirman la existencia de grandes disparidades entre los países de las Américas. La situación que guardan los seis países ubicados en el estrato de necesidad muy alta, con respecto a los países ubicados en el nivel de baja necesidad, reflejan la gran heterogeneidad en la composición de los países de la Región. Por otro lado, la necesidad muy alta y alta en salud tiende a estar distribuida en más de la mitad de los países de la Región, es decir abarca 18 de los 35 países estudiados.

Asimismo, en la subregión de la América Latina y en particular, en el área de Centroamérica se concentra un porcentaje elevado de países con gran necesidad en salud.

CUADRO 2.67

INDICE DE NECESIDADES EN SALUD POR ESTRATO Y NIVEL DE NECESIDAD  
35 PAISES, REGION DE LAS AMERICAS, ALREDEDOR DE 1984

RANGO	PAIS	INDICE DE NECESIDAD	ESTRATO	NIVEL DE NECESIDAD
1.	Haití	-1.99099		
2.	Bolivia	-1.73970		
3.	Guatemala	-1.47766	-0.9 y menos	MUY ALTA
4.	Nicaragua	-.95835		
5.	Rep. Dominicana	-.94612		
6.	Honduras	-.90169		
7.	Perú	-.86259		
8.	El Salvador	-.81263		
9.	México	-.58185		
10.	Ecuador	-.46888		
11.	Belice	-.45763		
12.	Paraguay	-.41158	-0.01 a -0.8	ALTA
13.	Brasil	-.34489		
14.	Colombia	-.24551		
15.	Jamaica	-.14477		
16.	Santa Lucía	-.14153		
17.	Venezuela	-.12300		
18.	Guyana	-.09605		
19.	S. Cristóbal-Nieves	.05305		
20.	Grenada	.13136		
21.	Suriname	.13669		
22.	San Vicente	.22198		
23.	Panamá	.25583		
24.	Costa Rica	.28410		
25.	Dominica	.31675	0.0 a 0.9	MEDIA
26.	Antigua	.51301		
27.	Chile	.56817		
28.	Trinidad y Tobago	.57354		
29.	Montserrat	.67439		
30.	Bahamas	.70838		
31.	Argentina	.76423		
32.	Islas Vírgenes (UK)	.83791		
33.	Barbados	1.29690		
34.	Canadá	2.57252	1.0 y más	BAJA
35.	USA	2.79661		

CUADRO 12

PESOS ASIGNADOS A LOS 20 INDICADORES EN LA CONSTRUCCION  
DEL INDICE DE NECESIDAD DE SALUD, (COMPONENTE PRINCIPAL)  
REGION DE LAS AMERICAS, ALREDEDOR DE 1984

INDICADOR	PESO/PONDERACION
1. Población 15	-.09239
2. Vacuna Polio	.4327
3. Egreso-Hospitalario	.07542
4. Vacuna DPT	.05733
5. Aux. Enfermería	.08090
6. Enfermeras	.08943
7. Gasto en Salud per Capita	.08121
8. Población 60	.09221
9. % Crec. Población	-.07477
10. Fecundidad	-.08918
11. % Alfabet.	.06780
12. Esperanza-Vida	.08330
13. Mort. Infantil	-.07413
14. Mort. 04 Años	-.06569
15. Bajo Peso/Nacer	-.04743
16. Camas Hospital	.08231
17. Médicos	.05923
18. Dentistas	.05933
19. PNB	.08397
20. Gasto-Salud/PNB	.03225

La varianza explicada por el primer componente (Indice de Necesidad en Salud) representó el 46.3 por ciento de la varianza total.

**ENFOQUES Y METODOS PARA LA  
DETERMINACION DE NECESIDADES**

Table 3-2 Needs Assessment Approaches and Methods

Approach	Method	Information Processing Function	Measurement Expertise Needed	Time and Resources Needed
Indicator approach Health indicators	Analysis of statistics on life expectancy, morbidity, mortality, and disability	Compilation	Moderate to high	Moderate
	Social indicators	Compilation	Moderate to high	Moderate
	Extrapolation/assumption	Compilation and integration	Moderate	Minimal
Survey approach	Analysis of service utilization or rates under treatment	Compilation	Moderate	Moderate
	Survey sample of labor and service facilities	Compilation and development	Moderate	Moderate
	Survey sample of general population	Development	High	Extensive
	Survey sample of service or provider population	Development	High	Moderate
Consensus-reaching approach	Community forum Nominal group Key informants Delphi technique	Integration Development Development Development and integration	Low Moderate Moderate Moderate	Moderate Minimal Minimal Moderate
	Community impressions	Integration Development, compilation, and integration	Moderate	Minimal

Sources: Adapted from "Need Identification and Program Planning in the Community Context" by L.M. Siegel, C.C. Attkisson, and L.G. Carson in *Evaluation of Human Service Programs*, ed. C.C. Attkisson, with permission of Academic Press, Inc., ©1978, p. 226; from *Guide to Health Needs Assessment: A Critique of Available Sources of Health Care Information* by L.W. Chambers, C.A. Woodward, and C. Dok with permission of the Canadian Public Health Association, ©1980, p. 32, and from *Determining Health Needs* by Robin E. MacStravic with permission of Health Administration Press, ©1978, p. 268, as reprinted in "Program Planning in a Small Community Health Care Setting" by Carol Clemenhagen and Francois Champagne with permission of *Health Care Management Review* 7, no. 1 (Winter 1982), ©1982, 47-55.

Fuente:  
Dever, Alan  
Epidemiology in Health Services Management.  
Aspen Publication, Rockville, 1984. p.59

**CONDICIONES SENTINELA PROPUESTAS  
POR RUTSTEIN Y COL.**

Table 2-7 Sentinel Health Events\*

Condition	Unneces- sary Disease <sup>1</sup>	Unneces- sary Dis- ability <sup>1</sup>	Unneces- sary Un- timely Death <sup>1</sup>	Notes <sup>2</sup>
Cholera	P		P T	
Typhoid fever	P		P T	
Baculism	P		P	
Tuberculosis (all forms)			T	
Silicotuberculosis	P	P	P T	P—Occupational
Plague			T	
Diphtheria	P		P T	
Rubella	P	P	P T	Disability in off-spring
Congenital syphilis	P	P	P T	
Major complications of syphilis	P T	P T	P T	
Gonococcal infections			T	
Malignant neoplasm of lip	F		P T	P—Pipe smoking and sun exposure
Malignant neoplasm of larynx	P		P T	P—Cigar and cigarette smoking
Malignant neoplasm of trachea, bronchus, and lung	P		P	P—Cigarette smoking and occupational exposure
Malignant neoplasm of bladder	P		P	P—Aniline dyes and cigarettes
Malignant neoplasm of eye		T	T	Genetic—screening and treatment
Thyroid carcinoma	P		P	P—Radiation exposure
Endemic goiter	P			Iodine deficiency
Nutritional deficiencies	P	P T	P T	Not associated with neoplasia and malabsorption
Pericious anemia		T	T	
Pulmonary heart disease	P	P	P	P—Occupational and environmental exposure
Chronic bronchitis, emphysema	P	P	P	P—Occupational and environmental exposure
Appendicitis			T	
All maternal deaths			P	

\*This listing is intended only as an illustration.

1. P denotes prevention; T denotes treatment.

2. The symbols P or T in the notes indicate that the prevention or treatment is limited to the circumstances described.

Source: Reprinted from "Measuring the Quality of Medical Care: A Clinical Method," by David D. Rutstein, M.D., William Berenberg, M.D., Thomas C. Chalmers, M.D., Charles G. Child III, M.D., Alfred P. Fishman, M.D., and Edward B. Perrin, Ph.D., with permission of *The New England Journal of Medicine*, ©1976, 294: 585.

**Fuente:**

**Rutstein, D., W. Berenberg, et al. 1976:  
"Measuring the Quality of Medical Care:  
A Clinical Method". New England Journal  
of Medicine. 294:582-588**

**PROCEDIMIENTOS SIMPLES UTILIZADOS EN LA  
DETERMINACION DE NECESIDADES DE SALUD**

**Dr. C. Castillo Salgado**

A continuación se presentan los pasos de un procedimiento comunmente utilizado en la determinación de necesidades de salud.

- 1. La selección de unidades de análisis**
  
- 2. La definición de variables**

**ejemplo :**

- % de la población económicamente activa que percibe ingresos inferiores al salario mínimo.**
- % de población rural incomunicada.**
- % de población analfabeta de 10 o más años.**
- % de población de 15 años y mas sin estudios primarios completos.**
- % de viviendas sin agua entubada.**
- % de viviendas con uno y dos cuartos.**
- % de viviendas sin drenaje.**
- Tasa de mortalidad general.**
- Tasa de mortalidad infantil.**
- Tasa de mortalidad perinatal.**
- % población menor de 15 años.**
- % nacimientos de mujeres menores de 18 años.**
- % población mayor de 65 años.**

### 3. Fórmula y algoritmo usado:

a. Se asume que todas las variables ( $X_1, X_2, \dots, X_n$ ) tienen un valor. La media ( $m$ ) es calculada de acuerdo a la fórmula siguiente:

$$m = \frac{X_1 + X_2 + \dots + X_n}{n}$$

La desviación estándar ( $s$ ) es calculada con la fórmula siguiente:

$$s = \sqrt{\frac{\sum (X - m)^2}{n - 1}}$$

b. Los valores o puntajes de " Zeta " ( $Z$ ) de las  $X$ s son calculados de la siguiente manera :

$$Z_i = (X_i - M) / s \quad i = 1, \dots, n ;$$

Si los valores de Zeta encontrados son:

	<	-3.5	entonces el puntaje asignado =	-4
-3.5	= < Z <	-2.5	=	-3
-2.5	= < Z <	-1.5	=	-2
-1.5	= < Z <	-0.5	=	-1
-0.5	= < Z <	0.0	=	0
0.0	= < Z <	0.5	=	0
0.5	= < Z <	1.5	=	1
1.5	= < Z <	2.5	=	2
2.5	= < Z <	3.5	=	3
	Z >	3.5	=	4

### 4. Asignación de los puntajes de necesidad

El valor del puntaje final de cada unidad es la suma de los puntajes asignados a cada variable.

## INDICADORES POR COMUNIDAD A

	COMUNIDADES AUT.	AUTOVALORACION EDO SALUD "MALO"	TASA SARAMPION (X 100,000) 1988	TETANOS 1988	ESPERANZA DE VIDA (AMBOS SEXOS)	ESPERANZA DE VIDA (HOMBRES)	AÑOS DE VIDA PERDIDA TASA X 1000 ( < 55 AÑOS )	MORTALIDAD INFANTIL	DIAS DE RESTRICION ACTIV. PRINCIPAL X 100 HAB (16-44) MASC.	DIAS-CAMA/100 HAB
1.	6.3	37.71	0.06	74.46	71.36	45.41	10.18	19.8	22.9	
2.	5.2	16.53	0.16	76.54	73.71	39.44	8.78	13.7	13.2	
3.	5.0	3.20	0.59	74.81	71.03	46.40	11.29	4.5	9.1	
4.	6.7	26.51	0.00	75.76	72.60	49.29	10.94	16.1	14.1	
5.	7.2	672.39	0.19	74.24	71.23	46.95	8.66	14.5	15.9	
6.	3.9	2.92	0.00	75.46	71.41	40.90	8.51	17.8	9.9	
7.	7.7	23.21	0.12	75.60	73.21	43.57	8.61	8.6	9.7	
8.	7.1	67.10	0.06	76.34	73.49	42.80	7.37	26.2	16.2	
9.	7.5	20.09	0.09	76.88	73.90	34.92	6.78	43.1	17.2	
10.	7.5	64.95	0.24	74.87	72.02	39.37	8.04	15.0	17.3	
11.	8.2	33.25	0.10	74.82	71.61	46.66	10.31	35.8	17.8	
12.	7.4	11.13	0.17	75.23	72.01	47.71	10.73	23.8	14.7	
13.	5.5	41.71	0.09	76.27	72.65	38.72	7.77	16.9	16.9	
14.	6.5	4.93	0.09	74.67	71.55	42.84	9.07	17.2	15.4	
15.	6.7	9.96	0.00	75.33	71.75	46.40	11.29	17.5	9.8	
16.	5.3	10.98	0.00	74.97	71.04	42.96	9.97	61.9	29.6	
17.	5.2	15.58	0.00	74.76	71.59	47.71	11.37	3.4	7.3	
X	6.4	62.47	0.697	75.36	72.12	43.65	9.39	19.75	14.52	
S	1.19	158.36	2.401	0.77	0.952	3.94	1.47	15.04	6.08	

26311

## INDICADORES POR COMUNIDAD A

	COMUNIDADES AUT.	PERSONAS CON LIMITACIONES PERMANENTES TASA X 100,000	T.B. (X 100,000) 1988	TOSFERINA 1988 (TASA X 100,000)	% DISCAPACITADOS ACTIVOS-PARA DOS (6-64)	NACIDOS VIVOS DE MADRES - 20 AÑOS % TOTAL NACIDOS VIVOS	# ODONTOLOGOS/ 10,000	PORCENTAJE DE POBLACION SIN SERVICIOS HIGIENICOS EN LA VIVIENDA (FAM)	% HOGARES SIN SERVICIOS HIGIENICOS EN VIVIENDA	MORTALIDAD PERINATAL
1.	11.224	14.89	68.86	9.68	8.28	1.45	5.56	11.16	14.13	
2.	6.904	27.79	19.12	5.68	3.86	1.95	3.23	4.36	9.04	
3.	6.465	31.96	12.70	9.01	7.88	1.98	5.88	5.93	13.93	
4.	6.742	21.52	25.59	6.72	6.31	1.74	2.49	7.72	20.09	
5.	13.390	30.68	62.75	7.11	10.44	1.45	3.20	4.54	9.42	
6.	14.393	18.82	9.32	2.74	6.33	1.94	7.20	9.81	11.88	
7.	6.240	35.58	30.65	5.75	5.63	1.37	14.30	16.65	10.45	
8.	7.963	20.27	52.23	5.94	4.85	1.00	10.45	18.63	10.59	
9.	9.380	20.87	39.75	7.11	4.86	2.01	1.74	1.07	7.46	
10.	9.236	21.02	10.52	7.29	5.36	1.98	1.24	1.70	9.30	
11.	9.170	13.66	7.53	9.59	7.24	0.78	14.36	24.34	12.43	
12.	9.084	24.68	4.14	6.42	10.62	1.13	16.66	21.29	11.89	
13.	11.576	17.68	44.45	6.64	4.61	2.96	1.69	1.76	8.40	
14.	10.995	22.83	7.01	2.57	7.86	0.85	3.89	2.13	11.34	
15.	3.410	24.36	44.47	2.97	3.54	1.97	1.10	0.37	9.98	
16.	9.999	5.36	19.17	7.15	3.95	2.25	0.86	2.41	12.54	
17.	6.676	41.18	6.31	4.93	4.95	2.02	2.93	2.27	12.10	
X	8.991	23.12	26.73	6.31	6.26	1.69	5.69	8.01	11.46	
S	2.810.8	8.59	21.68	2.14	2.16	0.56	5.14	7.73	2.90	

26311

NECESIDADES ESENCIALES  
EN MEXICO  
*Situación actual y perspectivas al año 2000*

5

GEOGRAFIA DE LA MARGINACION



## 2. METODOLOGIA

No existe una sola forma de medir el concepto de marginación aquí adoptado, ni acuerdo en cuanto al tipo de indicadores que deban utilizarse. Algunos estudios subrayan más los aspectos económicos de la marginación, otros dan más importancia a los sociales. Tampoco existe una única manera de combinar los indicadores para obtener una sola expresión del grado de marginación. Por último, no existe un concepto único, universalmente aceptado.

Consecuentemente, tres necesidades metodológicas de este trabajo fueron: a) determinar un sistema de indicadores adecuados al concepto de marginación adoptado, que pudieran elaborarse en los niveles estatal y municipal; b) seleccionar un método estadístico capaz de sintetizar el comportamiento de los indicadores seleccionados, y jerarquizar las unidades geográficas estudiadas de acuerdo con su grado de marginación; c) adoptar un concepto de región (o zona) aplicable a la delimitación de regiones, de zonas y núcleos. Los indicadores y el método estadístico empleados en el cálculo de los índices de marginación se describen en este capítulo.

En relación con el concepto de región se analizaron los principales conceptos existentes: región homogénea, región nodal y región plan, los cuales no son necesariamente excluyentes entre sí. Tratando de hacer una síntesis de ellos, se puede decir que el primero agrupa un conjunto de unidades geográficas (estados, municipios, distritos, etc.) lo más homogéneas posible a partir de uno o varios indicadores de tipo físico-geográfico, económico, social o cultural. La región nodal delimita su área a partir de un centro de población importante, denominado polo de atracción, sobre el cual gira la actividad económica del resto de las localidades del área. La región plan define un territorio con propósitos de programación a partir del criterio básico de la coherencia administrativa que debe existir entre la problemática del área y la presencia de instituciones capaces de llevar a cabo las tareas que se programen.<sup>1</sup>

Como podrá apreciarse más adelante, la regionalización de múltiples niveles adoptada en este documento combina los conceptos primero y tercero de región; esto es, homogeneidad en cuanto al nivel de marginación y coherencia administrativa que garantice la posibilidad de ejecución de los programas.

### INDICADORES UTILIZADOS

En concordancia con lo señalado respecto al concepto de marginación adoptado, se

<sup>1</sup> Para abundar sobre los conceptos de región véanse: H. Richardson, *Elements of regional economics*, Londres, Penguin Modern Economics, 1970; y R. Arias Hernández, *Una aproximación al enfoque integral de la región económica*, México, Instituto Nacional Indigenista, 1978.

consideró que la marginación se expresa en: 1) bajos niveles de ingreso de la población económicamente activa; 2) altos niveles de subempleo; 3) altos porcentajes de población rural agrícola; 4) incomunicación de las localidades; 5) alimentación inadecuada; 6) bajos niveles de escolaridad; 7) bajos niveles de salud y dificultad de acceso a los servicios médicos; 8) viviendas inadecuadas y sin servicios, y 9) escaso acceso a otros satisfactores como calzado, radio y televisión.

Para medir estas características se construyeron 19 indicadores divididos con propósitos de ordenamiento en: generales, de alimentación, de educación, de salud, de la vivienda y sus servicios, y de otras necesidades:

a) *Generales*

1. Porcentaje de población económicamente activa (PEA) que percibe ingresos inferiores a 1 000 pesos mensuales.
2. Porcentaje de PEA subempleada.<sup>2</sup>
3. Porcentaje de población rural.<sup>3</sup>
4. Porcentaje de PEA en el sector agropecuario.
5. Porcentaje de población rural incomunicada.

b) *Alimentación*

6. Porcentaje de población que consume leche dos o menos días a la semana.
7. Porcentaje de población que consume carne dos o menos días a la semana.
8. Porcentaje de población que consume huevo dos o menos días a la semana.

c) *Educación*

9. Porcentaje de población analfabeta de 10 y más años de edad.
10. Porcentaje de población de 15 años y más sin primaria completa.<sup>4</sup>

d) *Salud*

11. Tasa de mortalidad general.<sup>5</sup>
12. Tasa de mortalidad preescolar.<sup>6</sup>
13. Habitantes por médico.

e) *Vivienda y sus servicios*

14. Porcentaje de viviendas sin agua entubada.<sup>7</sup>
15. Porcentaje de viviendas de uno y dos cuartos.<sup>8</sup>
16. Porcentaje de viviendas sin electricidad.
17. Porcentaje de viviendas sin drenaje.

f) *Otras necesidades*

18. Porcentaje de población de un año y más que no usa calzado.<sup>9</sup>
19. Porcentaje de viviendas sin radio ni televisión.

Los indicadores 1 a 4, 6 a 10 y 14 a 19 fueron construidos a partir de los datos del IX Censo General de Población y Vivienda. La fuente de los indicadores 11, 12 y 13

<sup>2</sup> Aquella que trabaja nueve o menos meses al año.

<sup>3</sup> La que habita localidades menores de 2 500 habitantes.

<sup>4</sup> En el caso de municipios se tomó la población de 12 años y más debido a la no disponibilidad de información para 15 y más años de edad.

<sup>5</sup> Defunciones totales entre población total, por 1 000.

<sup>6</sup> Defunciones de niños de 1 a 4 años de edad entre la población respectiva por 1 000.

<sup>7</sup> No disponen de agua entubada ni dentro ni fuera de la vivienda; ni tienen acceso a llave pública o hidrante. En el caso de entidades federativas la variable se refiere a las viviendas que además de estar en la situación anterior carecen de drenaje.

<sup>8</sup> Este indicador es una aproximación burda, pero la única disponible, del hacinamiento a nivel municipal.

<sup>9</sup> Población que no usa zapatos ni huaraches.

fue *Estadísticas vitales de los Estados Unidos Mexicanos*, Secretaría de Salubridad y Asistencia (SSA), México, 1970. El indicador 5 se construyó a partir de datos de la Secretaría de Asentamientos Humanos y Obras Públicas (SAHOP) "Inventario de caminos rurales por construir", 1979 (hojas de computadora).<sup>10</sup>

Una aclaración importante es que el indicador 2, porcentaje de población rural, y el 3, porcentaje de PEA en el sector agropecuario, han sido considerados no por ser reflejo de marginalidad, sino porque según se ha argumentado, una buena parte de la población rural ha quedado al margen de los beneficios del desarrollo existiendo, consecuentemente, una alta asociación entre el grado de ruralización y de dedicación a las actividades agropecuarias, y los niveles de pobreza y marginalidad.

El resto de las variables están directamente relacionadas con las condiciones de vida de la población y, en especial, con los mínimos de bienestar en materia de alimentación, educación, salud y vivienda; se incluyen indicadores muy importantes de carácter general como nivel de ingresos y población rural incomunicada. Asimismo, se trató de cubrir la información sobre otras necesidades de la población, encontrándose únicamente datos sobre el uso de calzado y la disponibilidad de radio y televisión.

Aunque se estima que los 19 indicadores cubren un amplio espectro de lo que es la marginación, es necesario señalar que la información estadística en el nivel municipal presenta algunas deficiencias en calidad y en disponibilidad. Ambas situaciones se observan en la información sobre los aspectos de salud de la población, en la que tanto las tasas de mortalidad general y preescolar de la población, así como la cantidad de habitantes por médico reflejan sólo de manera indirecta la salud de la población, encontrándose además problemas de subregistro de muertes en algunos municipios. En el caso de la alimentación, aunque la información es medianamente confiable, el consumo de leche, carne y huevo, expresado en número de días de la semana en que se consume tales alimentos, ofrece también sólo un reflejo indirecto del nivel de nutrición de la población mexicana.

Por otra parte, inicialmente se pensó incluir algunos indicadores relacionados con el grado de dispersión de la población tales como la densidad demográfica por km<sup>2</sup>, o el porcentaje de localidades pequeñas respecto al total de localidades. Sin embargo, después de realizar algunas pruebas de asociación entre varios indicadores de dispersión y el nivel de marginación de las entidades federativas, se encontró una correlación muy baja y en ocasiones negativa, por lo que se decidió no incluir este tipo de indicadores. En el anexo 7 se muestran los indicadores de dispersión que fueron ensayados y la explicación del porqué no fueron considerados dentro de los indicadores finales.

No obstante, una vez calculados los índices de marginación, el uso de algunos indicadores de dispersión será importante para conocer la densidad de la población marginada y definir las prioridades de atención.

#### EL METODO ESTADISTICO

Para calcular los índices de marginación de las unidades geográficas estudiadas (en-

<sup>10</sup> Los datos de los 19 indicadores en los niveles de entidades, secciones agrícola y no agrícola de éstas, regiones, municipios y zonas y núcleos marginados, se presentan en los anexos 1, 2, 3, 4 y 6 respectivamente.

tidades federativas, regiones, municipios, zonas y núcleos), se utilizó el método de componentes principales.<sup>11</sup>

El uso de este método se recomienda cuando se quiere analizar un fenómeno que está asociado a un buen número de indicadores, y se desea obtener una sola variable que represente el comportamiento de los indicadores. En nuestro caso, el método sintetizó el comportamiento de los 19 indicadores señalados anteriormente en una sola variable denominada "Índice de marginación".

Debido al tipo de indicadores utilizados, el valor del índice de marginación se mueve en sentido directo al grado de marginación; esto es, a mayor valor del índice mayor marginación, y viceversa.

Si los indicadores están medidos en unidades diferentes (como en el caso que aquí se presenta) tienen que estandarizarse con base en sus medias y desviaciones estándar. Los valores estandarizados de los indicadores en cada unidad geográfica se ponderan con un conjunto de números. La suma ponderada de los valores de las variables de cada unidad geográfica es igual al índice de marginación.<sup>12</sup>

Es necesario aclarar que los índices de marginación sólo son comparables entre unidades que pertenezcan a un mismo nivel geográfico; esto es, el índice del municipio *x* es comparable al del municipio *y*, pero no es comparable con los del estado, región, zona o núcleo *z*. Lo mismo ocurre con los de estados, regiones, estratos y zonas. Son comparables entre sí pero no con los de otros niveles geográficos. Esto se debe a que en el cálculo de los índices de marginación de los diferentes niveles geográficos, se usan ponderadores específicos, óptimos para cada nivel.

Con el propósito de ofrecer al lector un cuadro general comparativo del grado de marginación de entidades federativas, de regiones y de zonas y núcleos, el anexo 6 contiene índices de marginación para cada una de estas unidades geográficas, recalculados utilizando en todos los casos los ponderadores obtenidos en el análisis municipal. Los índices que así resultan son diferentes a los presentados en el texto, pero estrictamente comparables en distintos niveles. Asimismo, el anexo proporciona la información estadística utilizada en el cálculo de dichos índices.

<sup>11</sup> Varios autores, *Statistical package for the social sciences*, Nueva York, McGraw-Hill, 2a. ed., 1975.

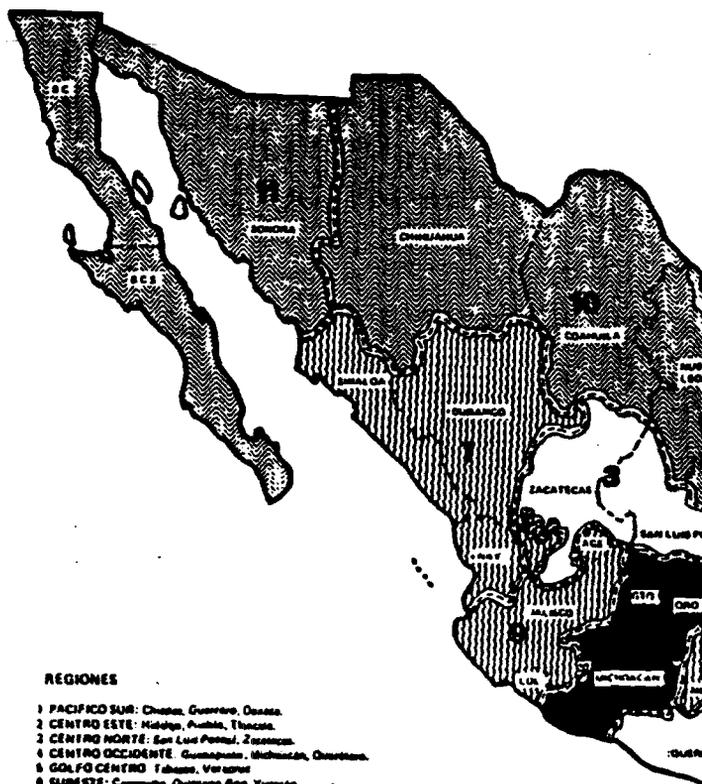
<sup>12</sup> La fórmula es la siguiente:

$$IM_j = \sum_{i=1}^{19} FPI_i \frac{X_{ij} - \bar{X}_i}{d_i} \quad j = 1 \dots N$$

en donde

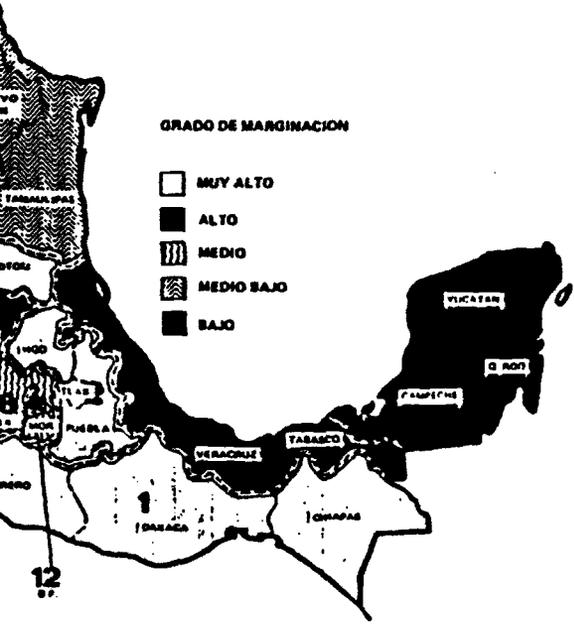
- IM<sub>j</sub> = Índice de marginación de la unidad geográfica *j*
- X<sub>ij</sub> = Valor de indicador *i*, de la unidad geográfica *j*
- $\bar{X}_i$  = Valor de la media del indicador *i*
- d<sub>i</sub> = Desviación estándar del indicador *i*
- FPI = Factor de ponderación del indicador *i*
- N = 32 en el caso de entidades, 12 para las regiones, 1 848 para los municipios y 111 para las zonas y núcleos marginados.

**MAPA 2**  
**REGIONES SEGUN GRADOS DE MARGINACION**



**REGIONES**

- 1 PACIFICO SUR: Chetumal, Guerrero, Oaxaca.
- 2 CENTRO ESTE: Hidalgo, Puebla, Tlaxcala.
- 3 CENTRO NOROCCIDENTE: San Luis Potosí, Zacatecas.
- 4 CENTRO OCCIDENTE: Guanajuato, Michoacán, Querétaro.
- 5 GOLFO CENTRO: Tlaxcala, Veracruz.
- 6 SURESTE: Campeche, Quintana Roo, Yucatán.
- 7 PACIFICO CENTRO: Durango, Nayarit, Sinaloa.
- 8 CENTRO: México, Morelos.
- 9 OCCIDENTE: Aguascalientes, Colima, Jalisco.
- 10 NOROCCIDENTE: Coahuila, Chihuahua, Nuevo León, Tamaulipas.
- 11 PACIFICO NOROCCIDENTE: Baja California, Baja California Sur, Sonora.
- 12 DISTRITO FEDERAL: Cd. de México, Delegaciones.



CUADRO 5  
PONDERACION DE LOS INDICADORES DE MARGINACION. NIVEL REGIONAL

Indicadores	Ponderaciones *
1. Bajos ingresos de la PEA	.86
2. Subempleo	-.17
3. Población rural	.94
4. Ocupación agrícola	.95
5. Incomunicación rural	.70
6. Subconsumo de leche	.97
7. Subconsumo de carne	.92
8. Subconsumo de huevo	.83
9. Analfabetismo	.93
10. Población sin primaria	.96
11. Mortalidad general	.58
12. Mortalidad preescolar	.86
13. Habitantes por médico	.83
14. Viviendas sin agua entubada	.95
15. Hacinamiento	.92
16. Viviendas sin electricidad	.97
17. Viviendas sin drenaje	.90
18. No disponibilidad de radio ni T. V.	.73
19. Población que no usa calzado	.96

\* La varianza explicada por la primera componente representó el 73.7% de la varianza total. Esta cifra es mayor que la observada en el nivel de entidades federativas (63.3%) porque además de que una mayor agregación permite compensar posibles deficiencias en la información de los indicadores a nivel de entidad, la regionalización "separó" a los estados más marginados de los menos marginados, de tal forma que los indicadores cobraron mayor fuerza discriminatória.  
Fuente: Elaboración de Coplamar con base en diversas fuentes. Véase el capítulo 2, indicadores utilizados.

sin electricidad, que tienen valores medios nacionales de 41.4, 39.4, 43.3, 38.9 y 41.1%, y que se elevan a 63.3, 62.7, 77.5, 57.7 y 61.9%, respectivamente, en el estrato de regiones de *muy alta* marginación.

#### ANÁLISIS DE RESULTADOS

##### *Comparación con estudios sobre el grado de desarrollo estatal*

Aunque con propósitos un tanto diferentes a los del presente trabajo, existen otros estudios que intentan jerarquizar las entidades federativas del país.<sup>4</sup>

<sup>4</sup> Véase por ejemplo: L. Unikel, L. y E. Victoria, "Medición de algunos aspectos del desarrollo socio-

CUADRO 6  
INDICES DE MARGINACION POR REGIONES AGRUPADAS EN ESTRATOS

Regiones	Indice de marginación	Estratos	Grado de marginación
1. <u>Pacífico Sur.</u> Chiapas, Guerrero, Oaxaca.	22.70		
2. <u>Centro Este.</u> Hidalgo, Puebla, Tlaxcala.	13.40	10.00 y más	MUY ALTO
3. <u>Centro Norte.</u> San Luis Potosí, Zacatecas.	10.38		
4. <u>Centro Occidente.</u> Guanajuato, Michoacán, Querétaro.	6.80		
5. <u>Golfo Centro.</u> Tabasco, Veracruz.	5.92	0.01 a 9.99	ALTO
6. <u>Sureste.</u> Campeche, Quintana Roo, Yucatán.	5.80		
7. <u>Pacífico Centro.</u> Durango, Nayarit, Sinaloa.	- 1.00		
8. <u>Centro.</u> México, Morelos.	- 1.06	- 0.01 a -9.99	MEDIO
9. <u>Occidente.</u> Aguascalientes, Colima, Jalisco.	- 6.66		
10. <u>Norte.</u> Coahuila, Chihuahua, Nuevo León, Tamaulipas.	-13.07		
11. <u>Pacífico Norte.</u> Baja California, Baja California Sur, Sonora.	-14.98	-10.00 a -19.99	MEDIO BAJO
12. <u>Distrito Federal.</u> Ciudad de México y Delegaciones.	-28.22	-20.00 y menos	BAJO

FUENTES: Elaboración de Coplamar con base en diversas fuentes. Véase el capítulo 2, indicadores utilizados.

económico de las entidades federativas de México, 1940-1960", *op. cit.*; K. Appendini, D. Murayama y R. M. Domínguez, "Desarrollo desigual en México, 1900 y 1960", *op. cit.*; J. Casimir, *Aspectos sociales de las desigualdades regionales*. Plan Lerma Asistencia Técnica, Guadalajara, 1973; J. Boltvinik y R. Passah, *La asignación de recursos públicos a la agricultura en México, 1959-1976*, *op. cit.*; Comisión Nacional de los Salarios Mínimos, *Niveles de desarrollo económico de los municipios y entidades federativas de México, 1970*, México, inédito, 1975.

**TALLER SOBRE ESTRATIFICACION DE LA MALARIA**

**FAVOR NO REPRODUCIR  
SIN LA AUTORIZACION  
DE LA ORGANIZACION  
PANAMERICANA DE LA  
SALUD (OPS)**

**MATERIALES DE TRABAJO**

**I**

**MEDICION DEL RIESGO DE ENFERMAR O MORIR  
(CONCEPTOS BASICOS)**

**(OPS) ORGANIZACION  
PANAMERICANA DE LA  
SALUD**

**NOTA**

**VERSION BORRADOR.**

**FAVOR DE NO REPRODUCIR O CITAR SIN LA AUTORIZACION DE LA  
ORGANIZACION PANAMERICANA DE LA SALUD**

## MEDICION DEL RIESGO DE ENFERMAR O MORIR DE MALARIA

### 1. Introducción

Los presentes materiales tienen como propósito ofrecer las bases metodológicas para el apoyo del proceso de la Estratificación Epidemiológica de la Malaria en los países de América Latina y el Caribe. La aplicación del método epidemiológico para el análisis de situaciones, eventos y factores relacionados con el proceso de la transmisión de la malaria así como de las medidas de intervención y control de la misma, representa una parte importante del contenido epidemiológico\* de estos materiales.

Entre las medidas de mayor importancia en la investigación epidemiológica relacionada con la transmisión de la malaria que aquí se presentan se encuentran: (a) el riesgo absoluto o incidencia; (b) riesgo relativo, y; (c) riesgo atribuible. Estas medidas son fundamentales ya que permiten la cuantificación del exceso de riesgo de enfermar o morir prematuramente de malaria, que tienen tanto individuos como grupos sociales debido a sus condiciones de vida y la exposición a uno o más determinados factores de riesgo.

\* EPIDEMIOLOGIA es la disciplina que "estudia la distribución y determinantes de eventos o situaciones relacionadas con la salud en poblaciones definidas y la aplicación de este estudio en el control de los problemas de salud". (Last:42)

## 2. Medición del Riesgo de Enfermar y de Morir

### 2.1 Concepto de factor de riesgo

Como se enuncia en la introducción, en el campo de la investigación epidemiológica, la medición del riesgo de enfermar o de morir de cualquier daño a la salud, como malaria, implica el uso del concepto epidemiológico de factor de riesgo. Se ha denominado como factor de riesgo a toda característica o circunstancia de una persona o grupo que está asociada con un incremento en la probabilidad de padecer o desarrollar un particular proceso mórbido.

De acuerdo a los textos contemporáneos de epidemiología y expuesto en el diccionario internacional de epidemiología editado por John Last (1), el término "factor de riesgo" se usa para expresar las siguientes tres consideraciones:

1. "un atributo o exposición que se asocia con una probabilidad mayor de desarrollar un resultado específico, tal como la ocurrencia de una enfermedad; este atributo no necesariamente constituye un factor causal";
2. "un atributo o exposición que aumenta la probabilidad de la ocurrencia de una enfermedad u otro resultado específico";
3. "un determinante que puede ser modificado por alguna forma de intervención, logrando disminuir la probabilidad de una enfermedad u otro daño específico a la salud; para evitar confusión, esta connotación debe ser referida como factor de riesgo modificable".

El conocimiento de la historia natural de un gran número de enfermedades, el avance logrado por la investigación epidemiológica de las enfermedades crónicas (infecciosas y no infecciosas), así como el reconocimiento de la naturaleza múltiple-causal de las enfermedades, han llevado a un reconocimiento de múltiples fuerzas que operan en el proceso

de salud-enfermedad. Estas fuerzas son llamadas factores de riesgo. Es importante mencionar que una de las características mas relevantes de los factores de riesgo radica en que son características o atributos que son observables o identificables antes de la ocurrencia del fenómeno mórbido que predicen. Por lo tanto, permiten el máximo uso de los conceptos de prevención y promoción a la salud.

Los factores de riesgo pueden ser aquellos que caracterizan al individuo, la familia, el grupo social, la comunidad o el ambiente. La presencia o combinación de uno o mas factores de riesgo en los mismos individuos o grupos sociales, incrementa la probabilidad de desarrollar un daño a la salud.

Por daño a la salud\* se entiende a todos los posibles cambios negativos en el estado de salud que son consecuencia de las condiciones de vida y de la vulnerabilidad y exposición a determinados factores de riesgo. Los daños a la salud mas estudiados son la enfermedad y la muerte. Sin embargo, las lesiones, los accidentes y las incapacidades son también importantes daños a la salud.

Es importante hacer notar que en diversas situaciones, una característica puede ser identificada tanto como un daño como un factor de riesgo. Por ejemplo la desnutrición puede ser reconocida como un daño a la salud específico, así mismo puede ser vista como un factor de riesgo de otros subsecuentes daños a la salud. La definición precisa de los daños y factores de riesgo que van a ser investigados es de gran importancia para la conducción de la investigación epidemiológica.

---

\*La salud es un estado de COMPLETO bienestar físico, mental y social, y no solamente la AUSENCIA de afecciones o enfermedades" (Constitución de la OMS).

Los factores de riesgo pueden ser específicos para un daño en particular. Por ejemplo, la presencia de criaderos de mosquitos transmisores cercanos a la vivienda puede ser un factor de riesgo específico para el desarrollo de la malaria. Por otro lado, puede ocurrir que un mismo factor de riesgo incremente la probabilidad de tener distintos o varios daños a la salud. Un ejemplo de esto lo constituye la pobreza, analfabetismo, agua contaminada, etc. cuya presencia se asocia a diversos padecimientos: malaria, gastroenteritis o desnutrición.

La importancia epidemiológica que puede tener un factor de riesgo en particular, depende básicamente de tres aspectos:

- (a) del grado de asociación y de determinación que tenga con el daño a la salud que se este investigando;
- (b) de la frecuencia que dicho factor de riesgo tenga en la comunidad; y
- (c) de la posibilidad que exista de prevenirlo o controlarlo.

Durante la década de los 80's, diversos países de América Latina han experimentado un incremento acentuado y continuo de la incidencia de la malaria. Por tal motivo, se ha reconocido la necesidad de realizar un número cada vez mas extenso de estudios epidemiológicos y sociales que aborden la dinámica de la transmisión de la malaria con una metodología en que la medición y evaluación de la distribución de los distintos factores de riesgo de enfermar o de morir de malaria de los distintos grupos sociales sirva de base a los programas de intervención (prevención y control) de los servicios de salud. Así, es fundamental el poder otorgar

especial atención a la definición epidemiológico-social de los factores de riesgo y daños que se busca estudiar.

Los factores de riesgo pueden ser clasificados, entre otros, en:

- (1) biológicos (ej. ciertos grupos de edad);
- (2) ambientales (ej. falta de adecuado sistema de disposición de excretas, abastecimiento deficiente de agua);
- (3) económicos (ej. bajos ingresos, desempleo);
- (4) sociales-culturales (ej. algunos grupos étnicos, bajo nivel de educación, grupos humanos desplazados);
- (5) relacionado con los servicios de atención a la salud (ej. cobertura insuficiente, baja calidad de la atención); y
- (6) de comportamiento (ej. fumar, drogadicción).

En términos generales, se considera que los factores de riesgo asociados causalmente con un daño a la salud permiten acciones de salud múltiples, incluyendo las de prevención primaria. Aquellos factores de riesgo asociados no causalmente pueden ser considerados para actividades de prevención secundaria y son de gran auxilio en la identificación de los grupos humanos con mayor riesgo de tener el daño a la salud estudiado.

Para poder distinguir o reconocer los individuos o grupos humanos con mayor riesgo de enfermar o de morir prematuramente de un daño a la salud estudiado, se hace uso de diversas medidas epidemiológicas de riesgo. Estas medidas son: (a) el riesgo absoluto; (b) el riesgo relativo; y (c) el riesgo atribuible. La revisión de estos conceptos, así como la medición de los mismos son cubiertas en el apartado siguiente.

## 2.2 Riesgo Absoluto, Riesgo Relativo y Riesgo Atribuible

Basándose en el conocimiento acerca de las características presentes en los individuos y grupos humanos que experimentan el daño a la salud estudiado, se inicia una búsqueda metodológica cuyo objetivo primario es poder calcular la probabilidad que los individuos con ciertas características experimenten el daño a la salud en el futuro, y de esta forma poder tomar las medidas preventivas necesarias. Esta búsqueda implica la noción de la distribución del riesgo de enfermar o morir en poblaciones definidas.

Recordando el concepto de incidencia, podemos decir que incidencia es el número de nuevos casos de una enfermedad, durante un período definido y en una población determinada. La incidencia es una medida del riesgo que tiene una población dada de enfermar. Mide el paso de estar sano al de estar enfermo. Incidencia es pues, sinónimo de riesgo absoluto. Esta medida del riesgo absoluto o incidencia es la medida básica para el cálculo del Riesgo Relativo y del Riesgo Atribuible.

El Riesgo Relativo es una medida de la probabilidad de experimentar un daño a la salud, que tienen aquellos individuos y grupos humanos con un determinado factor de riesgo en comparación con aquellos que no tienen ese factor.

El Riesgo Relativo es una razón de riesgos absolutos. Es decir, es una razón en que en el numerador se tiene a la incidencia del grupo expuesto a un factor de riesgo y en el denominador se tiene la incidencia del grupo NO expuesto a tal factor. Esta medida nos indica que tanto mayor

es el riesgo de enfermar en aquellos con la característica o factor de riesgo con respecto a los que no la tienen (2). Así, por ejemplo, si el riesgo relativo asociado con la presencia de un factor de riesgo es de 3.25, esto significa que la probabilidad para desarrollar o tener la enfermedad es 3.25 veces mas alta en los que tienen dicho factor que en los que no la tienen. (ver cuadro 1.1)

Una característica importante del riesgo relativo es el hecho que mide la FUERZA DE ASOCIACION entre un factor de riesgo y un determinado daño a la salud. De ahí que esta medida sea ampliamente usada en la búsqueda de la causalidad o determinación de las enfermedades y de su etiología.

El riesgo relativo así como el riesgo atribuible son medidas de asociación del exceso de riesgo debido a la exposición o presencia de un factor de riesgo determinado.

El riesgo atribuible mide el exceso de riesgo (incidencia) que puede ser atribuido a la exposición a un factor (por ejemplo, vivienda incompleta, desnutrición, tabaquismo, etc.). Se calcula como una diferencia (substracción) de dos tasas de incidencia. Es decir, a la tasa de incidencia del grupo expuesto al factor de riesgo se le subtrae o resta la tasa de incidencia del grupo no expuesto a dicho factor. El exceso observado en las tasas es el exceso de riesgo debido al factor en cuestion. Usualmente, con la idea de facilitar la interpretación del riesgo atribuible, esta medida se calcula como porcentaje. Así, por ejemplo en el cálculo presentado en el cuadro 1.1, se puede interpretar que 69.25% es el exceso de riesgo de malaria que tiene el grupo expuesto a

viviendas cercanas a criaderos. Este es el llamado riesgo atribuible en los expuestos. Un concepto de mayor importancia en salud pública es el denominado riesgo atribuible en la población. En esta medida, se toma en cuenta la prevalencia del factor de riesgo en toda la población y su interpretación estaría en términos del beneficio potencial que se esperaría en la comunidad (no solo en el grupo expuesto) de disminución del riesgo si la exposición a tal factor se redujera en la población, dado que el factor fuera causal.

El riesgo atribuible en la población total es una medida de asociación influenciada por la prevalencia del factor en la población total. El riesgo atribuible en la población mide el descenso porcentual en el número de casos o muertes que pueden evitarse si el factor causal fuera eliminado o neutralizado totalmente.

A continuación se presentan las fórmulas y cálculos necesarios para estimar el riesgo relativo y el riesgo atribuible.

$$RR = \frac{IE}{INE}$$

$$\text{RIESGO RELATIVO} = \frac{\text{Incidencia de los expuestos al Factor}}{\text{Incidencia de los No expuestos al Factor}}$$

$$\text{RIESGO ATRIBUIBLE} = \text{Incidencia de expuestos} - \text{Incidencia en los No expuestos}$$

$$RAE = IE - INE$$

PORCENTAJE DE Incidencia de expuestos -  
RIESGO ATRIBUIBLE =  $\frac{\text{Incidencia de No expuestos}}{\text{Incidencia de expuestos}}$  X 100  
EN LOS EXPUESTOS  
$$\% \text{ RAE} = \frac{\text{IE} - \text{INE}}{\text{IE}} \text{ X } 100$$

PORCENTAJE DE Incidencia en la población total -  
RIESGO ATRIBUIBLE =  $\frac{\text{Incidencia de No expuestos}}{\text{Incidencia en el total de la población}}$  X 100  
EN LA POBLACION  
$$\% \text{ RAP} = \frac{\text{Ipt} - \text{INE}}{\text{Ipt}} \text{ X } 100$$

Una formula alternativa es la siguiente:

PORCENTAJE DE  
RIESGO ATRIBUIBLE =  $\frac{P (\text{RR}-1)}{1 + P (\text{RR}-1)}$  x 100  
EN LA POBLACION

donde: P es igual a la proporción del factor de riesgo en la población;  
RR es igual a riesgo relativo.

Utilizando la información del cuadro 1.1, se presentan los cálculos de las medidas de riesgo previamente referidas.

Cuadro 1

Relación entre casos de Malaria y Vivienda Cercana a Criaderos  
de Mosquitos Transmisores en la localidad X, 1988.

Característica	Casos de Malaria		Total
	S I	N O	
Vivienda a < 500 mts de los Criaderos	202	2016	2218
Vivienda a > 500 mts de los Criaderos	81	2811	2892
Total	283	4827	5110

Tasa de incidencia en expuestos  $202/2218 = 91.07$  por 1,000

Tasa de incidencia en No expuestos  $81/2892 = 28.00$  por 1,000

$$\text{RIESGO RELATIVO} = \frac{91.07}{28.00} = 3.25$$

$$\text{RIESGO ATRIBUIBLE} = 91.07 - 28.00 = 63.07$$

(en expuestos)

$$\% \text{ RIESGO ATRIBUIBLE} = \frac{91.07 - 28.00}{91.07} \times 100 = 69.25 \%$$

(en expuestos) 91.07

$$\% \text{ RIESGO ATRIBUIBLE} = \frac{0.434 (3.25-1)}{1 + 0.434 (3.25-1)} \times 100 = 49.4\%$$

EN LA POBLACION

donde P = 2218/5110

RR = 3.25

Según los resultados presentados en el ejemplo anterior, para esta población y en el tiempo dado, el riesgo de enfermar de malaria es 3.25 veces mayor para los individuos que viven a menos de 500 metros de los criaderos de vectores, comparados con los que viven mas alejados de dichos criaderos. Para los expuestos (los que viven cercanos de los criaderos), hasta el 69.25% de su problema de malaria se explicaría por este factor de riesgo. Por otro lado, hasta el 49.4% del problema de malaria en esta población se resolvería si dicho factor de riesgo es causal y fuera eliminado.

Es importante destacar que hay otras posibles explicaciones para los resultados, y que la presencia de otros factores concomitantes puede interferir en la actuación del factor de riesgo en la determinación de la ocurrencia de malaria. Para la correcta determinación de los resultados, dichos factores deben ser reconocidos y tomados en cuenta en el análisis.

REFERENCIAS

Last, J.M. (ed) A Dictionary of Epidemiology. Second Edition. New York, Oxford University Press. 1988.

NO. 1000000000  
SERIAL 0000000000  
FEB 10 1988  
FARMACIA I  
SAUD (OPS)

1525I  
Dic.90

FAVOR NO REPRODUCIR  
SIN LA AUTORIZACION  
DE LA ORGANIZACION  
PANAMERICANA DE LA  
SALUD (OPS)

A N E X O I

M A L A R I A N D A :

SITUACION SOCIODEMOGRAFICA E INDICADORES

MALARIOMETRICOS BASICOS.

**NOTA:**

LA INFORMACION BASICA PRESENTADA EN ESTE ANEXO NO REPRESENTA A NINGUN PAIS PARTICULAR O REAL. EL PROPOSITO DE LA INCLUSION DE ESTA INFORMACION ES FACILITAR EL PROCESO DE ENSEÑANZA-APRENDIZAJE ASI COMO LA REALIZACION DE LOS EJERCICIOS CONTEMPLADOS EN LOS MATERIALES

FAVOR NO REPRODUCIR  
SIN LA AUTORIZACION  
DE LA ORGANIZACION  
PANAMERICANA DE LA  
SALUD (OPS)

## CONTENIDO

### INTRODUCCION A MALARIANDA

#### I. INTRODUCCION

#### II. CARACTERISTICAS GENERALES DE LAS REGIONES DE MALARIANDA.

##### A. REGION COSTA NORTE

- A.1 Temperatura
- A.2 Altitud
- A.3 Precipitación pluviométrica
- A.4 Población
- A.5 Actividad económica
- A.6 Infraestructura de Salud
- A.7 Situación de la Malaria

##### B. REGION NORESTE (SELVA)

- B.1 Temperatura
- B.2 Altitud
- B.3 Precipitación pluviométrica
- B.4 Población
- B.5 Vías de Comunicación
- B.6 Actividad económica
- B.7 Infraestructura de Salud
- B.8 Situación de la Malaria

##### C. REGION SURESTE (ALTIPLANO)

- C.1 Temperatura
- C.2 Altitud
- C.3 Precipitación pluviométrica
- C.4 Población
- C.5 Actividad económica
- C.6 Infraestructura de Salud
- C.7 Situación de la Malaria

##### D. REGION COSTA SUR

- D.1 Temperatura
- D.2 Altitud
- D.3 Precipitación pluviométrica
- D.4 Población
- D.5 Actividad económica
- D.6 Infraestructura de Salud
- D.7 Situación de la Malaria

**LISTADO DE CUADROS DE MALARIANDA**

**CUADRO 1. DISTRIBUCION DE LA POBLACION Y DENSIDAD DE POBLACION POR REGIONES, MALARIANDA. 1988.**

**CUADRO 2. DISTRIBUCION DE LA POBLACION Y SUPERFICIE, AREAS URBANA Y RURAL, POR MUNICIPIOS EN AREAS MALARICAS DE MALARIANDA, 1988.**

**CUADRO 3. DISTRIBUCION DE POBLACION URBANA Y RURAL, POR REGION Y MUNICIPIO, AREA MALARICA, MALARIANDA. 1988.**

**CUADRO 4. DISTRIBUCION DE MUESTRAS DE SANGRE EXAMINADAS Y LA POSITIVAS SEGUN ESPECIE PARASITARIA POR REGION Y MUNICIPIO, MALARIANDA, 1988.**

**CUADRO 5. DISTRIBUCION DE LA POBLACION, POR REGION Y MUNICIPIO, AREAS MALARICAS DE MALARIANDA, 1984-1988.**

**CUADRO 6. DISTRIBUCION DE LOS CASOS DE MALARIA POR REGION Y MUNICIPIO, MALARIANDA, 1984-1988.**

## INTRODUCCION A MALARIANDA

### I. INTRODUCCION

Malarianda es un país de 1 millón de habitantes, con una superficie de 42,931 kilómetros cuadrados. Está dividido en cinco regiones, cada una con un perfil geográfico y sociopolítico distintos. Estas regiones son: 1) Costa Norte, 2) Noreste (Selva), 3) Sureste (Altiplano), 4) Costa Sur, y 5) Oriente (Cordillera).

En 1988 se estimaba que Malarianda tenía alrededor del 40 por ciento de su población habitando en áreas originalmente maláricas, con condiciones socioeconómicas, ambientales, culturales y políticas problemáticas. Según estimaciones oficiales, de la población total, existen alrededor de 200,000 habitantes sin acceso permanente a los servicios de salud. De estos, al menos 54 por ciento se ubican en áreas donde existe transmisión activa de malaria o riesgo de que ésta ocurra.

El área potencialmente malárica del país comprende la región de la selva, parte de las regiones costeras norte y sur, y parte del altiplano.

Tradicionalmente, el problema de malaria estaba a cargo del Servicio Nacional de Erradicación de la Malaria (SNEM). Los servicios generales de salud se limitaban a dar tratamiento a un porcentaje limitado de casos graves o complicados. A partir de 1987, se inició un proceso de descentralización en el que el SNEM fue incorporado formalmente a los servicios generales de salud. La integración se está dando de manera gradual y lenta.

Como en otros países del continente, en Malarianda se observan intensos movimientos migratorios internos, de una región a otra y del campo a la ciudad. En la regiones norte y noreste se observa además migración entre los países fronterizos y Malarianda.

### II. CARACTERISTICAS GENERALES DE LAS REGIONES DE MALARIANDA.

#### A. REGION COSTA NORTE

A.1 Temperatura: oscila entre los 20-32 grados centígrados, humedad elevada.

A.2 Altitud: máxima de 200 mts.

A.3 Precipitación pluviométrica: elevada, con picos máximos de 300-500 mm. entre enero y abril.

- A.4 Población: predominantemente rural (60%), con tendencia a la urbanización. Intensa migración proveniente del resto de las regiones y países vecinos realizada de manera incoordinada.
- A.5 Actividad económica: área de implementación de grandes proyectos agrícolas y de explotación de mineral, especialmente extracción de oro de aluvión en los ríos. Esta región cuenta con un puerto marítimo de pequeña escala que se ubica en la desembocadura del Río Verde y que permite comunicación con la región de la selva. Actividad pesquera moderada.
- A.6 Infraestructura de Salud: se encuentra en expansión pero no al ritmo del crecimiento poblacional. La distribución geográfica de los servicios de salud y vías de acceso no son homogéneas en toda la región.
- A.7 Situación de la Malaria: la transmisión malárica ha sido un problema de salud pública durante varias décadas. El control de la malaria se ha planteado a largo plazo. Actualmente se realizan operaciones de rociamiento intra-domiciliario con DDT y tratamiento de los casos con cloroquina y primaquina.

Esta región registra el IPA mas elevado del país; en los últimos años ha mostrado un incremento importante, doblándose en los últimos cuatro años. A partir de hace tres años se ha intensificado la búsqueda de casos y el tratamiento, incrementándose el acceso al diagnóstico parasitológico y doblando el número de recursos humanos para las acciones de rociamiento. No se ha ampliado la capacidad instalada de atención a los casos graves. Se observa una mayor proporción de *P. falciparum* y un aumento en la resistencia de estas infecciones a las drogas antimaláricas. El número de casos de malaria registrados se ha triplicado entre 1984 y 1988.

## B. REGION NORESTE (SELVA)

- B.1 Temperatura : 22-32 grados centígrados durante todo el año.
- B.2 Altitud: máxima 200 mts. Esta región es el área selvática del país.
- B.3 Precipitación pluviométrica: lluvias durante todo el año, con los niveles mas altos en los meses de enero-abril 400-600 mm, y los mas bajos en junio-agosto 20-40 mm.
- B.4 Población: El 38% de la población reside en zonas urbanas. Intensa migración a las ciudades en los últimos 5 años, incrementando las áreas periféricas de las mismas y la formación de barrios "miseria" de las ciudades. El 62% restante lo constituye la población rural dispersa a lo largo de los ríos y en proyectos de colonización con desforestación de la selva.
- B.5 Vías de comunicación: deficientes, durante una época del año existe aislamiento de los poblados con el resto de las otras regiones.

Dos ríos y sus afluentes cruzan gran parte de la región. Estos ríos son utilizados como vías de comunicación y transporte.

- B.6 Actividad económica: área de subdesarrollo económico y social. Actividad de extracción de madera, agricultura de subsistencia, algunos proyectos de colonización para incentivar la agricultura. Existe además, tráfico de drogas y contrabando.
- B.7 Infraestructura de salud: precaria, concentrada en los centros poblacionales mas grandes. Existen algunas unidades móviles (en barcos) para asistir a poblaciones localizadas en las riberas de los ríos.
- B.8 Situación de la Malaria: Esta región ha tenido durante varias décadas alta incidencia de malaria. Después de haber logrado una reducción de la IPA con medidas ambientales y la campaña de erradicación, en los años 80's se observa un incremento sostenido de la malaria. El control se realiza con una gran red de trabajadores voluntarios que recolectan muestras de sangre y distribuyen tratamiento a nivel domiciliario. Hay dificultad para el diagnóstico y referencia a los servicios de atención para los casos graves. En esta región se encuentra un instituto de investigación de enfermedades transmisibles, interesado en el estudio del comportamiento de la malaria en la región.

#### C. REGION SURESTE (ALTIPLANO)

- C.1 Temperatura: temperaturas mínimas de 2 grados centígrados en la montaña durante junio-julio, hasta 38 grados centígrados en diciembre-febrero, en las áreas mas bajas. Clima sub-tropical, con grandes cambios climáticos.
- C.2 Altitud: el 75% del territorio se ubica entre los 200-800 mts sobre el nivel del mar y el resto de 800-1200 mts sobre el nivel del mar.
- C.3 Precipitación pluviométrica: baja precipitación pluviométrica con picos en noviembre y enero de 100-400 mm y bastante seco en los meses de junio-julio 4-12 mm.
- C.4 Población: 65% de la población radica en áreas urbanas. Urbanización acelerada y desordenada. Envejecimiento relativo de la población.
- C.5 Actividad económica: desarrollo económico acentuado, industrialización rápida, actividades de exportación. Proyectos agropecuarios con mecanización. Mejor infraestructura y mejores indicadores socio-económicos.

- C.6 Infraestructura de salud: los servicios generales de salud se ubican en las áreas urbanas y disponen de recursos modernos y de tecnología avanzada. Así mismo, existe una gran red privada de servicios de salud. Los centros de atención primaria tienen poca capacidad resolutive.
- C.7 La situación de la Malaria: con la mejoría de las condiciones socioeconómicas de la población y la campaña de erradicación de la malaria, se logró interrumpir la transmisión en todos los municipios, excepto dos de ellos. En los últimos cinco años, la transmisión de la malaria se reinstaló en una tercera parte del territorio. Actualmente el 28% de la población vive en áreas maláricas.

#### D. REGION COSTA SUR

- D.1 Temperatura: clima tropical, oscila entre 21.5 y 33 grados centígrados, caliente entre diciembre y marzo, fresco en los meses de junio-julio.
- D.2 Altitud: 85% de la superficie se encuentra entre los 0-800 mts, el resto entre 800-1200 mts.
- D.3 Precipitación pluviométrica: 300-500 mm entre marzo-mayo. De 2-5 mm entre noviembre-diciembre. El 30% de la región es área semidesértica.
- D.4 Población: 51% de la población es urbana y el resto rural. Se observa migración hacia la ciudad y hacia la Región Costa Norte.
- D.5 Actividad económica: industria pesquera (langosta y camarón), textil, artesanía. Agricultura algodón, caña de azúcar y frutales. Ganadería predominantemente bovino y caprino; explotación de petróleo y sal marina. Tiene el puerto marítimo mas importante del país. Gran concentración de la riqueza, grandes conflictos por la tenencia o propiedad de la tierra. Grandes grupos poblacionales marginales sin acceso a bienes y servicios. La región tiene los indicadores socio-económicos mas precarios del país.
- D.6 Infraestructura de salud: gran red de centros de servicios de atención primaria tanto en la zona rural como urbana, con poca capacidad resolutive. Existen planes de regionalización y descentralización de la atención médica. Se mantiene la concentración de recursos y servicios en las sedes de los municipios mas grandes.
- D.7 Situación de la Malaria: esta región constituía en el pasado, una área malárica; con los esfuerzos de la campaña de erradicación, muchas áreas se convirtieron en áreas libres de malaria. En los últimos cinco años algunos casos de malaria fueron diagnosticados y caracterizados como probables casos importados.

E. REGION ORIENTE (CORDILLERA)

- E.1 Temperatura: clima frío y seco durante la mayor parte del año, oscila entre 4 y 26 grados centígrados. Las temperaturas mas bajas se observan en los meses de mayo-agosto.
- E.2 Altitud: 35% de la superficie se encuentra entre los 500-800 mts sobre el nivel del mar, el restante 65% entre 800-1500 mts.
- E.3 Precipitación pluviométrica: alta precipitación en diciembre y enero de 200-260 mm y baja en los meses restantes (20-80mm).
- E.4 Población: 55% de la población radica en áreas rurales, el resto reside en núcleos urbanos pequeños.
- E.5 Actividad económica: desarrollo económico incipiente. Ganadería de pequeña escala (carne, lácteos y lana); agricultura de subsistencia. Minas en desactivación y desempleo creciente.
- E.6 Infraestructura de salud: pocas unidades de servicios de salud con algunos centros de atención primaria, sin suficiente personal de salud. La población usualmente demanda los servicios de la región vecina (altiplano), mas desarrollada.
- E.7 La situación de la Malaria: durante las últimas décadas no se han reportado casos autóctonos de malaria en la región.

CUADRO 1

DISTRIBUCION DE LA POBLACION Y DENSIDAD DE POBLACION  
POR REGIONES, MALARIANDA. 1988.

Región	Población Area Malárica	Población Area no Malárica	Población Total	Densidad Población hab/km2
REGION COSTA NORTE	45350	13059	58409	21.91
REGION NORESTE (SELVA)	122386	---	122386	19.12
REGION SURESTE (ALTIPLANO)	97934	254410	352344	20.18
REGION COSTA SUR	150005	181590	331595	25.13
REGION ORIENTE (CORDILLERA)	---	155676	155676	16.96
TOTAL GENERAL	415675	604735	1020410	23.77

Fuente: Censo de Población. Proyecciones 1981-1989. Malarianda 1985.

**CUADRO 2**  
**DISTRIBUCION DE LA POBLACION Y SUPERFICIE, AREAS URBANA Y**  
**RURAL, POR MUNICIPIOS EN AREAS MALARICAS DE MALARIANDA, 1988**

Nombre Munici- pio	Población del Area Malárica			Superficie en Km <sup>2</sup>		
	Total	Urbana	Rural	Urbana	Rural	Total
<b>REGION COSTA NORTE</b>						
San Jose	13994	5578	8366	47	480	527
Las Rosas	10880	4352	6528	35	800	835
Golondrina	20476	8461	12285	68	640	708
<b>Total</b>	<b>45350</b>	<b>18391</b>	<b>26959</b>	<b>150</b>	<b>1920</b>	<b>2070</b>
<b>REGION NORESTE (SELVA)</b>						
Aguales	3819	1451	2368	10	255	265
Romeros	1257	478	779	4	230	234
Río Verde	19011	7224	11787	58	2150	2208
San Miguel	46868	17809	29059	142	2505	2647
Tarama	2371	901	1470	7	230	237
Monte de Sal	49060	18643	30417	149	660	809
<b>Total</b>	<b>122386</b>	<b>46506</b>	<b>75880</b>	<b>370</b>	<b>6030</b>	<b>6400</b>
<b>REGION SURESTE (ALTIPLANO)</b>						
Santa Rosa	17929	11654	6275	81	1065	1146
Los Cazos	17107	11120	5987	77	581	658
Santo Domingo	22667	14734	7933	109	1220	1329
Villa Union	23788	15462	8326	107	370	477
Marquez de Río	9969	6479	3490	51	684	735
Pueblo Viejo	6474	4208	2266	36	472	508
<b>Total</b>	<b>97934</b>	<b>63657</b>	<b>34277</b>	<b>461</b>	<b>4392</b>	<b>4853</b>
<b>REGION COSTA SUR</b>						
Los Gatos	8140	4151	3989	36	107	143
Los Moros	6724	3429	3295	27	78	105
San Mateo	70836	36126	34710	289	1833	2122
El Paso	10510	5360	5150	43	130	173
San Bartolome	11427	5828	5599	42	180	222
Santa Lucía	42368	21608	20760	160	340	500
<b>Total</b>	<b>150005</b>	<b>76502</b>	<b>73503</b>	<b>597</b>	<b>2668</b>	<b>3265</b>
<b>TOTAL AREAS MALARICAS</b>	<b>415675</b>	<b>205056</b>	<b>210619</b>	<b>1578</b>	<b>15010</b>	<b>16588</b>

Fuente: Estadísticas de Salud. Programa de Malaria. Ministerio de Salud. Malarianda 1989.

**CUADRO 3. DISTRIBUCION DE LA DENSIDAD DE POBLACION URBANA Y RURAL, POR REGION Y MUNICIPIO, AREA MALARICA, MALARIANDA. 1988.**

Nombre Munici- pio	Población Area Ma- larica	Densidad Poblacional (hab/km <sup>2</sup> )		
		Gene- ral	Zona Urbana	Zona Rural
<b>REGION COSTA NORTE</b>				
San José	13994	26.5	120.0	16.1
Las Rosas	10880	13.0	186.5	8.2
Golondrina	20476	28.9	124.4	19.2
<b>Total</b>	<b>45350</b>	<b>21.9</b>	<b>122.6</b>	<b>14.1</b>
<b>REGION NORESTE (SELVA)</b>				
Aguales	3819	14.4	145.1	9.3
Romeros	1257	5.4	119.5	3.4
Rio Verde	19011	8.6	124.5	5.5
San Miguel	46868	17.7	124.7	11.6
Tarama	2371	10.0	124.8	6.4
Monte de Sal	49060	60.6	124.8	46.2
<b>Total</b>	<b>122386</b>	<b>19.1</b>	<b>125.7</b>	<b>13.6</b>
<b>REGION SURESTE (ALTIPLANO)</b>				
Santa Rosa	17929	15.6	143.9	5.9
Los Cazos	17107	26.0	144.4	10.3
Santo Domingo	22667	17.1	135.0	6.5
Villa Unión	23788	49.9	144.0	22.5
Marquez de Rio	9969	13.6	127.1	5.1
Pueblo Viejo	6474	12.7	117.7	4.8
<b>Total</b>	<b>97934</b>	<b>20.2</b>	<b>138.1</b>	<b>7.8</b>
<b>REGION COSTA SUR</b>				
Los Gatos	8140	56.9	116.9	37.1
Los Moros	6724	64.0	127.0	42.1
San Mateo	70836	33.4	124.9	18.9
El Paso	10510	60.8	124.6	39.5
San Bartolome	11427	51.5	140.0	31.3
Santa Lucia	42368	84.7	134.4	60.6
<b>Total</b>	<b>150005</b>	<b>45.9</b>	<b>129.1</b>	<b>27.6</b>
<b>TOTAL AREAS</b>				
<b>MALARICAS</b>	<b>415675</b>	<b>25.1</b>	<b>129.9</b>	<b>14.1</b>

Fuente: Estadísticas de Salud. Programa de Malaria. Ministerio de Salud. Malarianda 1989.

**CUADRO 4. DISTRIBUCION DE MUESTRAS DE SANGRE EXAMINADAS Y POSITIVAS SEGUN ESPECIE PARASITARIA, POR REGION Y MUNICIPIO, MALARIANDA, 1988.**

Nombre Municipio	Población Area Malarica	Muestras de Sangre Examinadas	Positivas	Tipo de Plasmodio		
				falc.	vivax	f+v
<b>REGION COSTA NORTE</b>						
San Jose	13994	7215	3305	1680	1615	10
Las Rosas	10880	1145	71	38	32	1
Golondrina	20476	2692	290	130	160	0
<b>Total</b>	<b>45350</b>	<b>11052</b>	<b>3666</b>	<b>1848</b>	<b>1807</b>	<b>11</b>
<b>REGION NORESTE (SELVA)</b>						
Aguales	3819	793	219	122	96	1
Romeros	1257	484	189	64	124	1
Rio Verde	19011	1033	152	45	106	1
San Miguel	46868	4964	1191	716	461	14
Tarama	2371	279	75	31	43	1
Monte de Sal	49060	3681	566	306	251	9
<b>Total</b>	<b>122386</b>	<b>11234</b>	<b>2392</b>	<b>1284</b>	<b>1081</b>	<b>27</b>
<b>REGION SURESTE (ALTIPLANO)</b>						
Santa Rosa	17929	1027	218	88	130	0
Los Cazos	17107	756	17	5	12	0
Santo Domingo	22667	523	10	3	6	1
Villa Unión	23788	596	12	6	6	0
Marquez de Rio	9969	627	7	3	4	0
Pueblo Viejo	6474	308	31	10	21	0
<b>Total</b>	<b>97934</b>	<b>3837</b>	<b>295</b>	<b>115</b>	<b>179</b>	<b>1</b>
<b>REGION COSTA SUR</b>						
Los Gatos	8140	72	1	0	1	0
Los Moros	6724	31	1	1	0	0
San Mateo	70836	3743	5	3	2	0
El Paso	10510	204	2	1	1	0
San Bartolome	11427	166	2	1	1	0
Santa Lucia	42368	521	4	1	3	0
<b>Total</b>	<b>150005</b>	<b>4737</b>	<b>15</b>	<b>7</b>	<b>8</b>	<b>0</b>
<b>TOTAL AREAS MALARICAS</b>						
	<b>415675</b>	<b>30860</b>	<b>6368</b>	<b>3254</b>	<b>3075</b>	<b>39</b>

Fuente: Estadísticas de Salud. Programa de Malaria. Ministerio de Salud. Malarianda 1989.

**CUADRO 5. DISTRIBUCION DE LA POBLACION, POR REGION Y MUNICIPIO, AREAS MALARICAS DE MALARIANDA, 1984-1988.**

Nombre Munici- pio	P o b l a c i ó n ( á r e a m a l á r i c a )				
	1988	1987	1986	1985	1984
<b>REGION COSTA NORTE</b>					
San Jose	13994	13083	12231	11434	10690
Las Rosas	10880	10370	9884	9649	9418
Golondrina	20476	19953	19430	18929	18438
<b>Total</b>	<b>45350</b>	<b>43406</b>	<b>41545</b>	<b>40012</b>	<b>38546</b>
<b>REGION NORESTE (SELVA)</b>					
Aguales	3819	3706	3597	3490	3387
Romeros	1257	1187	1148	1091	1050
Rio Verde	19011	18326	17841	17281	16744
San Miguel	46868	45229	43791	42121	40948
Tarama	2371	2293	2238	2145	2095
Monte de Sal	49060	47816	46690	45425	44277
<b>Total</b>	<b>122386</b>	<b>118557</b>	<b>115305</b>	<b>111552</b>	<b>108501</b>
<b>REGION SURESTE (ALTIPLANO)</b>					
Santa Rosa	17929	17481	17147	16620	16409
Los Cazos	17107	16865	16625	15881	15606
Santo Domingo	22667	22145	22029	21786	21411
Villa Unión	23788	23357	22914	22519	21989
Marquez de Rio	9969	9760	9565	9355	9189
Pueblo Viejo	6474	6412	6332	6289	6239
<b>Total</b>	<b>97934</b>	<b>96020</b>	<b>94612</b>	<b>92450</b>	<b>90743</b>
<b>REGION COSTA SUR</b>					
Los Gatos	8140	7982	7828	7674	7516
Los Moros	6724	6591	6461	6333	6197
San Mateo	70836	69431	68057	66702	65374
El Paso	10510	10323	10130	9957	9787
San Bartolome	11427	11267	11109	10954	10791
Santa Lucía	42368	41694	41046	40355	39683
<b>Total</b>	<b>150005</b>	<b>147288</b>	<b>144631</b>	<b>141975</b>	<b>139348</b>
<b>TOTAL GENERAL</b>	<b>415675</b>	<b>405271</b>	<b>396093</b>	<b>385989</b>	<b>377138</b>

\* Area potencialmente malárica.

Fuente: Estadísticas de Salud. Programa de Malaria. Ministerio de Salud. Malarianda 1989.

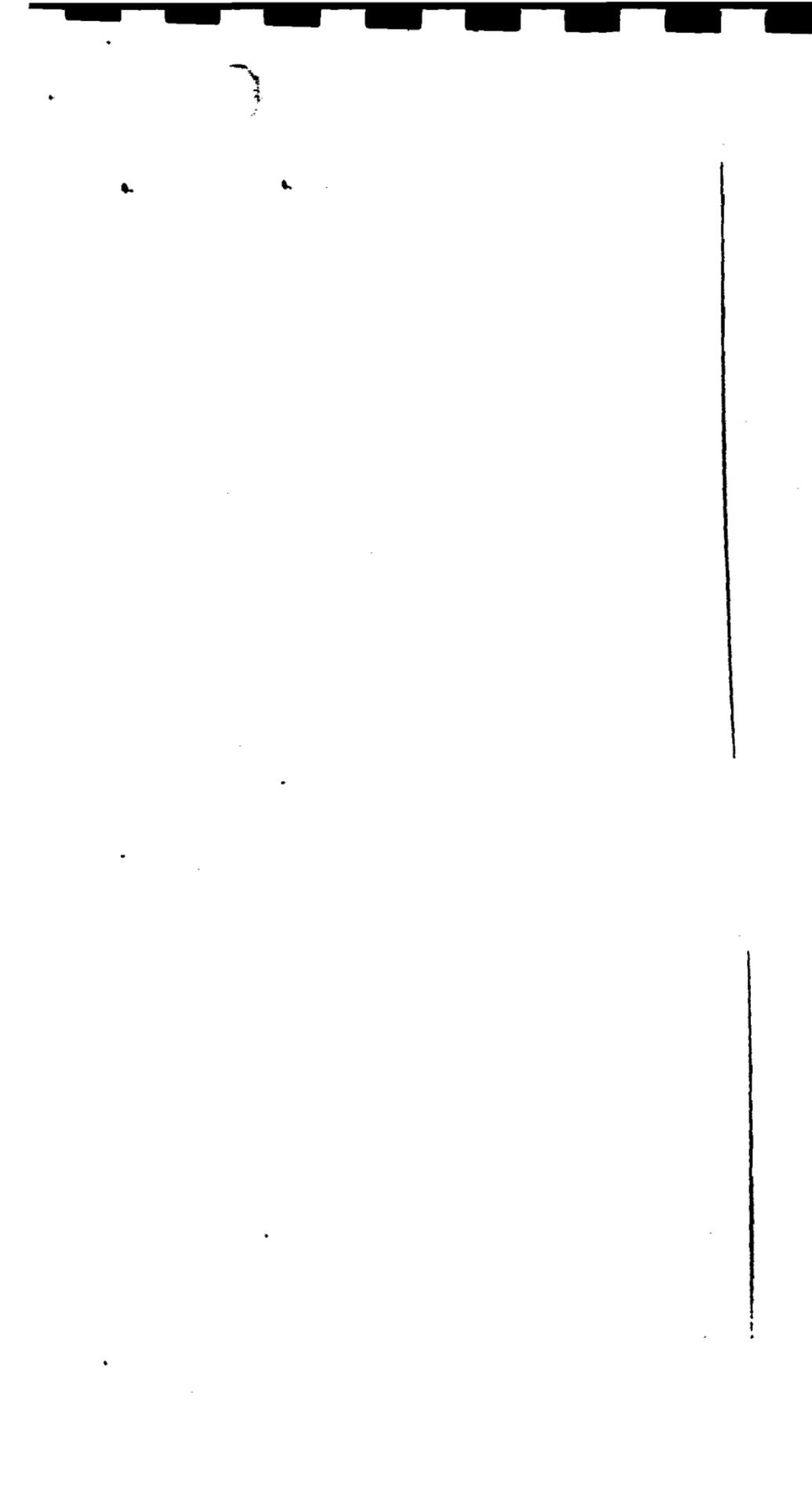
**CUADRO 6. DISTRIBUCION DE LOS CASOS DE MALARIA POR REGION Y MUNICIPIO, MALARIANDA, 1984-1988.**

Nombre Munici- pio	C a s o s d e				
	1984	1985	1986	1987	1988
<b>REGION COSTA NORTE</b>					
San José	1045	1432	2687	3019	3305
Las Rosas	89	85	61	70	71
Golondrina	118	149	195	267	290
<b>Total</b>	<b>1252</b>	<b>1666</b>	<b>2943</b>	<b>3356</b>	<b>3666</b>
<b>REGION NORESTE (SELVA)</b>					
Aguales	76	123	121	105	219
Romeros	184	192	183	181	189
Rio Verde	140	110	91	119	152
San Miguel	771	1254	1390	1223	1191
Tarama	55	51	79	80	75
Monte de Sal	184	226	253	390	566
<b>Total</b>	<b>1410</b>	<b>1956</b>	<b>2117</b>	<b>2098</b>	<b>2392</b>
<b>REGION SURESTE (ALTIPLANO)</b>					
Santa Rosa	197	187	132	152	218
Los Cazos	1	14	15	18	17
Santo Domingo	-	-	1	15	10
Villa Unión	-	2	12	11	12
Marquez de Río	-	-	-	4	7
Pueblo Viejo	13	21	20	28	31
<b>Total</b>	<b>211</b>	<b>224</b>	<b>180</b>	<b>228</b>	<b>295</b>
<b>REGION COSTA SUR</b>					
Los Gatos	2	1	-	-	1
Los Moros	1	-	1	2	1
San Mateo	4	2	7	4	5
El Paso	2	1	-	-	2
San Bartolomé	-	1	-	1	2
Santa Lucía	4	3	2	6	4
<b>Total</b>	<b>13</b>	<b>8</b>	<b>10</b>	<b>13</b>	<b>15</b>
<b>TOTAL GENERAL</b>	<b>2886</b>	<b>3854</b>	<b>5250</b>	<b>5695</b>	<b>6368</b>

Fuente: Estadísticas de Salud. Programa de Malaria. Ministerio de Salud. Malarianda 1989.

---

**Anexo I**  
**Ficticia: Situación de salud y recursos**  
**para la prestación de servicios de salud**  
**en una republica imaginaria**



**Ficticia: Situación de salud y recursos  
para la prestación de servicios de salud  
en una república imaginaria**

**Contenido**

<b>I. Introducción a Ficticia</b>	
a) Mapa de Ficticia (figura 1) .....	193
<b>II. Perfil sociodemográfico</b>	
a) Población por edad, sexo y región (cuadro 1) .....	194
b) Distribución de la población por región (cuadro 2) .....	195
c) Descripción de las áreas urbanas en las regiones (cuadro 3) .....	196
d) Características regionales con relación a algunos aspectos de importancia sociomédica (cuadro 4) .....	197
e) Tasas de fecundidad general por región y por tipo de asentamiento de población (cuadros 5a y 5b) .....	198
f) Tasas de fecundidad específicas para la edad en las áreas urbanas marginales (cuadro 6) .....	198
<b>III. Situación general de salud</b>	
a) Estimaciones de algunas estadísticas vitales (cuadro 7) .....	199
b) Variaciones regionales en las tasas de mortalidad (cuadro 8) .....	200
c) Variaciones de las tasas de mortalidad según tipo de asentamiento de población (cuadro 9) .....	200
d) Principales causas de mortalidad y morbilidad .....	201
e) Tasas de mortalidad específica por peso al nacer (cuadro 10) .....	202
f) Variaciones en las tasas de mortalidad perinatal por peso al nacer, según región y tipo de asentamiento de la población (cuadro 11) .....	202
g) Causas de muerte perinatal y neonatal tardía (cuadro 12) .....	203
h) Causas de muerte en niños con peso al nacer de 2.500 gramos o más (cuadro 13) .....	203
i) Diferencias regionales en las causas de muerte perinatal (cuadro 14) .....	204
j) Percepción pública de los problemas de salud .....	204
<b>IV. Recursos para la prestación de servicios de salud</b>	
a) La atención a la salud en Ficticia .....	205
b) Sistemas de prestación de servicios de salud .....	205
c) Algunas características de las fuentes de recursos humanos para la salud (cuadro 15) .....	207
d) Estructura administrativa de los servicios de salud en Ficticia (figura 2) .....	208
e) Número de establecimientos y personal del servicio de salud gubernamental por tipo de establecimiento para cada región (cuadro 16) .....	209

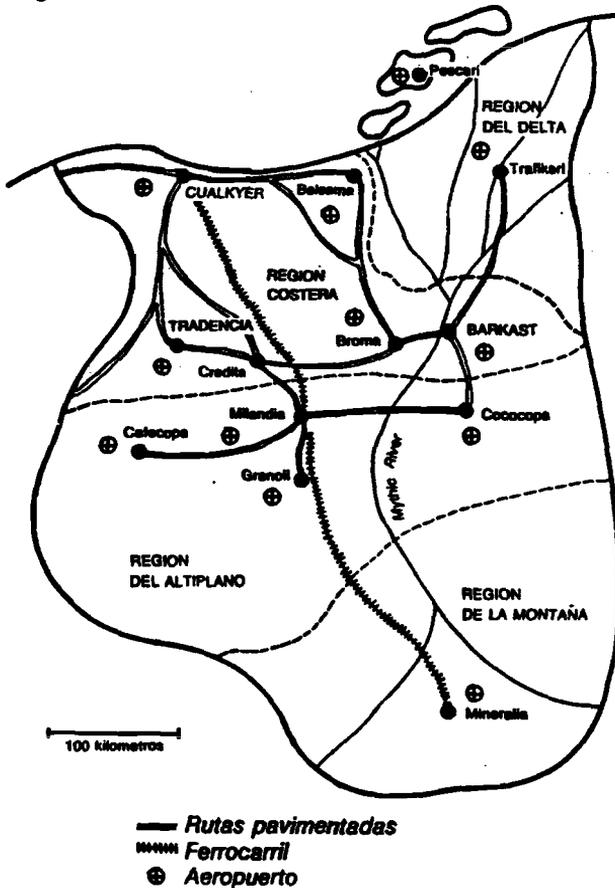
---

g) Resultados del cuestionario sobre actividades de salud suministrado por los directores regionales de salud (cuadro 18) .....	210
h) Gasto gubernamental total y específico por sectores (cuadro 19) .....	213
i) Gasto en salud en Ficticia por regiones (cuadro 20).....	214
j) Gasto en salud por programa (cuadro 21) .....	215
k) Metas operativas de salud (cuadro 22) .....	215
V. Información específica sobre riesgos perinatales .....	216
a) Relación entre la mortalidad perinatal y la talla materna (cuadro 23) .....	216
b) Relación entre la mortalidad perinatal y algunas características sociodemográficas y de salud (cuadro 24) .....	217
c) Distribución porcentual del puntaje de riesgo del período <i>preparto</i> . Primer nivel de atención (cuadro 25) .....	223
d) Distribución porcentual del puntaje de riesgo del período <i>preparto</i> . Primer nivel de atención (cuadro 26) .....	224
e) Distribución porcentual del puntaje de riesgo del período <i>intraparto</i> . Primer nivel de atención (cuadro 27). .....	225
f) Distribución porcentual del puntaje de riesgo del período <i>posparto</i> . Primer nivel de atención (cuadro 28) .....	226
g) Distribución de los nacimientos totales y las muertes perinatales y su clasificación por puntaje de riesgo para el período <i>preparto</i> . Primer nivel de atención (cuadro 29) .....	227
h) Puntaje de riesgo del período <i>preparto</i> en centros de salud (cuadro 30) .....	228
i) Puntaje de riesgo del período <i>intraparto</i> en centros de salud (cuadro 31) .....	229
j) Puntaje de riesgo del período <i>posparto</i> en centros de salud (cuadro 32) .....	230
k) Distribución de los nacimientos totales y las muertes perinatales y su clasificación por puntaje de riesgo para el período <i>preparto</i> en centros de salud (cuadro 33) ..	231
l) Grupos de riesgo según tipo de asentamiento de población (cuadro 34) .....	232
m) Grupos de riesgo según lugar del parto (cuadro 35).....	232

## I. Introducción

Ficticia es un país de 16 millones de habitantes, (Censo de 1981), con una extensión de 245.000 km<sup>2</sup>. Está dividido en cuatro regiones, cada una con distinta geografía, clima y características socioculturales: a) Delta; b) Costera; c) Altiplano; d) Montaña. Los cuadros y la información que se presentan a continuación describen el país y sus regiones.

Figura 1



## II. Perfil Sociodemográfico

Cuadro No. 1  
Población de Ficticia según edad,  
sexo y región (Censo de 1981)\*

Región	Sexo	Grupos de edad				Total
		0-4	5-14	15-44	≥45	
Delta	M	97	139	197	61	494
	F	93	133	211	69	506
	Total	190	272	408	130	1,000
Costera	M	659	1,149	2,351	955	5,114
	F	616	1,078	2,204	998	4,886
	Total	1,275	2,227	4,555	1,953	10,000
Altiplano	M	324	490	840	333	1,987
	F	317	483	842	371	2,013
	Total	641	973	1,682	704	4,000
Montaña	M	93	146	219	67	525
	F	87	134	194	60	475
	Total	180	280	413	127	1,000
Ficticia	M	1,173	1,924	3,607	1,416	8,120
	F	1,173	1,828	3,451	1,488	7,880
	Total	2,286	3,752	7,058	2,904	16,000

\* Miles de habitantes

Cuadro 2  
Distribución y densidad de la población  
para cada región de Ficticia en 1981

Distribución y densidad de la población	Delta		Costera		Altiplano		Montaña		Ficticia	
	No.	Hab.	No.	Hab.	No.	Hab.	No.	Hab.	No.	Hab.
Superficie (km <sup>2</sup> )	33.000		70.000		70.000		70.000		245.000	
Densidad de la población	29/km <sup>2</sup>		143/km <sup>2</sup>		57/km <sup>2</sup>		14/km <sup>2</sup>		64/km <sup>2</sup>	
Tipo de localidad	No.	Hab.	No.	Hab.	No.	Hab.	No.	Hab.	No.	Hab.
Ciudades de más de 500.000 habitantes	0	--	3	2.040.000	0	--	0	--	3	2.040.000
Ciudades de 100.000—500.000 habitantes	2	211.000	3	1.250.000	4	1.000.000	1	102.000	10	2.563.000
Poblados de 15.000—100.000 habitantes	5	280.000	30	1.800.000	12	700.000	3	130.000	50	2.910.000
Aldeas de menos de 15.000 habitantes	104	509.000	1.894	4.910.000	926	2.300.000	512	768.000	3.436	8.487.000
Total	111	1.000.000	1.930	10.000.000	942	4.000.000	516	1.000.000	3.499	16.000.000

Ficticia

**Cuadro No. 3**  
**Descripción de las áreas urbanas**  
**en las regiones de Ficticia, 1981**

<i>Nombre</i>	<i>Población</i>	<i>Características distintivas</i>	<i>Actividad económica</i>
Gualkyer	900.000	capital nacional, principal puerto, aeropuerto internacional	industrias de aluminio, cemento, fósforos, jabón y cerveza
Tradencia	540.000	centro agrícola	comercio
Barkast	512.000	sitio de una presa importante	industria hidroeléctrica

*Ciudades de 100.000 a 500.000 Habs.*

<i>Nombre</i>	<i>Población</i>	<i>Región</i>	<i>Características distintivas</i>
Pescara	101.000	Delta	centro pesquero
Trafikara	110.000	Delta	puerto (tráfico de la laguna)
Balsama	390.000	Costera	puerto
Broma	450.000	Costera	industria textil
Credita	410.000	Altiplano	centro de actividad bancaria agrícola
Cafecopa	225.000	Altiplano	centro cafetalero
Cocncopa	175.000	Altiplano	agricultura
Granolf	300.000	Altiplano	agricultura
Hilandia	300.000	Altiplano	_____
Mineralia	102.000	Montaña	centro minero

**Cuadro 4**  
**Características regionales con relación a algunos aspectos**  
**de importancia sociomédica en Ficticia**

	<i>Delta</i>	<i>Costera</i>	<i>Altiplano</i>	<i>Montaño</i>
1 Altitud (m)	0-30	30-100	100-2.000	2.000-5.000
2 Temperatura (C)	+30 to +45	+25 to +38	+15 to +30	-10 to +25
3 Precipitación pluvial (cm <sup>3</sup> )	180-200	70-150	40-80	10-20
4 Medio de transporte predominante	bote, autobús	auto, autobús	autobus, animales	a pie, animales
5 Geografía	islas pantanos ríos aldeas	granjas ríos selva tropical ciudades	mesetas colinas poblados aldeas	montañas valles altos
6 Población	esparcida, pescadores, contrabandistas	densa, jornaleros, empleados públicos, agricultores	densidad media, agricultores, ganaderos, nómadas	escasa, cazadores, guerrilleros, mineros
7 Industria	pesca, contrabando	agricultura, comercio, industria pesada	agricultura, ganadería, industria ligera	minería (asbestos)
8 Agua (% de la población rural con agua)	14	44	20	20
9 Tasa de alfabetismo en mujeres (%)				
a) urbana	44	48	35	33
b) rural	21	39	22	13
10 Proporción de niños de 6 a 11 años que asisten a la escuela	40%	85%	60%	20%
11 Comunicaciones receptores de radio por 1.000 hab. Teléfonos por 1.000 hab.	150 0,5	300 30	175 10,0	145 0,1
12 Caminos pavimentados (km)	100	3.000	300	..
13 Caminos transitables todo el año (km)	250	5.400	3.500	200
14 Ferrocarril		-----De Mineralía a Gualkyer-----		
15 Pistas de aterrizaje	2	5	4	1

**Cuadro 5a**  
**Tasas de fecundidad general\* por región**

<i>Región</i>	<i>Tasa</i>
Delta	218
Costera	112
Altiplano	193
Montaña	217

**Cuadro 5b**  
**Tasas de fecundidad general\* por tipo de  
 asentamiento de población**

<i>Asentamiento de población</i>	
Rural	170
Áreas urbanas marginadas	170
Urbano	65
Ficticia	144

\*Tasa calculada por cada 1.000 mujeres de 15 a 44 años de edad.

**Cuadro 6**  
**Tasas de fecundidad específicas para  
 la edad en las áreas urbanas  
 marginadas de Ficticia, 1981\***

<i>Grupo de edad de la madre</i>	<i>Tasa</i>
15-17	200
18-19	200
20-24	174
25-29	148
30-34	148
35-44	148
Todas las edades (15-44)	170

\*Tasas calculadas por cada 1.000 mujeres del grupo de edad correspondiente.

### III. Situación general de salud

**Cuadro 7**  
**Estimaciones de algunas estadísticas**  
**vitales de Ficticia\***

Indicador	Delta	Costera	Altiplano	Montaña	Ficticia
Esperanza de vida al nacer	50	59	49	47	55
Tasa cruda de mortalidad	17	9	15	21	12
Tasa cruda de natalidad	46	25	41	42	31
Tasa de mortalidad infantil	113	81	99	139	95
Tasa de mortalidad neonatal	69	47	58	81	55
Tasa de mortalidad postneonatal	44	34	41	58	40
Mujeres de 15 a 44 años	211.000	2.204.00	842.000	194.000	3.451.000
Tasa general de fecundidad	218	112	193	217	144
Nacimientos anuales	46.020	247.080	162.860	42.130	498.090
Muertes anuales de niños					
Menores de un mes	3.190	11.520	9.430	3.430	27.570
De 1 a 12 meses	2.030	8.450	6.750	2.460	19.690
Total en menores de 1 año	5.220	19.970	16.180	5.890	47.260

\*Estimaciones basadas en cifras oficiales de 1981.

**Cuadro 8**  
**Tasas de mortalidad en menores de un año**  
**por región de Ficticia (1981)**

<i>Región</i>	<i>Tasas de mortalidad</i>			
	<i>Perinatal</i>	<i>Neonatal</i>	<i>Neonatal Tardia</i>	<i>Postneonatal</i>
Delta	63,9	69,3	31,1	44,1
Costera	52,4	46,6	20,9	34,2
Altiplano	53,5	57,9	24,9	41,4
Montaña	70,8	81,4	38,5	58,4
Ficticia	55,4	55,4	24,6	39,5

**Cuadro 9**  
**Variaciones en las tasas de mortalidad en menores de un**  
**año en Ficticia (1981)**  
**según tipo de asentamiento de población**

<i>Asentamiento de población</i>	<i>Tasas de mortalidad</i>			
	<i>Perinatal</i>	<i>Neonatal</i>	<i>Neonatal</i>	<i>Postneonatal</i>
Rural	57,8	59,5	27,3	41,2
Áreas urbanas marginadas	78,2	69,3	28,0	46,4
Urbano	43,0	42,2	17,8	33,8
Ficticia	55,4	55,4	24,6	39,5

### Principales causas de morbilidad y mortalidad

La información sobre mortalidad y morbilidad en Ficticia es relativamente escasa. Se cuenta, sin embargo, con datos obtenidos a partir de una encuesta "ad hoc" efectuada como parte de los esfuerzos del Programa Ampliado de Inmunizaciones (PAI) y del Programa de Control de Enfermedades Diarreicas (PCED). Se estimaba, antes del comienzo del PAI, que la mortalidad entre los niños menores de cinco años era de 50 por mil y que el 30% de las muertes se debía a enfermedades evitables por medio de la vacunación. Al parecer, tras la instalación del PAI se produjo una reducción del 60% en las muertes por estas causas y la mayor parte de las muertes restantes se debe al sarampión. Desafortunadamente no se emprendieron, como parte del programa, acciones intensivas de inmunización antitetánica de las mujeres embarazadas.

La información revelada por el PCED señala que aproximadamente una tercera parte de las muertes en menores de cinco años se debía a las enfermedades diarreicas y que, tras la instalación

del programa, basado esencialmente en soluciones de rehidratación oral, se redujo esa cuota de defunciones en un 50%.

Las estimaciones de la mortalidad infantil revelan una cifra de 100 por mil. Las causas más importantes de muerte entre los menores de cinco años, fuera de las ya mencionadas, para las que se considera que no hay información son: enfermedades respiratorias, desnutrición y accidentes.

Una encuesta de mortalidad en un país semejante a Ficticia arrojó resultados en el sentido de que los padecimientos respiratorios, los accidentes, la desnutrición y otras enfermedades neonatales causaron el 33% de las muertes de menores de cinco años.

Con la información de diversas fuentes se ha estimado que la mortalidad entre los niños menores de cinco años había descendido a 34 por mil al cabo de un lustro de haberse iniciado los programas.

Un dato adicional de importancia para los programas de salud materno-infantil es que el 15% de los recién nacidos en Ficticia tiene al nacer un peso menor de 2.500 g.

A continuación se ofrece información sobre costos aproximados para el tratamiento y/o prevención de algunos problemas de salud de importancia en Ficticia.

#### Principales causas de morbilidad y mortalidad

ENFERMEDAD	PREVENCIÓN		TRATAMIENTO	
	Costo*	Efectividad	Costo*	Efectividad
Tétanos	1,00—3,00	95%	muy caro	40%
Malaria	5,50 años/persona (Droga) Mosquito control	alta problemática	2,00	alta
TBC	1,00—2,00	mediana (al nacimiento)	50—200	alta
Diarrea	alto (s. ambiental)	alto	1,00 (oral) 7,500 (I.V.)	alta alta
Sarampión	1,15	95-100%**	100—400 (hospital)	mediano
Pertusis	3,12 (DPT)	80%** (tres dosis)	alto	baja
Polio	3,30	alta	100—1.000	muy bajo
Difteria	3,12 (DPT)	alta (tres dosis)	100—1.000	bajo moderado
Infección respiratoria baja	—	—	2,00—20,00	alta
Infecciones de la piel	(higiene)	alta	15,00	moderado

\* Dólares (USA). Aproximado

\*\*Requiere cadena fría

**Cuadro 10**  
**Tasas de mortalidad específica en menores de un año por peso al nacer**

Peso al nacer	Incidencia porcentaje	Tasas de mortalidad		
		Perinatal	Neonatal temprana	Neonatal tardía
<2.500 g	14,3	261,7	158,0	73,5
≥2.500 g	85,7	20,7	12,2	18,4

En vista de que el peso al nacer no se registró en 2.665 de los 51.110 nacimientos, las tasas arriba mencionadas fueron calculadas a partir de un total de 48.455 nacimientos.

**Cuadro 11**  
**Tasas de mortalidad perinatal específica por peso al nacer por región y tipo de asentamiento de población Ficticia 1981**

Región	<1.500 g	Peso al nacer 1.500-2.499 g	≥2.500 g	Total
Delta	604	116	25	64
Costera	582	100	19	52
Altiplano	589	114	20	54
Montaña	539	124	26	71
<i>Tipo de asentamiento de población</i>				
Rural	553	111	22	58
Urbano	613	97	16	43
Área urbana marginada	684	122	27	78
Ficticia	581	109	21	55

**Cuadro 12**  
**Causas de muerte perinatal y neonatal**  
**en Ficticia (1981)**

Causas de muerte	Número de muertes	
	Perinatales	Neonatales tardías
Asociadas con el parto	1.156	147
Anomalías congénitas	304	89
Tétanos neonatal	298	773
Diarrea	131	71
Trastornos respiratorios	27	37
Traumatismos obstétricos	24	4
Causas mal definidas	892	105
Todas las causas	2.832	1.226

**Cuadro 13**  
**Causas de muerte perinatal y neonatal en niños**  
**con peso al nacer  $\geq$  2.500g**  
**Ficticia, 1981**

Causa de muerte	No. de defunciones	
	Perinatales	Neonatales tardías
Asociadas con el parto	347	14
Anomalías congénitas	148	34
Tétanos neonatal	232	631
Diarrea	38	23
Enf. respiratorias	12	25
Traumatismos obstétricos	22	4
No especificada o mal definida	58	28
TOTAL	857	759

El tétanos neonatal fue el responsable del 27% de las muertes perinatales en niños con peso al nacer  $\geq$  2.50

**Cuadro 14**  
**Diferencias regionales en las causas**  
**de muerte perinatal**

Causa de muerte	Delta	Costera	Región Altiplano	Montaña	Ficticia
Asociadas con el parto	119	528	368	141	1.156
Anomalías congénitas	29	157	94	24	304
Tétanos neonatal	38	110	111	39	298
Diarrea	19	73	32	7	131
Problemas respiratorios	0	9	11	7	27
Trauma	2	12	7	3	24
Causas mal definidas	95	443	268	86	892
Todas las causas	302	1.332	891	307	2.832

N.B. Un estudio de algunas enfermedades padecidas por las mujeres embarazadas de Ficticia, en 1981, reveló lo siguiente:

1. El 3,3% de las mujeres padecía malaria.
2. La anemia estuvo presente en el 14,7% de las embarazadas, y estaba distribuido predominantemente en las regiones del Delta y de la Montaña.
3. La presencia de la enfermedad hipertensiva del embarazo fue del 3,2%, con distribución predominantemente rural.
4. La diabetes estuvo presente en el 8,8% de los embarazos de Ficticia. La región Costera tuvo una frecuencia más elevada que las otras. Igual acentuación en las áreas urbanas marginadas.

### Percepción pública de los problemas de salud

En 1979, una encuesta llevada a cabo por la Universidad reveló que la población consideraba que los principales problemas de salud maternoinfantil eran: niños pequeños y débiles que mueren poco después de haber nacido, tétanos neonatal, diarrea y enfermedades infecciosas como la poliomielitis y el sarampión.

La malaria, las infecciones del tracto respiratorio bajo y la desnutrición también fueron percibida como importantes. En general el público no estaba bien informado sobre las medidas de control de los distintos problemas de salud, a excepción de las inmunizaciones, ni reconocía el valor de la atención prenatal como medida protectora de la salud del feto.

La opinión de la gente, especialmente de las madres, con respecto a la utilidad de los centros de salud, estaba fuertemente dividida. Muchas personas dijeron que habían tenido una buena experiencia en el centro de salud, mientras que otras se quejaron de tener que esperar demasiado tiempo, de poca amabilidad de parte del personal y de que la atención no era efectiva.

Existen creencias muy firmes con relación al tipo de alimentación más apropiado para los niños y las embarazadas. Hay prácticas tradicionales ampliamente aceptadas para el tratamiento de las diarreas, la fiebre y otros padecimientos, y para la atención del embarazo.

En el pasado las campañas publicitarias han sido eficaces por períodos cortos.

## IV. Recursos para la prestación de servicios de salud

### La atención a la salud en Ficticia

El gobierno de Ficticia está firmemente dispuesto a impulsar el desarrollo de la atención primaria en salud. Aún no se ha señalado la manera precisa en que esto va a llevarse a cabo. El gobierno ha formulado un "Plan Quinquenal" de carácter nacional que incluye varias metas a alcanzar en 1990. Entre las metas de salud se cuentan algunas relacionadas con el control de enfermedades diarreicas, las inmunizaciones en la infancia, el agua potable, la planificación familiar y los servicios de salud maternoinfantiles. Se ha efectuado las asignaciones de recursos a cada programa para los próximos cinco años y se ha designado coordinadores para el de inmunizaciones, el de control de enfermedades diarreicas y para el de planificación familiar.

En las regiones Costera y del Altiplano se ha puesto de manifiesto el interés por esquemas innovadores de prestación de servicios de salud. Por su parte, el gobierno central considera que las regiones menos desarrolladas, la del Delta y la de la Montaña, deben recibir especial atención en futuras actividades de desarrollo.

Entre las metas del plan quinquenal se cuentan las siguientes:

- Construir de 30 nuevos centros en áreas descubiertas, incluyendo un hospital (en la región del Altiplano) y 29 centros de salud. Esto permitirá la cobertura de 1.625.000 habitantes más.
- Renovar el 50% de las instalaciones actuales para que proporcionen mejores servicios y den más apoyo a los trabajadores de salud de las aldeas.
- Duplicar el número de aldeas atendidas por trabajadores de salud (aumentaría de 1.350 a 2.700 el número de trabajadores lo cual garantizaría el acceso local a la atención a la salud de 3.400.000 personas más).
- Hacer que las inmunizaciones rutinarias estén disponibles para todos los niños del país.
- Ampliar el alcance de los programas de planificación familiar.
- Mejorar las prácticas de cuidados de los niños, especialmente mediante el estímulo de la lactancia materna.
- Proveer a todas las instalaciones prestadoras de servicios y a todos los trabajadores de las aldeas de sales de rehidratación oral.
- Aumentar el acceso al agua potable mediante la perforación de nuevos pozos y el mejoramiento de los existentes para poder llegar a servir al 50% de la población.

### Sistemas de prestación de servicios de salud

**Sistema de atención rural:** En las áreas rurales, Ficticia tiene una serie de trabajadores, en general no profesionales, con los cuales se trata de llevar a cabo acciones de salud. No parece haber superposición de actividades entre este sistema y los servicios clásicos de salud (hospitales, clínicas, etc).

**Cuadro 21**  
**Gasto en salud por programa,**  
**Ficticia, 1979-1980\***

	AÑOS	
	1979	1980
<i>Gasto hospitalario</i>		
Salarios	10.854	12.555
Suministros	1.206	945
<b>Total</b>	<b>12.060</b>	<b>13.500</b>
<i>Gasto en servicio comunitarios</i>		
Adiestramiento de recursos humanos		
Salarios	680	740
Instituciones	171	185
Programas de prevención y control		
Maternoinfantil (PAI, nutrición, control de enfermedades diarreicas)	6.812	7.400
Tuberculosis	1.703	1.850
Malaria	4.162	3.700
Saneamiento ambiental	3.502	4.625
<b>Total</b>	<b>17.030</b>	<b>18.500</b>

\*En miles de dólares.

**Cuadro 22**  
**Metas operativas de salud**

- 1—Todas las embarazadas que asisten a los hospitales o centros de salud para someterse al control prenatal, deben ser sometidas a una prueba para estudio de diabetes al 2°, 6°, y 8° meses de embarazo.
- 2—Todas las mujeres serán pesadas en la consulta prenatal.
- 3—Toda mujer embarazada debe tener por lo menos una consulta con el médico.
- 4—El número aconsejable de controles prenatales, en un embarazo normal, es de 8.
- 5—Todas las mujeres deben ser sometidas a un examen para verificación de hemoglobina; la anemia se define cuando hay menos de 8 gs. de Hgb por 100 ml.
- 6—Todas las mujeres atendidas en los hospitales de la región de la Costa, recibirán leche en polvo para lactancia, sin costo alguno para ellas.
- 7—Todos los partos deben ser atendidos por personal legalmente habilitado, según la legislación vigente.
- 8—La presión arterial será tomada cuando sea necesario.
- 9—A todas las mujeres embarazadas les serán practicados estudios para investigar malaria.
- 10—Todas las embarazadas recibirán suplemento alimenticio.

## V. Información específica sobre riesgos perinatales

**Cuadro 23**  
Relación entre la mortalidad perinatal  
y la talla materna

Indicador	Talla materna (cms)		Total
	<155 cms	≥155 cms	
Muertes perinatales (1)	1.617	1.215	2.832
Sobrevivientes perinatales (2)	15.496	32.782	48.278
Total (3) = 1 + 2	17.113	33.997	51.110
Tasa de mortalidad perinatal* (1) / (3)	94,5	35,7	55,4

\*Por mil nacidos vivos

El riesgo relativo de experimentar una muerte perinatal entre quienes tienen una talla ≥ 155 cm. fue de 2,6 con respecto a quienes tienen una talla de 155 cms (RR =  $\frac{94,5}{35,7} = 2,6$ ). Aun cuando el 33,5% de las madres median menos de 155 cms., el 57,1% de las muertes perinatales ocurrió en la categoría de talla baja.

**Cuadro 23a**  
La información anterior se presenta aquí en forma  
similar a la que se usará en el cuadro 24

Características	Muertes perinatales			Total	x <sup>2</sup> y Riesgo relativo
	<155 cm	≥155 cm	Total		
Talla materna	<155 cm	1.617	15.496	17.113	x <sup>2</sup> = 751 RR = $\frac{94,5}{35,7} = 2,6$
	≥155 cm	1.215	32.782	33.997	
	Total	2.832	48.278	51.110	
Talla materna	<145 cm	334	1.966	2.300	x <sup>2</sup> = 371 RR = $\frac{145,2}{51,2} = 2,8$
	≥145 cm	2.498	46.312	48.810	
	Total	2.832	48.278	51.110	

**Cuadro 24**  
**Relación entre mortalidad perinatal y algunas**  
**características sociodemográficas y de salud**

Características		Muertes perinatales	Vivos	Total	$\chi^2$ y Riesgo relativo
Agua contaminada	Sí	1.311	13.914	15.225	
	No	1.521	34.364	35.885	$\chi^2 = 390$
	Total	2.832	48.278	51.110	$RR = \frac{86,1}{42,4} = 2,0$
Atención del parto	No profesional	1.448	20.832	22.280	
	Otras	1.384	27.446	28.830	$\chi^2 = 69$
	Total	2.832	48.278	51.110	$RR = \frac{65,0}{48,0} = 1,4$
Atención prenatal	Sí	2.056	25.891	27.947	
	No	776	22.387	23.163	$\chi^2 = 388$
	Total	2.832	48.278	51.110	$RR = \frac{73,6}{33,5} = 2,2$
Atención prenatal	No o tercer trimestre	2.263	30.743	33.006	
	Sí	569	17.535	18.104	$\chi^2 = 308$
	Total	2.832	48.278	51.110	$RR = \frac{68,6}{31,4} = 2,2$
Toxemia	Sí	388	1.248	1.636	
	No	2.444	47.030	49.474	$\chi^2 = 1.067$
	Total	2.832	48.278	51.110	$RR = \frac{237,2}{49,4} = 4,8$

Características		Muertes perinatales	Vivos	Total	$\chi^2$ y Riesgo relativo
Complicaciones durante el embarazo	Sí	1.865	13.378	15.243	
	No	967	34.900	35.867	$\chi^2 = 1.860$
	Total	2.832	48.278	51.110	$RR = \frac{122.4}{27.0} = 4.5$
Complicaciones previas al embarazo	Sí	584	6.019	6.603	
	No	2.248	42.259	44.507	$\chi^2 = 158$
	Total	2.832	48.278	51.110	$RR = \frac{88.4}{50.5} = 1.8$
Complicaciones médicas durante el parto	Sí	976	4.367	5.343	
	No	1.856	43.911	45.767	$\chi^2 = 1.846$
	Total	2.832	48.278	51.110	$RR = \frac{182.7}{40.6} = 4.5$
Complicaciones obstétricas durante el parto	Sí	1.588	10.647	12.235	
	No	1.244	37.631	38.875	$\chi^2 = 1.700$
	Total	2.832	48.278	51.110	$RR = \frac{129.8}{32.0} = 4.1$
Número de controles prenatales	<3	1.503	25.768	27.271	
	≥3	1.329	22.510	23.839	$\chi^2 = 0.1$
	Total	2.832	48.278	51.110	$RR = \frac{55.1}{55.7} = 1.0$

## Ficticio

Características		Muertes perinatales	Vivos	Total	$\chi^2$ y Riesgo relativo
Trabajo de la madre fuera de la casa	Sí	561	9.720	10.281	
	No	2.271	38.558	40.829	$\chi^2 = 0,2$
	Total	2.832	48.278	51.110	$RR = \frac{54,6}{55,6} = 1,0$
Fumar durante embarazo	Sí	362	6.156	6.518	
	No	2.470	42.122	44.592	$\chi^2 = 2,4$
	Total	2.832	48.278	51.110	$RR = \frac{55,5}{55,4} = 1,0$
Presentación al parto	Nalga	375	1.173	1.548	
	Otras	2.457	47.105	49.562	$\chi^2 = 1,065$
	Total	2.832	48.278	51.110	$RR = \frac{242,2}{49,6} = 4,9$
Ocupación del padre	No pro- fesional	2.518	37.880	40.398	
	Profesional	314	10.398	10.712	$\chi^2 = 176$
	Total	2.832	48.278	51.110	$RR = \frac{62,3}{29,3} = 2,1$
Estado civil	No casada	991	8.854	9.845	
	Casada	1.841	39.424	41.265	$\chi^2 = 477$
	Total	2.832	48.278	51.110	$RR = \frac{100,7}{44,6} = 2,3$
Paridad	3 o más	950	13.641	14.591	
	Menos de 3	1.882	34.637	36.519	$\chi^2 = 37$
	Total	2.832	48.278	51.110	$RR = \frac{65,1}{51,5} = 1,3$

Características		Muertes perinatales	Vivos	Total	$\chi^2$ y Riesgo relativo
Paridad	4 o más	607	6.683	7.290	
	Menos de 4	2.225	41.595	43.820	$\chi^2 = 126$
	Total	2.832	48.278	51.110	RR = $\frac{83,3}{50,8} = 1,6$
Número de hijos vivos	2 o más	586	9.878	10.464	
	Menos de 2	1.201	20.021	21.222	$\chi^2 = 0$
	Total	1.787	29.899	31.686*	RR = $\frac{56,0}{56,6} = 1,0$
Número de hijos vivos	3 o más	297	3.989	4.286	
	Menos de 3	1.490	25.910	27.400	$\chi^2 = 15$
	Total	1.787	29.899	31.686*	RR = $\frac{69,3}{54,4} = 1,3$
* El número de niños vivos no se supo en 19.424 casos (incluyendo 1.045 muertes perinatales).					
Educación materna	Analfabetas	1.668	21.019	22.687	
	Otras	1.164	27.259	28.423	$\chi^2 = 256$
	Total	2.832	48.278	51.110	RR = $\frac{73,5}{41,0} = 1,8$
Distancia de la casa al centro (km)	Cinco o menos	1.289	18.547	19.836	
	Más de cinco	1.543	29.731	31.274	$\chi^2 = 57$
	Total	2.832	48.278	51.110	RR = $\frac{65,0}{49,3} = 1,3$
Tipo de embarazo	Múltiple	303	1.049	1.352	
	Unico	2.520	47.154	49.674	$\chi^2 = 757$
	Total	2.823	48.203	51.026*	RR = $\frac{224,1}{50,7} = 4,4$

\*\*El tipo de embarazo no se supo en 9 casos de muertes perinatales y en 75 nacidos vivos.

Características		Muertes perinatales	Vivos	Total	$\chi^2$ y Riesgo relativo
Edad materna (años)	19 o menos	778	8.513	9.291	
	Más de 19	2.044	39.567	41.611	$\chi^2 = 174$
	Total	2.822	48.080	50.902*	$RR = \frac{83,7}{49,1} = 1,7$

Edad materna	17 o menos	306	3.350	3.656	
	18 o más	2.516	44.730	47.246	$\chi^2 = 60$
	Total	2.822	48.080	50.902*	$RR = \frac{83,7}{53,3} = 1,6$

\* La edad materna no se supo en 208 casos, 10 de los cuales fueron muertes perinatales.

Grupo étnico	B	2.025	20.164	22.189	
	A	807	28.114	28.921	$\chi^2 = 963$
	Total	2.832	48.278	51.110	$RR = \frac{91,3}{27,9} = 3,3$

Sexo del niño	M	1.639	24.989	26.628	
	F	1.189	23.283	24.472	$\chi^2 = 41$
	Total	2.828	48.272	51.100**	$RR = \frac{61,6}{48,6} = 1,3$

\*\* Hay 10 casos sin información sobre el sexo del niño; cuatro entre las muertes perinatales y seis entre los nacidos vivos.

Región	Montaña	307	4.032	4.339	
	Otras	2.525	44.246	46.771	$\chi^2 = 21$
	Total	2.832	48.278	51.110	$RR = \frac{70,8}{54,0} = 1,3$

Sector	Rural	1.824	29.731	31.555	
	Otras	1.008	18.547	19.555	$\chi^2 = 9$
	Total	2.832	48.278	51.110	$RR = \frac{57,8}{51,5} = 1,1$

Características		Muertes perinatales	Vivos	Total	$\chi^2$ y Riesgo relativo
Peso nacer (g)	<2.500	1.819	5.131	6.950	
	$\geq 2.500$	857	40.638	41.495	$\chi^2 = 6.629$
	Total	2.676	45.769	48.445*	$RR = \frac{261,7}{20,7} = 12,7$

\* Peso al nacer desconocido en 2.665 casos

Historia de aborto o muerte fetal						
Historia de aborto o muerte fetal	Sí	511	6.349	6.860		
	No	2.321	41.929	44.250	$\chi^2 = 55$	
	Total	2.832	48.278	51.110	$RR = \frac{74,5}{52,5} = 1,4$	

Historia de partos prematuros						
Historia de partos prematuros	Sí	280	2.662	2.942		
	No	2.552	45.616	48.168	$\chi^2 = 94$	
	Total	2.832	48.278	51.110	$RR = \frac{95,2}{53,0} = 1,8$	

**Cuadro 25**  
**Distribución porcentual del puntaje de riesgo\*1**  
**del período *preparto*, primer nivel de atención**  
**Ficticia, 1981**

Puntaje de riesgo	Mortalidad perinatal (Tasa por 1.000 nacimientos)	Muertes perinatales		Nacimientos		Porcentajes acumulados invertidos	
		Número	Porcentaje acumulado	Número	Porcentaje acumulado	Nacimientos	Muertes perinatales
0	6.5	16	0.6	2.455	4.8	100.0	100.0
1	18.3	67	2.9	3.667	12.0	95.2	99.4
2	24.6	149	8.2	6.045	23.8	88.0	97.1
3	39.4	325	19.7	8.239	39.9	76.2	91.8
4	50.8	450	35.6	8.859	57.3	60.1	80.3
5	60.4	536	54.5	8.872	74.6	42.7	64.4
6	88.1	556	74.1	6.314	87.0	25.4	45.5
7	100.2	392	88.0	3.914	94.6	13.0	25.9
8	117.5	230	96.1	1.957	98.5	5.4	12.0
9	137.8	89	99.2	646	99.7	1.5	3.9
10	128.8	17	99.8	132	100.0	0.3	0.8
11	500.0	5	100.0	10	100.0	0.0	0.2
<b>Total</b>	<b>55.4</b>	<b>2.832</b>		<b>51.110</b>			

\* Excluyendo la educación materna como factor de riesgo.

1 Todas las variables incluidas en el puntaje de riesgo recibirán igual valor.

**Cuadro 26**  
**Distribución porcentual del puntaje de riesgo\*1**  
**del período *preparto*. Primer nivel de atención**  
**Ficticia, 1981**

Puntaje de riesgo	Mortalidad perinatal (Tasa por 1.000 nacimientos)	<i>Muertes perinatales</i>		<i>Nacimientos</i>		<i>Porcentajes acumulados invertidos</i>	
		Número	Porcentaje acumulado	Número	Porcentaje acumulado	Nacimientos	Muertes perinatales
0	5,9	13	0,5	2.212	4,3	100,0	100,0
1	18,3	58	2,5	3.178	10,5	95,7	99,5
2	19,4	102	6,1	5.268	20,9	89,5	97,5
3	33,0	229	14,2	6.934	34,4	79,1	93,9
4	46,8	380	27,6	8.118	50,3	65,6	85,8
5	57,5	460	43,9	7.994	65,9	49,7	72,4
6	79,6	536	62,8	6.730	79,1	34,1	56,1
7	83,6	421	77,6	5.033	89,0	20,9	37,2
8	104,0	325	89,1	3.126	95,1	11,0	22,4
9	116,2	204	96,3	1.756	98,5	4,9	10,9
10	133,7	83	99,3	621	99,7	1,5	3,7
11	123,1	16	99,8	130	100,0	0,3	0,7
12	500,0	5	100,0	10	100,0	0,0	0,2
<b>Total</b>	<b>55,4</b>	<b>2.832</b>		<b>51.110</b>			

\* Incluyendo la educación materna como factor de riesgo.

1 Todas las variables incluidas en el puntaje de riesgo recibieron igual valor.

**Cuadro 27**  
**Distribución porcentual del puntaje de riesgo\*1**  
**del periodo intraparto. Primer nivel de atención**  
**Ficticia, 1981**

Puntaje de riesgo	Mortalidad perinatal (Tasa por 1.000 nacimientos)	<i>Muertes perinatales</i>		<i>Nacimientos</i>		<i>Porcentajes acumulados invertidos</i>	
		Número	Porcentaje acumulado	Número	Porcentaje acumulado	Nacimientos	Muertes perinatales
0	4,3	8	0,3	1.875	3,7	100,0	100,0
1	7,4	21	1,0	2.856	9,3	96,3	99,7
2	12,9	62	3,2	4.789	18,6	90,7	99,0
3	22,7	145	8,3	6.400	31,1	81,4	96,8
4	35,7	267	17,8	7.473	45,8	68,9	91,7
5	47,1	364	30,6	7.724	60,9	54,2	82,2
6	73,4	505	48,4	6.876	74,3	39,1	69,4
7	87,8	491	65,8	5.591	85,3	25,7	51,6
8	106,3	411	80,3	3.867	92,8	14,7	34,2
9	122,6	278	90,1	2.268	97,3	7,2	19,7
10	161,0	162	95,8	1.006	99,2	2,7	9,9
11	253,2	78	98,6	308	99,8	0,8	4,2
12	403,2	25	99,5	62	100,0	0,2	1,4
13	1.000,0	13	99,9	13	100,0	0,0	0,5
14	1.000,0	2	100,0	2	100,0	0,0	0,1
<b>Total</b>	<b>55,4</b>	<b>2.832</b>		<b>51.110</b>			

\* Incluyendo la educación materna como factor de riesgo.

1 Todas las variables incluidas en el puntaje de riesgo recibieron igual valor.

**Cuadro 28**  
**Distribución porcentual del puntaje de riesgo\*1**  
**del período *posparto*. Primer nivel de atención**  
**Ficticia, 1981**

Puntaje de riesgo	Mortalidad perinatal (Tasa por 1.000 nacimientos)	<i>Muertes perinatales</i>		<i>Nacimientos</i>		<i>Porcentajes acumulados invertidos</i>	
		Número	Porcentaje acumulado	Número	Porcentaje acumulado	Nacimientos	Muertes perinatales
0	4,3	8	0,3	1.843	3,6	100,0	100,0
1	5,4	15	0,8	2.797	9,1	96,4	99,7
2	8,0	37	2,1	4.639	18,2	90,9	99,2
3	17,5	109	6,0	6.245	30,4	81,8	97,9
4	25,5	184	12,5	7.221	44,5	69,6	94,0
5	40,7	306	23,3	7.515	59,2	55,5	87,5
6	63,7	435	38,6	6.830	72,6	40,8	76,7
7	90,2	507	56,5	5.620	83,6	27,4	61,4
8	117,1	473	73,2	4.041	91,5	16,4	43,5
9	131,4	330	84,9	2.511	96,4	8,5	26,8
10	173,6	205	92,1	1.181	98,7	3,6	15,1
11	279,4	140	97,1	501	99,7	1,3	7,9
12	434,1	56	99,0	129	99,9	0,3	2,9
13	655,2	19	99,7	29	100,0	0,1	1,0
14	1.000,0	8	100,0	8	100,0	0,0	0,3
<b>Total</b>	<b>55,4</b>	<b>2.832</b>		<b>51.110</b>			

\* Incluyendo la educación materna como factor de riesgo.

1 Todas las variables incluidas en el puntaje de riesgo recibieron igual valor.

**Cuadro 29**  
**Distribución porcentual de los nacimientos**  
**totales y de las muertes perinatales y su clasificación**  
**por puntaje de riesgo para el período preparto.**  
**Primer nivel de atención ficticia, 1981**

Riesgo relativo (puntaje)	Mortalidad perinatal (tasa por 1.000 nacimientos)	Nacimientos totales		Muertes perinatales		Riesgo relativo estandarizado
		Número	Porcentaje acumulado	Número	Porcentaje acumulado	
0	9.6	3.632	100.0	35	100.0	1.0
1	18.1	3.811	92.9	69	98.8	1.9
2	27.9	5.050	85.4	141	96.3	2.9
3	28.1	6.378	75.5	179	91.3	2.9
4	28.7	6.089	63.0	175	85.0	3.0
5	38.4	5.526	51.1	212	78.8	4.0
6	62.5	4.401	40.3	275	71.4	6.5
7	84.9	4.392	31.7	373	61.6	8.8
8	101.5	3.705	23.1	376	48.5	10.6
9	97.5	3.006	15.9	293	35.2	10.2
10	114.1	2.392	10.0	273	24.8	11.9
11	140.4	1.289	5.3	181	15.2	14.6
12	154.2	759	2.8	117	8.8	16.0
13	170.6	428	1.3	73	4.7	17.8
14	200.0	160	0.5	32	2.1	20.8
15	283.8	74	0.2	21	0.9	29.7
16	307.7	13	0.0	4	0.2	32.0
17	600.0	5	0.0	3	0.1	62.5
<b>Total</b>	<b>55.4</b>	<b>51.110</b>	<b>100.2</b>	<b>2.832</b>	<b>100.0</b>	

Cada variable recibió un valor basado en el riesgo relativo calculado para la mortalidad perinatal.

**Cuadro 30**  
**Distribución porcentual del puntaje de riesgo del período**  
**preparto en centros de salud**  
**Ficticia, 1981**

Puntaje de riesgo	Mortalidad perinatal (Tasa por 1.000 nacimientos)	Muertes perinatales		Nacimientos		Porcentajes acumulados invertidos	
		Número	Porcentaje acumulado	Número	Porcentaje acumulado	Nacimientos	Muertes perinatales
0	5,8	10	0,4	1.729	3,4	100,0	100,0
1	6,8	19	1,0	2.791	8,8	96,6	99,6
2	13,2	61	3,2	4.632	17,9	91,2	99,0
3	20,3	124	7,6	6.103	29,8	82,1	96,8
4	33,4	239	16,0	7.165	43,9	70,2	92,4
5	50,7	383	29,5	7.554	58,6	56,1	84,0
6	65,3	440	45,1	6.743	71,8	41,4	70,5
7	88,6	487	62,3	5.499	82,6	28,2	54,9
8	96,8	405	76,6	4.182	90,8	17,4	37,7
9	113,0	313	87,6	2.769	96,2	9,2	23,4
10	166,7	218	95,3	1.308	98,8	3,8	12,4
11	179,5	91	98,5	507	99,7	1,2	4,7
12	292,5	31	99,6	106	100,0	0,3	1,5
13	550,0	11	100,0	20	100,0	0,0	0,4
14	0,0	0	100,0	2	100,0	0,0	0,0
<b>Total</b>	<b>55,4</b>	<b>2.832</b>		<b>51.110</b>			

**Cuadro 31**  
**Distribución porcentual del puntaje de riesgo del período**  
**intraparto en centros de salud**  
**Ficticia, 1981**

Puntaje de riesgo	Mortalidad perinatal (Tasa por 1.000 nacimientos)	<i>Muertes perinatales</i>		<i>Nacimientos</i>		<i>Porcentajes acumulados invertidos</i>	
		Número	Porcentaje acumulado	Número	Porcentaje acumulado	Nacimientos	Muertes perinatales
0	4,4	7	0,2	1.589	3,1	100,0	100,0
1	4,8	12	0,7	2.516	8,0	96,9	99,8
2	7,0	29	1,7	4.116	16,1	92,0	99,3
3	10,9	61	3,8	5.585	27,0	83,9	98,3
4	25,0	166	9,7	6.644	40,0	73,0	96,2
5	36,0	247	18,4	6.856	53,4	60,0	90,3
6	51,3	336	30,3	6.547	66,2	46,6	81,6
7	69,1	392	44,1	5.677	77,3	33,8	69,7
8	103,2	456	60,2	4.420	86,0	22,7	55,9
9	126,2	429	75,4	3.399	92,6	14,0	39,8
10	142,9	303	86,1	2.121	96,8	7,4	24,6
11	190,0	196	93,0	1.032	98,8	3,2	13,9
12	264,2	107	96,8	405	99,6	1,2	7,0
13	381,6	58	98,8	152	99,9	0,4	3,2
14	548,4	17	99,4	31	100,0	0,1	1,2
15	777,8	14	99,9	18	100,0	0,0	0,6
16	1.000,0	2	100,0	2	100,0	0,0	0,0
<b>Total</b>	<b>55,4</b>	<b>2.832</b>		<b>51.110</b>			

**Cuadro 32**  
**Distribución porcentual del puntaje de riesgo del período**  
**posparto en centros de salud**  
**Ficticia, 1981**

Puntaje de riesgo	Mortalidad perinatal (Tasa por 1.000 nacimientos)	<i>Muertes perinatales</i>		<i>Nacimientos</i>		<i>Porcentajes acumulados invertidos</i>	
		Número	Porcentaje acumulado	Número	Porcentaje acumulado	Nacimientos	Muertes perinatales
0	2.7	2	0,1	737	1,4	100,0	100,0
1	4,1	8	0,4	1.970	5,3	98,6	99,9
2	3,7	12	0,8	3.226	11,6	94,7	99,6
3	6,4	30	1,8	4.667	20,7	88,4	99,2
4	11,9	71	4,3	5.966	32,4	79,3	98,2
5	21,5	138	9,2	6.416	45,0	67,6	95,7
6	30,3	201	16,3	6.640	58,0	55,0	90,8
7	51,4	305	27,1	5.937	69,6	42,0	83,7
8	81,2	419	41,9	5.161	79,7	30,4	72,9
9	108,1	430	57,1	3.978	87,5	20,3	58,1
10	136,8	410	71,5	2.997	93,3	12,5	42,9
11	178,3	321	82,9	1.800	96,8	6,7	28,5
12	231,8	216	90,5	932	98,7	3,2	17,1
13	334,9	147	95,7	439	99,5	1,3	9,5
14	416,7	70	98,2	168	99,9	0,5	4,3
15	576,9	30	99,2	52	100,0	0,1	1,8
16	1.000,0	14	99,7	14	100,0	0,0	0,8
17	800,0	8	100,0	10	100,0	0,0	0,3
<b>Total</b>	<b>55,4</b>	<b>2.832</b>		<b>51.110</b>			

**Cuadro 33**  
**Distribución porcentual de los nacimientos**  
**totales y de las muertes perinatales y su clasificación**  
**por puntaje de riesgo para el período *preparto***  
**en centros de salud Ficticia, 1981**

Riesgo relativo (puntaje)	Mortalidad perinatal (tasa por 1.000 nacimientos)	Nacimientos totales		Muertes perinatales		Riesgo relativo estandarizado
		Número	Porcentaje acumulado	Número	Porcentaje acumulado	
0	6.2	2.898	100.0	18	100.0	1.0
1	7.8	2.929	94.3	23	99.4	1.3
2	10.6	3.758	88.6	40	98.6	1.7
3	11.1	4.785	81.2	53	97.2	1.8
4	16.7	4.657	71.9	78	95.3	2.7
5	28.9	4.663	62.8	135	92.5	4.7
6	41.3	4.313	53.6	178	87.7	6.7
7	52.8	4.454	45.2	235	81.5	8.5
8	59.5	4.184	36.5	249	73.2	9.6
9	64.6	3.778	28.3	244	64.4	10.4
10	88.5	3.333	20.9	295	55.7	14.3
11	131.6	2.212	14.4	291	45.3	21.2
12	158.0	1.836	10.1	290	35.1	25.5
13	180.3	1.348	6.5	243	24.8	29.1
14	173.3	825	3.8	143	16.2	28.0
15	205.3	565	2.2	116	11.2	33.1
16	284.7	295	1.1	84	7.1	46.0
17	303.0	132	0.5	40	4.1	48.9
18	422.5	71	0.3	30	2.7	68.2
19	666.7	45	0.1	30	1.6	107.6
20	437.5	16	0.0	7	0.6	70.6
21	777.8	9	0.0	7	0.3	125.5
22	1.000.0	2	0.0	2	0.1	161.4
23	500.0	2	0.0	1	0.0	80.7
<b>Total</b>	<b>55.4</b>	<b>51.110</b>	<b>100.2</b>	<b>2.832</b>		

Cada variable recibió un valor basado en el riesgo relativo calculado para la mortalidad perinatal.

**Cuadro 34**  
**Grupos de riesgo según tipo de asentamiento de población**

Grupo de riesgo	Tipo de asentamiento de población			Total
	Rural	Áreas urbanas marginadas	Urbano	
Bajo	5.810	2.406	12.190	20.406
Medio	23.032	2.337	2.590	27.959
Alto	2.713	13	19	2.745
<b>Total</b>	<b>31.555</b>	<b>4.756</b>	<b>14.799</b>	<b>51.110</b>

El 8,6% de los embarazos en el sector rural se considera de alto riesgo, cuando sólo el 5,4% lo es en el nivel nacional.

**Cuadro 35**  
**Grupos de riesgo según lugar del parto**

Grupo de riesgo	Lugar del parto				Total
	Partera empírica	Enfermera obstétrica	Centros de salud	Hospital	
Bajo	4.438	1.915	5.134	8.919	20.406
Medio	15.899	2.449	6.324	3.287	27.959
Alto	1.943	172	555	75	2.745
<b>Total</b>	<b>22.280</b>	<b>4.536</b>	<b>12.013</b>	<b>12.281</b>	<b>51.110</b>

El 8,7% de los partos atendidos por parteras empíricas, 3,8% de los atendidos por enfermeras obstétricas, 4,6% de los ocurridos en centros de salud y 0,6% en hospitales, fueron el alto riesgo.

# PROBLEMS AND PROGRESS IN MEDICAL CARE

*Essays on Current Research*

SEVENTH SERIES

EDITED BY  
GORDON McLACHLAN

Published for the  
Nuffield Provincial Hospitals Trust  
by the Oxford University Press  
London New York Toronto

1972

# A taxonomy of social need

JONATHAN BRADSHAW

---

*One of the most crucial problems facing the social services is how to identify social need. This article attempts to provide a framework for clearer thinking about need.*

---

The concept of social need is inherent in the idea of social service. The history of the social services is the story of the recognition of needs and the organization of society to meet them. The Seebohm Report (1) was deeply concerned with the concept of need, though it never succeeded in defining it. It saw that 'The Personal Social Services are large scale experiments in ways of helping those in need'.

Despite this interest it is often not clear in a particular situation what is meant by social need. When a statement is made to the effect that a person or group of persons are in need of a given service, what is the quality that differentiates them—what definition of social need is being used?

The concept of social need is of particular interest to economists. They have a clearcut measure of 'effective demand': demand is 'effective' when people are prepared to back it by pecuniary allocation and ineffective or non-existent when they are not. This measure will not do for the social services, because there is normally no link between service and payment (though some economists think

indicate a high incidence of pathology in one area which are not present in another. Need established by this method is the gap between what services exist in one area and what services exist in another, weighted to take account of the difference in pathology. This is an attempt to standardize provision, but provision may still not correspond with need. The question still has to be asked—supply at what level? The statement that one area A is in need in comparison with another area B does not necessarily imply that area B is still not in need.

Comparative need used to define individuals in need can be illustrated by the following statements: 'this person X is in receipt of a service because he has the characteristics A-N. This person Z has also the characteristics A-N but is not receiving the service. Therefore Z is in need.' The difficulty in this situation is to define the significant characteristics. The method has been used by some local health authorities to compile a risk register of babies in need of special attention from the preventive services. Conditions which in the past have been associated with handicap such as forceps delivery, birth trauma, birth to older mothers, etc., are used as indicators to babies in special need. The definition is more commonly used in an *ad-hoc* way—a crude rule of precedence to assess eligibility for selective services provided by the personal social services.

Fig. 3.1 demonstrates diagrammatically the interrelation of the four definitions. Plus (+) and minus (-) denote the presence or absence of need by each of the foregoing definitions, i.e. + - - + is a need that is accepted as such by the experts, but which is neither felt nor demanded by the individual despite the fact that he has the same characteristics as those already being supplied with the service. Other examples of the twelve possible combinations are given. It will be noted that none of the circles in Fig. 3.1 are coterminous and the problem the policymaker has to face, is deciding what part of the total is 'real need'—that is need it is appropriate to try to meet.

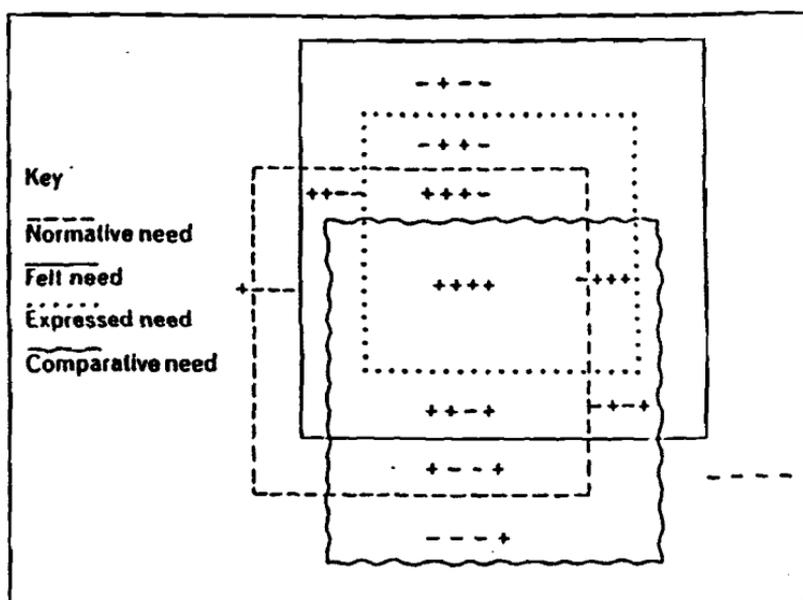


FIGURE 3.1

I. ++++

This is the area where all definitions overlap, or (using an analogy from studies of intelligence tests) the 'g' factor of need. An individual is in need by all definitions and so this is the least controversial part of need.

2. ++-+

Demand is limited by difficulties of access to a service. Although the individual is in need by all other definitions he has not wanted to, or been able to, express his need. Difficulties of access may be due to a stigma attached to the receipt of a service, geographical distances that make it difficult to claim, charges which are a disincentive to take up, administrative procedures that deter claimants or merely ignorance about the availability of the service. Demand must also vary according to how intense is the felt need.

11. - + - +

A need that is not postulated by the experts but is felt, not expressed, but is supplied. People feel a need to make contributions for social benefits and the need is met by insurance stamps, but many experts feel it would be simpler to finance these benefits wholly through taxation.

12. - - - -

Absence of need by all definitions.

To illustrate how this could be used by research-workers and policymakers it might be useful to outline a hypothetical situation. The taxonomy will be discussed in relation to housing need, but there is no significance in this choice—the discussion is equally relevant to any other area of need. A local housing authority is concerned about the housing position of the elderly in their area. They wish to have assessment of the need for public sector housing for this age-group. A research-worker is therefore commissioned to do a study of the housing need. The first problem the research-worker has to face is the question of what constitutes housing need? He can either make a decision as to what he himself believes housing need to be, or he can produce information on the amount of need under each section of the taxonomy and allow the policymaker to decide what part of the total they regard as 'real need'. The research-worker decides to take the latter course of action. This will provide the maximum information with the minimum number of value-judgements. In order to produce a figure for each section of the taxonomy, he must first decide on the amount of need under each of the four separate definitions.

*Normative need.* It has already been pointed out that there is no one definition of normative need. Let us assume that the local housing authority are laying down the norms in this situation and they would agree that old persons living

in homes lacking any of the basic amenities and old persons living in overcrowded accommodation are in need by their standards. An estimate of the number of persons in this situation could be obtained by a sample survey.

*Felt need.* An estimate of the degree of felt need can be obtained by means of the same sample survey by asking the respondents whether they are satisfied with their present housing and if not whether they would like to move. Ignoring the problems inherent in exploring people's attitudes on such a delicate question and remembering that their attitudes will be affected by their knowledge of alternative housing opportunities, as well as their fears about the upheaval of the move, another measure of need is obtained.

*Expressed need.* The local housing authority's waiting-lists provide the measure of expressed need in this context. It is at the same time the easiest measure of need to obtain and the most inadequate. On the one hand the list may be inflated by persons who have resolved their housing problem since they applied for the housing and yet who have not withdrawn their application, and on the other hand the list may underestimate expressed need if certain categories are excluded from the waiting-list; there may be a residence qualification, applications from owner-occupiers may not be accepted unless they are overcrowded, and persons who have refused the first offer may also be excluded. All these exclusions mean that the waiting-list is not an adequate measure of expressed need but because it is the only one available, it is used as another measure of need.

*Comparative need.* The measure of comparative need is more difficult to obtain. It would entail investigating the characteristics of elderly persons already in public sector housing and then through a sample survey obtaining an estimate of the number of persons in the community (not in public sector housing) who have similar characteristics. As the

local housing authority's norms have been taken for the measure of normative need, and as the local housing authority is responsible for choosing their tenants, it is likely that in this example the characteristics of tenants will be similar to those norms and thus the measure of comparative need will be very similar (though not necessarily identical) to the measure of normative need.

The research-worker has now produced four separate but interrelated measures of need. By sorting he is able to put a figure against each of the permutations of the four measures. For instance:

+ + - + This will consist of persons whose houses are overcrowded or lack basic amenities, who want to move but who are not on the council waiting-list and yet who are 'as deserving as' other residents in council accommodation.

- + - - This will consist of persons whose housing is considered satisfactory by local authority standards, who are not on the council waiting-list, and are not in need when compared with other residents in council property and yet who want to move.

So now the policymaker is presented with a picture of 'total need' for public sector housing in their area. He is now able to use the taxonomy to clarify his decision-making. Instead of housing being allocated on the basis of either first come first served, or whether the old person is articulate, energetic, and knowledgeable enough to get on the housing waiting-list, it can now be allocated on the basis of explicit priorities. No longer is the local authority providing houses 'to meet need' but rather providing houses to meet certain specific conditions of need.

Thus the policymaker can do one of two things. Either he can decide that certain categories of the total (say + + + +, + + - +, + - - +, - + + +, + + + -) constitute 'real need' and plan to provide enough housing for the numbers in these groups, or secondly if it is found that

need is very large and his resources are limited, he can decide that certain categories of need should be given priority. For instance, he may decide that category + + - + : those in need who have not applied for help (the iceberg below the waterline) should be given priority over category - + + + : those in need on all bases except that they are already adequately housed on a normative measure.

The policymaker can now return to the research-worker. Having made his priorities explicit he could ask the research worker to carry out a detailed study of the 'real need' categories to ascertain their aetiology so that in future they may be more easily identified and the services explicitly designed to get at and help them. The research worker could also use the taxonomy as a framework for monitoring the effects on need of technical advances, demographic change, changes in the standard of living, and improvements in the services.

This taxonomy may provide a way forward in an area where precise thinking is needed for both theoretical and practical reasons. Without some further classification much social policy must remain a matter of political hunches and academic guesswork. The taxonomy provides no easy solutions either for the research-worker or the policymaker. The research-worker is still faced with difficult methodological problems and the policymaker has still to make complex decisions about which categories of need should be given priority. But the taxonomy may help to clarify and make explicit what is being done when those concerned with the social services are studying or planning to meet social need.

#### REFERENCES

1. **SERMONS REPORT (1968).** *Report of the Committee on Local Authority and Allied Personal Services*, Cmnd. 3703 (London: HMSO).
2. **LAMBERT, ROYSTON (1964).** *Nutrition in Britain 1950-60*, Occasional Papers in Social Administration no. 6 (London: Bell).
3. **TOWNSEND, PETER, and WEDDERBURN, DOROTHY (1965).** *The Aged in the Welfare State*, Occasional Papers in Social Administration no. 14 (London: Bell).

1527I  
Dic./90

**TALLER SOBRE ESTRATIFICACION DE LA MALARIA**

**MATERIALES DE TRABAJO**

**IV**

**EVALUACION DE LAS INTERVENCIONES**

**BASADAS EN EL ENFOQUE**

**EPIDEMIOLOGICO DE RIESGO**

**NOTA**

**VERSION BORRADOR**

**FAVOR DE NO REPRODUCIR O CITAR SIN LA AUTORIZACION DE LA  
ORGANIZACION PANAMERICANA DE LA SALUD**

**FAVOR NO REPRODUCIR  
SIN LA AUTORIZACION  
DE LA ORGANIZACION  
PANAMERICANA DE LA  
SALUD (OPS)**

**EVALUACION DE LAS INTERVENCIONES DE CONTROL DE MALARIA  
BASADAS EN EL ENFOQUE EPIDEMIOLOGICO DE RIESGO.**

**I. Conceptos básicos sobre indicadores de evaluación.**

El proceso de evaluación de las intervenciones de control de la malaria basadas en el enfoque epidemiológico de riesgo constituye la base fundamental para reconocer su impacto en los niveles de transmisión malárica de la población.

El diccionario de epidemiología define a la evaluación como "Un proceso que tiene como objeto la determinación sistemática y objetiva de la relevancia, efectividad e impacto de actividades a la luz de sus objetivos" (Last: 34).

El objetivo fundamental del proceso de evaluación es identificar en qué medida la estrategia basada en el enfoque epidemiológico de riesgo es una alternativa mejor que la que orienta las intervenciones tradicionales. Por lo tanto, los tres posibles resultados de la evaluación son:

1. Que NO exista cambio en los niveles de transmisión malárica de la comunidad, según lo reflejado por los indicadores seleccionados para la evaluación.

2. Que exista un cambio negativo en los niveles de transmisión malárica, es decir, que exista un deterioro de las condiciones de salud, manifestado por el aumento de la incidencia de malaria.
  
3. Que exista un cambio favorable en los niveles de transmisión malárica en la comunidad, el cual se refleje a su vez en el descenso de la incidencia de malaria y aumento en los niveles de salud de la población.

Un aspecto fundamental del proceso de evaluación de las intervenciones basadas en el enfoque epidemiológico de riesgo es la posibilidad de reconocer si el IMPACTO positivo o negativo, o la ausencia de impacto, es consecuencia de la estrategia de estratificación basada en el enfoque de riesgo y no de otros factores. Este problema metodológico puede controlarse únicamente a través de un cuidadoso diseño de evaluación y de una selección adecuada de los indicadores de evaluación.

En términos generales la evaluación puede realizarse a través de tres tipos de indicadores:

- a) indicadores de estructura
- b) indicadores de proceso
- c) indicadores de impacto o resultado

Los indicadores de estructura son aquellos que reflejan información cuantitativa sobre la infraestructura de los servicios de salud, su nivel de organización, los tipos de servicios disponibles para la población y los recursos con que se cuenta. El tipo de evaluación que puede efectuarse con estos indicadores presupone que la calidad de la atención a la salud y los niveles de salud dependen de los recursos disponibles.

Los indicadores de proceso incluyen información sobre la manera en que los servicios se adecuan a las necesidades de la población o sobre los pasos seguidos en el diagnóstico, el tipo de tratamiento, las condiciones y las formas de acceso al sistema de atención a la salud, los patrones de referencia y el tipo de servicios ofrecidos en los distintos niveles de atención. En resumen, el proceso se relaciona con la forma de organizar y articular los distintos componentes del sistema de salud en la prestación de servicios e intervenciones de salud.

Los indicadores de impacto o resultado son utilizados para evaluar el nivel de salud alcanzado como consecuencia de la intervención. Buscan identificar las variaciones y cambios en los niveles de enfermedad, invalidez o muerte en la comunidad, producidos por la nueva estrategia basada en el enfoque de riesgo. La evaluación de este tipo de indicadores requiere en ocasiones períodos largos de tiempo, ya que los cambios en las tasas de mortalidad o morbilidad se expresan con mayor claridad al mediano plazo. Estos indicadores son fundamentales para evaluar el impacto de las intervenciones en la transmisión malárica de la población.

2. **Importancia de una lineal basal de información.**

Se entiende por línea basal el conjunto de datos fundamentales reunidos con anterioridad a la intervención, a manera de información mínima, que refleje las condiciones de salud de la población en áreas maláricas antes de la puesta en práctica de la estrategia de estratificación epidemiológica y que permita medir los cambios y efectos producidos por la estrategia basada en el enfoque de riesgo.

Los indicadores de la evaluación deben utilizar la información de la línea basal con el fin de verificar los cambios que han ocurrido en el área de estudio. La selección de indicadores guarda una relación estrecha con el diseño del estudio para la evaluación. A su vez el diseño de estudio reflejará la existencia o no existencia de grupos control y el tipo de información básica que será utilizada.

3. **Diseño del estudio de evaluación**

En la literatura sobre metodología de evaluación de los servicios de salud se mencionan diversos diseños posibles para el estudio de evaluación de nuevas intervenciones. A continuación se presentan algunos de dichos diseños de investigación.

a) Un diseño de corte transversal único. En este diseño, todas las mediciones son realizadas en un punto en el tiempo posterior a la intervención.

b) Un diseño antes/después, sin área(s) de comparación. Este diseño se basa en encuestas de corte transversal, la primera realizada antes y la segunda realizada después de la intervención. En este diseño, la premisa fundamental de la que se parte es que las diferencias observadas en los indicadores antes y después es debida a los efectos de la(s) intervención(es) que ocurrieron durante dicho intervalo de tiempo.

c) Un diseño antes/después, con áreas de comparación, donde la información puede ser tomada de las estadísticas existentes, recolectadas rutinariamente. Este diseño implica la medición de las variables en dos grupos de población, uno al cual se introdujo la nueva estrategia de intervención y en otro al cual se mantiene la intervención tradicional. Así mismo, requiere dos mediciones, una antes de la intervención y otra después de que la intervención ha sido introducida. El grupo de comparación debe ser lo mas semejante al grupo en estudio, difiriendo solo en el tipo de intervención realizada.

d) un diseño antes/después en el área de intervención con estudios de corte transversal y usando los datos nacionales como área de comparación.

Estos diseños de estudios de evaluación como todos los estudios observacionales pueden estar sujetos a problemas de validez y confiabilidad. Debido a ésto, la generalización de los resultados dependerá de la magnitud y dirección de los sesgos introducidos o encontrados en los mismos. Así mismo, deben considerarse los aspectos éticos relevantes hacia la población que se está estudiando.

Para la evaluación de una nueva estrategia de intervención, la situación ideal sería tener una o mas áreas de comparación donde no se introduzcan cambios en el sistema de atención e intervención. Tanto en las áreas de estudio como en las de comparación, deberían evaluarse los niveles de transmisión malárica en los grupos de alto y bajo riesgo así como documentarse las diferencias encontradas.

Si existen diferencias estadísticamente significativas en los niveles de transmisión malárica en los grupos de alto y bajo riesgo entre las áreas de estudio y comparación, podría decirse que la nueva estrategia modificó la situación malárica y de salud en el área de estudio, y si tal cambio fue positivo, podemos sugerir que la nueva estrategia de intervención es mas efectiva que la anterior.

#### 4. Análisis de datos y conclusiones

A partir de los aspectos señalados anteriormente, se puede afirmar que los principales componentes del proceso de evaluación son los siguientes:

- a) el objeto de estudio en la evaluación es una actividad o intervención que se introduce en un programa o servicio;
- b) presupone la existencia de objetivos previamente establecidos;
- c) implica un proceso de medición; y
- d) requiere de evidencia que muestre si el cambio observado es debido a las actividades o intervenciones realizadas.

Consistente con lo anterior, en todo estudio de evaluación deben formularse dos preguntas centrales. En la primera de ellas debe plantearse si durante el curso de la implementación de la nueva estrategia, programa o servicio hubo cambios significativos en los niveles de salud de la población cubierta por tales acciones. La segunda pregunta fundamental es reconocer si dichos cambios y/o el logro de los objetivos planteados pueden ser atribuidos al programa o intervención planeada. Ambas preguntas deben ubicarse en un período de tiempo definido, esto significa que el proceso de evaluación requiere de un lapso de tiempo que corresponde al punto en que el posible cambio puede ser observado, medido y analizado.

Los cambios observados en los grupos de estudio y control pueden ser medidos mediante pruebas de significancia estadística. Las mas comunes son pruebas de Chi-cuadrado, análisis de varianza, regresión múltiple lineal y regresión logística. Una discusion extensa de cada una de ellas puede ser encontrada en textos de bioestadística y epidemiología avanzada.

La idea general del análisis estadístico es documentar cuidadosamente todas las diferencias y conclusiones que puedan obtenerse de los datos. El análisis de los datos permite derivar conclusiones acerca del impacto alcanzado por la nueva estrategia de riesgo. Hay que tener gran precaución en la interpretación de los datos. Lo que a primera vista puede parecer un juicio simple sobre la diferencia entre dos series de datos, es en realidad un proceso complicado. Los métodos estadísticos usados en el análisis deben ser trabajados con sumo cuidado. La participación en el proceso de evaluación deberá ser amplia e incluir al grupo de estudio, al personal de salud, y de administración, al personal con poder en las decisiones políticas, y a los representantes de la comunidad. Cada uno de estos grupos tiene intereses propios que influyen en la interpretación de los datos. Cuando se realiza una evaluación como un proceso de grupo, es difícil sesgar los resultados. Por otra parte, la participación amplia ayudará a facilitar la aplicación de los resultados. El estudio no debe perder su valor científico en general y las interpretaciones del estudio deben ser aceptadas como parte de un proceso socio-político tendiente al control de la malaria y al desarrollo de los niveles de salud.

### EJERCICIO

1. Formule una hipótesis de cambio que pueda ser usada en la evaluación de la estrategia de riesgo en el área malaria en estudio.
2. Mencione el diseño de estudio de evaluación que utilizaría en la evaluación de las intervenciones basadas en el enfoque de riesgo. Justifique su selección. Discuta convenientes y limitaciones del mismo.
3. Mencione tres indicadores para medir elementos de estructura, proceso e impacto relacionados con la estrategia de estratificación epidemiológica de la malaria.
4. Con el fin de evaluar la nueva estrategia de intervención en malaria, use su conocimiento general de Malarianda y sus respuestas a los distintos ejercicios para describir las áreas de estudio y de comparación que usted elegiría para esta evaluación.
5. Prepare los cuadros de contingencia necesarios para cada uno de los indicadores mencionados en el ejercicio dos.
6. Para cada uno de los cuadros preparados sugiera que tipo de comparaciones y pruebas estadísticas aplicaría.

7. Discuta que otras variables pueden estar asociadas con el resultado a ser observado. Señale algunas fuentes de error y discuta las probables limitaciones acerca de la generalización de los resultados.

RECEIVED  
ADMINISTRACION  
DE LA UNIVERSIDAD  
DE LOS ANGELES  
1970

Administracion de la Universidad de Los Angeles

15251

Dic./90

TALLER SOBRE  
ESTRATIFICACION DE LA MALARIA

FAVOR NO REPRODUCIR  
SIN LA AUTORIZACION  
DE LA ORGANIZACION  
PANAMERICANA DE LA  
SALUD (OPS)

MATERIALES DE TRABAJO

II

"LA INVESTIGACION EPIDEMIOLOGICA EN MALARIA"

( CONCEPTOS BASICOS )

FAVOR NO REPRODUCIR  
SIN LA AUTORIZACION  
DE LA ORGANIZACION  
PANAMERICANA DE LA  
SALUD (OPS)

NOTA :

VERSION BORRADOR

FAVOR DE NO REPRODUCIR O CITAR SIN LA AUTORIZACION DE LA  
ORGANIZACION PANAMERICANA DE LA SALUD

**LISTA DE CUADROS**

Cuadro 1	7
Cuadro 2	10
Cuadro 3	13
Cuadro 4	15
Cuadro 5	25
Cuadro 6	25
Cuadro 6-A	26
Cuadro 7	27
Cuadro 8	29
Cuadro 9	33

**EJERCICIOS**

Ejercicio 1	20
Ejercicio 2	29
Ejercicio 3	35

# I N D I C E

UNIDAD DE INVESTIGACION  
SIN LA APOYABILIDAD  
DE LA ORGANIZACION  
PANAMERICANA DE LA  
SALUD (OPS)

	Página
I. <u>Introducción</u>	1
II. <u>La Investigación Epidemiológica</u>	
2.1 Concepto	2
2.2 El Razonamiento Epidemiológico	2
2.3 Etapas de la Investigación Epidemiológica	3
III. <u>Tipos de Diseño de Estudios Epidemiológicos</u>	
3.1 Introducción	5
3.2 Diseño de Estudios Experimentales	5
3.3 Diseños de Estudios de Observación	6
3.4 Tipos de Estudios de Observación	8
3.4.1 E. de Corte Transversal	9
3.4.2 E. Retrospectivos (Casos y Controles)	9
3.4.3 E. Prospectivos	12
IV. <u>Investigación Epidemiológica de Malaria</u>	
4.1 Generalidades	16
4.2 Definición de Propósitos y Objetivos	17
4.3 Formulación de Hipótesis de Investigación	18
4.4 Selección del Diseño de Estudio Epidemiológico	21
4.5 Análisis de las relaciones entre Malaria y los probables Factores de Riesgo	22
4.5.1 Introducción	22
4.5.2 Prueba del Chi-Cuadrado	22
4.5.3 Cálculo de la Prueba del Chi-Cuadrado	26
4.6 Cálculo del Riesgo Relativo y Riesgo Atribuible	31
4.7 Preparación del Reporte Final de la Investigación	36
Bibliografía	37

## INVESTIGACION EPIDEMIOLOGICA DE LA MALARIA

NO SE REPRODUCE  
SIN LA AUTORIZACION  
DE LA ORGANIZACION  
PANAMERICANA DE LA  
SALUD (OPS)

### 1. Introducción

La finalidad de esta segunda parte de los materiales de trabajo es proporcionar los elementos básicos de la investigación epidemiológica usada en apoyo al proceso de Estratificación Epidemiológica de Areas Maláricas en países de América Latina y el Caribe. Los materiales ofrecen una introducción general a los conceptos e instrumentos epidemiológicos que facilitan el reconocimiento de las relaciones existentes entre la malaria y los diversos factores de riesgo que participan y determinan la transmisión malárica por medio del mosquito vector en los diversos-- grupos humanos.

Es importante destacar que el conocimiento de los principios y métodos epidemiológicos no es campo exclusivo de técnicos o profesionales dedicados a la investigación epidemiológica. Este conocimiento es fundamental para todos los integrantes del equipo de salud. El estudio epidemiológico permite el reconocimiento de los determinantes de la distribución diferencial de los riesgos de enfermar o morir en los distintos grupos sociales que componen una población o comunidad determinada. Así mismo, facilita y fundamenta las opciones y alternativas de intervención existentes para la prestación de los servicios de salud. Los resultados de la investigación epidemiológica aportan elementos importantes para el desarrollo de infraestructuras capaces de apoyar a los sistemas locales de salud en la prevención y el control de la malaria bajo la estrategia de atención primaria en salud.

## 2. La Investigación Epidemiológica

### 2.1 Concepto

La investigación epidemiológica implica el estudio de la frecuencia y distribución de los determinantes del proceso de salud-enfermedad, en grupos de población humana. Debido a que los daños a la salud no se distribuyen al azar en toda la población, la búsqueda de las características, eventos o factores que influyen y determinan dicha distribución es parte fundamental de la investigación epidemiológica.

La información epidemiológica referente al grupo poblacional con una frecuencia mayor de una enfermedad o con un mayor riesgo de contraerla es útil a los servicios de salud, ya que indica en cuales segmentos de la población deben concentrar sus actividades, así como el tipo de intervenciones que deben ser realizadas.

### 2.2 El razonamiento epidemiológico

La epidemiología es considerada como una serie ordenada de razonamientos relacionados con la observación de un daño a la salud que ocurre en grupos específicos de la población y cuyo objeto es poder dilucidar y determinar los factores causales de dicho daño a la salud.

El método epidemiológico es primariamente comparativo. El estudio de los daños a la salud se realiza comparando diferentes tiempos, lugares o grupos poblacionales. En el estudio de un daño a la salud, la presencia de ciertas características o la exposición a determinados factores en un grupo poblacional (con el daño) se compara con la exposición o presencia de dichos factores en otro grupo poblacional (sin el daño).

La determinación de las vinculaciones, relaciones y asociaciones entre factores y daños a la salud, así como la formulación de inferencias basadas en dicha determinación es parte central del proceso del razonamiento epidemiológico.

Es importante mencionar que si bien la medición de los eventos epidemiológicos representa una parte básica de la investigación, no es suficiente para la formulación de las inferencias epidemiológicas acerca de la población. Al respecto, el epidemiólogo americano W. Frost (1936) establecía estas consideraciones de la siguiente manera:

"La epidemiología, en cualquier tiempo, es algo más que el total de sus hechos establecidos. Abarca su agrupación ordenada en cadenas de inferencia que rebasan más o menos los límites de la observación directa. Tales cadenas, dispuestas bien, guían la investigación hacia los hechos del futuro; las mal hechas retardan el progreso".

### 2.3 Etapas de la investigación epidemiológica

Una de las consideraciones fundamentales de la investigación epidemiológica es la definición del marco teórico o conceptual que permite dar dirección y sentido a cada una de las etapas de la investigación que a continuación se mencionan. Así, por ejemplo, el enfoque epidemiológico de riesgo tiene como marco de referencia central, la búsqueda de la eliminación de las desigualdades en la distribución del riesgo de enfermar y morir de los distintos grupos humanos. El concepto y uso de la noción de factor de riesgo se articula con el concepto de equidad. En este sentido, los riesgos diferenciales de enfermar o morir

son conceptualizados como estructuras determinadas por las condiciones de vida de los distintos grupos socioeconómicos. Dentro de estas condiciones, destacan los problemas de salud y los factores de riesgo que los determinan, así como el acceso de los distintos grupos a los servicios de salud. Así, los conceptos de equidad y de eliminación de las desigualdades en el riesgo de enfermar o morir constituyen un substrato importante en el marco de referencia de la investigación epidemiológica de riesgo.

Para lograr establecer las relaciones de determinación entre el daño a la salud seleccionado y sus factores de riesgo, se deben considerar diversos pasos o etapas en la investigación epidemiológica, entre los que se destacan:

**ETAPAS BASICAS DE LA INVESTIGACION EPIDEMIOLOGICA:**

- (a) el planteamiento del problema;
- (b) la formulación de los objetivos específicos e hipótesis de investigación del estudio epidemiológico;
- (c) la selección de un diseño de estudio que sea apropiado para demostrar las hipótesis de investigación formuladas;
- (d) la definición y selección e la población en estudio y del método de muestreo adecuado;
- (e) en caso de tratarse de un estudio muestral, el cálculo del tamaño de la muestra requerido;
- (f) la elaboración del listado de datos e información que se recolectara, así como de los cuadros de contingencia a ser usados en el análisis de la información;
- (g) el análisis de las relaciones entre daños a la salud y los factores de riesgo estudiadas, a través de pruebas de significancia estadística y el cálculo del riesgo relativo, riesgo atribuible y de otras medidas relevantes;
- (h) la preparación de los perfiles de riesgo que describan los grupos de alto riesgo así como la distribución de riesgo en la comunidad; y

- (i) la preparación y difusión del informe final de los resultados de la investigación de forma tal, que puedan ser utilizados para el diseño de estrategias de intervención por parte de los servicios de salud.

Cada una de estas etapas será descrita brevemente en el apartado cuarto, en el contexto de la investigación epidemiológica en malaria.

A continuación se presentan los diseños de investigación epidemiológica mas comunmente utilizados en el estudio de las enfermedades y sus determinantes.

### 3. Tipos de Diseño de Estudios Epidemiológicos

#### 3.1 Introducción

En epidemiología, en general, los estudios de investigación pueden ser divididos en dos grandes grupos: el estudio experimental y el estudio observacional. Una de las diferencias que existe entre estos dos tipos de estudio radica en que en el estudio experimental, el equipo de investigación puede especificar y controlar las condiciones bajo las cuales se ha de conducir el estudio, mientras que en el estudio observacional no se controlan esas condiciones.

#### 3.2 Diseños de Estudios Experimentales

En epidemiología, el estudio experimental es denominado Ensayo Clínico Controlado.

#### ENSAYO CLINICO CONTROLADO:

"Estudio en que las condiciones del mismo están bajo el control directo del investigador. La población es seleccionada para el estudio controlado de una

intervención o régimen en donde los efectos son medidos comparando el resultado de la intervención en el grupo experimental con el resultado de otra intervención en un grupo control. Con el propósito de evitar sesgos, tanto los miembros del grupo experimental como del grupo control deben ser comparables excepto en el régimen que se les ofrece. En un ensayo clínico aleatorio controlado, la asignación de los individuos al grupo experimental o grupo control es a través de la aleatorización o azar".  
(Last:35).

En los ensayos clínicos controlados se estudia la eficacia y seguridad de intervenciones terapéuticas o preventivas. La eficacia de medicamentos, vacunas, procedimientos quirúrgicos y el impacto de intervenciones como la fluorización del agua en la prevención de la caries dental, la aplicación de insecticidas y la eliminación de criaderos de anofelinos para el control de malaria, son ejemplos de este tipo de estudio. En los estudios de población humana, por consideraciones éticas, no siempre es factible llevar a cabo los estudios experimentales, por lo que la investigación debe realizarse a través de los estudios de observación.

### 3.3 Diseños de Estudios de Observación

Como se ha indicado previamente, en los estudios de observación, el equipo de investigación no crea las condiciones o manipula las variables participantes, sino que los resultados son observados bajo condiciones o situaciones "naturales" o sin manipulación.

#### ESTUDIO DE OBSERVACION:

"Es aquel realizado en situaciones en que los cambios o diferencias en una característica son estudiados en relación a cambios o diferencias en otra característica, sin la intervención del investigador". (Last:72).

En los estudios de observación el equipo de investigación realiza el estudio bajo las situaciones, tal como ellas se presentan. Los resultados obtenidos se relacionan a los distintos grados de exposición del factor o factores estudiados tal como éstos se manifestaron "naturalmente". Los individuos o grupos sociales tienen atributos o características propias o están expuestos a un factor, la exposición a un factor no es dada o creada por el investigador. Por ejemplo: el efecto de la edad avanzada, habitación sin paredes, pobreza, tipo de ocupación y su vinculación con el desarrollo de una enfermedad u otro daño a la salud.

Una fuente de sesgo o problema en los estudios de observación está dada por el hecho de que los grupos observados pueden diferir en otras características además del factor específico que se estudia. Si estas características son relevantes en la cadena de causalidad, pueden originar confusión y afectar las conclusiones si no se realiza el ajuste o control de las características.

Los datos recolectados en un estudio de observación pueden clasificarse bajo la forma de un cuadro de contingencia. Usualmente, este cuadro es cuadrangular, como lo indica el cuadro 1.

Cuadro 1

Clasificación de los datos recolectados en estudios epidemiológicos de observación.

		<u>DANO A LA SALUD</u>		
		Presente	Ausente	TOTAL
Exposición Al Factor	SI	A	B	A + B
	NO	C	D	C + D
TOTAL		A+C	B + D	A+B+C+D

### 3.4 Tipos de Estudios de Observación

Los dos eventos relevantes que sirven para clasificar y reconocer el tipo de estudio epidemiológico que se está diseñando son:

1. Un daño a la salud, por ejemplo: enfermedad, incapacidad, lesiones o muerte.
2. La exposición a uno o varios factores de riesgo.

La forma como estos dos eventos ocurren en el tiempo y la manera de observarlos metodológicamente, permite caracterizar el tipo de diseño de estudio epidemiológico de que se trata.

Así, los estudios analíticos de observación pueden dividirse en tres tipos: 1) estudios transversales; 2) estudios retrospectivos, y 3) estudios prospectivos. En los estudios transversales y los retrospectivos, la selección de los individuos se basa en la presencia o ausencia de la enfermedad u otro daño a la salud, mientras que en el prospectivo, la selección de los individuos se hace de acuerdo con la exposición a un factor, agente o condición. Un estudio transversal se diferencia de uno retrospectivo porque en el transversal, la presencia del daño y la exposición al factor se obtienen simultáneamente, es decir, no se puede reconocer cuál de los dos eventos ocurre primero en el tiempo; dicho de otra forma, no se puede demostrar si la exposición ocurrió previamente al daño. En el estudio retrospectivo, la exposición o característica ha ocurrido o existido en el pasado, previamente a la presencia del daño (Lilienfeld y L.:174).

### 3.4.1 Estudios de Corte Transversal

Los estudios de corte transversal, también llamados estudios de prevalencia, son aquellos en que la información sobre la presencia de los factores de riesgo y los daños a la salud se obtiene simultáneamente, lo cual generalmente impide reconocer la relación de tiempo entre una posible asociación causal de uno de ellos sobre el otro. Este diseño no permite diferenciar si la presencia del factor de riesgo es anterior o posterior o simultáneo al daño observado. Permite, sin embargo establecer si existe una asociación estadística entre ellos.

#### ESTUDIO DE CORTE TRANSVERSAL:

ESTUDIO QUE EXAMINA LA RELACION ENTRE ENFERMEDAD Y OTRAS VARIABLES DE INTERES TAL COMO EXISTEN EN UNA POBLACION DADA Y EN UN MOMENTO DETERMINADO. ESTA RELACION SE EXPRESA EN TERMINOS DE PREVALENCIA DE LA ENFERMEDAD EN LOS SUBGRUPOS CON O SIN LAS VARIABLES DE INTERES. LA SECUENCIA TEMPORAL DE CAUSA Y EFECTO NO PUEDE NECESARIAMENTE SER DETERMINADA. (LAST:24)

Este tipo de diseño es el sencillo y generalmente es utilizado para obtener un perfil de la población y para la determinación de la prevalencia de factores de riesgo o de un daño a la salud en dicha población.

Los estudios de corte transversal, en general representan una muestra tomada de toda la población; se realizan en un periodo único, breve y bien delimitado; y se emplean para describir los atributos de quienes han experimentado un daño a la salud y aquellos que no lo han sufrido.

### 3.4.2 Estudio Retrospectivo (Casos-Controles)

En el estudio retrospectivo, también denominado de casos-controles, el investigador inicia con la determinación de los casos y los controles.

Siguiendo el cuadro 2, se seleccionan A + C casos, los cuales van a ser comparados con B + D controles. En este cuadro de 2 x 2, los casos (A+C) y los controles (B+D) representan respectivamente el total de las columnas. Posteriormente, tanto los casos como los controles son asignados a las celdillas de las hileras, que representan la exposición o no exposición retrospectiva al factor de riesgo estudiado.

**ESTUDIO RETROSPECTIVO:**

ESTUDIO QUE SE INICIA CON LA IDENTIFICACION DE PERSONAS CON LA ENFERMEDAD DE INTERES, Y UN GRUPO APROPIADO DE PERSONAS SIN LA ENFERMEDAD. LA RELACION ENTRE UN FACTOR O ATRIBUTO Y LA ENFERMEDAD ES EXAMINADA COMPARANDO LOS CASOS Y NO CASOS CON RESPECTO A QUE TAN FRECUENTE ESTA PRESENTE EL ATRIBUTO EN CADA UNO DE DICHS GRUPOS.  
(Last:15).

**Cuadro 2**  
**DISEÑO DEL ESTUDIO RETROSPECTIVO (Casos-Controles).**

		DAÑO A LA SALUD		
		Presente CASOS	Ausente CONTROLES	TOTAL
Exposición SI Al Factor	SI	A	B	
	NO	C	D	
TOTAL		A+C	B + D	A+B+C+D

- 
- 1: EL ESTUDIO SE INICIA CON:  
LA SELECCION DE CASOS Y NO CASOS (CONTROLES).
  - 2: EL ANALISIS SE REALIZA MIDIENDO LA EXPOSICION  
PREVIA A FACTORES DE RIESGO EN CASOS Y EN CONTROLES.
  - 3: EL OBJETIVO ES COMPARAR LA PROPORCION DE CASOS  
EXPUESTOS A UN FACTOR DE RIESGO CON LA PROPORCION DE  
CONTROLES EXPUESTOS AL MISMO FACTOR DE RIESGO.
-

El análisis del estudio retrospectivo se realiza comparando las tasas de exposición al factor de riesgo de los casos con las tasas de exposición de los controles. Si mediante pruebas de significancia estadística se encuentra que la frecuencia de individuos con el factor de riesgo es mas elevada en los casos que en los controles, y la diferencia es estadísticamente significativa, puede suponerse que hay una asociación estadística entre la enfermedad y el factor estudiado.

Para medir el grado o fuerza de la asociación en los estudios retrospectivos, se han elaborado diversos métodos estadísticos. Una de las medidas más importantes es la llamada RAZON DE DISPARIDAD, también denominada razón de productos cruzados o razón de momios. Esta medida es una razón en donde el numerador incluye el momio (Odds) de los casos expuestos al factor de riesgo y en el denominador el momio (Odds) de los controles expuestos al mismo factor de riesgo.

Es importante destacar que la razon de disparidad puede ser un estimador del riesgo relativo si se llenan dos supuestos: a) la frecuencia de la enfermedad en la población es muy baja (enfermedades raras); b) los casos y los controles estudiados deben ser representativos, respectivamente, de los casos y de los "no casos" que existen en la población (Lilienfeld y L: 188).

De acuerdo al cuadro 2 las fórmulas de las medidas usadas en el diseño retrospectivo son las siguientes:

$$\text{Tasa de Exposición de los Casos} = \frac{A}{A + C}$$

$$\text{Tasa de Exposición de los Controles} = \frac{B}{B + D}$$

$$\text{RAZON DE DISPARIDAD} = \frac{A / C}{B / D}$$

Donde A/C es el momio (Odds) de los casos expuestos al factor y B/D es el momio (Odds) de los controles expuestos al factor.

La razón de disparidad es también denominada razón de productos cruzados porque es equivalente a:

$$\frac{A D}{B C}$$

---

### 3.4.3 Estudio Prospectivo

El estudio prospectivo es también denominado estudio de cohorte, estudio longitudinal o estudio de incidencia. El diseño de este estudio permite el cálculo de incidencia y del riesgo relativo.

#### **ESTUDIO PROSPECTIVO:**

ESTUDIO EN EL CUAL SE IDENTIFICAN SUBGRUPOS DE POBLACION EXPUESTOS Y NO EXPUESTOS A UN FACTOR, QUE HIPOTETICAMENTE INFLUYE EN LA PROBABILIDAD DE QUE OCURRA UNA ENFERMEDAD U OTRO DAÑO A LA SALUD. IMPLICA LA OBSERVACION DE UN NUMERO SUFICIENTE DE PERSONAS-AÑOS, QUE PERMITA GENERAR TASAS DE INCIDENCIA CONFIABLES EN LOS SUBGRUPOS DE POBLACION. (Last:20).

En el estudio prospectivo, al inicio del estudio, usualmente, ha ocurrido la exposición al factor de riesgo, mientras que la enfermedad aun no ocurre. Tanto el grupo de individuos expuestos como el de no expuestos están libres de enfermedad. El investigador inicia el estudio con la selección de los individuos expuestos al factor de riesgo, que

corresponden en el cuadro 3 al total A + B expuestos y de individuos no expuestos C + D. Ambos grupos se siguen durante un período especificado de tiempo. Al final de este período de seguimiento, se observa el número de individuos expuestos que desarrollaron la enfermedad, así como el número de no expuestos que se enfermaron.

Cuadro 3

DISEÑO DEL ESTUDIO PROSPECTIVO (Cohorte).

- 1: EL ESTUDIO SE INICIA CON:  
LA SELECCION DE EXPUESTOS Y NO EXPUESTOS AL FACTOR DE RIESGO.
  
- 2: EL ANALISIS SE REALIZA MIDIENDO LA PRESENCIA DEL DAÑO A LA SALUD  
EN EXPUESTOS Y NO EXPUESTOS AL FACTOR DE RIESGO ESTUDIADO.
  
- 3: EL OBJETIVO ES COMPARAR LA INCIDENCIA DEL DAÑO A LA SALUD ENTRE  
LOS EXPUESTOS A UN FACTOR DE RIESGO CON LA INCIDENCIA ENTRE LOS  
NO EXPUESTOS AL MISMO FACTOR DE RIESGO.

	<u>DAÑO A LA SALUD</u>		
	Presente	Ausente	TOTAL
EXPUESTO AL FACTOR DE RIESGO	A	B	A + B
NO EXPUESTO AL FACTOR	C	D	C + D
	TOTAL		A+B+C+D

En el análisis, se compara la tasa de incidencia del grupo expuesto con la tasa de incidencia del grupo no expuesto. Debido a que este diseño permite estimar la incidencia del daño a la salud en los grupos de expuestos y no expuestos, se puede calcular directamente el riesgo relativo. Al poder medir directamente el riesgo absoluto y relativo de que se desarrolle un daño a la salud en una población, este diseño es muy poderoso para probar hipótesis sobre la relación entre los factores de riesgo y daños a la salud.

A continuación, se presentan las fórmulas mas comúnmente usadas en este diseño de estudio epidemiológico:

$$\begin{aligned} \text{Tasa de Incidencia en los Expuestos} &= \frac{A}{A + B} \\ \text{Tasa de Incidencia en No Expuestos} &= \frac{C}{C + D} \\ \text{RIESGO RELATIVO} &= \frac{A / A + B}{C / C + D} \end{aligned}$$

Las ventajas y desventajas de cada uno de los tres diseños observacionales se describen en el cuadro 4.

Cuadro 4  
COMPARACION DE LAS VENTAJAS Y DESVENTAJAS DE LOS PRINCIPALES  
TIPOS DE DISEÑO DE INVESTIGACION EPIDEMIOLOGICA

<u>TIPO DE DISEÑO</u>	<u>VENTAJAS</u>	<u>DESVENTAJAS</u>
<b>CORTE TRANSVERSAL:</b>	<ol style="list-style-type: none"> <li>1. Sencillo y barato.</li> <li>2. Permite determinar la prevalencia de una enfermedad o de un factor de riesgo</li> <li>3. Permite la descripción de la población.</li> </ol>	<ol style="list-style-type: none"> <li>1. No sustenta inferencias de causalidad.</li> <li>2. No se puede determinar incidencia ni el riesgo relativo.</li> </ol>
<b>RETROSPECTIVO:</b>	<ol style="list-style-type: none"> <li>1. Barato.</li> <li>2. Corto tiempo de estudio.</li> <li>3. Menor cantidad de individuos.</li> <li>4. Adecuado para enfermedades raras.</li> <li>5. Se identifica fácilmente a los casos.</li> </ol>	<ol style="list-style-type: none"> <li>1. Posibilidad de sesgos</li> <li>2. No se puede determinar la incidencia.</li> <li>3. El riesgo relativo es aproximado.</li> <li>4. Difícil selección de los controles.</li> <li>5. No es adecuado para Factores de Riesgo poco prevalentes.</li> </ol>
<b>PROSPECTIVO (Cohorte):</b>	<ol style="list-style-type: none"> <li>1. Menos posibilidad de sesgo</li> <li>2. Permite la determinación del riesgo relativo</li> <li>3. Permite la determinación de incidencia.</li> <li>4. Puede combinar el estudio de varios daños a la salud.</li> </ol>	<ol style="list-style-type: none"> <li>1. Requiere un tiempo de estudio prolongado.</li> <li>2. Frecuentemente costoso.</li> <li>3. No es adecuado para enfermedades de baja frecuencia.</li> <li>4. Requiere un mayor número de sujetos.</li> <li>5. Adolece del problema de pérdida de sujetos durante el seguimiento.</li> </ol>

Fuente: Modificado de OMS-OPS. Manual sobre el enfoque de riesgo en la atención Maternoinfantil. Paltex No. 7, Washington, D.C. 1986:55-56

#### 4. Investigación epidemiológica en malaria

##### 4.1 Generalidades

La necesidad de una investigación epidemiológica en malaria puede variar mucho no sólo de un país a otro, sino fundamentalmente de una localidad a otra. Las tasas de incidencia y mortalidad por malaria varían tanto geográficamente como por la composición social de las poblaciones. Por ejemplo, en 1986 aunque sólo el 3.8% de la población de la Región de las Américas estaba localizada en la Subregión de Centro America, Panamá y Belice, en ella se produjeron 123,000 casos de malaria, representando el 20% del total de casos para todas las Américas. Dentro de esta subregión, Costa Rica y Panama tuvieron solamente el 2% de los casos, mientras que el 98% se concentraba en los cinco países restantes (OPS 1988:5). Al interior de dichos países, podrían ser identificadas aquellas localidades y grupos poblacionales que contienen el mayor número de casos de malaria y en que la probabilidad de la ocurrencia de este padecimiento es mayor que en otras localidades o grupos humanos. Es de vital importancia establecer, en tales localidades y grupos humanos, el conjunto de características y factores, que explican el aumento de la probabilidad de enfermar o morir de malaria.

El reconocimiento de los factores de riesgo de la malaria, y de la importancia relativa de ellos, en cada una de las localidades y grupos humanos en un período determinado, es posible mediante estudios de investigación epidemiológica.

La investigación epidemiológica busca la recolección y agrupación de la información sobre los factores de riesgo asociados con un

mayor riesgo de enfermar de malaria, estableciendo un conocimiento mas completo de la situación epidemiológica de la malaria. Así mismo, facilita el establecimiento de los perfiles de riesgo de la población y posibilita la selección de intervenciones dirigidas específicamente hacia la modificación o eliminación de los mismos en las localidades y grupos de mayor riesgo.

Los tres supuestos básicos que siguen la secuencia del razonamiento de una investigación de malaria basada en el enfoque epidemiológico de riesgo pueden ser enunciados de la forma siguiente:

- (1) Dentro de las áreas maláricas seleccionadas, es posible identificar mediante indicadores basados en factores de riesgo, determinados subgrupos de la población que tienen una mayor probabilidad de enfermar de malaria que el resto de la población.
- (2) Dichos subgrupos representan una proporción substancial de la morbilidad por malaria que ocurre en la población total.
- (3) Es posible aumentar la efectividad y eficacia de los servicios generales de salud en el manejo de las intervenciones preventivas y de control de la malaria por medio de la redefinición de tipos de intervenciones y de recursos hacia:
  - a. los factores de riesgo causales identificados; y
  - b. los individuos y grupos con mayor riesgo de contraer malaria.

Estas tres premisas facilitan la formulación de los objetivos e hipótesis de investigación que establezcan las relaciones entre malaria y los diversos factores de riesgo en estudio.

#### 4.2 Definición de propositos y objetivos

En todas las situaciones en las que sea necesario realizar un estudio epidemiológico de riesgo en malaria, es fundamental que los propósitos y objetivos de la investigación estén definidos racionalmente.

Uno de los objetivos generales de la investigación epidemiológica en malaria es la identificación, en áreas maláricas, de aquellos individuos y grupos sociales cuyas condiciones de vida social, económica y ambiental los condicionan a un riesgo muy elevado de enfermar o morir de malaria.

#### 4.3 Formulación de hipótesis de investigación

La investigación epidemiológica se conduce para responder ciertas interrogantes específicas o hipótesis relacionadas con la frecuencia, distribución y determinantes de la malaria en los diversos grupos sociales.

Así, un paso inicial del estudio epidemiológico es la formulación de preguntas relevantes que puedan ser objeto de investigación. Tales preguntas son presentadas en forma de una o varias hipótesis. En tales hipótesis se expresa la probable relación existente en la población en estudio, entre la exposición a uno o varios factores de riesgo y la presencia de un daño a la salud, en este caso malaria.

##### HIPOTESIS DE INVESTIGACION:

Una suposición o conjetura que se plantea para explicar ciertos hechos o eventos y se emplea como base para llevar a cabo una investigación mediante la cual se busca demostrarla o refutarla.  
(OMS-OPS, 1986:47)

Hipótesis bien formuladas permitirán una mejor estructuración de la investigación. Un ejemplo del planteamiento de una hipótesis de investigación epidemiológica en malaria podría ser la siguiente:

"Los individuos de la localidad "x" que no tengan una fuente regular de atención médica, tendrán una morbilidad por malaria mayor que aquellos que tengan dicha fuente regular de servicios de salud"

El número de hipótesis dependerá de los objetivos y alcance de la investigación así como del número de factores de riesgo estudiados. La

mayoría de las hipótesis de investigación son redactadas en forma negativa. Esta manera de enunciar las hipótesis es denominada Hipótesis Nula.

**HIPOTESIS NULA:**

La hipótesis nula señala que los resultados en un estudio, experimento o prueba, no son diferentes de los que hubiera ocurrido como resultado sólo del azar. Es la hipótesis que establece que la variable independiente bajo estudio no tiene asociación o efecto sobre la variable dependiente". Last, 71.

La razón básica de la transformación de las hipótesis de investigación en hipótesis nulas es debido a que las Pruebas de Significancia Estadística utilizadas en el análisis de las hipótesis solo pueden rechazar (o no rechazar) la hipótesis nula cuando miden la probabilidad de que los resultados obtenidos en un estudio hayan ocurrido por azar.

La formulación de una hipótesis nula correspondiente a la hipótesis de investigación previamente enunciada es la siguiente:

"No hay diferencia en la morbilidad por malaria entre los individuos de la localidad "x" que no tengan una fuente regular de atención médica, y aquellos que tengan dicha fuente regular de servicios de salud"

La formulación de las hipótesis de investigación facilita la selección del diseño de estudio, el tipo de metodología a ser empleada, de las variables que se incluirán y de la forma en que los datos deben ser analizados. Así mismo, evita confusión, pérdida de tiempo y la recolección innecesaria de datos.

**EJERCICIO 1:**

El problema de malaria ha sido definido como prioritario en Malarianda,\* por lo que un grupo de investigadores de los servicios generales de salud de dicho país, ha decidido realizar una serie de estudios epidemiológicos de riesgo de la malaria en las regiones maláricas del país. En el Anexo I\* se presenta la información general del país así como de las cuatro regiones maláricas del país. Usando la información que tiene a su disposición, inicie su propuesta de investigación con las siguientes actividades:

**PREGUNTA 1.** Seleccione de las cuatro regiones maláricas de Malarianda, aquella que puede ser definida de mayor prioridad. Así mismo, identifique dentro de la misma, a la localidad o localidades en que se iniciará la investigación, estableciendo los criterios utilizados o razonamiento seguido en la selección de la región y localidad ó localidades.

**PREGUNTA 2.** Formule los objetivos generales de la investigación epidemiológica de riesgo de malaria.

**PREGUNTA 3.** Enuncie seis hipótesis de investigación que se desean demostrar, a fin de estudiar el papel de seis probables factores de riesgo que pudieran explicar el aumento en la incidencia de la malaria en las localidades seleccionadas.

**PREGUNTA 4.** Formule las hipótesis nulas correspondientes a cada una de las hipótesis de investigación previamente enunciadas. :

**PREGUNTA 5.** Elabore un listado preliminar de la información que requiere ser recolectada para probar las hipótesis de investigación.

\* Anexo I: Malarianda: Situación Sociodemográfica e Indicadores Malariométricos Básicos. La información presentada en este Anexo no representa a ningún país particular o real. El propósito es facilitar el proceso de enseñanza-aprendizaje y la realización de los ejercicios contemplados en los materiales.

#### 4.4 La Selección del Diseño de Estudio.

Como toda investigación científica, el estudio epidemiológico de malaria debe realizarse conforme a una estructura, plan o diseño. El diseño de un estudio permite estructurar el marco teórico de referencia con las hipótesis de investigación. Así mismo, permite la selección del número y tipo de características o factores a estudiar, así como de la forma de organizar, analizar e interpretar la información recolectada.

##### DISEÑO DE INVESTIGACION

Diseño de investigación contiene los procedimientos y métodos, predeterminados por el equipo de investigación, los cuales serán utilizados en la conducción de un proyecto de investigación.

(Last:92).

La selección y planeación del diseño de la investigación implica la identificación de las variables o características a estudiar, la definición de los procedimientos de selección de los grupos a ser investigados, así como del plan de aleatorización, si el diseño es experimental. Como se ha señalado en el capítulo 3, los diseños de investigación epidemiológica de uso mas extendido son los estudios de:

- (a) corte transversal; (b) retrospectivo (casos y controles);
- (c) prospectivo (cohorte) y (d) los ensayos clínicos controlados.

El examen de las hipótesis y objetivos de la investigación ayuda a tomar la decisión acerca de que tipo de diseño de estudio debe ser seleccionado. Así mismo, debe tomarse en cuenta las limitaciones de información y de recursos existentes. Finalmente, debe considerarse si el diseño de investigación podrá ofrecer la estructura necesaria para poder analizar adecuadamente las preguntas fundamentales del estudio.

#### 4.5 Análisis de las relaciones entre Malaria y los probables factores de riesgo estudiados.

##### 4.5.1 Introducción

Una vez seleccionado el diseño del estudio epidemiológico, se procede a la conducción del mismo. Así mismo, recolectada y depurada toda la información epidemiológica, ésta debe ser desglosada en sus respectivos cuadros de contingencia. Posteriormente se procede al cálculo de las medidas fundamentales del estudio y de la aplicación de la prueba estadística seleccionada para el análisis de las hipótesis del estudio.

El análisis e interpretación de los estudios epidemiológicos parten de los resultados obtenidos en las pruebas de significancia estadística. Una de las más utilizadas en el campo epidemiológico y que ejemplifica el uso de los cuadros de contingencia es la PRUEBA DEL CHI-CUADRADO.

##### 4.5.2 Prueba del Chi-Cuadrado

En la investigación epidemiológica, una estrategia analítica en la interpretación de los datos es el uso de una prueba estadística que permita reconocer si existe una asociación estadística entre la Malaria y el conjunto de factores riesgo postulados.

La asociación estadística indica el grado de probabilidad que tiene un valor observado, o una serie de valores, o de las diferencias entre dos valores de ser producto del azar.

#### **SIGNIFICANCIA ESTADÍSTICA:**

**LOS METODOS ESTADISTICOS PERMITEN ESTIMAR LA PROBABILIDAD DEL GRADO DE ASOCIACION IGUAL O MAYOR AL OBSERVADO QUE EXISTE ENTRE LAS VARIABLES INDEPENDIENTES Y LAS DEPENDIENTES BAJO LA SUPOSICION DE LA HIPOTESIS NULA. DE ESTA ESTIMACION, BAJO UNA MUESTRA DE UN TAMAÑO DADO, PUEDE SER EXPRESADA LA "SIGNIFICANCIA" ESTADISTICA DE UN RESULTADO. USUALMENTE EL NIVEL DE SIGNIFICANCIA ESTADISTICA ES EXPRESADO POR EL VALOR "P". (Last, 100).**

Si bien existe un gran número de pruebas de significancia estadística, la selección de alguna de ellas depende en gran medida del tipo de hipótesis de investigación formulada y de la medición estadística que se desee. En el apartado 4.3 se señala que en todas las pruebas de significancia estadística se incluyen las llamadas "hipótesis nulas". La elaboración de las "hipótesis nulas" es un requisito indispensable debido a que las pruebas estadísticas solo pueden ser usadas con las hipótesis de nulidad. Como se ha mencionado previamente, estas hipótesis plantean que no existen diferencias en los parámetros de población (promedio, proporciones, etc.) entre los dos o mas grupos comparados. Las pruebas estadísticas miden la probabilidad de que los resultados obtenidos en un estudio hayan ocurrido por azar (ver definición de valor de "P"). Si esto es extremadamente improbable, podemos entonces rechazar la hipótesis de "no diferencia" o hipótesis de nulidad y "aceptar" la hipótesis alternativa, o sea, la hipótesis propuesta en la investigación.

Debido a que el tema central de la investigación epidemiológica de estos materiales es el problema de malaria y se postula entre las hipótesis de investigación que el tipo de vivienda sin cerramientos externos es uno de los probables factores sociales de riesgo que explican un incremento en el riesgo de desarrollar malaria, debemos establecer primeramente la existencia de una asociación estadística entre la malaria y la exposición a este factor de riesgo. Una prueba muy utilizada en la determinación de esta asociación estadística es la llamada Prueba del "Chi-Cuadrado.

PRUEBA DEL CHI-CUADRADO.

La Prueba del Chi-Cuadrado es un método estadístico que permite determinar si dos o más frecuencias o proporciones son distintas estadísticamente o si un conjunto de proporciones difiere significativamente de una distribución teórica esperada. (Last, 18)

Esta prueba permite rechazar o no rechazar la hipótesis nula, es decir permite establecer la asociación o la falta de asociación entre la Malaria y cada uno de los postulados factores de riesgo que se estudian. Indica la probabilidad de que los resultados obtenidos se deban al azar. La anotación usada para registrar esta probabilidad es el Valor de "P". El valor de "P" expresa la probabilidad de que una diferencia tan grande como la observada podría haber ocurrido solamente por azar. El valor de "P" universalmente aceptado como nivel de aceptación de la significancia estadística es  $P < 0.05$ . Sin embargo la selección de este nivel es arbitrario y solo debe servir como guía en la interpretación de los resultados.

EL VALOR "P":

LA LETRA "P", SEGUIDA DE LA ABBREVIACION "n.s." (NO SIGNIFICANTE) O DEL SIMBOLO "<" (MENOS QUE), ASI COMO DE UNA NOTACION DECIMAL TAL COMO 0.01, 0.05, ES UNA EXPRESION DE LA PROBABILIDAD DE QUE LA DIFERENCIA OBSERVADA PODRIA HABER OCURRIDO POR AZAR SI LOS GRUPOS FUERAN REALMENTE SIMILARES. (Last,75).

Usando el cuadro 5, se tienen cuatro celdas, marcadas con las letras a, b, c, d, las cuales contienen el número de individuos que poseen las dos características al mismo tiempo. Así, por ejemplo, la celda a, expresa el número de individuos con el factor de riesgo que desarrollaron Malaria (el daño a la salud que se está estudiando). Las celdas b, c y d, expresan también otras relaciones importantes. La fórmula de la Prueba del Chi-Cuadrado es la siguiente:

La suma de las cuatro celdas de =  $\frac{(\text{valor observado} - \text{valor esperado})^2}{\text{valor esperado}}$

**Cuadro 5**  
Relación entre Malaria y participación social en la lucha antimalárica de los habitantes de la localidad X, Malarianda, 1988.

Factor Riesgo	Daño a la Salud MALARIA		Total
	Presente	Ausente	
Renuencia a la colaboración Antimalárica	A	B	A + B
Colaboración antimalárica	C	D	C + D
Total	A + C	B + D	N

Cálculo del valor esperado para cada celda:

VALOR ESPERADO =  $\frac{(\text{total de la columna}) \times (\text{total de la hilera})}{\text{total de casos (N)}}$

Para la celda "A", el valor esperado se calcula así:

VALOR ESPERADO DE "A" =  $\frac{(A+C) \times (A+B)}{N}$

FORMULA DEL CHI-CUADRADO = Suma de  $\frac{(\text{Observado} - \text{Esperado})^2}{\text{Esperado}}$

para cada celda.

**Cuadro 6**  
Relación entre Malaria y participación social en la lucha antimalárica de los habitantes de la localidad X, Malarianda, 1988.

Característica	M A L A R I A		Total
	SI	NO	
Renuencia a la colaboración antimalárica	152	1313	1465
Colaboración antimalárica	502	7911	8413
Total	654	9224	9878

**4.5.3 Cálculo de la Prueba del Chi-Cuadrado.**

Un ejemplo del cálculo de esta prueba se presenta a continuación. Tomando en cuenta los datos del cuadro 6, se calculan los valores esperados para cada celda suponiendo que la hipótesis nula es cierta. Usando la fórmula correspondiente, se obtuvieron los siguientes valores esperados:

**VALORES ESPERADOS:**

**GELDA A** :  $654 \times 1465/9878 = 97$     **GELDA C**:  $654 \times 8413/9878 = 557$

**GELDA B** :  $9224 \times 1465/9878 = 1368$     **GELDA D**:  $9224 \times 8413/9878 = 7856$

Cuadro 6-A  
Resultado del Cálculo de los valores esperados del Cuadro 6  
que serán usados para la prueba del Chi-cuadrado.

Factor Riesgo	Daño a la Salud MALARIA		Total
	Presente	Ausente	
Renuencia a la colaboración antimalárica	97	1368	1465
Colaboración Antimalárica	557	7856	8413
Total	654	9224	9878

Una vez obtenidos los valores esperados para cada celda, se procede al cálculo de la diferencia entre lo observado y lo esperado. Esta diferencia se eleva al cuadrado y se divide por el valor esperado. Este cálculo se realiza para las cuatro celdas y los resultados obtenidos se suman. En nuestro ejemplo, el resultado obtenido del valor de la prueba del Chi-Cuadrado fue 39.21. Este valor se compara con los valores

tabulados en el cuadro 7 para conocer el nivel de significancia (valor de P) alcanzado. Así, la probabilidad de encontrar un valor del Chi-Cuadrado con un grado de libertad de 2.71 o más, es de 10% y la de encontrar un valor de 3.84 o mayor es de 5 %. En nuestro ejemplo, como el valor del Chi-Cuadrado fue de 39.21, el nivel de significancia estadística fue menor de 0.001. Esto significa que la asociación entre el factor de riesgo y la malaria es improbable que sea debida al azar (menos de 0.001), ó lo que es lo mismo, la posibilidad de rechazar incorrectamente la hipótesis nula es inferior a 1 por 1000.

Pasos seguidos para obtener el valor de la Prueba del Chi-Cuadrado:

$$X^2 = \frac{(O - E)^2}{E} \text{ sumando para cada celda.}$$

$$X_2 = \frac{(152-97)^2}{97} + \frac{(1313-1368)^2}{1368} + \frac{(502-557)^2}{557} + \frac{(7911-7856)^2}{7856}$$

$$X_2 = 86.185 + -52.788 + -49.569 + 55.38$$

$$X_2 = 39.21$$

$$p < 0.001 \text{ n.s.}$$

Cuadro 7

Valores del nivel de significancia estadística para la Prueba del Chi-Cuadrado con un grado de libertad\*.

Quando el valor de Chi-Cuadrado es	El Nivel de Significancia (valor de P)
1.64	0.20
2.71	0.10
3.84	0.05
6.63	0.01
10.83	0.001

\* El cuadro de 2 x 2 contiene siempre datos con un grado de libertad.

Fuente: Snedecor G, Cochran W.: Métodos Estadísticos. Compañía Editorial Continental, México, 1980 pag.47

Obviamente que encontrar una asociación estadística no es suficiente para demostrar el gradiente de determinación o causalidad entre los factores de riesgo estudiados y la Malaria. Se debe determinar si la asociación estadística encontrada no fue resultado de un error o sesgo, es decir, que se trate de una asociación artificial o espuria. Así mismo, para la determinación de la causalidad, además de los criterios de la fuerza de asociación, deben ser considerados otros criterios no estadísticos, tales como especificidad, temporalidad y consistencia de la asociación así como la credibilidad biológica.

**EJERCICIO 2**

Siguiendo un diseño de estudio prospectivo se realizó el estudio epidemiológico de riesgo de la malaria. El cuadro 8 muestra el número de individuos que presentaron alguno de los primeros seis factores de riesgo estudiados así como en los que se observó la presencia o no de malaria.

**Cuadro 8**

**NUMERO DE INDIVIDUOS CON MALARIA QUE PRESENTARON ALGUNO DE LOS  
PRIMEROS SEIS FACTORES DE RIESGO ESTUDIADOS, MALARIANDA 1988.**

<u>Factor de Riesgo</u>	<u>Número de Individuos</u>	<u>M a l a r i a</u>	
		<u>Si</u>	<u>No</u>
1. ocupación campesino	433	52	381
2. ingreso familiar en línea de pobreza	654	70	584
3. tipo de vivienda sin cerramientos externos	598	121	477
4. escolaridad < 6 años	302	33	269
5. vivienda sin baño intradomiciliario	505	59	446
6. presencia de criaderos a menos de 500 m	714	190	524
Total de individuos en la muestra	9878	633	9245

**PREGUNTA 1:**

Utilizando la información del cuadro 8 complete los cuadros de contingencia para cada uno de los seis factores de riesgo.

**PREGUNTA 2:**

Calcule el valor del Chi-Cuadrado para cada uno de los seis cuadros elaborados.

**PREGUNTA 3:**

Comente los resultados obtenidos en el ejercicio anterior, señalando cuáles de los factores de riesgo tienen una mayor significancia estadística.

#### 4.6 Cálculo del RIESGO RELATIVO y RIESGO ATRIBUIBLE.

Una vez concluidas las pruebas de significancia estadística, el equipo de investigación cuenta con un listado de factores asociados estadísticamente con la malaria. El paso siguiente consiste en cuantificar el grado o fuerza de cada una de las asociaciones encontradas. La importancia de un factor de riesgo en particular, depende del grado de asociación y de determinación que tenga con la malaria, de la frecuencia que dicho factor de riesgo tenga en la comunidad, y de la posibilidad que exista de prevenirlo o controlarlo.

Por lo tanto, la tarea siguiente es calcular el riesgo relativo para cada uno de los factores de riesgo que se encontraron estadísticamente significativos. Recordemos que en nuestra investigación, el riesgo relativo es una medida de la probabilidad de experimentar malaria, que tienen aquellos individuos con alguno de los factores de riesgo estudiados en comparación con aquellos que no tienen tales factores.

Esta medida nos indica que tanto mayor es el riesgo de enfermar de malaria en aquellos con la característica o factor de riesgo con respecto a los que no la tienen.

El riesgo relativo mide la FUERZA DE ASOCIACION entre la malaria y cada uno de los postulados factores de riesgo.

El riesgo relativo así como el riesgo atribuible son medidas de asociación del exceso de riesgo debido a la exposición o presencia de un factor de riesgo determinado.

El riesgo atribuible mide el exceso de riesgo de malaria que puede ser atribuido a la exposición a un factor, por ejemplo, la exposición a una vivienda sin cerramientos externos.

El riesgo atribuible en la población total es una medida de asociación influenciada por la prevalencia del factor en la población total. El riesgo atribuible en la población mide el descenso porcentual en el número de casos de malaria que pueden evitarse si el factor causal fuera eliminado o neutralizado totalmente.

Las fórmulas y cálculos necesarios para estimar el riesgo relativo y el riesgo atribuible son las siguientes:

$$\text{RIESGO RELATIVO} = \frac{\text{Incidencia de los expuestos al Factor}}{\text{Incidencia de los No expuestos al Factor}}$$

$$\text{RIESGO ATRIBUIBLE EN LOS EXPUESTOS} = \frac{\text{Incidencia de expuestos} - \text{Incidencia de los No expuestos}}{\text{Incidencia de los No expuestos}}$$

$$\text{PORCENTAJE DE RIESGO ATRIBUIBLE EN LOS EXPUESTOS} = \frac{\text{Incidencia de expuestos} - \text{Incidencia de No expuestos}}{\text{Incidencia de expuestos}} \times 100$$

$$\text{PORCENTAJE DE RIESGO ATRIBUIBLE EN LA POBLACION} = \frac{\text{Incidencia en el total de la población} - \text{Incidencia de No expuestos}}{\text{Incidencia en el total de la población}} \times 100$$

Una fórmula alternativa es la siguiente:

$$\text{PORCENTAJE DE RIESGO ATRIBUIBLE EN LA POBLACION} = \frac{P (RR-1)}{1 + P (RR-1)} \times 100$$

donde P es igual a la proporción del factor de riesgo en la población; RR es igual a riesgo relativo.

En el cuadro 9 se presentan los datos referentes a otro de los factores de riesgo estudiados en la comunidad X de Malarianda: la renuencia de los habitantes en participar en la lucha antimalárica. A partir de los datos aquí contenidos, se presentan los cálculos del Riesgo Relativo y Riesgo Atribuible.

**Cuadro 9**  
Relación entre Malaria y participación social en la lucha antimalárica de los habitantes de la localidad X, Malarianda, 1988

Característica	M A L A R I A		Total
	SI	NO	
Renuencia a la colaboración antimalárica	152	1313	1465
Colaboración antimalárica	502	7911	8413
<b>Total</b>	<b>654</b>	<b>9224</b>	<b>9878</b>

Tasa de incidencia en expuestos  $152/1465 = 103.75$  por 1,000  
Tasa de incidencia en No expuestos  $502/8413 = 59.67$  por 1,000

$$\text{RIESGO RELATIVO} = \frac{103.75}{59.67} = 1.74$$

$$\text{RIESGO ATRIBUIBLE (en expuestos)} = 103.75 - 59.67 = 44.08$$

$$\% \text{ RIESGO ATRIBUIBLE EN EXPUESTOS} = \frac{103.75 - 59.67}{103.75} \times 100 = 42.49 \%$$

$$\% \text{ RIESGO ATRIBUIBLE EN LA POBLACION} = \frac{0.148 (1.74-1)}{1 + 0.148 (1.74-1)} \times 100 = 9.89\%$$

Según los resultados presentados en el ejemplo anterior, para esta población y en el tiempo dado, el riesgo de enfermar de malaria es 1.74 veces mayor en aquellos individuos que muestran renuencia a la lucha antimalárica, comparados con aquellos que participan en la lucha antimalárica. Para los expuestos (los que muestran renuencia al programa), 42.49% de su problema de malaria se explicaría por este factor de riesgo.

Por otro lado, 9.89% del problema de malaria en toda la comunidad se resolvería si dicho factor de riesgo fuera eliminado. Es importante destacar que existen otras posibles explicaciones para estos resultados, y que la presencia de otros factores concomitantes puede interferir en la actuación del factor de riesgo en la determinación de la ocurrencia de malaria. Para la correcta determinación de los resultados, dichos factores concomitantes deben ser reconocidos y tomados en cuenta en el análisis.

En resumen, el riesgo relativo y el riesgo atribuible son medidas que permiten identificar y reconocer el aporte que cada uno de los factores de riesgo estudiados tiene en el riesgo de desarrollar malaria. Así mismo, ofrecen información acerca de la proporción de malaria que se controlaría si cada uno de los factores de riesgo causales se eliminaran de la comunidad. Estas medidas permiten al equipo de salud calcular el nivel esperado de reducción de la malaria que se alcanzaría si el programa de salud se orientara específicamente a la modificación de los factores de riesgo causales estudiados. El equipo encargado de la implementación de las acciones esperaría que la reducción de los niveles de malaria fuera proporcional al exceso de riesgo eliminado. Así mismo, como resultado de la información y conocimiento epidemiológico local adquirido, nuevas estrategias de control de la malaria pueden emerger mejorando los objetivos operativos de la eliminación de aquellos factores de riesgo que determinan la distribución local de la malaria.

**EJERCICIO 3**

**PREGUNTA 1:**

Calcule el riesgo relativo de los factores de riesgo de las hipótesis de investigación seleccionadas previamente.

**PREGUNTA 2:**

Usando el riesgo relativo, ordene los factores de riesgo según su importancia y relación con la malaria.

**PREGUNTA 3:**

Calcule el riesgo atribuible para cada uno de los factores de riesgo listados en las preguntas anteriores.

**PREGUNTA 4:**

Estructure una nueva lista de factores de riesgo, jerarquizándolos de acuerdo al riesgo atribuible poblacional.

**PREGUNTA 5:**

Discuta las diferencias encontradas en las dos listas de factores de riesgo. Señale la importancia que tiene el riesgo relativo y el riesgo atribuible para un individuo en particular y para la comunidad.

**PREGUNTA 6:**

Cuáles de los factores de riesgo pueden modificarse?

- a) En el sector salud
- b) Intersectorialmente

**PREGUNTA 7:**

Señale la importancia de la investigación epidemiológica y de la medición de estas dos medidas de riesgo para la programación de las intervenciones para el control de la malaria.

**4.7 Preparación del reporte final de la investigación**

El equipo de investigación debe preparar una serie de informes y documentos para publicación y divulgación de las conclusiones de la investigación dirigidas tanto a los programas de salud como a las organizaciones sociales y comunitaria y la comunidad científica.

En la elaboración de los informes o reportes técnicos deben articularse claramente las hipótesis de la investigación con el diseño del estudio, el tipo de muestreo y métodos de recolección de la información, las técnicas de análisis y la interpretación de los resultados. Deben destacarse las implicaciones que los resultados tienen en la formulación de estrategias de intervención para el control de la malaria.

RECEIVED  
MAY 15 1968  
MEXICO CITY  
INSTITUTO VENEZOLANO DE LA  
SALUD (IVS)

B I B L I O G R A F I A

Mauser J. y Bahn A.: Epidemiología. Editorial Interamericana, Mexico, 1977.

Lilienfeld A. y Lilienfeld D.: Fundamentos de Epidemiología. Fondo Educativo Interamericano, Mexico, 1983.

Last J.M. (editor): A Dictionary of Epidemiology. Oxford University Press-I.E.A., New York, 1983.

Kahn H.A.: An Introduction to Epidemiologic Methods. Oxford University Press. New York, 1983.

OMS-OPS. Manual sobre el enfoque de riesgo en la atención maternoinfantil. Serie Paltex No. 7. Washington, D.C. 1986.

Schlesselman J.J.: Case-Control Studies: Design, Conduct, Analysis. Oxford University Press, New York, 1982.

Fleiss J.L.: Statistical Methods for Rates and Proportions. 2nd Edition. Wiley, New York, 1981.

Kleinbaum D.G., Kupper L.L., Morgenstern H.: Epidemiologic Research. Lifetime Learning Publications, Belmont, Calif., 1982.

Camel F.: Estadística Médica y de Salud Pública. Universidad de los Andes, Mérida, Venezuela, 1979.

Fox J.P., Hall C.E., Elveback L.R.: Epidemiology: Principles and Methods. Little, Brown & Co., Boston, 1970.

# Manual sobre el enfoque de riesgo en la atención maternoinfantil

Serie PALTEX para ejecutores de programas de salud No. 7

ORGANIZACION PANAMERICANA DE LA SALUD  
Oficina Sanitaria Panamericana, Oficina Regional de la  
ORGANIZACION MUNDIAL DE LA SALUD

1986

## Capítulo IV

### Identificación de los problemas prioritarios de salud

Introducción		Sección 4.3	Posibles fuentes de información ¿están disponibles todos los datos requeridos?
Sección 4.1	Revisión de las políticas y programas relevantes existentes.	Sección 4.4	Algunos métodos para obtener información complementaria
Sección 4.2	Información requerida para establecer el diagnóstico de salud y las prioridades. La disponibilidad de la información.	Sección 4.5	Presentación de la información.
		Sección 4.6	El establecimiento de prioridades en los problemas de salud maternoinfantil.

#### Objetivos de aprendizaje

Al finalizar este capítulo, los participantes deberán ser capaces de:

- Analizar las políticas y programas existentes de una manera sistemática.
- Analizar la información existente y desarrollar información adicional utilizando técnicas rápidas de evaluación.
- Cuantificar y colocar en orden de prioridades los problemas de salud.
- Conciliar dichas prioridades con las de la comunidad y con el marco político social respectivo.

#### Introducción

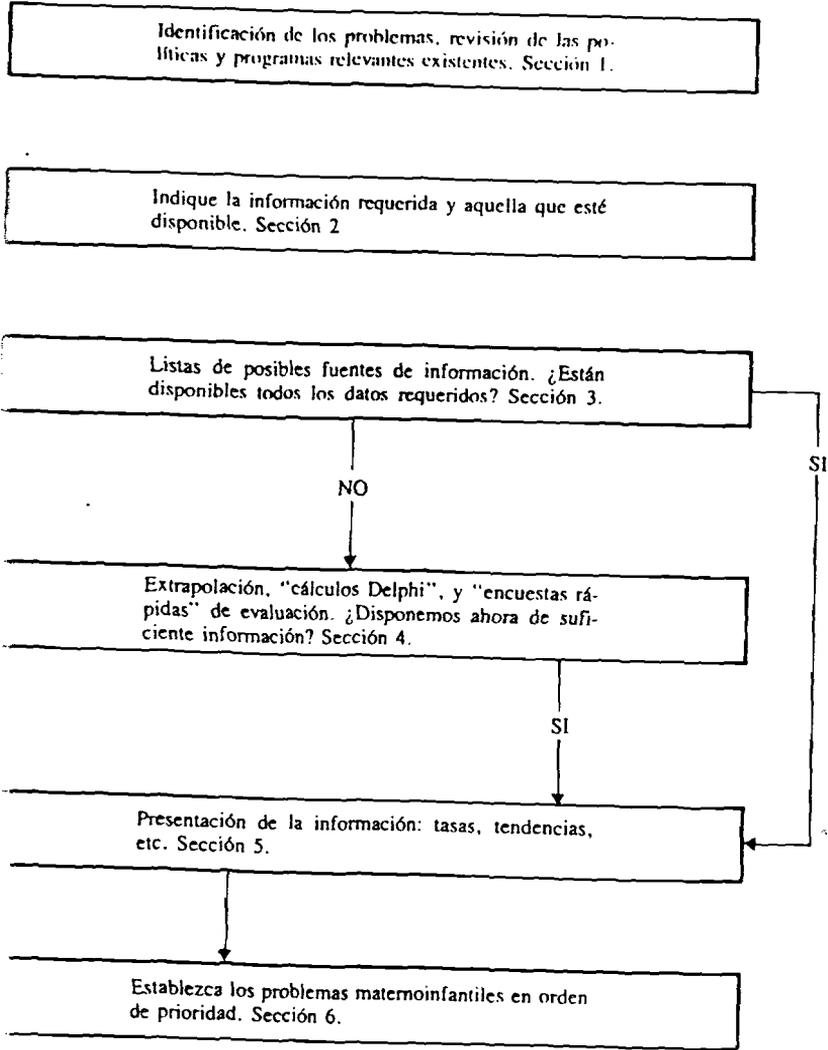
El análisis de los problemas de salud de madres y niños, así como la determinación de

prioridades, siguiendo el concepto de condiciones de vida de los grupos humanos y los factores de riesgo, constituyen los elementos esenciales para la programación y la asignación de recursos, de acuerdo con los requerimientos de las estrategias regionales. Este enfoque tiene los siguientes fundamentos:

La meta de "Salud para todos en el año 2000" y las estrategias regionales reconocen la existencia de desigualdades entre los diferentes grupos humanos y definen como un objetivo principal la eliminación de dichas desigualdades (equidad).

La existencia de grupos humanos postergados puede reconocerse a través de las desigualdades en las condiciones de vida de los distintos grupos socioeconómicos. Dentro de estas condiciones están los problemas de salud y

Cuadro 1. Esquema del proceso de identificación de los problemas de salud y el establecimiento de prioridades



los factores de riesgo que condicionan y determinan esos problemas, así como el acceso de los diferentes grupos a los servicios de salud.

Los vacíos de información que existen no constituyen impedimentos para el análisis, ni para la programación de acciones. El estado actual del conocimiento es suficiente para iniciar la formulación de intervenciones mínimas para afrontar los problemas prioritarios de salud maternoinfantil.

Las estrategias acordadas exigen la articulación intersectorial para elevar los niveles de vida y controlar los factores de riesgo que condicionan los problemas específicos de salud.

La determinación de prioridades y la asignación de recursos son esencialmente *actos políticos* que se definen y perfeccionan en términos políticos. Las consideraciones de tipo técnico y administrativo contribuyen a enriquecer y apoyar la información necesaria para la toma de la decisión política, pero no a su adopción.

Se ha indicado ya que el concepto y uso de la noción de factor de riesgo trasciende su connotación biológica orientada hacia conductas clínicas. Se requiere la identificación de los grupos humanos de acuerdo con sus condiciones de vida (social, económica y ambiental y de salud), así como su exposición particular a los diversos factores de riesgo que esas condiciones determinan. Este enfoque de la planificación y programación de la salud, constituye un marco de referencia para todas las acciones de salud.

El cuadro 1 esquematiza las etapas de revisión de la información existente y del establecimiento de prioridades de salud, tal como se discuten en este capítulo.<sup>1</sup>

<sup>1</sup> Para mayor información sobre este tema refiérase a: Bainbridge, J. y Sapire, S. *Health Project Management—A Manual of Procedures for Formulating and Implementing Health Projects*. Offset Publication 12, Geneva, WHO, 1974, págs. 39-40. *Setting Priorities for Health Development* (draft). WHO/HMD Documents, Geneva, 1980.

## Sección 4.1 Revisión de las políticas y programas relevantes existentes

### Resumen

La identificación de los problemas de salud de madres y niños requiere que la estrategia nacional de atención maternoinfantil/planificación familiar —presente y futura— se revise en relación con las políticas y programas existentes. Muchos creen que esto es bastante conocido y que no es necesaria una investigación especial; sin embargo, la experiencia ha mostrado que es más útil preparar un resumen de políticas y programas relevantes en forma ordenada y sistemática, sobre la base de las condiciones de vida de los grupos humanos y los factores de riesgo derivados de éstos.

Debe hacerse notar que en la región de las Américas las políticas y los planes nacionales de salud, frecuentemente comprenden los problemas prioritarios de salud maternoinfantil. Este conocimiento es útil, pero requiere la verificación de los criterios seguidos en la adopción de esas prioridades, así como la información utilizada para su establecimiento.

La relevancia de estos problemas no está siempre explícita en las políticas nacionales formales. En algunos casos los entendimientos no escritos son más importantes que la política escrita por lo que deben considerarse en este estudio. Se destaca que la identificación de problemas debe realizarse en un contexto nacional específico y tomando en consideración la situación política y social. Las fuentes de información constituyen un elemento esencial en esta etapa. Algunas fuentes valiosas podrían ser:

- Las estipulaciones concernientes al derecho a la atención de la salud dentro de las constituciones nacionales o federales.
- El plan nacional de salud. La mayoría de los países de América Latina y del Caribe ha adoptado estos planes de salud a mediano plazo. En muchos de ellos se establecen políticas y estrategias pertinentes al grupo maternoinfantil. Lo importante

es verificar si estos planes están en operación y si realmente se están aplicando las políticas.

- Los programas especiales en el área de salud maternoinfantil.
- Las asignaciones presupuestarias nacionales con énfasis en el área maternoinfantil.
- El plan nacional de desarrollo. Es importante identificar la forma como considera a los sectores sociales (subsidiarios, complementarios o constitutivos del desarrollo mismo), el papel del estado y la participación social, así como la referencia particular a los grupos humanos postergados y, dentro de ellos, a algunos sectores del grupo maternoinfantil.
- Los compromisos internacionales (Organización Mundial de la Salud, WHO); regionales (Organización Panamericana de la Salud, OPS; Estrategias y objetivos regionales) y subregionales (Grupo Andino) adquiridos por el gobierno.

En concreto, el análisis de las políticas de salud debe proporcionar información sobre:

- El reconocimiento nacional de la importancia de los problemas socioeconómicos, ambientales y de salud de los grupos humanos postergados, tanto en áreas rurales como urbanas.
- La prioridad asignada a madres y niños dentro de estos grupos.
- La posibilidad de que el enfoque integral de riesgo constituya un elemento clave para identificar problemas prioritarios, los cuales deben ser la base para la formulación de programas de salud maternoinfantil y la asignación de recursos.
- El posible impacto de la investigación actual y de los análisis consiguientes en los ajustes de política y estrategias en salud maternoinfantil.

En este momento sería útil que el equipo de investigación elaborara un resumen de las políticas y programas relevantes, en el que figure

un análisis sobre dicha relevancia, en relación con los factores de riesgo y la posible reorientación de la planificación de salud maternoinfantil. Debería hacerse notar las políticas o programas en conflicto, o aquellas que sean obviamente inapropiadas o contradictorias. Además, tomar decisiones sobre cuáles son las políticas dominantes y dirigir el énfasis a sus objetivos, y determinar qué programas merecen mayor atención durante investigaciones subsiguientes.

### Ejercicio 1

- (i) De acuerdo con la información de ficticia en sus metas operativas de salud (véase cuadro 22) dé un ejemplo de una política que usted piense que puede ser cambiada como resultado de la investigación basada en el enfoque de riesgo.
- (ii) Dé un ejemplo de una política que usted cree que *no* puede ser cambiada como resultado de la investigación con base en el enfoque de riesgo.
- (iii) Dé dos ejemplos de metas o de programas, o la falta de ellos, capaces de entorpecer la identificación de problemas prioritarios y la nueva programación basada en el concepto de riesgo.
- (iv) Sobre la base de su propia experiencia, describa una política *no escrita* que podría influir en el reconocimiento de los problemas de los grupos humanos postergados y dentro de éstos, del grupo maternoinfantil.

## Sección 4.2 Información requerida para establecer el diagnóstico de salud y las prioridades. Su disponibilidad

### Resumen

La información requerida debe referirse específicamente a los propósitos del estudio para que se obtengan los datos verdaderamente necesarios, y así ayudar a la selección de los problemas prioritarios. La información puede ser cuantitativa y cualita-

tiva; ambos tipos son útiles para decidir prioridades. A pesar que algunos datos pueden ser incompletos, no deben desecharse ya que pueden ser muy útiles si son enriquecidos y se interpretan cuidadosamente.

Se ha preparado una propuesta de estudio que describe en términos generales las características de los problemas de salud y los enfoques de la investigación. Esta propuesta de estudio ha sido aceptada por lo que ahora va a desarrollarse el protocolo del proyecto.

Los grupos involucrados en el estudio deben determinar ahora qué problemas específicos de salud han de considerarse. ¿Cuáles son esos problemas? ¿Qué información se requiere para su reconocimiento?

Dentro de lo posible la información debe ser selectiva y concretarse a los fines del estudio. En este capítulo se describe su recopilación y análisis y su propósito es seleccionar y asignar prioridad a los problemas de salud de madres y niños.

#### **4.2.1 Criterios para la selección de problemas prioritarios siguiendo el enfoque de riesgo**

Los siguientes criterios pueden ser aplicados a cualquier grupo de edad y sexo y, por lo tanto, son pertinentes al grupo materno-infantil.

##### **4.2.1.1 Ubicación espacial de los grupos humanos postergados**

Se considera como grupo humano postergado a un conjunto de individuos con oportunidades y condiciones de vida inferiores a las de otros grupos que también integran una sociedad dada. Se caracteriza habitualmente por la distribución espacial de la pobreza y la insatisfacción de las necesidades básicas.

##### **4.2.1.2 Identificación y definición de las condiciones de vida**

Desde el punto de vista operativo los factores que permiten identificar las condiciones de vida

se refieren al acceso a los servicios públicos, la sanidad ambiental, la vivienda adecuada, la educación, el empleo y el ingreso. La estrecha relación entre estos factores permite la sustitución de unos por otros, de acuerdo con la información disponible. Para los fines del sector salud, es fundamental que los factores seleccionados sean lo suficientemente sensibles para caracterizar las condiciones de vida.

##### **4.2.1.3 Identificación y análisis de los factores de riesgo**

Esta identificación y el examen correspondiente se basan en los análisis de los dos puntos anteriores sobre los grupos humanos postergados y las condiciones de vida, así como en la identificación de los problemas de salud y la accesibilidad a los servicios de salud. En el capítulo correspondiente al "Concepto y enfoque de riesgo en salud materno-infantil" se mencionan tanto los factores de riesgo biológicos, como los socioeconómicos, los ambientales y de accesibilidad a los servicios.

##### **4.2.1.4 Identificación de los problemas de salud pública**

En esta área es importante definir los problemas actuales, los potenciales y los previsibles, así como analizar la inferencia de los factores de riesgo derivados de las condiciones de vida de los distintos grupos humanos. Tomando en cuenta estos criterios, el examen de los problemas para la identificación y establecimiento de prioridades, puede comprender los siguientes elementos:

###### **a) Magnitud del problema**

Las medidas de magnitud de los problemas de salud deben ser expresados en tasas específicas de morbi-mortalidad por edad, sexo y/o causa. Aunque no siempre la información está disponible, generalmente es posible lograr una aproximación bastante exacta a través de los datos reco-

### b) Gravedad del problema

Aunque la gravedad del problema de salud es con frecuencia (y popularmente) la base del establecimiento de prioridades, esta idea es compleja. La gravedad o seriedad de una enfermedad o muerte puede medirse en términos del número de años de expectativa de vida perdida, la pérdida económica para la comunidad, la dependencia impuesta como consecuencia del problema de salud, y el pesar o la duración del dolor experimentado. Para medir muchas de estas variables probablemente no existan datos y la importancia que se atribuye a cada una variará de comunidad en comunidad.

Los cálculos de gravedad deben hacerse después de una discusión con el personal local y con frecuencia la base para la toma de decisiones la constituyen, tanto las medidas de mortalidad y morbilidad como la experiencia del personal de salud.

### c) Posibilidad de prevenir el problema

Se relaciona con los beneficios (reducción del problema) obtenido de cada acción o recurso que utilizemos para su control. Algunos problemas de salud (por ejemplo, el tétanus neonatorum) son simples y fáciles de controlar. Otros como (la espina bífida) son costosos y difíciles, si no imposibles de solucionar.

### d) Importancia para la comunidad

Una orientación sobre cuáles son las principales áreas de interés local puede obtenerse a través de los líderes del pueblo, las autoridades regionales, los planes nacionales de salud, el ministerio de salud, agencias voluntarias, o de otras fuentes.

Sin embargo, lo importante en esta etapa es determinar si las prioridades locales están en conflicto con las prioridades que surgen de los cálculos de la magnitud, gravedad, y prevención del problema de salud.

### e) TENDENCIA EN EL TIEMPO

Los problemas de salud no son estáticos

por lo que debe tomarse en cuenta los cambios que con el tiempo se producen en ellos, así como las transformaciones derivadas de los cambios socioeconómicos y del ambiente. Algunos problemas de salud—como la mortalidad por sarampión—pueden estar declinando rápidamente, mientras que otros pueden ir en aumento o, en el mejor de los casos, mantenerse estáticos durante cierto número de años. Un examen de las tendencias del pasado generalmente mostrará si el problema está aumentando o disminuyendo y, entonces frecuentemente, es posible prever las tendencias futuras. Las tendencias deben tomarse en consideración en el establecimiento de prioridades. A los problemas de salud que parecen estar aumentando debe dársele una mayor prioridad que a aquellos que están estáticos o declinando.

## 4.2.2 Tipos de información

La información puede ser de dos tipos básicos: cualitativa o cuantitativa. Ambos tipos son importantes para nuestro propósito.

### 4.2.2.1 Información cualitativa

La información cualitativa es de gran utilidad como antecedente, así como para decidir la gravedad de un problema de salud.

La información cualitativa también ayuda a tener un mayor entendimiento de los cambios temporales y de los posibles instrumentos de control. Nos permite una mejor percepción de los problemas políticos, humanos, sociales, económicos, y ambientales en el país donde ocurren sus interrelaciones. Esta clase de información nos indica algo sobre el funcionamiento del sistema de atención en salud. También ayuda a identificar las relaciones intersectoriales que pueden utilizarse para sugerir asignación de recursos. Si se toma un problema de salud como ejemplo (tal como la mortalidad infantil) se pueden hacer preguntas y presentar respuestas para construir un modelo.

## Ejemplo

"¿Por qué es esta una preocupación?"  
 "Porque . . ."! Entonces la pregunta a ser formulada es: "¿Cuál es el resultado de esto?"

Para ayudar en una revisión sistemática, las respuestas pueden ser aparecer en un diagrama. Las entradas en un diagrama son luego revisadas y puede estudiarse sus interrelaciones. Este análisis formal o en detalle nos lleva a una revisión sistemática de los problemas locales de salud, sus causas y consecuencias, su seriedad y sus interrelaciones. Este análisis también puede dar ideas en relación con los problemas que necesitan más documentación y también puede proporcionarnos información sobre los posibles factores de riesgo<sup>2</sup> que deberían tomarse en consideración para la definición de problemas prioritarios e incluirse en el estudio epidemiológico (véase Capítulo V).

Mediante el método Delphi, modificado o discutido por consenso (ver 4.4.), el grupo tratará ahora de identificar las relaciones de causa y efecto que existen entre los varios problemas de salud en estudio. La relación entre los problemas puede variar en cuanto a su intensidad. Puede ser útil establecer diferencias entre aquellas variables que tienen asociaciones fuertes y débiles.

Este análisis cualitativo del problema ayuda a enfocar la atención del equipo de investigación en los problemas de salud más importantes,

particularmente en aquéllos que pueden ser modificados. También permite identificar aquellos factores extrasectoriales que son, en general, reconocidos por la comunidad como de alta prioridad.

### 4.2.2.2 Información cuantitativa

Este tipo de información, aunque en general presente, raramente está disponible en la forma requerida. Aún más, si se le encuentra, con frecuencia no será de la calidad deseada. Sin embargo, es posible que sea suficiente para el grado de decisión que se necesita en esta etapa.

Información cuantitativa referente a los siguientes aspectos que en general están disponibles.

- enfermedades y eventos específicos, por ejemplo embarazo y parto, accidentes;
- síntomas específicos, por ejemplo fiebre, diarreas, enfermedades del aparato respiratorio superior;
- eventos específicos, por ejemplo muerte, incapacidad temporal, incapacidad permanente;
- los efectos sociales de las enfermedades y las condiciones específicas—días de trabajo o de escuela perdidos, ganancias o productividad reducida, grado de desarrollo rural;
- enfermedades y condiciones que están asociadas con factores ambientales como agua contaminada, tratamiento inadecuado de aguas negras y desperdicios; vivienda inadecuada; prevalencia de pestes y contaminación ambiental;
- enfermedades y condiciones que contribuyen a dificultades de los servicios de salud, o que son afectadas por éstos, como son: escasez en la capacidad de la clínica para atender pacientes externos, escasez de diversos tipos de personal específico y falta de cobertura del personal técnico específico.

<sup>1</sup> Una alta tasa de mortalidad infantil (TMI) y en niños menores de dos años, puede ser una preocupación local no sólo porque la vida de un niño es preciosa, sino porque mueren más niños que niñas. Un resultado, por lo menos ante los ojos occidentales, puede ser el rechazo a la planificación familiar en este país. Por lo tanto, una reducción en la TMI contribuirá también a reducir las tasas de natalidad.

<sup>2</sup> Los factores de riesgo encontrados en una alta TMI pueden deberse a la escasez y contaminación del agua, bajo peso al nacer, partos muy seguidos, falta de conocimientos sobre rehidratación, poco tiempo de amamantamiento, etc.

Es probable que no exista información exacta disponible a no ser la referente a muertes. Sin embargo, una búsqueda cuidadosa puede revelar información mínima que permita construir los siguientes parámetros:

Casos	El número de casos de determinada enfermedad en un año.
Incidencia	El número de casos nuevos de una enfermedad o evento durante determinado período en una población específica.

Para que la incidencia sea comparable de área a área y de tiempo a tiempo, se debe calcular la tasa de incidencia.

Tasa de incidencia	El número de casos nuevos de una enfermedad durante un determinado período en una población específica dividido por el número de individuos a riesgo de contraer dicha enfermedad en la población estudiada (multiplicado por una constante: 1.000 ó 100.000).
--------------------	--

Prevalencia	El número total de casos de enfermedades o personas enfermas, en una población específica, sin distinción entre casos antiguos y nuevos.
-------------	--

En forma similar, al comparar las prevalencias de área a área o de tiempo a tiempo, se calcula una tasa.

Tasa de prevalencia	El número de casos de enfermedades, o personas enfermas en una población específica dividido por el número de individuos en riesgo de contraer la enfermedad en la población estudiada (multiplicado por 1.000 y 100.000).
---------------------	--

Muertes	El número de muertes ocurridas durante un período dado
---------	--

(generalmente un año) en una población específica.

Las muertes se convierten en:

Tasa de mortalidad	El número de muertes ocurridas en un año dividido por la población total.
Tasas de mortalidad por una causa específica	El número de muertes debidas a una enfermedad específica dividido por la población en riesgo de contraer esa enfermedad.
Tasas de mortalidad por sexo específico	Se calcula en forma análoga.

Muertes debidas a determinada causa:

Tasas de letalidad	El número de muertes debidas a enfermedades específicas divididas por el número total de casos de dicha enfermedad.
Consultas	El número de contactos realizados durante un año por parte del servicio de salud debido a la enfermedad.
Admisiones	El número de personas hospitalizadas durante un año padeciendo de la enfermedad.
Días-cama	El número de admisiones en un año multiplicado por el promedio de días de hospitalización.

Algunas de estas medidas serán necesarias a fin de establecer el orden de prioridades de nuestro problema sanitario.

#### 4.2.2.3 Información disponible

En la vida real, la recopilación de datos no es sencilla. Mucha de la información que se nece-

sita no está disponible. Aquella que sí lo está, con frecuencia no es la que más se necesita.

Sin embargo, antes de establecer un sistema de recopilación de información, debe realizarse una búsqueda exhaustiva en todas las fuentes posibles de información. En muchas o en la mayoría de las situaciones hay más información disponible que la aparente. Los estudios especiales, las agencias voluntarias, las universidades, las instituciones gubernamentales no sanitarias, tales como los ministerios de planificación o del interior, y los sistemas de registros de atención de salud, pueden con frecuencia proporcionar datos útiles. Esta información, interpretada con precaución, debe también considerarse como fuente potencial de datos útiles.

## Ejercicio 2

- (i) En el texto se mencionaron <sup>CINCO</sup> ~~seis~~ criterios generales para la identificación de los principales problemas de salud (magnitud, gravedad, posibilidad de prevención e importancia para la comunidad). Para cada uno de ellos dé un ejemplo de un *indicador* que pueda cuantificarlos.
- (ii) Prepare una lista de datos básicos que usted necesitaría para formular los indicadores mencionados en el ejercicio 2(i), tomando la *mortalidad perinatal* como ejemplo.
- (iii) Revise el documento con la información de Ficticia y compare la información presentada en él con la lista de datos preparada en este ejercicio. Liste la información que usted necesitaría y que no está disponible en Ficticia (o que por lo menos no figura en el documento que se le dio).

## Sección 4.3. Posibles fuentes de información. ¿Están disponibles todos los datos requeridos?

### RESUMEN

Después de decidir que la información será requerida, se procederá a identificar las fuentes posibles de información. Con un poco de imaginación y búsqueda paciente, con frecuencia puede encontrarse una cantidad sorprendente de datos que no son usados.

Esta sección se ocupa de fuentes de datos estadísticos en *general* y de la información faltante en *particular*.

Al identificar las fuentes de datos no sólo se debe recurrir a fuentes conocidas sino que se debe realizar una búsqueda de información a través de contactos personales, cartas a instituciones e individuos y búsquedas en las bibliotecas, si posible, revisando en detalle la literatura publicada. Puede darse el hallazgo de informes de encuestas limitadas con datos poco usuales que han sido publicados, tales como tasas de muertes fetales, mortalidad de niños menores de dos años y peso al nacer. También debe buscarse documentación relacionada con actividades de investigación que se estén llevando a cabo y que pueden proveer literatura actualizada.

Deberían considerarse todas las fuentes de información.

Por ejemplo, las estadísticas gubernamentales y los informes *ad hoc*, reportes (periódicos o *ad hoc*) de instituciones, organizaciones privadas y profesionales, agencias de colaboración internacional (bilaterales y multilaterales); publicaciones científicas, revistas médicas u otras, perfiles de salud de país, planes y programas nacionales de salud, proyectos de estudiantes, etc., así como registros conservados localmente de hospitales, clínicas, agencias públicas y privadas y registros de nacimientos y muertes.

Puede encontrarse información diferente sobre el mismo tema en la misma fuente u otras

fuentes, por lo cual debe darse consideración cuidadosa a la fuente que ha de utilizarse. Si la misma información procede de distintos orígenes, uno podría pensar que es correcta, pero puede ser que todos estén citando la misma información (correcta o incorrectamente). Siempre busque la información original y utilícela luego de evaluarla.

### Ejercicio 3

- i) Sugiera posibles fuentes de donde pueda obtenerse la información faltante que se menciona en el ejercicio 2(iii).
- ii) Si usted llega a la conclusión que alguna de la información requerida no está disponible en estas fuentes, mencione tres alternativas para obtenerla.

## Sección 4.4 Algunos métodos para obtener información complementaria

### RESUMEN

A pesar de todos los esfuerzos, en muchas situaciones alguna información necesaria no estará disponible. Esta sección describe los siguientes métodos que contribuyen a obtener la información faltante: extrapolación, análisis de información anecdótica, cálculos Delphi y encuestas rápidas de evaluación.

#### 4.4.1 Extrapolación

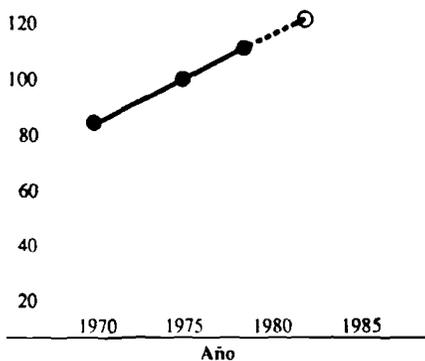
Algunos datos están disponibles solamente durante pocos años. Al continuar el patrón observado, pueden calcularse los valores para los próximos años. La forma más fácil de hacerlo es colocando los datos en una gráfica contra el tiempo. Se dibuja una línea a través de los puntos y se extiende o extrapoliza obteniendo información para los años previos o siguientes. Para cuantificar estas extrapolaciones puede recurrirse a métodos de regresión simple, usando calculadoras de mano. Debe recordarse que

estos cálculos están basados en el principio que los factores encontrados en otros años permanecerán estables.

#### Ejemplo

Usted sabe que en 1970 hubo 85 nacimientos en un pueblo, en 1975, 100 y en 1978, 110. Coloque esta información en una gráfica y dibuje una línea a través de los puntos.

Número de nacimientos



Al extenderse la línea (extrapolación) se puede obtener una estimación aproximada del número de nacimientos esperados para 1982.

#### 4.4.2 Información anecdótica

Los miembros de mayor edad en la comunidad, incluyendo a los profesionales de salud, pueden generalmente hablar con conocimiento sobre los nacimientos de niños durante los últimos 10 años más o menos. Con frecuencia se puede confirmar estos recuerdos a través de otras fuentes. A esta información se le puede llamar anecdótica, esto es, que se ha basado en impresiones más que en hechos o cifras. Tal información puede ser muy útil pero debe ser interpretada con precaución. Esta información anecdótica puede usarse también para evaluar, en líneas generales, la confiabilidad de registros de dudosa calidad.

#### 4.4.3 Inferencias de otra información

Con frecuencia se puede inferir la magnitud de algunas tasas no disponibles basada en información relacionada. Por ejemplo, si se pretende saber la tasa de mortalidad materna, ésta puede inferirse de la tasa de mortalidad entre mujeres de 15-45 años en los últimos años, suponiendo que las causas perinatales y el aborto son responsables de la mayoría de las muertes en ese grupo de edad. Se puede inferir también que la mortalidad infantil es alta si se registra un número desproporcionado de niños entre el total de muertes en los últimos años. Una alta tasa de mortalidad entre niños menores de cinco años generalmente está acompañada de una alta mortalidad posneonatal (las causas son similares).

Finalmente la observación, que es uno de los métodos más útiles para obtener información cuando faltan datos, es raramente descrito y formalizado. Si prácticamente cada niño que usted ve en sus viajes al distrito está obviamente malnutrido, puede concluir que la desnutrición en los niños es un problema importante aunque nunca se haya registrado este hecho en las historias clínicas. Puede ser que se le discuta la exactitud de su observación. Su respuesta a la persona debería ser que proporcione información más exacta sobre el tema.

#### 4.4.4 Estimaciones basadas en el método Delphi

El método Delphi consiste en una agrupación y análisis organizado de las "opiniones informadas" sobre un tema. Se puede preguntar a 10 personas expertas en el tema: "¿Cuál es en su opinión la tasa de mortalidad infantil en Ficticia. Mencione un valor que usted considere que se acerca a la realidad". Todas las respuestas se analizan y se hace una nueva pregunta al mismo grupo de personas basada en las respuestas previas. "La tasa de mortalidad infantil en Ficticia se estima entre 130 y 250 por 1.000. Basándose en esta información, ¿cuál es la cifra que usted cree que se acerca lo más posible al valor real?"

Las respuestas a esta segunda pregunta usualmente convergen más que las de la anterior. El proceso puede repetirse varias veces en caso que sea necesario. Al final puede hacerse una estimación basándose en el consenso del grupo de expertos. El proceso puede realizarse a través de discusión de grupo y generalmente es posible llegar a un acuerdo sobre una buena aproximación.

#### 4.4.5 Encuestas rápidas<sup>1</sup>

En la mayoría de los países la información estadística requerida para identificar los problemas prioritarios en salud materno-infantil no es muy completa o exacta. La llamada encuesta de evaluación rápida puede ayudar a proporcionar información adicional a la obtenida por los enfoques descritos previamente. La información requerida puede ser, por ejemplo, mortalidad materna por edad y paridad o el uso del sistema de atención en salud por parte de mujeres embarazadas. Los datos se necesitan rápidamente. Debe llevarse a cabo una encuesta que busque los hechos en el pasado, ya que un seguimiento de los eventos futuros tomaría mucho tiempo (y sería difícil y costoso).

La encuesta depende de la memoria de los miembros de las familias, comadronas, profesionales de salud, de las mujeres con influencia en el pueblo y de cualquier información escrita disponible. Nosotros probablemente buscaríamos información hasta tres años atrás, ya que la memoria a veces falla. Necesitaríamos establecer un acercamiento amistoso con la población y con los trabajadores sanitarios. La encuesta podría utilizar una lista de verificación muy simple que busque información a través de en-

<sup>1</sup> Para mayor información sobre este tema refiérase a:

*Taking Soundings for Development and Health*, Document FHE/8.2 Rev. 1. WHO/Family Health Division, 1219 Geneva 27, Switzerland.

W. Lutz: *Planning and Organizing a Health Survey*, Geneva, IEA/WHO, 1981.

trévistas. *En primer lugar*, los hechos sobre los eventos que puedan ser confirmados (nacimientos y muertes maternas), *en segundo lugar*, información sobre las mujeres que dieron a luz en el pueblo durante el período correspondiente, su edad (aproximada), su paridad, etc., y en *tercer lugar*, preguntas sobre el uso de la atención en salud que darán información menos exacta.

Además, los registros clínicos pueden revisarse sistemáticamente. Aun cuando los registros no sean completos, la información obtenida a través de entrevistas puede ser confirmada. Un análisis rápido del patrón de asistencia a una clínica puede revelar información útil.

Pero hay que tener cuidado en llegar a conclusiones sobre todas las mujeres del área basándose en el pequeño grupo que asista al centro de salud. Por lo tanto, debe tratarse siempre de interrelacionar la información obtenida de la población general con aquella producida por hospitales y/o centros de salud.

### Ejercicio 4

- (i) Sobre la base de una serie de encuestas y otras investigaciones realizadas, se han hecho las siguientes estimaciones sobre la mortalidad en Ficticia.

Año	Tasa de Mortalidad Infantil (TMI)	Tasa de Mortalidad Posneonatal (TMPN)
1961	132	79
1972	120	68
1975	115	63
1981	103	55

- (ii) Use una gráfica para estimar la tasa de mortalidad infantil y posneonatal para 1985 y 1990 si la tendencia actual continúa.

- (iii) Calcule las tasas de mortalidad neonatal para 1961-1981.

- (iv) Utilizando la información anterior, des-

criba la tendencia actual en mortalidad neonatal, posneonatal e infantil en Ficticia.

- (iv) ¿Cuál es el valor de estas tendencias para identificar más concretamente los problemas de salud en madres y niños en Ficticia? ¿Cuál de estos problemas de salud parecería ser el más importante?
- (v) El Ministerio de Salud de Ficticia ha informado que la tasa de mortalidad infantil es de 45 por 1.000 nacidos vivos en el contexto de información hospitalaria sobre una incidencia de bajo peso al nacer del 10%. ¿Cómo se relacionan estas cifras entre sí y con los datos obtenidos en los ejercicios (i) y (ii)? Discuta las implicaciones de las discrepancias.

Recuerde que en áreas con mortalidad infantil  
 ≈ 45 por 1.000 nacidos vivos:

	Porcentaje
Posneonatal	35
Neonatal	65
Muertes > 2.500 g	25
Muertes < 2.500 g	75

y que cifras "aceptables" de mortalidad neonatal específicas por peso al nacer son:

- > 2500 ≈ 2.5 a 5.0 por 1.000 nacidos vivos
- < 2500 ≈ 90 - 100 por 1.000 nacidos vivos

- (vi) El grupo de investigación en Ficticia ha decidido realizar una encuesta de evaluación rápida para obtener información sobre la edad y paridad de las madres y los niños que murieron antes de llegar a un año. Describa cómo organizaría esa encuesta, cuyo objetivo es obtener información que permita calcular la mortalidad materna específica por edad y paridad, y la mortalidad infantil, neonatal y posneonatal específica por edad y paridad, en una área rural de Ficticia.

## Sección 4.5. Presentación de la información

### Resumen

Los datos deben presentarse en forma clara y correcta a fin de que puedan utilizarse adecuadamente. Esta sección da algunas definiciones y explica procedimientos para la presentación de la información. Particularmente, se pone énfasis en que las tasas, proporciones y tendencias que se utilizan para describir eventos, con frecuencia son mal presentadas y mal interpretadas.

### Introducción<sup>1</sup>

Los datos numéricos que usted ha recopilado deben ser útiles y podrá expresarlos en una forma que le permita hacer comparaciones entre los grupos de población o área y en diferentes épocas; entre enfermedades o eventos; o entre cualesquiera de ellos en otra forma que se pueda usar como medida de magnitud.

Cuando se haya recopilado todos los datos necesarios, ya sea de lo que estaba disponible, o a través de encuestas especiales, extrapolación o cálculos al azar, el equipo de investigación estará listo para presentar la información en una forma que pueda ser interpretada inteligentemente.

#### 4.5.1 Números absolutos

Algunas veces es suficiente describir a las personas y a los eventos en números absolutos. El número de individuos con un problema de salud particular puede ser igual en dos poblaciones diferentes, pero si el número total de indivi-

duos en cada población fuera diferente, las proporciones y, por lo tanto, la importancia del problema sería también diferente. Si estuviéramos interesados en la cantidad de casos atendidos en una clínica, se requiere conocer el número absoluto de personas que asistieron. Pero si se está interesado en la cobertura de los servicios de atención en salud se requiere conocer tanto el número de personas que fueron atendidas por lo menos una vez, como el número total de personas en el área.

#### 4.5.2 Razones, proporciones y tasas

Una forma más útil en que se puede apreciar la información es por medio de los números relativos, es decir un número que exponga la magnitud de un evento en términos de otro. Esto generalmente se expresa en forma de fracción donde el numerador es el total de eventos estudiados y el denominador es la población de individuos a quienes pudo haberle ocurrido el evento. Este tipo de fracción—donde el numerador está incluido en el denominador—es conocido como *proporción*. Una medida a la que por razones prácticas podemos recurrir con frecuencia, es la fracción, donde el numerador *no* se incluye en el denominador. Un ejemplo de esto es la mortalidad materna. Lógicamente, las muertes maternas deberían relacionarse con los embarazos, es decir que el numerador debe contener el número de mujeres que mueren por causas relacionadas con el embarazo, el parto y el puerperio. La población a la cuál puede ocurrir este evento (población a riesgo) es obviamente el total de mujeres embarazadas. (Si no está embarazada no puede morir por causas debidas a la maternidad.) Sin embargo, generalmente es imposible obtener este denominador y tenemos que usar una medida sustituta que, en este caso, es el número de nacidos vivos. Este tipo de fracción no es una proporción, sino una razón, o sea, la relación entre las muertes maternas y los nacimientos.

Muchas, aunque no todas, de las medidas que estaremos usando en el análisis de enfoque de

<sup>1</sup> Para una mayor información sobre este tema refiérase a:

*Descriptive Statistics—Tables and Rates, Ratios, Proposals and Indices—Graphs and Charts*. Publicado por el Departamento de Salud, Educación y Bienestar de los Estados Unidos, Servicios de Salud, Centros para el Control de las Enfermedades (CDC), Atlanta, Georgia, 30033, EUA.

riesgo, representan eventos, tales como la mortalidad infantil, que ocurre en un tiempo dado. Esto significa que nuestra fracción tiene que convertirse en una *tasa*—el número de eventos que ocurren en una población dada en determinado tiempo, multiplicado por una constante (por ejemplo, el número de muertes de niños de uno a cuatro años entre todos los niños de la misma edad en un año). Cuando se reportan tasas hay que definir el tiempo (un año), la población (todos los niños de uno a cuatro años) y el lugar (país, departamento, municipio) y la constante utilizada ( $\times 100$ ; 1.000, etc.).

Una fuente común de error surge cuando el numerador y el denominador no tienen el mismo origen. Esto puede presentarse, por ejemplo, cuando hay casos de una enfermedad que son atendidos en un centro de salud (que también atiende casos de pueblos vecinos) en cuya circunstancia estos casos se deben relacionar únicamente con la población del lugar donde está localizado el centro de salud.

#### 4.5.3 Tasas brutas

Las tasas expresadas en términos de población total se describen como *brutas*. No ofrecen información sobre las diferencias, por ejemplo, en la edad y sexo entre las poblaciones comparadas. Las tasas brutas, sin embargo, resumen las características de la población en una sola cifra. Estas tasas pueden calcularse con un mínimo de información; por ejemplo, el número total de muertes y el número total en la población—sin tomar en cuenta otras características en el grupo. Esto también es una desventaja porque puede dar lugar a equivocaciones. Ello significa, sin embargo, que dos poblaciones con la misma mortalidad, pero con diferente estructura de edad, tendrán tasas brutas de mortalidad diferentes. Por lo tanto, las tasas brutas no deben usarse para comparar poblaciones que tienen estructuras distintas (por ejemplo, edad, paridad, raza, sexo).

#### 4.5.4 Tasas ajustadas o estandarizadas

Para comparar poblaciones de distinta estructura, pueden utilizarse diferentes procedimientos. Uno de los más utilizados consiste en preparar tasas ajustadas o estandarizadas. Este procedimiento permite comparar poblaciones de distinta estructura (por ejemplo, paridad) sin la distorsión que esto introduce en las tasas comparadas (por ejemplo, mortalidad perinatal).

Generalmente se emplean dos tipos de ajuste: directo e indirecto. Supóngase que deseamos comparar la población A con la B en términos de la mortalidad perinatal. Sabemos que la distribución de edad y paridad en las dos poblaciones son distintas. También se sabe que la edad y la paridad materna están asociadas con la mortalidad perinatal.

Si se conoce las tasas específicas de mortalidad perinatal por edad y paridad de ambas poblaciones, se puede utilizar el método directo de ajuste. Este método supone que las dos áreas tienen la misma distribución de edad y paridad que una población tomada como de referencia. Las tasas de mortalidad perinatal son luego calculadas para cada una de las poblaciones estandarizadas usando la distribución de la población de referencia. Cuando no se conocen las tasas específicas por edad y paridad de cada población (cosa que sucede frecuentemente), se puede usar el método indirecto de ajuste. Este método solamente utiliza la distribución de edad y paridad de las dos poblaciones a comparar y la tasa de mortalidad perinatal específica por edad y paridad de una población de referencia. El objeto es calcular el número de muertes perinatales *esperado* en cada población si ellas tuvieran las distribuciones de población de referencia. Un paso final compara este número *esperado* con el *observado* en cada población.<sup>1</sup>

<sup>1</sup> Para una descripción más detallada de estos métodos véase:

Kahn, H. *An Introduction to Epidemiologic Methods*. Oxford University Press. Capítulo 5, p. 63, 1983.

#### 4.5.5 Tasas específicas

Cuando el problema que estamos investigando sólo comprende un sector específico de la población, es más útil tratar con las tasas específicas por edad, sexo o causa específica de morbilidad y mortalidad. En nuestro ejemplo, la tasa de mortalidad calculada como el número de muertes de niños de uno a cuatro años, dividido por el total de niños de uno a cuatro años a la mitad del año, nos da una tasa de mortalidad específica para ese grupo de edad. Esta tasa se denomina *tasa de mortalidad de 1-4 años*. La mortalidad de mujeres de 15 a 49 años es una tasa específica de sexo y edad. El número de muertes por tuberculosis por cada 1.000 personas es una tasa específica de mortalidad por causa. Finalmente, la tasa de mortalidad perinatal por peso al nacer es una de las tasas específicas por causa, de uso más definido en la programación de salud maternoinfantil.

#### 4.5.6 Incidencia y prevalencia

Mientras que las tasas de mortalidad son útiles para señalar las prioridades de los problemas de salud midiendo las pérdidas de vida, no nos dicen nada sobre el estado de salud de aquellos que no mueren. Las tasas de incidencia y prevalencia son medidas que se utilizan con frecuencia para cuantificar la importancia relativa de las enfermedades, tales como diarrea, anemia o desnutrición. La tasa de incidencia se define como el número de casos de enfermedades que comienzan (o de personas que se enferman) durante un determinado período en una población específica, dividido por el número de individuos en esa población que estuvieron a riesgo de contraer la enfermedad durante ese período estudiado.

Para un ejemplo del método indirecto de ajuste puede verse:

Villar, J.; Belizan, J.M. y Fischer, P. *Epidemiologic Observations of the Relationship between Calcium intake and Eclampsia. Int. J. Gynecol. Obstet.*, 1983, 21:271-278.

En lugar de enfermedades podemos usar la incidencia para otros eventos que deseamos medir como, por ejemplo, accidentes infantiles.

La incidencia nos da una idea de la probabilidad que tienen los individuos de esa población de contraer una enfermedad específica durante un período dado. En otras palabras, cuantifica la probabilidad de pasar del estado de salud a de enfermedad o muerte.

#### Ejemplo

Cincuenta y ocho niños de uno a cuatro años tuvieron sarampión en 1981. El número total de niños de uno a cuatro años en la población de la cual son originarios estos niños, es de 700. La tasa de incidencia de sarampión es entonces:

$$\frac{58}{700} \times 1.000 = 83 \text{ por mil}^1 \text{ niños de 1 a 4 años en 1981.}$$

En comparación con la incidencia que mide la frecuencia de un suceso (nuevo) a través del tiempo, la prevalencia es una visión de la situación de diagnóstico prevaleciente en sola instancia en el tiempo. ¿Cuántas mujeres del pueblo están anémicas ahora? La tasa de prevalencia se define como el número de casos de la enfermedad, o personas enfermas (anteriores y nuevas) en una población específica, dividido por el número de individuos en esa población en ese momento dado.

#### Ejemplo

Si 75 de las 640 mujeres embarazadas que asisten a la clínica prenatal están anémicas, la prevalencia de anemia en el embarazo en esa clínica es:

$$\frac{75}{640} \times 100 = 12\% \text{ (11,71\%)}^1$$

<sup>1</sup> Las tasas pueden expresarse en porcentajes.

Es un error común utilizar el término incidencia como una expresión general de frecuencia, o presentar una cifra que de hecho es prevalencia. Esto es importante porque la incidencia describe el riesgo de contraer una enfermedad en una población mientras que la prevalencia describe el riesgo de padecer dicha enfermedad en este momento.

### 4.5.7 Tendencias

Una de las preguntas cruciales a formularse en la selección de problemas prioritarios es: "¿Qué está sucediendo a través del tiempo? El problema analizado, ¿está incrementándose, es constante, o tiende a disminuir?"

A fin de tener alguna medida de los cambios que ocurren a través del tiempo, necesitamos tener nuestros datos en cierto orden para que puedan realizarse comparaciones en determinados momentos,

#### Ejemplo

El número de casos de sarampión en niños de uno a cuatro años era de 58 en 1981; con una población de 700 la tasa de incidencia en esos niños era calculada en 83 por mil. En 1978 había 40 casos entre 500 niños (incidencia de 80 por mil) y en 1975 había 30 casos entre 400 niños (incidencia de 75 por mil).

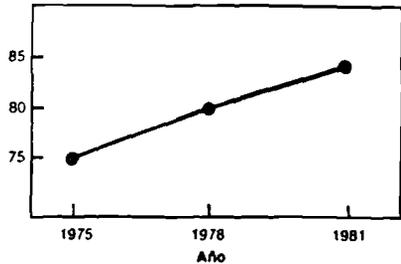
Puede observarse que la tasa de incidencia va en aumento.

1975 - 75 por mil  
 1978 - 80 por mil  
 1981 - 83 por mil

Las tendencias en diferentes épocas pueden ser elaboradas calculando tasas o algunas veces aun comparando los datos brutos en diferentes momentos. Hay varias maneras de lograr que las tendencias a través del tiempo surjan de manera más clara.

A. La forma más simple es haciendo un diagrama de los datos, ya sean estos datos brutos o tasas, o por medio de una gráfica para hacer una evaluación visual de las tendencias.

Incidencia de Sarampión



B. Una forma alternativa sería expresar todas las cifras como porcentajes del primer año. Tales porcentajes son conocidos como *números indicadores*.

Este método permite formarse un juicio sobre la rapidez con que aumenta (o disminuye) el problema. En nuestro caso puede notarse que la incidencia de sarampión ha tenido un aumento del 11% en seis años.

En esta sección únicamente se ha descrito la forma en que los datos brutos pueden convertirse en información útil. La decisión sobre *qué* datos usar, y *cómo* presentarlos, es crucial para

#### Ejemplo:

Año	Incidencia de sarampión		Número indicador
1975	75	$75 / 75 \times 100 =$	100
1978	80	$80 / 75 \times 100 =$	107
1981	83	$83 / 75 \times 100 =$	111

el desarrollo del estudio. Generalmente se tiende a pedir *mucha información* cuando en realidad debe existir un objetivo para cada dato que se busca. Antes de recopilar y procesar los datos es útil escribir el uso que se piensa dar a la información recogida. Preparar los esqueletos de los gráficos y cuadros que se pretenden construir es un método de gran utilidad.

**Ejercicio 5**

- (i) El equipo de Ficticia decidió obtener más información sobre la mortalidad en niños menores de un año, a través de una encuesta rápida de evaluación. La siguiente información pudo obtenerse a través de los ayudantes de salud del pueblo:

Durante los dos últimos años, los trabajadores sanitarios del pueblo habían atendido a 775 mujeres. De éstas, 730 dieron a luz hijos vivos. Las otras tuvieron muerte fetal. De los nacidos vivos, el 20% de los niños eran pequeños. Aunque no habían sido pesados, se piensa que su peso al nacer era de 2.500 gm o menos. De estos niños pequeños 24 murieron antes de cumplir el primer mes. De los niños grandes, 53 murieron antes de cumplir un año y 28

de ellos antes de cumplir un mes. En total, murieron 95 niños antes de llegar a su primer cumpleaños. Dos mujeres murieron durante el parto. La causa de muerte de los niños que murieron en el primer mes, el período neonatal, se registró en las tarjetas de la siguiente forma:

Causa de muerte	Número de muertes de Niños	
	pequeños	grandes
Tétanos	3	9
Otras infecciones	4	8
Ictericia	2	0
Niño anormal	1	1
Lesiones al nacer	1	7
Prematuro	10	1
Otras	0	1
Desconocidas	3	1
Todas las causas	24	28

Presente esta información en forma de cuadro.

- (ii) Use la información anterior para calcular los indicadores de salud que se mencionan en el cuadro que sigue. (La tasa de mortalidad fetal se da como ejemplo.)

Problema de salud	Tipo de medida	Cifras a emplearse
Mortalidad fetal	Tasa <del>Razón</del>	$\frac{775 - 730}{775} \times 1.000 = 60 \times 1.000$ nacidos vivos
Bajo peso al nacer ( $\leq 2.500$ g)		
Mortalidad neonatal		
Mortalidad posneonatal		
Mortalidad infantil		
Mortalidad por tétanos		
Mortalidad materna	Tasa Razón	
Mortalidad neonatal por otras infecciones que no sea tétanos en todos los recién nacidos		
Mortalidad neonatal de $\geq 2.500$ g		

- (iii) ¿Cuáles son las diferencias entre razones y tasas?
- (iv) Calcule las tasas de mortalidad neonatal y posneonatal específicas por peso al nacer ( $2.500 \text{ g} \leq 2.500 \text{ g}$ ) ¿Cómo puede interpretarse estos resultados? ¿Cómo esta información ayuda a establecer áreas de atención perinatal prioritarias?

Nota: Las cifras esperadas de mortalidad neonatal y posneonatal para  $\geq 2.500 \text{ g}$  para países desarrollados son

Mortalidad neonatal	Posneonatal
2.5 a 5.0/1.000	1.5 a 3.0/1.000

## Sección 4.6. El establecimiento de prioridades en los problemas de salud materno-infantil

### Resumen

Frecuentemente cuando se identifican los problemas prioritarios para definir las acciones, se introducen sesgos o prejuicios personales.

Los prejuicios o la subjetividad del personal pueden reducirse aplicando una serie de criterios en la selección de prioridades. Aquí se describen algunas formas para mejorar este procedimiento. Los procedimientos descritos deben utilizarse con precaución y sentido común, ya que sólo constituyen herramientas pero no proveen respuestas infalibles.

### Introducción

El grupo debe ahora decidir cuál es el problema de salud más importante para atacarlo utilizando la estrategia de riesgo. Al decidir las prioridades relativas tenemos que ser prácticos, empleando la información disponible. Existe cierto número de principios que pueden orientarnos:

- Como es obvio, los problemas más comu-

nes deberían tener mayor prioridad que los más raros, haciendo a un lado otras consideraciones (frecuencia).

- Los problemas de salud más serios deberían tener mayor prioridad que los menos serios (severidad).
- Los problemas de salud que pueden prevenirse con mayor facilidad deberían tener una prioridad más alta que aquellos cuya prevención es más difícil (posibilidad de prevención).
- Aquellos con tendencias crecientes a través del tiempo deben tener mayor prioridad que los estáticos, y éstos a su vez mayor prioridad que los que están disminuyendo (tendencia temporal).
- Debemos respetar y considerar los intereses, prioridades y preferencias de la comunidad (interés de la comunidad).

El material que se presenta a continuación representa un esquema para facilitar el listado técnico de prioridades. Sin embargo como se expresó anteriormente, la decisión final de la selección de prioridades es política.

Se ha propuesto métodos cuantitativos que comprenden desde los extremadamente simples, como una simple lista de los problemas y los criterios, hasta aquellos más complejos que permiten calcular la cantidad de "días de vida sana" que pueden ser atribuidos a distintos problemas de salud.<sup>1</sup> Debe recordarse, sin embargo, que estos métodos no son completamente objetivos y dependen de la selección subjetiva de los criterios que se van a usar, los puntajes y los pesos relativos que se le den a cada uno.

<sup>1</sup> Ghana Health Assessment Project Team: A Quantitative Method of Assessing the Health Impact of Different Diseases in less Developed Countries. Int. J. of Epidemiology. 10,17, 1981.

### 4.6.1 Un ejemplo de procedimientos cuantitativos para establecer prioridades

Este procedimiento otorga un peso relativo a cada criterio así como a la relación de éste con un problema o enfermedad específico.

El cuadro 2 presenta tres problemas específi-

cos y su puntuación relativa basada en los cinco criterios utilizados. Además, estos cinco criterios han sido clasificados entre sí, otorgándoles un puntaje relativo. Por ejemplo, puede considerarse que la frecuencia y la seriedad son más importantes que la tendencia de la enfermedad.

La matriz que se presenta muestra este procedimiento y los cálculos requeridos.

Cuadro 2

Problema	Criterio	Puntaje máximo posible	Puntaje asignado
Sarampión	Magnitud	10	5
	Gravedad	10	5
	Posibilidad de prevención	10	10
	Interés local	10	3
	Tendencias en el tiempo	10	2
Diarrea	Magnitud	10	4
	Gravedad	10	9
	Posibilidad de prevención	10	5
	Interés local	10	6
Accidentes de tránsito que involucran niños	Tendencias en el tiempo	10	4
	Magnitud	10	2
	Gravedad	10	9
	Posibilidad de prevención	10	4
	Interés local	10	7
	Tendencias en el tiempo	10	7

Criterio para la ubicación de prioridades

	Magnitud	Gravedad	Prevenible	Interés local	Tendencia	Puntaje total
Valores	4	4	2	3	1	
Problemas de salud						
Sarampión	$5 \times 4 = 20$	$5 \times 4 = 20$	$10 \times 2 = 20$	$3 \times 3 = 9$	$2 \times 1 = 2$	71
Diarrea	$4 \times 4 = 16$	$9 \times 4 = 36$	$5 \times 2 = 10$	$6 \times 3 = 18$	$4 \times 1 = 4$	84
Accidentes de tránsito que involucran niños	$2 \times 4 = 8$	$9 \times 4 = 36$	$4 \times 2 = 8$	$7 \times 3 = 21$	$7 \times 1 = 7$	80

---

**Pero no debe olvidarse que:**

Cualquier matriz es óptima según los números que contenga. El sentido común tiene que prevalecer. Si los resultados del análisis de su matriz van en contra del sentido común debería revisar tanto la matriz como su método y sus datos.

---

**Ejercicio 6**

- (i) Basándose en su experiencia y en la información provista para Ficticia haga una lista de los problemas prioritarios de salud para el país.
- (ii) Utilice el método Delphi para establecer una lista corta de prioridades.
- (iii) Utilice el procedimiento presentado para establecer una lista complementaria de las prioridades.
- (iv) Utilice *toda* la información para establecer las prioridades que desearía resolver o controlar con la nueva estrategia de atención maternoinfantil.

**Referencias**

1. Bainbridge, J. y Sapire, S. *Health Project Management—A Manual of Procedures for formulating and implementing health projects*. Offset Publication No. 12. WHO-Geneva, 1974, pág. 39-40.
2. *Descriptive Statistics, Tables and Rates, Ratios Proposals and Indices—Graphs and Charts*. Publicado por el Departamento de Salud, Educación y Bienestar de los Estados Unidos, Servicios de Salud, Centros para el Control de las Enfermedades (CDC), Atlanta, Georgia 30033, EUA.
3. Ghana Health Assessment Project team: A Quantitative Method of Assessing the Health Impact of Different Diseases in less Developed

---

Countries. *Int. J. of Epidemiology* 10, 17, 1981.

4. Khan, H. *An Introduction to Epidemiologic Methods*. Oxford University Press, Capitulo 5, pág. 63, 1983.

5. *Setting Priorities for Health Development* (Borrador). WHO/HMD Documents, Geneva, 1980.

6. *Taking Soundings for Development and Health*, Document FHE/8.2/Rev. 1, WHO Family Health Division, 1219 Geneva 27, Switzerland.

7. Villar, J., Belizan, J. M. y Fisher, P. Epidemiologic observations of the Relationship between calcium intake and Eclampsia. *Int. J. Gynecol. Obstet.*, 1983, 21: 271-278.

8. W. Lutz: *Planning and Organizing a Health Survey*, Geneva, IEA/WHO. 1981.

# **Manual sobre el enfoque de riesgo en la atención maternoinfantil**

Serie PALTEX para ejecutores de programas de salud No. 7

**ORGANIZACION PANAMERICANA DE LA SALUD  
Oficina Sanitaria Panamericana, Oficina Regional de la  
ORGANIZACION MUNDIAL DE LA SALUD**

1986

## Capítulo V

### Los instrumentos epidemiológicos empleados en el enfoque de riesgo

Sección 5.1	Definición de los objetivos del estudio	Sección 5.7	Interpretación de los datos
Sección 5.2	Selección de diseño del estudio	Sección 5.8	La cuantificación del riesgo en individuos y la elaboración de estrategias de intervención
Sección 5.3	Selección del método de muestreo	Sección 5.9	La cuantificación del riesgo en la comunidad
Sección 5.4	Selección del tamaño de la muestra	Sección 5.10	Presentación de datos y preparación de informes
Sección 5.5	Preparativos para el análisis de los datos	Apéndice:	Modelo de una propuesta de investigación
Sección 5.6	Conceptos generales sobre manejo de los datos		

#### Objetivos de aprendizaje

#### Objetivos específicos

##### Objetivo general

El objetivo general es planificar, ejecutar e interpretar,—en cooperación con otros miembros de un equipo de investigación, incluyendo un estadístico y/o un epidemiólogo,—un estudio epidemiológico en el que se establezcan relaciones entre daños a la salud (enfermedad y/o muerte) y factores de riesgo.

Al finalizar este capítulo, los participantes deberán ser capaces de:

1. Formular correctamente los objetivos e hipótesis de un estudio epidemiológico de riesgo.
2. Seleccionar un modelo de diseño de estudio que sea apropiado para demostrar las hipótesis que se formulen de acuerdo con el objetivo anterior.

3. Definir la población en estudio y participar en la selección del método de muestreo adecuado.
4. Determinar el tamaño aproximado de la muestra requerido en un estudio epidemiológico de riesgo.
5. Preparar un listado de los datos que se compilarán y un modelo que describa las tabulaciones cruzadas que se ha de obtener en el análisis.
6. Participar en el diseño, prueba e instrumentación de sistemas de recolección de datos y en la formulación de planes para su codificación y procesamiento.
7. Determinar la posible naturaleza causal de uno o más factores de riesgo y seleccionar entre ellos los que pueden ser modificados.
8. Describir las relaciones entre factores de riesgo y daños a la salud desde el punto de vista del riesgo relativo.
9. Preparar perfiles de riesgo que describan los grupos de personas de alto riesgo así como la distribución del perfil en la comunidad, para determinar los límites de inclusión de los puntajes de riesgo.
10. Determinar el riesgo atribuible en la población de los factores de riesgo presuntamente causales, con el fin de calcular el probable efecto de la intervención proyectada.
11. Presentar los resultados de los estudios epidemiológicos en forma tal que, junto con la información sobre el sistema de salud, puedan ser utilizados para diseñar estrategias de intervención.

los objetivos de la investigación permitirá seleccionar el diseño del estudio más adecuado para probar la(s) hipótesis de trabajo.

### 5.1.1 Definición del área de investigación.

En esta etapa el grupo encargado del estudio habrá identificado los principales problemas de salud que existen en el país o la región. También tendrá un buen conocimiento sobre las políticas y programas nacionales de salud materno-infantil. La siguiente acción es la preparación del protocolo para el estudio. La primera tarea consiste en especificar las áreas de investigación, incluir la formulación de hipótesis y fijar los objetivos de la investigación.

El tema central de la investigación es el problema prioritario de salud identificado. De ahí que sea importante definir el tópico de la investigación desde el punto de vista de la diferencia entre la situación actual y la situación que debería o podría existir; éste último aspecto indica lo que se considera que podría lograrse al finalizar la intervención. El término daño a la salud se usa para señalar un hecho o problema de salud que puede ser medido. Los daños a la salud seleccionados habitualmente como puntos terminales en los estudios epidemiológicos, son la enfermedad, la incapacidad o la muerte.

Un daño es mensurable en categorías o escalas generales que incluyen distintos niveles de salud, tales como resolución completa de un problema de salud, mejoría, ausencia de cambios, deterioro o fracaso. Entre las categorías de interés figuran por ejemplo, el peso al nacer, la ocurrencia de gastroenteritis o la muerte perinatal. Como ejemplo de categorías más específicas puede citarse la anemia o la hipertensión. En la sección 5.5 se examina con más detalle los diversos tipos de daños.

## Sección 5.1 Definición de los objetivos del estudio

### Resumen

El tema central de la investigación es el problema de salud que se va a investigar. Las hipótesis de investigación girarán en torno a las relaciones que existen entre los factores de riesgo y los daños a la salud (enfermedad y/o muerte). La definición de

### Ejemplo

Planteamiento de problemas de salud usando un acontecimiento final como daño. En *Ficticia* la tasa de mortalidad

perinatal en niños que pesan más de 2.500 gramos al nacer es de 25 por mil nacidos y bien podría ser de entre 5 y 10 por mil nacidos. En este ejemplo el daño a la salud a medir es la mortalidad perinatal.

Después de este planteamiento es preciso examinar los problemas de salud identificados y, al menos parcialmente, cuantificados. La selección definitiva de los problemas a ser investigados se basará en:

- \* trabajos previamente publicados que incluyan datos nacionales y locales (ver capítulo IV);
- \* nuevas ideas de miembros del equipo de trabajo o de otros investigadores;
- \* participación de la comunidad;
- \* recursos disponibles.

Esta etapa de formulación de preguntas, hipótesis y objetivos específicos, requiere un abordaje sistemático.

### 5.1.2 Formulación de la hipótesis de investigación

Una vez identificados los temas de la investigación (por ejemplo, mortalidad perinatal, o mortalidad materna) es preciso formular las hipótesis de investigación. Una hipótesis de investigación se define como una suposición o conjetura que se plantea para explicar ciertos hechos o eventos y se emplea como base para llevar a cabo una investigación mediante la cual se busca demostrarla o refutarla. Debe incluir, en términos muy específicos, lo que se desea averiguar (por ejemplo, cuantificar la asociación existente entre el consumo de energía de la madre y el crecimiento del feto). Deberá evitarse el empleo de planteamientos generales y de una redacción ambigua (ej. "queremos encontrar la manera de mejorar la salud de madres y lactantes"). En investigaciones sobre enfoque de riesgo, interesa particularmente la relación existente entre los factores de riesgo y los daños a la salud. Debe indicarse claramente la manera

en que se va a cuantificar y medir dichos daños (por ejemplo, tasas de mortalidad específicas por edad, prevalencia o incidencia de una enfermedad específica), así como precisar los factores de riesgo que se sospecha guardan relación con tales daños a la salud y la forma en que serán definidos y cuantificados (prevalencia, distribución en subgrupos de la población, etc.).

Un ejemplo de planteamiento de una hipótesis general de investigación con el enfoque de riesgo podría ser el siguiente:

*"Es posible identificar—mediante el uso de indicadores de riesgo—grupos bien definidos de mujeres embarazadas cuyos hijos tendrán una morbilidad y una mortalidad perinatal más altas que las del resto de las mujeres embarazadas".*

A partir de esto debe formularse las hipótesis específicas del estudio. El número de hipótesis de investigación debe ser limitado, ya que es poco probable que un solo estudio permita resolver un gran número de preguntas. Además, es preciso tener en consideración los recursos que están disponibles.

La mayoría de las hipótesis de investigación se redacta en la forma negativa, llamada la hipótesis nula. Aunque esto puede parecer una elaboración compleja, las pruebas estadísticas utilizadas para confirmar los resultados de la investigación parten de la base de que la hipótesis está planteada en estos términos. Tales pruebas miden la probabilidad de que los resultados obtenidos en un estudio hayan ocurrido por azar. Si esto es extremadamente improbable, podemos entonces rechazar la hipótesis nula y aceptar la alternativa, o sea, la hipótesis propuesta en la investigación. Así, la formulación de una hipótesis nula correspondiente a la planteada previamente sería así:

*"No es posible identificar—mediante el uso de indicadores de riesgo—grupos definidos de mujeres embarazadas cuyos hijos tendrán una morbilidad y una mortalidad perinatal más altas que las del resto de las mujeres embarazadas".*

Las hipótesis de investigación también deben ser específicas. Un ejemplo de una hipótesis específica es:

*"Los hijos de las mujeres embarazadas que hayan tenido cinco o más embarazos previos, tendrán una morbilidad y una mortalidad perinatal más altas que las de las mujeres con menos embarazos anteriores".*

La formulación de hipótesis específicas es útil por varios motivos:

- A) Facilita la selección del diseño del estudio, del tipo de metodología, de las variables que se incluirán y de la manera en que deberán analizarse los datos.
- B) Ayuda a organizar y analizar los datos.
- C) Ayuda a decidir si las preguntas de la investigación son apropiadas y si han sido expresadas con claridad.
- D) Estimula el desarrollo de hipótesis conexas.
- E) Evita confusión, pérdida de tiempo y la recolección innecesaria de datos.

### 5.1.3 Determinación de los objetivos específicos de la investigación

Una vez que se haya seleccionado el tema de la investigación y se haya formulado las hipótesis, el siguiente paso consiste en enumerar los objetivos específicos de la investigación.

Se entiende aquí por objetivos específicos, aquellas tareas o actividades necesarias para la obtención de la información a usarse para probar las hipótesis.

La lista de objetivos específicos debe ser exhaustiva y detallada abarcando todos los que estén implícitos en el diseño del estudio. Cualquier omisión puede llevar a la obtención de datos engañosos o incompletos y al fracaso de una parte importante del estudio.

#### Ejercicio 1

Un grupo de investigadores en Ficticia ha decidido que el estudio se hará sobre el problema de la mortalidad perinatal. El problema

ha sido definido en términos generales y el protocolo del estudio será formulado con más detalle: es preciso definir las hipótesis generales y específicas del estudio, así como las correspondientes hipótesis nulas de la investigación.

Mediante la información que tiene a su disposición, programe en la forma siguiente:

- (i) Formule una hipótesis general de riesgo sobre la que pueda basarse todo el estudio epidemiológico.
- (ii) Enumere cinco hipótesis específicas de investigación que desee demostrar, a fin de usar el enfoque de riesgo en la solución del problema de la mortalidad perinatal.
- (iii) Formule la hipótesis nula correspondiente a cada una de las hipótesis de investigación previamente mencionadas.
- (iv) Indique los objetivos específicos de la investigación que se derivan de cada hipótesis.
- (v) Enumere la información que desea recolectar y compilar a fin de alcanzar estos objetivos y probar las hipótesis.

## Sección 5.2 Selección del diseño del estudio

### Resumen

Para estudiar un problema de salud puede emplearse varios tipos de diseños de investigación. Se describe aquí diseños en los que se tiene control sobre la exposición al factor de riesgo, (experimental), en los que no se tiene (de observación), en los que se compila los datos retrospectivamente, en un corte transversal o en forma prospectiva. Se revisa también la manera en que se debería analizar los datos. La selección del diseño se verá afectada por la hipótesis que se va a demostrar; la información necesaria para probar la hipótesis y para alcanzar los objetivos específicos de la investigación; los recursos disponibles (tiempo, dinero, personal, tecnología) y la facilidad con que se puede establecer contacto con la población en estudio.

### 5.2.1 La secuencia del razonamiento epidemiológico y la selección del diseño del estudio

La selección del diseño de investigación depende, en primer término, del nivel del razonamiento epidemiológico que se pueda alcanzar. En la figura 1 se presentan los distintos niveles del razonamiento y se establece cuál es la relación adecuada entre el diseño de estudio y el nivel correspondiente. Frecuentemente el investigador se encontrará en situaciones en que no tendrá ningún tipo de control en cuanto a la exposición de la población a ciertos factores de riesgo. Por consecuencia sólo podrá observar el evento en la población, mientras logra investigar la exposición. Los dos tipos de investigación epidemiológica de uso más común son los estudios de observación (descriptivos o analíticos) y los estudios experimentales (ensayos aleatorios controlados). En este último tipo es posible controlar la exposición al factor de riesgo y evaluar la efectividad de una intervención, bien sea profiláctica o terapéutica. La característica distintiva de los *ensayos controlados* es que el investigador recurre al azar para la selección de los grupos a investigar. Sin embargo, debido a que en los estudios con seres humanos existen diversas limitaciones de naturaleza ética que impiden la aplicación de éste método, el investigador frecuentemente utiliza los estudios de observación. En las secciones siguientes se examinará brevemente cada una de estas opciones de diseños de investigación.

### 5.2.2 Estudios de observación

La característica principal de los estudios de observación es que el investigador no tiene control sobre las condiciones de exposición de la población al factor de riesgo. Los estudios de observación pueden ser clasificados en descriptivos y analíticos (figura 1). Los descriptivos estudian la distribución de la enfermedad (tasa) de acuerdo con las características de la población (edad, sexo, raza etc.). Este tipo de estu-

dios es útil para examinar la magnitud de un problema en sectores de la población o para generar hipótesis que posteriormente puedan ser exploradas en estudios analíticos, cuyo objeto es probar hipótesis específicas de asociación entre factores de riesgo y daños a la salud.

Los estudios de observación analíticos pueden ser clasificados de acuerdo con su secuencia temporal en: a) de corte transversal b) retrospectivos o de casos y controles y c) prospectivos o de cohorte. A continuación se describen estos tres tipos de estudios de observación analíticos.

Las hipótesis de investigación también deben ser específicas. Un ejemplo de una hipótesis específica es:

*"Los hijos de las mujeres embarazadas que hayan tenido cinco o más embarazos previos, tendrán una morbilidad y una mortalidad perinatal más altas que las de las mujeres con menos embarazos anteriores"*.

La formulación de hipótesis específicas es útil por varios motivos:

- A) Facilita la selección del diseño del estudio, del tipo de metodología, de las variables que se incluirán y de la manera en que deberán analizarse los datos.
- B) Ayuda a organizar y analizar los datos.
- C) Ayuda a decidir si las preguntas de la investigación son apropiadas y si han sido expresadas con claridad.
- D) Estimula el desarrollo de hipótesis conexas.
- E) Evita confusión, pérdida de tiempo y la recolección innecesaria de datos.

### 5.1.3 Determinación de los objetivos específicos de la investigación

Una vez que se haya seleccionado el tema de la investigación y se haya formulado las hipótesis, el siguiente paso consiste en enumerar los objetivos específicos de la investigación.

Se entiende aquí por objetivos específicos, aquellas tareas o actividades necesarias para la obtención de la información a usarse para probar las hipótesis.

La lista de objetivos específicos debe ser exhaustiva y detallada abarcando todos los que estén implícitos en el diseño del estudio. Cualquier omisión puede llevar a la obtención de datos engañosos o incompletos y al fracaso de una parte importante del estudio.

#### Ejercicio 1

Un grupo de investigadores en Ficticia ha decidido que el estudio se hará sobre el problema de la mortalidad perinatal. El problema

ha sido definido en términos generales y el protocolo del estudio será formulado con más detalle: es preciso definir las hipótesis generales y específicas del estudio, así como las correspondientes hipótesis nulas de la investigación.

Mediante la información que tiene a su disposición, prosiga en la forma siguiente:

- (i) Formule una hipótesis general de riesgo sobre la que pueda basarse todo el estudio epidemiológico.
- (ii) Enumere cinco hipótesis específicas de investigación que desee demostrar, a fin de usar el enfoque de riesgo en la solución del problema de la mortalidad perinatal.
- (iii) Formule la hipótesis nula correspondiente a cada una de las hipótesis de investigación previamente mencionadas.
- (iv) Indique los objetivos específicos de la investigación que se derivan de cada hipótesis.
- (v) Enumere la información que desea recolectar y compilar a fin de alcanzar estos objetivos y probar las hipótesis.

## Sección 5.2 Selección del diseño del estudio

### Resumen

Para estudiar un problema de salud puede emplearse varios tipos de diseños de investigación. Se describe aquí diseños en los que se tiene control sobre la exposición al factor de riesgo, (experimental), en los que no se tiene (de observación), en los que compila los datos retrospectivamente, en un corte transversal o en forma prospectiva. Se revisa también la manera en que se debería analizar los datos. La selección del diseño se verá afectada por la hipótesis que se va a demostrar; la información necesaria para probar la hipótesis y para alcanzar los objetivos específicos de la investigación; los recursos disponibles (tiempo, dinero, personal, tecnología) y la facilidad con que se puede establecer contacto con la población en estudio.

### 5.2.1 La secuencia del razonamiento epidemiológico y la selección del diseño del estudio

La selección del diseño de investigación depende, en primer término, del nivel del razonamiento epidemiológico que se pueda alcanzar. En la figura 1 se presentan los distintos niveles del razonamiento y se establece cuál es la relación adecuada entre el diseño de estudio y el nivel correspondiente. Frecuentemente el investigador se encontrará en situaciones en que no tendrá ningún tipo de control en cuanto a la exposición de la población a ciertos factores de riesgo. Por consecuencia sólo podrá observar el evento en la población, mientras logra investigar la exposición. Los dos tipos de investigación epidemiológica de uso más común son los estudios de observación (descriptivos o analíticos) y los estudios experimentales (ensayos aleatorios controlados). En este último tipo es posible controlar la exposición al factor de riesgo y evaluar la efectividad de una intervención, bien sea profiláctica o terapéutica. La característica distintiva de los *ensayos controlados* es que el investigador recurre al azar para la selección de los grupos a investigar. Sin embargo, debido a que en los estudios con seres humanos existen diversas limitaciones de naturaleza ética que impiden la aplicación de este método, el investigador frecuentemente utiliza los estudios de observación. En las secciones siguientes se examinará brevemente cada una de estas opciones de diseños de investigación.

### 5.2.2 Estudios de observación

La característica principal de los estudios de observación es que el investigador no tiene control sobre las condiciones de exposición de la población al factor de riesgo. Los estudios de observación pueden ser clasificados en descriptivos y analíticos (figura 1). Los descriptivos estudian la distribución de la enfermedad (tasa) de acuerdo con las características de la población (edad, sexo, raza etc.). Este tipo de estu-

dios es útil para examinar la magnitud de un problema en sectores de la población o para generar hipótesis que posteriormente puedan ser exploradas en estudios analíticos, cuyo objeto es probar hipótesis específicas de asociación entre factores de riesgo y daños a la salud.

Los estudios de observación analíticos pueden ser clasificados de acuerdo con su secuencia temporal en: a) de corte transversal b) retrospectivos o de casos y controles y c) prospectivos o de cohorte. A continuación se describen estos tres tipos de estudios de observación analíticos.

**Figura 1**

**Relación entre la secuencia del razonamiento epidemiológico y el tipo de diseño del estudio**

Secuencia del razonamiento epidemiológico	Diseño del estudio
Nivel 1      Identificación y definición del problema de salud	De observación descriptivo * Descripción por persona, lugar, tiempo etc.
Nivel 2      Descripción del problema de salud	
Nivel 3      Búsqueda de la causalidad: Asociación estadística con factores de riesgo Verificación del sesgo Criterios de causalidad	De observación analítico * Corte transversal * Retrospectivo (casos- controles) * Prospectivo (cohorte)
Nivel 4      Identificación de medidas preventivas y de control	
Nivel 5      Búsqueda de la causalidad Determinación de la eficacia de una intervención (profiláctica y/o terapéutica)	Experimental * Ensayo clínico aleatorio controlado
Nivel 6      Evaluación de la intervención	
Redefinición del problema de salud (regresar al nivel 1)	

Figura 2

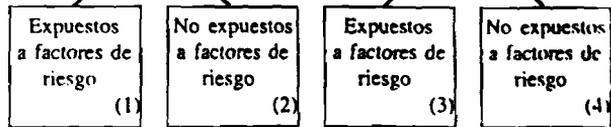
## Características básicas de los principales diseños de estudios epidemiológicos

## a) Estudio retrospectivo (casos-contrroles)

I. El estudio se inicia con la selección de:



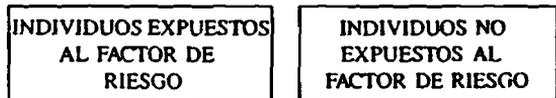
II. Se mide la exposición previa a factores de riesgo



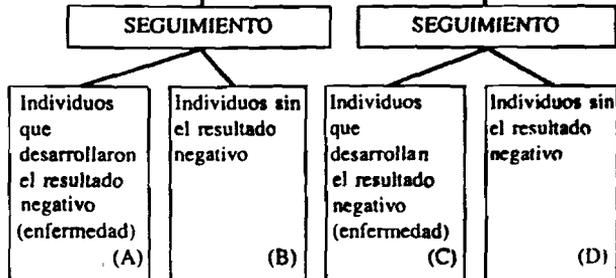
El objetivo de estos estudios es comparar la proporción de sujetos expuestos a los factores de riesgo entre los casos  $(1)/(1) + (2)$  con la proporción de sujetos expuestos a los factores de riesgo entre los controles  $(3)/(3) + (4)$ . Si la primera proporción es mayor que la segunda se puede inferir que hay una asociación probablemente causal.

## b) Estudio prospectivo (cohorte)

I. El estudio se inicia con la selección de:



II. Se mide la presencia del daño a la salud



El objetivo de estos estudios es comparar la incidencia del daño a la salud (enfermedad o muerte) entre los expuestos a los factores de riesgo  $(A)/(A) + (B)$  con la incidencia del daño a la salud entre los no-expuestos a los factores de riesgo  $(C)/(C) + (D)$ . Si la primera tasa de incidencia es mayor que la segunda se puede inferir que hay una asociación probablemente causal.

### a) Estudios de corte transversal

Por lo general un estudio de corte transversal es el de diseño más sencillo. Este tipo de diseño permite identificar la magnitud y la distribución de una enfermedad, de otros daños a la salud y de factores de riesgo en la población. Sin embargo, no permite identificar claramente asociaciones causales entre los daños a la salud y los factores de riesgo importantes o de interés. La característica esencial de este estudio es que la información sobre la presencia de los factores de riesgo y los daños a la salud se recaba simultáneamente, lo cual generalmente impide reconocer la relación temporal de una posible asociación entre ellos. Los estudios de corte transversal

- generalmente representan una muestra tomada de toda la población;
- se realizan en un período único, breve y bien delimitado;
- se utilizan para identificar la frecuencia de un daño a la salud o de un factor de riesgo; y
- se emplean para describir los atributos de quienes han experimentado un daño a la salud y aquellos que no lo han sufrido.

Pueden utilizar o no los registros de salud existentes. Los métodos usados para formular hipótesis de investigación, los procedimientos de muestreo, la recolección de información, el análisis de los datos y la interpretación son semejantes a los métodos empleados en cualquier tipo de estudio epidemiológico. El método para deducir sobre la naturaleza causal de la asociación es diferente, en el sentido de que generalmente el investigador sólo puede describir asociaciones entre los daños y los factores de riesgo presentes en el momento de realizar la encuesta.

Habitualmente no existe manera de asegurar, como podría hacerse mediante un estudio analítico, si el factor de riesgo era anterior, posterior o simultáneo al resultado observado. Un ejemplo podría ser el descubrimiento de una asociación

entre una elevación en la presión arterial sanguínea y el desprendimiento prematuro de la placenta. Ambos pueden estar asociados, pero un estudio de corte transversal no permitiría decir si el aumento en la presión arterial necesariamente condujo al desprendimiento prematuro de la placenta.

#### *Ventajas de los estudios de corte transversal*

Son relativamente sencillos, generalmente baratos y pueden hacerse con rapidez, utilizando información de fácil obtención. Pueden proporcionar rápidamente un perfil de la población y son especialmente útiles para determinar la prevalencia de factores de riesgo.

Existen formas de relacionar varios estudios de corte transversal, que permiten obtener conclusiones acerca de cambios dados en cierto período. En estos estudios, llamados de *corte transversal interconectados*, se puede decir, figurativamente, que se toman varias fotografías de las mismas personas en diferentes momentos. Cabe mencionar que son difíciles de realizar en zonas subdesarrolladas. Una serie de encuestas de corte transversal—no necesariamente de las mismas personas—permite medir cambios, con el paso del tiempo, en las tasas de prevalencia de daños a la salud. Este método, aún con limitaciones, es una aproximación a un estudio de cohorte o seguimiento.

#### *Desventajas de los estudios de corte transversal*

En los estudios de corte transversal frecuentemente no se puede llegar a deducciones con respecto a los factores de riesgo presentes en épocas anteriores al estudio. Por ejemplo, si se incluye mujeres que han dado a luz recientemente, sólo se puede sacar conclusiones válidas sobre factores de riesgo que están presentes en el momento de hacerse la entrevista y que se sabe estaban presentes antes del nacimiento—por ejemplo, estatura, situación socioeconómica, edad. No se puede derivar ninguna inferencia causal sobre factores que aparecen durante el embarazo (por ejemplo, hipertensión

del primer trimestre). Se puede hacer *conjeturas tentativas* sobre factores que están presentes en el momento de hacerse la encuesta y que se sospecha también estuvieron presentes durante la gestación (por ejemplo, proteinuria, anemia). Quizá el inconveniente más importante de los estudios de corte transversal es lo que se ha denominado fenómeno de supervivencia. Por definición, en el estudio sólo quedan incluidas aquellas personas que están vivas en el momento de la encuesta. Por ejemplo, si se pretende conocer la talla promedio en niños de un año de edad, aquellos con corta estatura que murieron no pueden ser incluidos en la muestra; si la mortalidad infantil es alta un número importante de niños va a quedar fuera de la muestra.

### b) Estudios retrospectivos (casos-contróles)

Estos estudios están destinados a explorar hipótesis de asociación entre factores de riesgo y daños a la salud. En ellos se seleccionan *casos* o sujetos con el daño a la salud bajo estudio, y *contróles* o sujetos que no poseen el daño a la salud bajo estudio. Interesa conocer la proporción de sujetos expuestos a los factores de riesgo en ambos grupos (figura 2.a).

En algunas oportunidades su diferenciación con los estudios de corte transversal puede ser difícil, dado que los estudios de casos y contróles pueden tener un componente de corte transversal, por ejemplo al estudiar los valores serológicos entre casos y contróles.

En los estudios de casos y contróles se puede recolectar una gran variedad de información de diferentes fuentes. Por ejemplo, información clínica, socioeconómica, del ambiente, la educación, los patrones de vida (*fumar*), el uso de drogas y medicamentos y la atención prenatal de los pacientes bajo estudio. Esta información puede extraerse de historias clínicas, encuestas especiales, entrevistas personales, etc.

#### *Ventajas de los estudios retrospectivos*

Los estudios retrospectivos son efectivos para estudiar hipótesis de manera relativamente sen-

cilla, rápida y económica si se los compara con los estudios de cohorte. Esto se debe a que permiten establecer una relación entre un daño a la salud y la exposición a factores de riesgo existentes o que incidieron en el pasado. Por ello puede deducirse una relación causal con mayor confianza que en un estudio de corte transversal. Los estudios de casos y contróles permiten además estudiar múltiples factores de riesgo.

La ventaja principal es que, de existir fuentes adecuadas de información recopilada rutinariamente, con el análisis retrospectivo no es necesario realizar nuevos trabajos de recolección en el terreno. Varios estudios de enfoque de riesgo realizados en países desarrollados se han basado en un análisis retrospectivo de registros prenatales y en certificados de nacimiento y defunción. Si el sistema de registro no es suficientemente bueno, un análisis retrospectivo de registros deficientes puede servir para obtener muchas indicaciones útiles para una ulterior investigación.

#### *Desventajas de los estudios retrospectivos*

El estudio retrospectivo tiene algunas de las desventajas de los estudios de corte transversal, así como otras adicionales, debido a que la información recabada en un estudio retrospectivo depende de la memoria de las personas y/o los registros de los servicios de salud. Algunas veces estos últimos no existen, o están tan incompletos que es imposible descubrir si durante el embarazo existía o no un determinado factor de riesgo.

Otra desventaja es que debido a la selección de los casos, los estudios no permiten establecer la incidencia de la enfermedad. Finalmente no son muy útiles para el estudio de la exposición a factores raros o inusuales.

### c) Estudios prospectivos (de cohorte)

Los estudios prospectivos se basan en la identificación y observación de *grupos bien definidos de sujetos que pueden experimentar en el futuro un daño a la salud* (figura 2.b). El riesgo de experimentar este daño a la salud, se puede

calcular tanto para quienes están expuestos a un atributo o factor de riesgo determinado como para quienes no lo están. En estos estudios es de interés que los sujetos sean representativos de la población a riesgo de experimentar un daño a la salud. No obstante, el criterio más importante es que los grupos bajo estudio (expuestos y no expuestos al factor de riesgo) sean comparables entre sí, con respecto a otras características que aquellas estudiadas. Las personas que son objeto del estudio forman una cohorte—un grupo bien definido, con una experiencia común—que es seguida en el tiempo para observar en ella una serie de acontecimientos o resultados (daños a la salud).

El método de recolección de datos durante todo el período del estudio debe ser confiable, de modo que se pueda obtener una referencia válida de la información posterior. Como generalmente se necesita mantener contacto con las personas que son objeto de estudio por períodos largos, es preciso diseñar un sistema para instrumentar dicho contacto a largo plazo. Aquellas personas que han quedado fuera del estudio por causa de migración u otras razones, deberán ser localizadas, de ser posible, para poder dejar constancia de su experiencia durante el período de observación.

#### *Ventajas de los estudios prospectivos (de cohorte)*

Permiten recopilar información sobre la exposición a los factores de riesgo en forma confiable. La asociación entre posibles factores de riesgo y el fenómeno bajo estudio tiene menos probabilidad de distorsión que en los estudios retrospectivos, dado que no dependen de la memoria o de los registros médicos. Los estudios de cohorte permiten estimar la incidencia del daño a la salud en el grupo expuesto y en el no expuesto al/los factor(es) de riesgo, y por lo tanto calcular directamente el riesgo relativo.

Dado que permiten medir directamente el riesgo de que se obtenga un daño a la salud, ofrecen un método poderoso y más directo de probar hipótesis sobre la relación entre factores

de riesgo y daños a la salud. Este diseño es útil tanto para estudiar resultados múltiples como para abordar problemas de exposición infrecuente al factor de riesgo postulado.

#### *Desventajas de los estudios prospectivos*

Cuando la incidencia del daño es relativamente baja (por ejemplo mortalidad perinatal o materna) se requiere un gran número de sujetos que serán seguidos durante un largo período para obtener conclusiones confiables. El seguimiento de un gran número de pacientes requiere una infraestructura compleja y costosa que quizá no sea factible sostener en un país en vías de desarrollo.

Si el problema de salud que se está investigando tiene una incidencia baja, es decir, si es bastante raro, se puede tener muchos más casos de personas sanas que los necesarios para el análisis y quizá muy pocos casos de personas enfermas. Este diseño es poco útil para daños o padecimientos de baja frecuencia.

Al igual que en los estudios de corte transversal interconectados, las personas que se retiran del estudio plantean un problema importante, puesto que sus características pueden ser diferentes de las de quienes permanecen en él. Si entre las personas que se retiran hay una proporción alta que tiene un riesgo elevado de presentar el daño, las tasas de incidencia de ese daño a la salud que se obtiene del estudio podrían ser más bajas de lo que son en realidad.

#### **d) Consideraciones generales**

En el cuadro 1 se presentan las ventajas y desventajas de los estudios de corte transversal, retrospectivos y prospectivos.

Un punto de importancia en el diseño de estudios es el que se refiere al precio que se paga por obtener un mayor número de eventos cuando la frecuencia del daño y/o del factor es baja. Al hacer la selección de los grupos a partir de la presencia o ausencia de un daño (estudios retrospectivos) o de la presencia o ausencia de la exposición a un factor de riesgo (estudios pros-

pectivos) ya no resulta posible derivar una medida de la incidencia del daño a la salud en el primer caso, ni una medida de la proporción del factor de riesgo en la población total en el segundo de los casos. Por lo tanto, debemos obtener esa información a partir de otras fuentes. Esto se puede apreciar con más claridad en el cuadro 2.

En un estudio retrospectivo, o de casos y controles, los grupos de estudio han sido escogidos conforme a la presencia o ausencia del daño. Por consiguiente la relación entre ambas expresada como razón,  $(a + c)/(b + d)$ , se distorsiona al no corresponder a la población general. Esto se debe a que se ha escogido un número

predeterminado del total de aquellos que tienen el daño y quienes no lo tienen, sin saber cuántas personas con el daño y cuántas sin él existen en la población general. A pesar de eso, es muy importante el hecho de que las razones  $a/(a + c)$  y  $b/(b + d)$  se mantienen constantes como las que existen en la población general y en consecuencia la razón de productos cruzados  $(a \cdot d)/(b \cdot c)$  en un estudio de casos y controles es la misma que la que se encontraría en la población general. Conviene recordar que  $a/(a + c)$  es la tasa de prevalencia del factor de riesgo entre los individuos que presentan el daño a la salud y  $b/(b + d)$  es la tasa de prevalencia del factor de riesgo entre los que no presentan dicho daño.

### Cuadro 1

Comparación de las ventajas y desventajas de los principales tipos de diseño de investigación epidemiológica

TIPO DE DISEÑO	VENTAJAS	DESVENTAJAS
CORTE TRANSVERSAL	<ol style="list-style-type: none"> <li>1. Sencillo y barato.</li> <li>2. Se determina la prevalencia de una enfermedad o de un factor de riesgo.</li> <li>3. Permite la descripción de la población.</li> </ol>	<ol style="list-style-type: none"> <li>1. No sustenta inferencias de causalidad.</li> <li>2. No se puede determinar el riesgo relativo.</li> </ol>
RETROSPECTIVO (CASOS Y CONTROLES)	<ol style="list-style-type: none"> <li>1. Barato.</li> <li>2. Corto tiempo de estudio.</li> <li>3. Menor cantidad de sujetos.</li> <li>4. Adecuado para enfermedades raras.</li> <li>5. Se identifica fácilmente a los sujetos (casos).</li> </ol>	<ol style="list-style-type: none"> <li>1. Posibilidad de sesgos en la obtención de la información.</li> <li>2. No se puede determinar la incidencia.</li> <li>3. La determinación del riesgo relativo es solamente aproximada.</li> <li>4. Puede ser difícil la selección e identificación del grupo control.</li> </ol>

PROSPECTIVO  
(COHORTE)

- |  |  |
|--|--|
| 1. Menos posibilidad de sesgo.   | 1. Requiere un tiempo de estudio más prolongado.                                 |
| 2. Se determina el riesgo relativo.  | 2. Frecuentemente costoso.   |
| 3. Se determina la incidencia.   | 3. No es adecuado para enfermedades de baja frecuencia.                          |
| 4. Puede combinar el estudio de varios daños a la salud (enfermedades o muerte). | 4. Requiere un mayor número de sujetos.  |
|  | 5. Adolece del problema de pérdida de sujetos de estudio durante el seguimiento. |

**Cuadro 2**

Representación esquemática de la relación entre la presencia o ausencia de un factor de riesgo y la presencia o ausencia de un daño a la salud

FACTOR DE RIESGO	DAÑO A LA SALUD		TOTAL
	PRESENTE	AUSENTE	
PRESENTE	a	b	a + b
AUSENTE	c	d	c + d
<u>TOTAL</u>	a + c	b + d	a + b + c + d

$$a + b + c + d = N$$

### 5.2.3 Estudios experimentales

La característica principal de un estudio experimental es que el investigador tiene control sobre las condiciones bajo las cuales la población está expuesta al factor de riesgo. De ahí que los estudios experimentales sean siempre de naturaleza prospectiva, pero con una característica adicional: la asignación de la población al grupo de estudio o al grupo control (expuestos, no expuestos) se hace al azar. Los grupos de estudio y control deben ser lo más posible comparables entre sí excepto en su exposición o no al factor que se desea estudiar.

Los estudios experimentales o ensayos clínicos aleatorios controlados se utilizan para evaluar estrategias de prevención primaria (por ejemplo prueba de vacunas o intervenciones nutricionales) o de prevención secundaria (por ejemplo intervenciones terapéuticas).

Lo ideal desde el punto de vista científico sería determinar la eficacia de todas las estrategias de intervención mediante el empleo de ensayos controlados, pero con frecuencia se contempla como algo poco viable tanto desde el punto de vista político como ético.

### 5.2.4 Selección del grupo de comparación

La selección adecuada de un grupo de comparación es una de las partes más importantes, y muchas veces difícil, en un estudio epidemiológico de observación analítico. Un grupo control es un instrumento de comparación empleado para evaluar el efecto de un factor o exposición.

Según el tipo de controles pueden derivarse determinadas inferencias en el estudio. Si se trata de un estudio de casos y controles, los controles expresan la proporción esperada de un factor de riesgo en un grupo control que es comparable al de los casos en todos los aspectos pertinentes, excepto en cuanto a que ninguno de sus integrantes ha experimentado el daño a la salud. En un estudio de cohorte, el grupo con-

trol (no expuesto) señala la incidencia esperada del daño a la salud de interés para el estudio en un grupo que es comparable al expuesto en todas las formas pertinentes, excepto en cuanto a que no ha estado expuesto al factor de riesgo o al tratamiento. La clave consiste en seleccionar controles que sean comparables con el grupo de estudio en cuanto a las características que pronostiquen tanto la exposición a la enfermedad, como la presencia de un factor de riesgo.

La pregunta fundamental es: ¿Cómo se define el grupo que se necesita para servir como base de comparación? Lo que hay que hacer es seleccionar un grupo que esté sujeto lo menos posible al sesgo de selección. Existen seis reglas prácticas para ello.

- a) *No escoger.* La mejor solución al problema de la comparación consiste en no escoger, es decir, tomar como grupo de estudio a todas las personas que tengan el daño a la salud y como controles a todas aquellas que no lo tengan.
- b) *Tomar una muestra sistemática de controles.* Cuando el número de posibles controles es demasiado grande, se toma una muestra sistemática de ellos (los principios y técnicas de muestreo serán discutidos en otras secciones).
- c) *En un estudio retrospectivo se debe seleccionar los casos (daño a la salud) y los controles, al mismo tiempo y en el mismo lugar.* Al seleccionar el grupo control al mismo tiempo y en el mismo lugar, se reducen las posibilidades de que algún factor externo afecte a uno y no al otro grupo y debilita la comparabilidad.
- d) *Tener cuidado con los registros.* Hacer una lista de los casos que han sido recolectados sistemáticamente.
- e) *Ejecutar el pareo individual de casos y controles sólo cuando sea necesario.* El pareo es la selección de un control particular para cada caso, de tal manera que

compartan una serie de características pre-determinadas.

El pareo por una característica particular elimina una diferencia entre los dos grupos y, por lo tanto, no se puede determinar si hay alguna variación entre los grupos como resultado de ese factor. Mientras más factores se decida incluir en el proceso de pareo de los grupos, más difícil será identificar un grupo control.

- f) *Estratificar en lugar de parear.* Si se necesita asegurar que los grupos de comparación sean comparables en cuestiones tales como edad, sexo y región, pero no obstante se quiere que el proceso resulte más fácil, puede estratificarse la muestra de controles. La estratificación permite que para cada estrato se compare, en el

análisis, el grupo de estudio con el grupo control.

### 5.2.5 ¿Cómo se selecciona el diseño de estudio más adecuado?

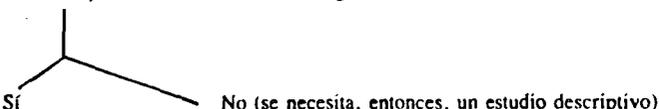
Es útil analizar en qué lugar de la secuencia del razonamiento epidemiológico está ubicada la problemática de estudio. En la mayoría de los casos se encontrará entre los niveles uno y cuatro de la figura 1.

Luego se debe examinar las hipótesis y objetivos de la investigación, lo que ayudara a determinar si se selecciona un enfoque descriptivo o uno analítico. En el diagrama que aparece a continuación (figura 3) se presentan algunas preguntas que son de utilidad para decidir entre estos dos enfoques.

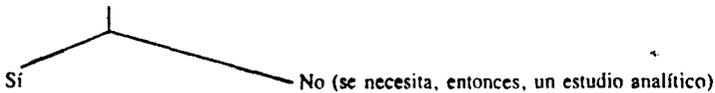
Figura 3

#### Preguntas a responder en la selección del diseño de estudio

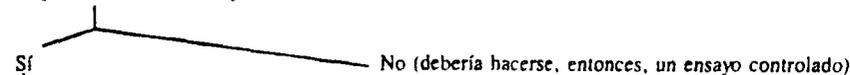
¿Se conoce la incidencia o prevalencia del factor de riesgo?



¿Está bien definida la naturaleza causal de la asociación entre el factor de riesgo y el padecimiento o acontecimiento de interés?



¿Se ha hecho un ensayo controlado?



Aplique los resultados del ensayo controlado

A continuación hay que determinar la viabilidad o factibilidad del diseño escogido en el escenario en cuestión. Algunas preguntas útiles pueden ser: ¿Se dispone de registros de salud? ¿Hay buenos caminos de acceso? ¿Se llevan a cabo los partos en las casas? Es necesario comparar el diseño deseado con los recursos disponibles e identificar las limitaciones, a fin de proponer soluciones.

Finalmente se deben examinar estas soluciones para ver si siguen siendo válidas desde el punto de vista científico, así como determinar si el estudio proporcionará la información necesaria para responder a la pregunta (hipótesis) inicial.

## Ejercicio 2

Aunque los investigadores pueden tener muchas hipótesis específicas que quieren demostrar, las cinco hipótesis de investigación en las que se centrará todo el documento de trabajo son:

1. La mortalidad perinatal en la región montañosa de Ficticia es más elevada que en las otras regiones del país.
2. La mortalidad perinatal de los productos de madres que son atendidas durante el parto por parteras empíricas sin adiestramiento, es mayor que la mortalidad perinatal entre las madres atendidas por parteras empíricas adiestradas.
3. La mortalidad perinatal de los productos de madres que padecen desórdenes hipertensivos del embarazo (DHE) o toxemia, es mayor que entre madres que no sufren de DHE.
4. La mortalidad perinatal de los productos de madres de 19 años de edad o menores, es mayor que entre madres de 20 a 35 años de edad.
5. La mortalidad perinatal de los productos de madres que tienen más de tres hijos vivos, es mayor que entre madres que tienen tres o menos hijos vivos.

(i) Tome en consideración las hipótesis de investigación para el estudio epidemiológico (daños a la salud—factores de riesgo) y describa el diseño del estudio que consideraría ideal para realizar esta investigación en Ficticia. Justifique su selección en términos de las ventajas y desventajas de los distintos diseños presentados en el texto.

(ii) Se está estudiando la relación entre mortalidad perinatal y sus factores de riesgo. Los recursos permiten realizar un estudio retrospectivo, que incluye 5.000 nacimientos, o un estudio prospectivo (de cohorte) de 1.000 mujeres embarazadas. ¿Cuál escogería y por qué? ¿Por qué no escogería el otro? Responda tomando en cuenta la cantidad de información y el costo del estudio.

(iii) Se estableció una reorganización de la atención perinatal en el hospital A, mientras que el hospital B mantuvo el sistema tradicional de atención perinatal. Para evaluar el impacto de la reorganización de la atención perinatal en el hospital A, se seleccionó al hospital B como grupo de comparación. De la información que se presenta en el cuadro 3 seleccione la medida de impacto que usaría para evaluar dicha reorganización y explique su selección. ¿Podría concluirse (con los datos del cuadro 3) que la diferencia en la mortalidad de ambos hospitales se debe a la reestructuración del sistema de atención? ¿Qué otros factores podrían explicar la diferencia?

**Cuadro 3**

Características de la población atendida en dos hospitales de Ficticia

Variable	Hospital A		Hospital B	
	No.	%	No.	%
<i>Edad de la madre (años)</i>				
10 - 17	45	1,1	678	19,3
18 - 19	117	2,9	698	19,9
20 - 29	2.911	71,1	1.821	51,9
30 - 34	658	16,1	144	4,1
35 +	346	8,4	154	4,4
Desconocida	18	0,4	14	0,4
<i>Grupo étnico A</i>	3.922	95,8	515	14,7
<i>Grupo étnico B</i>	173	4,2	2.994	85,3
<i>Solteras</i>	52	1,3	1.385	39,5
<i>Analfabetas</i>	420	10,3	1.859	53,0
<i>Población neonatal</i>				
<i>Peso al nacer (gms)</i>				
1.499	57	1,4	248	7,1
1.500 - 2499	231	5,6	482	13,7
2.500 +	3.804	92,9	2.767	78,9
Desconocido	3	0,1	12	0,3
<b>TOTAL</b>		100		100
<i>Número de muertes neonatales por peso al nacer</i>				
2.500	59	66	227	71
2.500 +	31	34	92	29
<b>TOTAL</b>	90	100	319	100
<i>Zonas:</i>				
— Rurales	455	11,1	258	7,4
— Urbano periférica	371	9,1	3.096	88,2
— Urbanas	3.269	79,8	155	4,4
<i>Operaciones cesáreas</i>	902	22,0	359	10,2
<i>Sin atención prenatal</i>	699	17,1	1.043	29,7
<i>Total de nacimientos</i>	4.095		3.509	

## Sección 5.3 Selección del método de muestreo

### Resumen

La población a partir de la cual se puede seleccionar una muestra de personas es la población objeto del estudio. La lista de nombres y direcciones de esa población es el marco de muestra del estudio. En los métodos de muestreo se emplean técnicas probabilísticas y no probabilísticas. El proceso de muestreo está sujeto a varios tipos de errores que pueden llevar a falta de representatividad. Este sesgo puede ser causado por una técnica de muestreo inapropiada o por la falta de respuestas al cuestionario.

### 5.3.1 Definición de la población bajo estudio y del marco de muestreo.

Un estudio epidemiológico tiene por objeto medir en una población definida, un daño a la salud y los factores de riesgo con él asociados. Es importante definir la población para la que se quiere generalizar los resultados (población objetivo) y la población a partir de la cual se obtendrá la información (población bajo estudio). ¿Se está estudiando la población de una región, de una aldea o de un país? ¿A toda la población o sólo a mujeres embarazadas y a niños pequeños? ¿Qué personas se quiere incluir para verificar la hipótesis de la investigación y para cumplir los objetivos del estudio? Estas son preguntas que deben formularse. Hay que decidir si se estudia a todas las personas de la población objetivo o bien si se toma una muestra de ella (población bajo estudio). Una muestra seleccionada correctamente tendrá características muy similares a las de la población objetivo.

La definición de la población objetivo de la encuesta debe ser explícita y precisa, y debe ajustarse a la realidad. Para definir la población bajo estudio es preciso describir la población de

la cual se va a tomar una muestra. Por ejemplo: ¿Se desea que los resultados sean válidos para todo un país, o sólo para una región? Dentro de la región geográfica seleccionada: ¿Quiénes son las personas que serán incluidas? ¿Todas las mujeres o las mujeres entre los 15 y los 49 años de edad? ¿Se incluirá aquellas mujeres que normalmente habitan en la zona pero que están ausentes temporalmente? ¿Deberá incluirse a todas aquellas personas que tienen el riesgo de desarrollar el daño que se esté investigando? Sin embargo, dependiendo de los objetivos del estudio, algunas personas tendrán que ser excluidas por razones de edad, sexo, lugar de residencia, etc.

Una población bajo estudio no necesariamente está constituida por personas. Si la investigación del riesgo se efectúa analizando registros médicos, entonces la población bajo estudio son los registros de las personas elegibles. Si se está interesado en comunidades enteras, entonces las comunidades componen la población en estudio.

Finalmente, después de definir la población bajo estudio, los investigadores deben decidir si aquella comprende a todos los sujetos de esa población o tan sólo a un grupo seleccionado de ésta. Con frecuencia la respuesta es obvia, como en los casos en que la población bajo estudio es tan grande y está tan dispersa que no se puede abarcar completamente, ya sea por falta de tiempo, recursos financieros o personal.

### Selección del marco de muestreo

Para determinar cómo se va a trabajar con la población bajo estudio y con el fin de escoger la muestra, es necesario contar con una lista de todas las unidades de posible selección, a la que se denomina marco de muestreo. Esta puede ser un censo, un registro electoral, los registros de un centro de salud o una lista de aldeas con su población aproximada. No obstante en algunos casos, por ejemplo si se usan admisiones a un hospital que están sucediendo continuamente,

esta lista no será necesaria ni estará disponible previamente.

En ciertos casos es necesario identificar y enumerar la población por separado antes de proseguir con el muestreo. Si el diseño del estudio incluye una investigación de poblaciones en zonas diferentes del país, el marco de muestreo se puede definir subdividiendo un mapa por medio de cuadrículas. En la práctica, frecuentemente surgen dificultades para definir adecuadamente los marcos de muestreo. El investigador debe ser extremadamente cuidadoso al describir la composición de la población bajo estudio en el marco de muestreo antes de proceder a obtener la muestra para la investigación

#### Relación entre el diseño del estudio epidemiológico y el tipo de muestreo

El tipo de muestreo en estudios epidemiológicos depende de la forma en que se interpretará los resultados. Por ejemplo, en estudios de ob-

servación descriptivos, es muy importante que la muestra sea representativa de la población objetivo, dado que la extrapolación o generalización de los resultados es la principal meta (prevalencia, tasas de mortalidad/morbilidad). En estudios de observación analíticos así como en los experimentales, dado que el principal objetivo es comparar grupos (riesgo relativo), no es necesario que la muestra sea representativa de la población objetivo. Se requiere principalmente que los grupos de casos y controles o grupos experimentales sean comparables entre sí o que, al menos, los sesgos sean similares en ambos grupos.

Finalmente, en los estudios de casos y controles las muestras deben obtenerse separadamente para casos y controles. Si la muestra se obtiene de la población general, es probable que si la enfermedad no es frecuente, no se incluya suficiente número de casos. La figura 4 resume la relación entre el tipo de estudio y algunas estrategias de muestreo recomendadas.

Figura 4:

#### Tipos de estudios epidemiológicos y posibles estrategias de muestreo

##### TIPO DE ESTUDIO

##### 1. Descriptivo

##### ESTRATEGIA DE MUESTREO

— muestra representativa de toda la población

##### 2. Analítico

##### Casos y controles

— muestreo separado de casos y controles

— estudiar todos los casos y tomar una muestra de los controles

##### Prospectivo o cohorte

— muestreo de expuestos y no expuestos separadamente

— muestreo de toda la población y luego clasificar en expuestos y no-expuestos

— estudiar todos los expuestos y muestreo de los no expuestos

### 5.3.2 Métodos de muestreo

#### A) Métodos de muestreo probabilístico

Una vez definido el marco de muestreo debe determinarse la forma de selección de las personas a estudiarse. Sería erróneo hacer una selección *arbitraria* recurriendo, por ejemplo, a voluntarios o personas que se encuentren presentes por casualidad. Los resultados obtenidos en esa forma no son fácilmente generalizables. La mejor forma de llegar a conclusiones válidas sobre la población objetivo, mediante los resultados del muestreo, es la selección de las personas de la población bajo estudio por *muestreo probabilístico*. El principio del muestreo probabilístico es que cada persona, u otra unidad de muestreo (por ejemplo, una casa, un registro, una familia), tenga una probabilidad conocida, generalmente igual, de ser incluida en el estudio.

Si se hace correctamente, los resultados obtenidos de dicha muestra serán representativos de toda la población objetivo, dentro de un margen de error que se puede calcular.

Se dispone de una variedad de métodos de muestreo diferentes. En los párrafos siguientes se ofrece una breve presentación de algunos de los métodos comúnmente utilizados. El proceso de muestreo está plagado de problemas por lo que se recomienda que desde un principio se consulte a profesionales con experiencia en muestreo.

##### a) Muestreo aleatorio simple

Cuando se usa el muestreo aleatorio simple cada persona incluida en el marco de muestreo tiene la misma probabilidad de ser seleccionada en la muestra. El muestreo aleatorio simple se puede lograr si: (i) a cada miembro en el marco de muestreo se le asigna un número exclusivo; y (ii) si las personas son seleccionadas mediante la obtención de números aleatorios en las tablas correspondientes. La mayoría de los libros de texto sobre métodos estadísticos tienen tablas de números aleatorios. También se pueden obtener

mediante programas especiales de computadora.

##### b) Muestreo sistemático

Por otro lado, una lista basada en la población (tal como un censo electoral), puede ofrecer la oportunidad de tomar una muestra de personas en el orden aproximado en que figuran en la lista. El investigador puede decidir tomar cada segunda o quinta persona de la fila de espera en el departamento de consulta externa, o del censo electoral.

Un método de selección de la muestra en esta forma consiste en tomar al azar la primera persona que es seleccionada y una vez seleccionada esta primera persona, tomar sistemáticamente cada *n*-ésima persona, de ahí en adelante. No obstante debe tenerse cuidado de que no haya patrones cíclicos de referencias de pacientes y que el orden de la lista sea razonablemente contínuo al azar.

##### c) Muestreo aleatorio estratificado

El muestreo aleatorio simple al que se ha hecho mención puede no satisfacer las necesidades del investigador. El método de muestreo aleatorio no garantiza la obtención de los números de personas con ciertas características que se requieren para el estudio. Si resulta importante estudiar subgrupos representativos de personas (por ejemplo, por grupo étnico, edad, paridad, historia obstétrica previa, etc.), es preciso formar cuadros de muestreo (*estratos*) para cada una de estas características y obtener muestras aleatorias al interior de cada estrato.

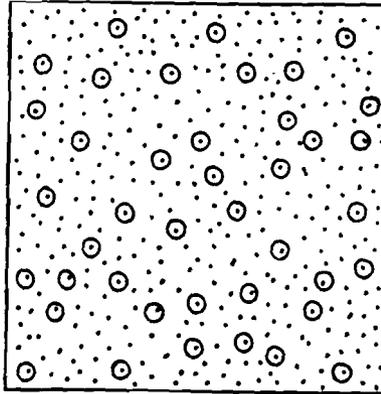
En caso necesario, debe corregirse la muestra para tener en cuenta las diferencias de tamaño de los estratos.

##### d) Muestreo por conglomerados

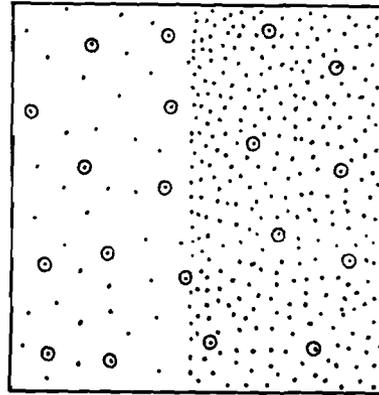
Puede ser difícil o imposible (debido a limitaciones de tiempo, dinero, personal, o por otras razones de logística) hacer un muestreo aleatorio, debido a que las personas se encuentran distribuidas sobre una amplia zona geográfica. Una solución consiste en escoger la muestra en

Figura 5: Esquema de muestreos probabilísticos

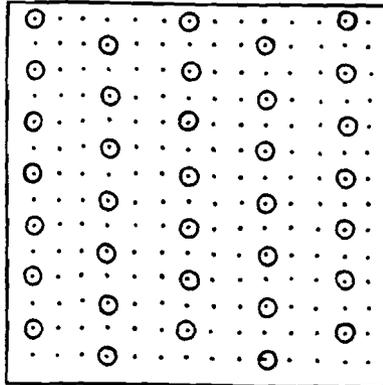
## MUESTREO ALEATORIO SIMPLE



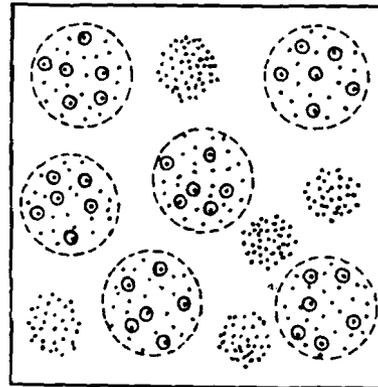
## MUESTREO ALEATORIO ESTRATIFICADO



## MUESTREO SISTEMÁTICO



## MUESTREO POR CONGLOMERADOS



dos etapas, obteniendo primero una muestra aleatoria simple, o estratificada por grupos o conglomerados de unidades de la población total. Estos conglomerados podrían abarcar aldeas, manzanas de casas en una calle, etc. Luego puede aplicarse los métodos de muestreo aleatorio simple o estratificado a las personas que se encuentran dentro de cada conglomerado seleccionado. Si se incluyen todas las personas en el conglomerado, el método se describe como de una sola etapa.

La figura 5 muestra una representación gráfica de los diferentes métodos de muestreo probabilístico presentados previamente.

#### B) Métodos de muestreo no probabilísticos: muestreo secuencial y por cuotas

Con frecuencia resulta conveniente, o necesario, tomar muestras de personas a medida que acuden al sitio de recolección de la información; por ejemplo, en una clínica de consulta externa. Este método puede describirse como secuencial.

El muestreo por cuotas es una manera de abreviar los pasos que se dan en el muestreo estratificado y se usa porque es más barato y más fácil de realizar; sin embargo, no es un método de muestreo probabilístico. Dado que no podemos calcular errores de muestreo, no tenemos ninguna indicación respecto a cuán representativa es la muestra. Se usa eficazmente en la investigación de opinión pública y, aunque puede tener éxito para echar un vistazo rápido y aproximado, no se recomienda para un estudio definitivo de factores de riesgo. Funciona de forma similar al muestreo estratificado en el sentido de que decidimos cuántas personas se quiere tener en cada estrato. En el muestreo por cuotas, el investigador entrevista a la mayor cantidad de personas que pueda encontrar dentro de cada categoría, hasta que llena su cuota.

#### 5.3.3 Problema del muestreo

En una situación ideal la muestra aleatoria debería ser una representación exacta de la po-

blación total; por ejemplo, si la prevalencia de la anemia en las mujeres embarazadas entre 30 y 34 años de edad es del 25% en la población en estudio, se desearía obtener una estimación de la prevalencia en la muestra cercana al 25%. Pero al estar trabajando con una muestra no siempre puede estimarse con plena exactitud la prevalencia en la población total, en parte por la variación aleatoria propia del muestreo y en parte por la influencia de los factores que se mencionan seguidamente.

#### A) Representatividad

Una muestra representativa es un subconjunto de la población cuya frecuencia de ciertos atributos es semejante a la frecuencia en la población total de la cual ha sido tomada. ¿Cómo decidir si las personas incluidas en la población bajo estudio son representativas de todas aquellas que tienen los factores de riesgo a los daños a la salud que se están estudiando? Por otra parte ¿cómo se asegura que aquellas personas seleccionadas porque no tienen el factor de riesgo o no han experimentado el daño, son verdaderamente representativas de todas aquellas que están exentas de factores de riesgo y/o daños? Estas son preguntas importantes que hay que contestar.

Si se produce una distribución desigual de una determinada característica entre la muestra y la población de la cual es tomada, la muestra ya no es representativa y los estudios de investigación basados en ella pueden resultar engañosos.

Cuando la falta de representatividad se debe simplemente al hecho de que se está tratando con una muestra en vez de con toda la población, se llama error aleatorio o de muestreo. Este error puede ser medido. El error de muestreo se reduce al aumentar el tamaño de la muestra y en él influye la prevalencia de la característica que se está estudiando. Por ejemplo, es máximo cuando la prevalencia ( $p$ ) es del 50%; desde este valor el error disminuye cuando ( $p$ ) tiende a cero o a 100%. (véase la sección 5.4 para más detalles).

**B) Error sistemático o sesgo**

No obstante, cuando la falta de representatividad se debe a otras razones (por ejemplo: la forma en que se hace una pregunta, la selección no aleatoria al escoger la muestra, una elevada tasa de falta de respuestas, etc.) se habla de sesgo. El sesgo es una diferencia sistemática que existe entre la población objetivo y la muestra, producido por un error técnico en la selección de personas o en la recolección de los datos. Por ejemplo, si la persona responsable de recolectar los datos es consciente del propósito del estudio, puede hacer preguntas de forma tal que obtenga una respuesta determinada que lleve, por ejemplo, a la demostración de una asociación entre un factor de riesgo y un daño a

la salud. El tipo y presentación de las preguntas y la interpretación de las respuestas puede ser afectado por las expectativas o prejuicios del investigador. Esto se conoce como sesgo del investigador.

El sesgo puede también intervenir en la manera en que se define la población bajo estudio. Si, por ejemplo, sólo se incluye mujeres embarazadas que asisten a clínicas prenatales, se sesga la muestra al excluir a aquellas mujeres, generalmente más pobres, que no asisten a esas clínicas. A éste se le llama error sistemático y es el tipo de error más importante, puesto que no se puede medir. Por lo tanto, es fundamental que se reduzca al mínimo o sea prevenido en la etapa de diseño del estudio.

**Ejemplo****Efecto del sesgo en la estimación de la tasa de morbilidad de un daño a la salud**

El sesgo aparece en una muestra durante varias etapas de un estudio epidemiológico:

- al definir poblaciones bajo estudio;
- en el muestreo y;
- al recopilar los datos.

Si en la población objetivo el 70% posee la característica A y el 30% la característica B, y la muestra contiene un 40% con A y un 60% con B, y además, el factor B está asociado con la prevalencia del daño a la salud, habrá un efecto de sesgo de la estimación de la frecuencia de personas afectadas por el daño. La magnitud del sesgo puede calcularse multiplicando los números de personas en la población objetivo y en la muestra por la probabilidad de que el resultado se encuentre presente por ejemplo en el 20% en quienes tienen el factor A y 50% en quienes está presente el factor B.

	<u>Factor A</u>	<u>Factor B</u>	<u>TOTAL</u>
Población objetivo	7.000	3.000	10.000
Muestra	400	600	1.000
Prevalencia del daño:	0,20	0,50	
Población objetivo:	$7.000 \times 0,2 + 3.000 \times 0,5$		= 290/1.000
Muestra:	$400 \times 0,2 + 600 \times 0,5$		= 380/1.000

Como puede verse, la muestra sesgada ha excedido, en más del 30%, la estimación de la tasa para un daño a la salud.

### C) Sesgos por falta de respuesta de los encuestados

Una elevada tasa de falta de respuestas es una fuente muy importante de sesgo que con frecuencia recibe poca atención al evaluarse la validez de las investigaciones. Las personas que rehúsan ser incluídas en un estudio, o que no pueden ser entrevistadas, con frecuencia son diferentes de aquellas que son fáciles de entrevistar y que acceden a cooperar. Así, por ejemplo, las mujeres que trabajan fuera del hogar posiblemente tengan menos hijos, tengan mayor o menor grado de escolaridad, o sean más o menos jóvenes que quienes trabajan en el hogar. El investigador encontrará que es más fácil entrevistar a quienes están en sus casas y, a menos que se haga grandes esfuerzos por entrevistar a las mujeres que están trabajando, éstas no estarán debidamente representadas y la muestra, por consecuencia estará sesgada. Algunos métodos de muestreo permiten a los investigadores substituir a las personas que no pueden ser entrevistadas. Sin embargo, aunque esto aumenta el tamaño de la muestra, el sesgo frecuentemente no puede ser corregido.

### Ejercicio 3

(i) Las cuatro regiones de Ficticia difieren considerablemente. No sólo son distintos los problemas de salud sino también las infraestructuras de tipo administrativo. En las partes urbanas de la zona costera, se registran todos los nacimientos y defunciones. En la mayoría de las aldeas de esa región, la mayor parte de la población es conocida por los servicios de salud. En la ~~zona~~ <sup>de Ficticia</sup> el registro es bastante bueno en las ciudades, pero muy deficiente en las zonas remotas. En las regiones del Delta y la Montaña el registro es extremadamente deficiente.

Si usted quisiera elaborar un cuadro representativo de la mortalidad perinatal en Ficticia: a) ¿qué método de muestreo probabilístico seleccionaría para cada una de las cuatro regiones?; b) ¿estratificaría su muestra por zonas ge-

ográficas?; c) en caso afirmativo ¿tomaría mismo número de personas en cada estrato ¿por qué?

(ii) Los investigadores decidieron elaborar un esquema que permita tomar una muestra representativa para estimar la mortalidad perinatal en Ficticia de la siguiente forma. Dado que las zonas urbanas de las regiones Costera y de ~~Delta~~ <sup>Delta</sup> tienen buenos sistemas de registros vitales, seleccionaron aleatoriamente una muestra de nacimientos y defunciones de ese registro. Las zonas rurales de las regiones Costera y de ~~Delta~~ <sup>Delta</sup> y la totalidad de las regiones de Montaña y el Delta, fueron muestreadas mediante conglomerados de las aldeas. En las ciudades de las regiones de la Montaña y el Delta, la muestra fue dividida en conglomerados y manzanas.

Discuta si el esquema de muestreo es adecuado para garantizar la representatividad de la muestra en Ficticia. ¿Qué sesgos puede haber introducido?

## Sección 5.4 Selección del tamaño de la muestra

### Resumen

El tamaño de la muestra mantiene relación directa con la confiabilidad y validez de las conclusiones que se esperan obtener. Depende de dos factores: (a) la prevalencia del resultado negativo o del factor de riesgo, según el caso, y (b) la precisión que se quiere obtener en los resultados.

### 5.4.1 Introducción a la selección del tamaño de la muestra

Una vez decididas las hipótesis, los objetivos y el tipo de diseño de la investigación, se enfrenta el problema de establecer el tamaño de muestra requerida. En general, puede ser difícil extraer conclusiones de una muestra muy pequeña. Por otra parte, una muestra demasiado grande puede ser una pérdida de recursos.

tiempo, tanto en la recolección como en el procesamiento de datos.

La selección del tamaño de la muestra depende del fin que se persigue con los resultados del estudio. La muestra debe ser lo suficientemente grande como para permitir sacar *conclusiones válidas*, relacionadas con la hipótesis. El que esto sea factible con una hipótesis determinada y con cierto tamaño de muestra depende, básicamente, de dos factores, de los cuales sólo uno de ellos es controlable: el grado de confianza que se desea tener en los resultados. El otro se refiere a la variación de los factores que serán medidos en la población.

Dado que la confianza que se tiene en las estimaciones de la población depende de la distribución que haya en los factores medidos, vale la pena usar medidas de tendencia central, que permitan describir la distribución de un factor en una población; las siguientes son *medidas de tendencia central* que ayudan a describir poblaciones.

#### 1) Medidas de tendencia central

##### a) *Media o promedio*

La media es el *valor promedio* de una serie de observaciones; se calcula dividiendo la suma de los valores individuales del grupo de observaciones entre el número total de observaciones en el grupo. Toda la información disponible se utiliza para obtener una sola cifra que describa la población. Una de las limitaciones de la media es que puede verse afectada por la *presencia*, en el grupo, de valores extremos, ya sea grandes o pequeños. No revela nada sobre la dispersión de las observaciones individuales alrededor de la media. Podría obtenerse valores medios similares o idénticos en grupos de observaciones que fuesen ya sea muy dispersos o muy agrupados.

##### b) *Mediana*

La mediana o *valor medio* de una serie de observaciones es aquel valor que divide el

total de observaciones (ordenadas de menor a mayor) en dos grupos de igual tamaño: es decir, que el 50% de la población está por encima y el 50% por debajo de la mediana. Dicho valor corresponde al percentil 50. Esto ofrece una medida de tendencia central que evita la distorsión que introducen en la *media* los valores extremos. La mediana tampoco proporciona ninguna información sobre la variabilidad de las observaciones.

##### c) *Modo*

El modo es el *valor que ocurre con más frecuencia*; puede estar representado por un número, grande o pequeño, de observaciones de un valor que es el que más se repite en la serie. No ofrece ninguna indicación acerca de la variación de las observaciones, o sobre los extremos de la distribución.

## Ejemplo

Compare los valores de las medidas de tendencia central (media, mediana, modo) en las dos muestras, A y B, que se presentan a continuación. Los datos corresponden a pesos al nacer en dos series de once valores cada una.

MUESTRA A		MUESTRA B	
Número de observación	peso al nacer (g)	Número de observación	peso al nacer (g)
1	1.700	1	1.500
2	2.000	2	1.800
3	2.200	3	2.100
4	2.500	4	2.400
5	2.700	5	2.500
6	3.200	6	3.100
7	3.300	7	3.400
8	3.300	8	3.500
9	3.300	9	3.700
10	3.400	10	4.000
11	4.000	11	4.000
media	2.873	media	2.909
mediana	3.200	mediana	3.100
modo	3.300	modo	4.000

Observe en qué grado las medidas de tendencia central pueden no ser representativas de todas las observaciones.

## B) Medidas de dispersión

*Variación y desviación estándar*

Una forma de medir la variabilidad o dispersión de las observaciones es la *variación*, o su transformación, luego de obtener la raíz cuadrada, en lo que se denomina *desviación estándar*. Se obtiene relacionando el valor de las observaciones individuales con la media global. Considérese el diagrama (figura 6) que muestra la distribución de once observa-

ciones ( $x_1, \dots, x_{11}$ ) obtenidas en un cuestionario.

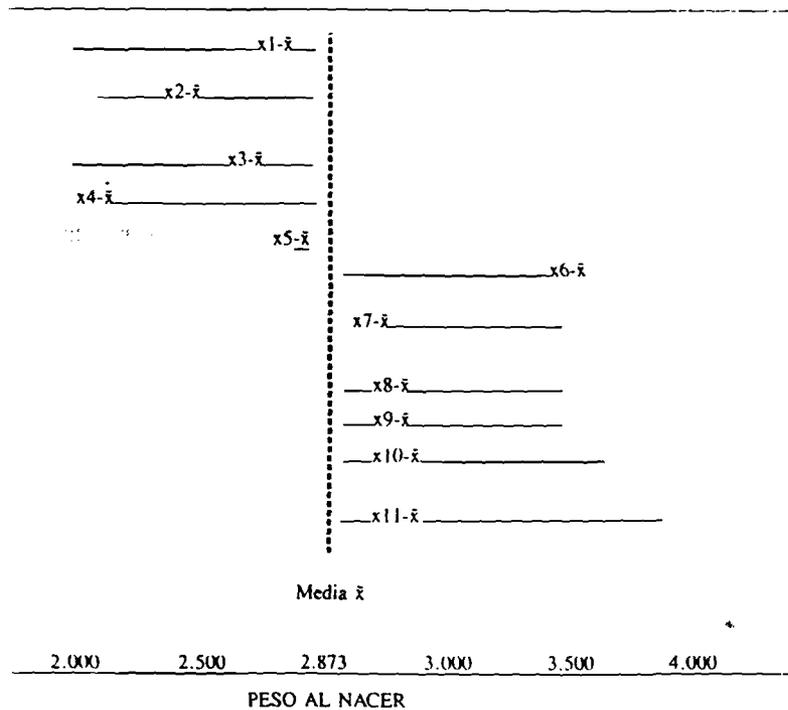
La desviación estándar se obtiene de la manera siguiente:

1. Se calcula las diferencias entre cada observación y la media total. Dado que sumar las diferencias negativas y positivas tenderían a cancelarse mutuamente, lo que se hace es elevar cada diferencia al cuadrado antes de efectuar su suma (se eliminan los signos).

- Las diferencias elevadas al cuadrado se suman para obtener la suma de las diferencias cuadráticas en torno a la media.
- La media de las diferencias cuadráticas se obtiene dividiendo la suma de las diferencias cuadráticas en torno a la media entre

- el número total de observaciones que hay en la muestra menos uno ( $n - 1$ ). El valor así obtenido es la *variancia*.
- La raíz cuadrada de la variancia es la *desviación estándar*.

Figura 6: Diagrama de la dispersión de un conjunto de valores de peso al nacer con relación a la medida ( $\bar{x}$ ) de la serie



Ejemplo

El conjunto de valores de peso al nacer ( $x$ ) para la muestra A que se presentó en el ejemplo anterior puede caracterizarse de acuerdo a su variabilidad, para lo cual se calcula la variancia y la raíz cuadrada de ésta: la desviación estándar.

Número de observación (n)	Peso al nacer (x)	(x- $\bar{x}$ )	(x- $\bar{x}$ ) <sup>2</sup>
1	1.700	- 1.173	1.375,929
2	2.000	- 873	762,129
3	2.200	- 673	452,929
4	2.500	- 373	139,129
5	2.700	- 173	29,929
6	3.200	+ 327	106,929
7	3.300	+ 427	182,329
8	3.300	+ 427	182,329
9	3.300	+ 427	182,329
10	3.400	+ 527	277,729
11	4.000	+ 1.127	1.270,129

$n = 11 \quad \bar{x} = 2.872,72$

$* \sum (x-\bar{x})^2 = 4\,961\,819$

Variancia =  $\frac{\sum(x-\bar{x})^2}{n-1} = \frac{4961819}{10}$   
 $= 496181,9$

Desviación estándar =  $\sqrt{\text{variancia}} = \sqrt{496181,9} = 704,4$  Una fórmula de trabajo para la desviación estándar es la siguiente:

$$DE = \sqrt{\frac{(\sum x)^2}{n} - \frac{\sum x^2}{n-1}}$$

C) La distribución normal y el uso de la desviación estándar

interesan (estatura, presión sanguínea, hemoglobina), tienen aproximadamente una distribución en forma de campana que se asemeja a la llamada *distribución normal*\*. No obstante, hay medidas que no se ajustan a la distribución normal, como en el caso del puntaje de App...

Se pueden describir y usar muchas distribuciones teóricas para identificar grupos o familias de diferentes tipos de variables. En biología, muchas de las variables continuas que

\* $\Sigma$  = Sumatoria

\*Recuérdese que el término "normal" no implica normalidad biológica.

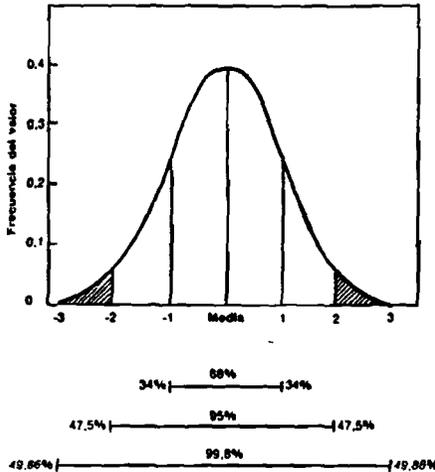
Si existe seguridad razonable de que la distribución de la variable que se está midiendo es semejante a una distribución normal, puede emplearse las propiedades de esta distribución para predecir la amplitud de los valores aplicados a ciertas proporciones específicas de la población en estudio.

En la figura 7 se indican las proporciones de la población incluidas dentro de la curva normal, describiéndose la amplitud mediante  $\pm 1$

desviación estándar (DE),  $\pm 2$  DE y  $\pm 3$  DE. Puede observarse que el 68.2% de la población está incluido dentro del rango de  $\pm 1$  DE, el 95.6% dentro del rango de  $\pm 2$  DE y el 99.8% por  $\pm 3$  DE. Obsérvese que son cifras aproximadas: por ejemplo, la cifra de 2 DE, correspondiente al 95% de la población, es una cifra redondeada del valor 1,96. En el resto de este documento de trabajo se usará la cifra 1,96.

Figura 7: Curva de distribución normal

Proporción de la población entre medidas de desviación estándar



#### D) Error estándar de la media

Se ha señalado que la distribución de las observaciones de diferentes variables a menudo se asemeja a una distribución simétrica o normal. Los valores *medios* obtenidos de las muestras de población también son similares a una distribución normal. Vale la pena destacar que *las medias de las muestras generalmente se*

*distribuyen normalmente, incluso si las observaciones individuales de las cuales se obtienen no lo son.* Se puede calcular una desviación estándar para una serie de medias (por ejemplo, las obtenidas mediante un muestreo repetido), de igual manera que para una serie de observaciones individuales.

De una muestra puede obtenerse una estimación de la desviación estándar de la media,

generalmente designada como el *error estándar de la media*. La variación en las medias dependerá del tamaño de la muestra y de la desviación estándar de las observaciones individuales.

Error estándar de la media (EEM) =

$$\frac{\text{Desviación estándar de la muestra}}{\sqrt{\text{Número de observaciones en la muestra}}} = \frac{DE}{\sqrt{n}}$$

### E) Intervalos de confianza

Ahora combíñese la información discutida sobre *errores estándar de la media* y la *distribución normal* para considerar el concepto de *intervalos de confianza*. Un intervalo de confianza es un intervalo dentro del cual se puede esperar que se encuentre un *parámetro* determinado (la media) verdadero, al fijar cierta confianza en que ello ocurra al efectuar muestras repetidas. Si se toma un número infinito de muestras, con un tamaño determinado, y se calculan las medias para cada una, es de esperar que en el 95% de esas muestras la media verdadera de la población estuviese dentro del rango de  $\pm 1,96$  errores estándar. Dicho de otro modo, se tiene confianza en que sólo en el 5%

de los intervalos, la media sea excluida de los intervalos. De esta manera se hace un planteamiento sobre la probabilidad de que surja una diferencia entre la media verdadera de toda la población y la media de la muestra que se obtuvo de ella.

Para ello, se obtiene el valor de lo que se denomina *precisión (d)*, que es la magnitud dentro de la cual se quiere permanecer a cada uno de los lados de la media. El valor de *d* (para un determinado nivel de confianza) se calcula multiplicando el error estándar de la media por el valor de  $z^*$  en que estamos interesados. Esta precisión entonces es sumada y restada a la media de la muestra obteniéndose un límite superior y un límite inferior de confianza.

\*El valor de  $z$  es el número de errores estándar que define el intervalo entre la media verdadera ( $\bar{X}$ ) y la media obtenida en la muestra, ( $x$ ). Se usa para transformar los valores observados a la escala de la curva normal, donde la media se vuelve cero y la desviación estándar uno. De esta forma se pueden usar las tablas de la distribución normal.

$Z$  es la variación estándar normal.

### Ejemplo

Media de la muestra ( $\bar{x}$ ) = 50

Tamaño de la muestra ( $n$ ) = 8

Suma de las diferencias cuadráticas en torno a la media = 175

Variación =  $175 / (\text{tamaño de la muestra} - 1) = 175 / (8 - 1) = 25$

Desviación estándar =  $\sqrt{\text{Variación}} = \sqrt{25} = 5$

Error estándar de la media (EEM) =  $\frac{5}{\sqrt{8}} = \frac{5}{2,83} = 1,77$

Precisión al 95% =  $d = \text{EEM} \cdot z = (1,77)(1,96) = 3,47$

Intervalos de confianza =  $\bar{x} \pm z (\text{EEM})$   
=  $\bar{x} \pm d$

Así, el intervalo de confianza del 95% =  $50 \pm 1,96 (1,77)$   
=  $50 \pm 3,47$   
= 46,53 a 53,47

En otras palabras, existe una confianza del 95% en que la media de la población sea mayor de 46,53 pero menor de 53,47.

### 5.4.2 Estimación del error estándar para una sola proporción

Ahora se examinará la tarea de estimar el tamaño de la muestra para un estudio, en el cual se va a determinar la proporción de la población con cierto daño o característica. Obviamente el investigador deseará evitar aquella situación en la que, por una parte, una muestra de tamaño inadecuado lleve a la detección de sólo uno o dos casos del problema que se está investigando y, por la otra, se tome una muestra innecesariamente grande, con las repercusiones en cuanto a costos, en la investigación de un problema relativamente común.

El (EEP) error estándar de una proporción, análogo al error estándar de una media, que fue descrito anteriormente, es una función del tamaño de la muestra ( $n$ ):

$$EEP = \sqrt{\frac{p(1-p)}{n}} \quad p = \text{proporción que se desea estimar}$$

Como se mencionó con anterioridad, el intervalo de confianza se representa mediante  $p \pm z$  (EEP), y la precisión de la estimación mediante  $d = z$  (EEP). Naturalmente  $p \pm d$  es igual al intervalo de confianza.

Observe que en general la precisión ( $d$ ) depende del tamaño de la muestra ( $n$ ) y no de la fracción de muestreo.

### 5.4.3 Pasos para determinar el tamaño de muestra necesario para un estudio de una sola proporción

Paso 1: Calcule la proporción del factor en la población (" $p$ ").

Paso 2: Utilizando las tablas de la distribución, determine el valor de " $z$ " que corresponda al nivel de confianza que usted desea. Para un nivel de confianza del 95%, las tablas de distribución normal

muestran un valor de  $z$  (la columna de la derecha) de 1.96. A continuación se presentan valores de  $z$  para cuatro diferentes niveles de confianza.

Nivel de confianza	Variación estándar normal o valor de "z" correspondiente (Dos colas)
90%	1.64
95%	1.96
99%	2.57
99.9%	3.29

90%	1.64
95%	1.96
99%	2.57
99.9%	3.29

Paso 3: Determine la precisión deseada ( $d$ ) que usted quiere obtener en su estudio. Recuerde que  $d = z$  (EEP).

Paso 4: Calcule el tamaño mínimo de la muestra mediante la fórmula siguiente:

$$n = \frac{z^2 [p(1-p)]}{d^2}$$

NOTA: El tamaño de la muestra ( $n$ ) para la proporción " $p$ " ha sido determinado de tal manera que la confianza en que el intervalo de confianza incluya la verdadera medida sea 95%.

### Ejemplo

Paso 1: Se determina la proporción,  $p = 0.20$ .

Paso 2: El nivel de confianza deseado es 95%, de modo que  $z = 1.96$

Paso 3: Se desea que la precisión sea de  $\pm 0.05$ . (EEP = 0.025, por lo que  $d = 0.05$ ).

Paso 4:  $n = \frac{(1.96)^2 (0.20)(1-0.20)}{(0.05)^2} = 246$

Por consiguiente, para que se pueda afirmar que la proporción buscada está en algún punto

entre 0.15 y 0.25 con un nivel de confianza del 95%, la muestra tiene que tener un *mínimo* de 246 observaciones.

*Nota:* El error estándar de la proporción (EEP) necesario para calcular  $d$  no se conoce en este caso porque no se tiene el número de observaciones ( $n$ ), que es lo que se desea obtener, ni el valor exacto de  $p$ . Por tanto, este valor debe estimarse sobre la base de las hipótesis y de la información obtenida por referencias bibliográficas. En aquellos casos en que se desconoce totalmente el valor de la proporción, se puede tomar 0.50 como el valor más conservador. En este caso la variancia es 0.25. ( $0.5 \times 0.5$ ).

#### 5.4.4 Estimación del tamaño de muestra necesario para un estudio de dos proporciones.

En la investigación de riesgos es de interés no sólo la prevalencia o incidencia de un factor de riesgo o una enfermedad, sino también la diferencia de tasas de incidencia o prevalencia entre dos subgrupos distintos de las poblaciones estudiadas. Por ejemplo, se desea saber si existe una diferencia real entre la prevalencia de anemia en aquellas mujeres cuyos hijos murieron al nacer y en aquellas cuyos niños sobrevivieron. En este caso son dos muestras separadas, y dos estimaciones muestrales separadas en cuanto a la prevalencia de un factor (anemia) en la población, cada una con su propio error estándar. Incluso si se encuentra una diferencia entre las dos tasas de prevalencia estimadas con las muestras, ¿cómo saber si la diferencia se debe simplemente a la superposición de la distribución de los errores estándar? Para averiguarlo se calcula el error estándar de la diferencia entre las dos proporciones. Esto depende de las dos tasas de prevalencia y de los dos tamaños muestrales.

En la práctica generalmente no se conoce las tasas de prevalencia en la población antes de hacer un estudio. Hay que hacer una conjetura

basada en el conocimiento del país, las condiciones prevalcientes en situaciones similares en otras áreas, en informes previos o simplemente en la intuición.

El investigador debe considerar una gran variedad de cuestiones al hacer la determinación del tamaño de la muestra. Las decisiones sobre el tamaño de la muestra *no* deben tomarse aisladamente de otros aspectos del estudio. El investigador debe examinar la hipótesis, los objetivos específicos y el diseño seleccionado. Estos deben relacionarse con conceptos estadísticos importantes que se presentan a continuación:

- $H_0$ : Hipótesis nula
- Error Alfa: Es la probabilidad de rechazar incorrectamente la hipótesis nula cuando es verdadera. Este valor, que convencionalmente se expresa como el valor  $p$  al indicar el resultado de una prueba estadística, se llama *Error Tipo I*, o *error Alfa*.
- Error Beta: Es la probabilidad de no rechazar incorrectamente la hipótesis nula cuando es falsa, es decir, cuando sí existe una diferencia real y se llama *Error Tipo II* o *Error Beta*.
- Poder =  $1 - \text{Beta}$ : Es la probabilidad que tiene el estudio de detectar una diferencia real de una magnitud determinada entre dos variables cuando se ha especificado la magnitud del error tipo I.
- $p_1$ , y  $p_2$ : Son las proporciones estimadas en los grupos de comparación, tanto de un factor de riesgo como de un daño a la salud, según el tipo de diseño del estudio.

## Cuadro 4

Relación entre la realidad y las decisiones que se toman con respecto a la veracidad de la hipótesis nula, basadas en los resultados del estudio

Decisión basada en los resultados del estudio	REALIDAD	
	$H_0$ falsa (existe una diferencia)	$H_0$ no falsa (no hay diferencia)
Rechazar $H_0$	Poder <u>1 - BETA</u>	Error tipo I <u>ALFA</u>
No Rechazar $H_0$	Error tipo II <u>BETA</u>	Confianza <u>1 - ALFA</u>

Si las dos proporciones que se desea estudiar (ej: incidencia en expuestos y no expuestos en estudios prospectivos) se expresan como  $p_1$  y  $p_2$ , por lo general, se expresa la hipótesis nula en la forma siguiente:  $H_0: p_2 - p_1 = 0$ , (no existe diferencia entre  $p_2$  y  $p_1$ ). La preocupación es por evitar cometer el error tipo I, es decir, evitar el rechazo de la hipótesis nula al señalar que existe una diferencia cuando en realidad no la hay. El nivel de confianza con que se acepta la hipótesis nula se expresa como 1 - Alfa. Un nivel de confianza mínimo es, por lo general, del 95%. También debe considerarse la probabilidad de no rechazar la hipótesis nula cuando en realidad existe una diferencia verdadera, es decir, Beta, o la probabilidad de cometer el error tipo II. La confianza con que correctamente se rechaza la hipótesis nula es el poder (1-Beta). En la literatura científica, en general, no se acepta un poder menor de 0.80, es decir, una probabilidad del 20% de señalar que no existe diferencia cuando en la realidad sí la existe.

La decisión de qué niveles de Alfa y Beta deben escogerse depende de la utilización que

se le dará a los resultados de la investigación. Para decisiones sobre programas se puede relajar los límites fijados tanto sobre Alfa como sobre Beta. Para decisiones que afectarán a personas, por ejemplo la realización de una operación cesárea, se desea aumentar la confianza con que se tomará la decisión y, por ende, reducir la probabilidad de cometer un error del tipo I o II.

## Ejemplo

## ¿Cómo no cometer errores del tipo I y II?

A un funcionario de salud de un distrito le gustaría seleccionar mujeres embarazadas y determinar quienes podrían necesitar una cesárea, ya que en su distrito sólo existe un centro donde realizar tal tipo de operación. Si dice que una mujer no necesitará una cesárea cuando en realidad la necesita (error tipo II), esto es mucho más grave para dicha mujer, que el decir que va a necesitar una operación cuando acaba por tener un parto vaginal normal (error tipo I). En el

primer caso (error tipo II), la mujer puede morir como resultado de no haber tenido la operación, mientras que en el segundo (error tipo I), habrá utilizado los recursos de un nivel más complejo de atención cuando no los necesitaba. La decisión que tome el funcionario de salud tendrá en cuenta las ventajas para las mujeres y las cotejará con el posible desperdicio de recursos de la comunidad.

#### 5.4.5 Pasos para estimar el tamaño de muestra necesario para un estudio de cohortes

Esta estimación es similar para estudios de casos y controles y para estudios de cohortes. Para simplificar la discusión aquí se presenta solamente para estudios de cohorte.

**Paso 1:** Calcular  $p_1$ , la tasa de incidencia en los expuestos y  $p_2$ , la tasa de incidencia en los no expuestos. La diferencia,  $p_1 - p_2$ , es la magnitud de la diferencia que se quiere comprobar si tiene importancia estadística.

**Paso 2:** Determinar la *variante estándar normal* ( $z$ ) de Alfa, que corresponde al nivel de error tipo I que se desea. Para simplificar las cosas, supóngase que siempre se usará una prueba de hipótesis de *dos colas*. Debe seleccionarse el valor correspondiente a  $z$  de Alfa en la tabla que aparece a continuación:

Nivel de confianza	Error-Alfa (nivel de significancia)	Valor $z$ de Alfa correspondiente
90 %	0,10	1,64
95 %	0,05	1,96
99,0 %	0,01	2,57
99,9 %	0,001	3,29

**Paso 3:** Determine la variante estándar normal,  $z$  de Beta, que corresponde al nivel de error tipo II que se desea y, consecuentemente, el poder de la prueba correspondiente. Seleccione el valor de  $z$  de Beta en la tabla que aparece a continuación:

Beta	Poder (1-Beta) correspondiente	$z$ de Beta
0,01	0,99	2,326
0,05	0,95	1,645
0,10	0,90	1,282
0,20	0,80	0,842

**Paso 4:** Calcule el tamaño de la muestra ( $n$ ), que es el tamaño de *cada* grupo a ser estudiado, es decir, que el tamaño de la muestra total es el doble del obtenido (se da por supuesto que los dos grupos son iguales).

$$\text{Si } \bar{p}^* = \frac{p_1 + p_2}{2}, \text{ Luego}$$

$$n = \frac{[z \text{ de Alfa } \sqrt{2 \bar{p}(1-\bar{p})} + z \text{ de Beta } \sqrt{p_1(1-p_1) + p_2(1-p_2)}]^2}{(p_1 - p_2)^2}$$

Recuerde que el tamaño obtenido es  $n$  para cada grupo, y que  $2n$  es el tamaño total de la muestra.

\* Nótese que  $\bar{p}$  es el promedio de las proporciones en las dos poblaciones que se están estudiando y no la proporción en la población global.

### Ejemplo

A un investigador le gustaría tener un poder del 90% para mostrar una diferencia entre una tasa de 0,10 en los expuestos y una tasa de 0,05 en los no expuestos, al hacer la prueba de hipótesis con un nivel de significación del 1% (error tipo I = 0,01).

$$p_1 = 0,10$$

$$p_2 = 0,05$$

z de Alfa = 2,576 z de Alfa es el valor correspondiente = 0,01 (dos colas)

z de Beta = 1,282 z de Beta es el valor correspondiente al = 0,10 (dos colas)

$$p = \frac{0,10 + 0,05}{2} = 0,075$$

$$n = \left[ \frac{2,576 \sqrt{2(0,075)(0,925)} + 1,282 \sqrt{0,10(0,90) + 0,05(0,95)}}{(0,10 - 0,05)^2} \right]^2$$

$$n = \frac{[0,9595 + 0,4754]^2}{(0,05)^2}$$

$$n = \frac{2,0589}{0,0025} = 824$$

Por consiguiente, en cada grupo se necesitan 824 observaciones con un tamaño total de la muestra de 1.648.

#### 5.4.6 Selección final del tamaño de la muestra

Las decisiones sobre el tamaño de la muestra no pueden tomarse en forma aislada de otros aspectos del estudio. Entre los factores principales que hay que considerar al tomar la decisión figuran:

1. Las hipótesis específicas, los factores de riesgo y el tipo de problema de salud a estudiar.
2. Estimaciones sobre la probable frecuencia de los factores de riesgo y del daño a la salud (el cálculo del tamaño de muestra debe usar la frecuencia más baja).

3. El tipo de prueba estadística que se va a realizar.

4. El grado de confianza que se adjudica a los resultados dependiendo de características y riesgos de la intervención propuesta.

5. Limitaciones logísticas: personal, recursos, conocimientos prácticos, instalaciones para el procesamiento de datos, etc.

Debe señalarse que el cálculo del tamaño de la muestra representa el número *mínimo* de observaciones necesarias para detectar la diferencia seleccionada. Sin embargo, en las investigaciones frecuentemente se pierden observaciones, lo que debe tomarse en cuenta para seleccionar el número *óptimo* de observaciones.

### Ejercicio 4

(i) Un estudio piloto ha demostrado que alrededor del 25% de las mujeres embarazadas están anémicas en el sexto mes de gestación. ¿Cuántas mujeres se necesitaría examinar para poder establecer (con una certeza del 95%) la proporción de mujeres anémicas dentro de una amplitud de más o menos 5%? ¿Cómo afectaría la duplicación de este número de mujeres a la precisión de su estimación? ¿Cómo se afectaría si el número se redujese a la mitad?

(ii) Posteriormente los investigadores continúan este proceso aplicándolo a todas las hipótesis escogidas para determinar el tamaño mínimo de la muestra necesario para demostrar las hipótesis. Habiendo decidido de qué manera se tomará la muestra, es preciso determinar su tamaño. Para los fines de este ejercicio, determine el tamaño de la muestra necesario para demostrar una de las cinco hipótesis señaladas por los investigadores. Esta hipótesis se relaciona con la región de la Montaña. Se cree que la tasa de mortalidad perinatal es de alrededor de 100 por 1.000 en zonas remotas. ¿De qué tamaño necesitaría ser la muestra de nacimientos para que el estudio pueda:

- a) calcular la proporción de niños que mueren durante el periodo perinatal, con un error de

precisión no mayor del 10 por mil, con un nivel de confianza del 95%; y

b) tener un poder del 95%, con un error tipo I de  $= 0.05$ , para detectar una diferencia de 50 por mil en la tasa de mortalidad perinatal de quienes nacieron en las zonas urbanas de la región de la Montaña, comparada con la de quienes nacieron en las zonas remotas, donde se supone que la tasa es de 100 por 1.000?

## Sección 5.5 Preparativos para el análisis de la información

### Resumen

Antes de realizar el análisis de los datos es preciso hacer planes detallados sobre la manera en que se resumirán, presentarán, (cuadros y gráficos) y analizarán. Antes de la realización del estudio principal, se considera pertinente analizar una pequeña muestra de los datos (estudio piloto) para evaluar características, calidad y formato de la información.

#### 5.5.1 Introducción

El proceso de formular las hipótesis, definir los objetivos de la investigación, y seleccionar la población objeto de estudio, el marco de muestreo y el método de muestreo, obliga a identificar una lista de variables, que son las características de personas y/o grupos de la población sobre las que es necesario recolectar información durante el estudio. Mediante el análisis de esas variables se podrá verificar cada hipótesis, es decir, se podrá decidir si la hipótesis es verdadera o falsa.

Hay que examinar ahora la manera en que se efectuará el análisis de los datos que se van a recolectar, a fin de asegurar que sus resultados adquieran una forma concluyente y útil que sirva de base para el diseño de una estrategia de intervención.

En este nivel de la investigación es de gran importancia que haya una estrecha colaboración entre los estadísticos y los expertos en computa-

ción. Cuando ya se ha recolectado todos los datos muchas veces es demasiado tarde para cambiar la forma en que fueron clasificados, agrupados o descritos. Con gran frecuencia los equipos de investigación solicitan estadísticos y/o epidemiólogos cuando los datos ya están recolectados y les preguntan: ¿qué podemos deducir de estos datos? Puede ahorrarse molestias, tiempo y dinero si se presta atención suficiente a la preparación de lo que se llama el plan de análisis. A continuación se describirá la forma en que se agruparán, tabularán y analizarán los datos. También se incluye una indicación de los métodos estadísticos que se usarán a fin de comprobar la validez de los resultados y de cuantificar las asociaciones.

Debe tratar de responderse a las siguientes preguntas:

- ¿cuál es la información que se necesita para probar cada hipótesis?
- ¿cuál es la mejor manera de recolectar los datos a fin de proporcionar esa información?
- ¿qué tabulaciones se necesitan?

#### 5.5.2 Tabulaciones de la relación entre los factores de riesgo y los daños a la salud estudiados

Una tabulación es un conjunto de datos dispuestos en hileras y columnas. Las tabulaciones se utilizan para presentar la frecuencia con que determinado acontecimiento ocurre en diferentes categorías o subdivisiones de una variable. Casi cualquier información cuantitativa puede organizarse en forma tabular. Con la ayuda de estas tabulaciones se preparan las gráficas y cuadros. Debe recordarse que los cuadros deberán explicarse por sí solos:

1. El título deberá ser claro y conciso. Debe responder a las preguntas de quién, qué, cuándo y dónde.
2. Cada hilera y cada columna deberán estar rotuladas con claridad y en forma concisa.

Debe indicarse las unidades específicas de medida para los datos. Las columnas deben estar separadas por líneas verticales.

3. Debe indicarse los totales de las hileras y las columnas.
4. Los códigos, abreviaturas o símbolos deberán explicarse en detalle en una nota al pie de página.

Si se refieren a la hipótesis de investigación basada en el enfoque de riesgo, se requiere la inclusión en el título de los siguientes tipos de información relacionada con los eventos a estudiar: la descripción del problema de salud prioritario desde el punto de vista de quién tiene el problema, dónde vive, cuándo ocurre el problema (hora, lugar, persona). Esto podría expresarse en términos de causa y mortalidad específica por edad, o tasas y cifras absolutas de morbilidad por zona geográfica, y rural o urbana. Esta es la descripción general. Pero se requiere ser específicos: decidir exactamente cuáles son los problemas que se van a describir y cómo se hará esa descripción. Si el problema es descrito como mortalidad perinatal, puede decidirse que hay que observar la mortalidad perinatal por edad de la madre, paridad o zona de origen. Un cuadro modelo puede ser similar al cuadro 5.

Este tipo de cuadro nos permite visualizar los números absolutos, y la *incidencia*. El factor de riesgo implícito en la mencionada hipótesis es la edad. Deberían prepararse cuadros similares para cada factor de riesgo con relación al daño a la salud estudiado. Dichas tabulaciones muestran cuán frecuente es el daño y cuán prevalente el factor de riesgo. Permiten analizar la relación entre determinado factor de riesgo y la presencia del daño de interés.

No todos los factores de riesgo son variables continuas, como la edad. La presencia de agua potable en la unidad familiar en que vive la mujer es una variable discreta, dicotómica. De ahí, que el cuadro modelo 6, similar al cuadro 5, sea mucho más sencillo.

Las variables dicotómicas permiten identificar los factores de riesgo en forma directa. Así, si se deduce que las mujeres sin abastecimiento de agua potable están expuestas a un riesgo mayor de sufrir una muerte perinatal, se concluiría que la ausencia de agua potable es un factor de riesgo.

### 5.5.3 Tabulaciones necesarias para dicotomizar las variables continuas.

En el caso de las variables continuas el método no es tan sencillo. El riesgo es una cuestión de grado. No es posible afirmar categóricamente que la edad por debajo de los 18 años es un factor de riesgo. Podría ser edad por debajo de los 16 o incluso edad por debajo de los 19. Se debe escoger el punto que a lo largo de la continuidad del riesgo servirá más eficazmente para definir un factor de riesgo. De las tasas de mortalidad perinatal específicas para la edad materna se puede sospechar que el riesgo aumenta por debajo de los 20 años, pero el cuadro 5 quizá no sea suficiente para permitir tomar una decisión práctica. Hay que hacer un desglose más detallado para las edades más jóvenes.

El cuadro 5, entonces, se transformaría en el cuadro 7.

Cuadro 5

Número de nacimientos y de muertes perinatales y tasa de mortalidad perinatal por edad de la madre. Zonas rurales de Ficticia, 1982

Edad de la madre (años)	Número total de nacimientos	Número de muertes perinatales	Tasa de mortalidad perinatal (por 1.000 nacimientos)
14 o menos			
15 - 19			
20 - 24			
25 - 29			
30 - 34			
35 o más			
TOTAL			

Cuadro 6

Número de nacimientos, número de muertes perinatales y tasa de mortalidad perinatal por presencia o ausencia de abastecimiento de agua. Zonas rurales de Ficticia, 1982

Abastecimiento de agua potable	Número total de nacimientos	Número de muertes perinatales	Tasa de mortalidad perinatal (por 1.000 nacimientos)
Presente			
Ausente			
TOTAL			

Cuadro 7

Número de nacimientos y de muertes perinatales y tasa de mortalidad perinatal por edad de la madre. Zonas rurales de Ficticia, 1982

Edad de la madre (años)	Número total de nacimientos	Número de muertes perinatales	Tasa de mortalidad perinatal (por 1.000 nacimientos)
12			
13			
14			
15			
16			
17			
18			
19			
20-24			
25-29			
30-34			
35 ó más			
TOTAL			

#### 5.5.4 Tabulaciones que demuestren la independencia de los factores de riesgo

Frecuentemente dos factores de riesgo están tan relacionados entre sí que, de hecho, miden el mismo riesgo. No están contribuyendo de forma independiente al daño. Por ejemplo, en muchos países la edad de la madre y el estado civil serían variables interdependientes: a medida que las mujeres tienen más edad, aumentan las probabilidades de que se casen. Si el estar soltera resulta ser un indicador significativo de riesgo, el factor real podría ser que la mujer es demasiado joven.

Por consiguiente al describir la influencia que tiene cada uno de los factores sobre el daño, se trata de saber si son independientes o no. Si son independientes, entonces el riesgo de tener un daño a la salud es mayor si ambos factores se hallan presentes. Si no son independientes, los dos factores están describiendo el mismo elemento de riesgo y el efecto de ambos combinados es menor del que si fuesen factores independientes.

La manera más sencilla de examinar el efecto de dos factores de riesgo para determinar si son independientes es hacer una tabulación doble. En el caso del estado civil y la edad, el modelo sería como el cuadro 8

Cuadro 8

Número de nacimientos y tasa de mortalidad perinatal (TMP) por edad y estado civil de la madre. Zonas rurales de Ficticia, 1982

Edad de la madre (años)	Mujeres solteras		Mujeres casadas	
	Nacimientos	TMP	Nacimientos	TMP
17 o menos				
18 - 19				
20 - 24				
25 - 29				
30 - 34				
35 o más				
TOTAL				

Cuadro 9

Tasa de mortalidad fetal, neonatal temprana y tardía y postneonatal específica por peso al nacer. Zonas rurales de Ficticia 1982

TASA	PESO AL NACER (g.)			
	1.500 ó menos	1.500-2.499	2.500-3.999	4.000 ó más
Mortalidad fetal*				
Mortalidad neonatal* temprana (menos de 7 días)				
Mortalidad neonatal tardía* (7 a 27 días)				
Mortalidad postneonatal*(28 A 365 días)				

\* por 1.000 nacidos vivos

### 5.5.5 Tabulaciones del peso al nacer como un indicador importante del nivel de salud

Una tabulación muy valiosa es la tasa de mortalidad específica por peso al nacer (cuadro 9). Proporciona información sobre la relación específica entre el peso al nacer y la edad al morir. Las muertes más evitables son las de lactantes con un peso al nacer mayor de 2.499 gramos y que han muerto en el período post-neonatal. Las muertes menos evitables son las de fetos de menos de 1.500 gramos. Los principales períodos de intervención pueden determinarse comparando las tasas esperadas u ópti-

mas, para cada una de las celdas del cuadro con las tasas de mortalidad observadas. De esta manera puede seleccionarse las estrategias de intervención prioritaria.

### 5.5.6 Tabulaciones de las categorías de riesgo

Una distribución de las categorías de riesgo en una población determinada identifica la frecuencia de las diferentes características y el punto de corte preseleccionado que deben usar los planificadores de salud en la estrategia de riesgo. El cuadro 10 presenta un esquema de cómo tabular esta información.

Cuadro 10

Tasas de mortalidad perinatal (TMP) por nivel de riesgo zonas rurales de Ficticia 1982

Nivel de riesgo de la madre	Porcentaje de nacimientos	Distribución de muertes (Porcentaje)	TMP
Bajo			
Mediano			
Alto			

En el cuadro 11 aparecen puntajes de riesgo de las madres, relacionados con la mortalidad perinatal por lugar de nacimiento. Una vez agrupados los puntajes de riesgos como bajos, medianos y altos, ello permitiría una mayor eficacia en la selección de pacientes, su referencia y el tipo de tratamiento. En resumen, esta información enfoca la distribución del riesgo y

evalúa los lugares de nacimiento. Se espera que las mujeres con elevado riesgo, que han sido atendidas por el médico en el hospital, tengan un resultado promedio mucho mejor desde el punto de vista probabilístico, que las mujeres de elevado riesgo que fueron atendidas por el personal auxiliar.

Cuadro 11

Muertes perinatales por lugar de nacimiento y personal de atención de acuerdo con el nivel de riesgo materno zonas rurales de Ficticia, 1982

Nivel de riesgo de la madre	Lugar de nacimiento			
	En la casa		En el hospital	
	Personal auxiliar	Enfermera-partera	Enfermera-partera	Médico
Alto				
Mediano				
Bajo				
Total				

Obviamente, no es fácil tabular todo contra todo, por lo que a medida que avanza el análisis tal vez haya que revisar las primeras tabulaciones seleccionadas. Es útil que todas las tabulaciones necesarias para el análisis sean discutidas suficientemente entre todos los miembros del grupo de estudio.

Puede prepararse un juego completo de tablas usando todas las variables expresadas como discretas; estas tablas de contingencia, muestran la forma en que se va a tabular cada una de las variables y debe incluirse en el protocolo de estudio. Una tabla de contingencia es el formato final en el cual se presentan los datos. Recuerde que las tablas debieran prepararse de tal modo que faciliten la demostración de las hipótesis y, de igual modo, el contenido de las tablas debe influir en la manera en que se procesen los datos. Cada tabla deberá ser independiente, o sea, el título deberá explicar exactamente su contenido. Cada columna e hilera deberá estar marcada claramente, y cada variable agrupada

de modo que se pueda obtener la información que proporciona la tabla. No deberá incluirse una tabla si no contribuye a demostrar la hipótesis o a un entendimiento general del problema.

### 5.5.7 Análisis preliminar de una muestra de datos

Tras haber examinado todos los pasos dados hasta ahora se tiene una idea clara de los objetivos del estudio, qué datos se están recolectando y por qué, y se ha llegado a un acuerdo respecto de las tabulaciones necesarias y de qué manera se van a analizar los datos.

Al analizar las muestras de los primeros datos recolectados, se puede probar y modificar el sistema de recolección y análisis de los datos. De esta manera, se observa la calidad y la dispersión de los valores que se obtendrán de los datos. Por ejemplo, en una encuesta de 3,000 madres, se examinaría las estadísticas básicas sobre el grupo de las primeras 300.

Dentro de esta submuestra se podría examinar:

- la integridad de los datos, su consistencia interna, y la proporción de datos que falte; y
- el rango y el tipo de información sobre diferentes variables, tales como:
  - edad de la madre;
  - nivel socioeconómico (clase social);
  - lugar de residencia;
  - tipo de atención de salud;
  - morbi-mortalidad del embarazo;
  - nacimientos únicos/múltiples;
  - nacimientos vivos/muertos;
  - presentación del feto;
  - incidencia de hipertensión; y
  - bajo peso al nacer.

### Ejercicio 5

- Para cada una de las cinco hipótesis de investigación específicas del ejercicio 2 prepare los cuadros de contingencia necesarios para relacionar los factores de riesgo con la mortalidad perinatal.

## Sección 5.6 Conceptos generales sobre manejo de los datos

### Resumen

Las variables que se incluyen en el estudio dependen de los objetivos de la investigación. Es necesario que se las defina con exactitud y que sean mensurables y asequibles. Hay que determinar también la escala de medición. Los datos pueden proceder de registros o entrevistas. Hay que preparar el cuestionario cuidadosamente y precodificarlo. Los entrevistadores deben recibir un adiestramiento para aplicar el cuestionario. Las fuentes de posibles errores son el bajo porcentaje de respuestas, los datos incorrectos y la variación del observador. Se debe contar con controles de consistencia interna de los datos y hay que probar previamente todas las partes del cuestionario.

Después de determinar el diseño del estudio y el tipo y tamaño de la muestra, debe definirse, con exactitud, la información que se desea reunir y el método seleccionado para hacerlo.

En la investigación sobre los problemas de salud materno-infantil en Ficticia se decidió que el daño a la salud prioritario es la mortalidad perinatal. Al enumerar las hipótesis concretas se señalaron diversos atributos de la madre como posibles factores de riesgo de ese daño a la salud. Resta ahora definir con mayor precisión tales factores, determinar la forma de medirlos y volver a examinar las razones de su inclusión en la investigación.

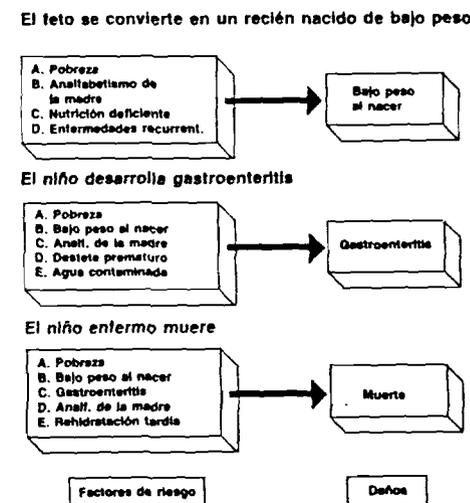
### 5.6.1 Definición de las variables del estudio

#### A) Medición de los daños a la salud

El daño a la salud escogido en Ficticia fue preciso. La mortalidad perinatal es un hecho relativamente fácil de definir y tiene una dimensión temporal bien delimitada. El niño muere en el útero, o durante los primeros siete días, si no es que sobrevive después de esa fecha. En cambio muchas otras mediciones de daños a la salud son más problemáticas. Por ejemplo, en el caso de la gastroenteritis, para que se pueda hablar de daño o resultado positivo o adverso hay que definir tanto la frecuencia de la manifestación como la severidad de la diarrea, así como la edad de los participantes. Luego se puede clasificar el hecho como daño a la salud o al estado óptimo de salud de la persona.

No siempre se descubre a primera vista una distinción neta entre los daños a la salud y los factores de riesgo. ¿Es la gastroenteritis un daño a la salud, o es un factor de riesgo para la muerte por gastroenteritis? Cuando se consideran esos problemas, es muy importante indicar con claridad lo que se entiende por un daño a la salud. La desnutrición es otro problema de salud que puede considerarse fácilmente bajo la perspectiva del enfoque de riesgo en la atención de salud materno-infantil, pero que presenta problemas similares de definición: puede ser un

FIGURA 8. FACTORES DE RIESGO Y LOS DAÑOS ASOCIADOS



daño a la salud, pero también es un factor de riesgo para la mortalidad en el preescolar. La figura 8 ejemplifica esta discusión.

#### B) Factores de riesgo

Antes de enumerar las variables que se usarán como factores de riesgo debe tenerse una idea clara de la razón por la que se desea incluir a cada uno de los factores de riesgo. ¿Qué se quiere indicar con la expresión factor de riesgo? Como se señala en el capítulo III, se le define como el atributo de una persona, o de un grupo de individuos, cuya presencia aumenta la probabilidad de experimentar un daño a la salud en comparación con aquellos que no lo presentan.

Los factores de riesgo pueden ser causas, o expresiones de un proceso. Si son solamente expresiones (pero no causas), se habla de los factores de riesgo como indicadores. Un aspecto importante es que se les puede observar o identificar antes de que se desarrolle el daño a la

salud: la enfermedad, la muerte, o la invalidez. Los factores de riesgo pueden ser aquellos que caracterizan a individuos, familias, grupos, comunidades o ambientes determinados. En numerosos estudios se ha demostrado que la primiparidad, la alta paridad, la gestación en edades reproductivas extremas, una muerte fetal previa y la desnutrición, son factores universales de riesgo perinatal. La presencia de uno de estos factores aumenta la probabilidad de que el embarazo tenga un desenlace insatisfactorio. La combinación de esos y otros factores en el mismo individuo aumenta, aún más, la probabilidad del daño a la salud. Por ejemplo, la multiparidad en madres que viven en la pobreza se asocia con un riesgo aumentado de muerte perinatal, mayor que el que acontece con la multiparidad por sí sola.

Algunos factores de riesgo pueden ser específicos de un hecho particular; por ejemplo, varios abortos provocados aumentan el riesgo de incompetencia cervical. Sin embargo, por lo

común un factor de riesgo (analfabetismo) aumenta la posibilidad de varios daños a la salud, (mortalidad) aunque quizá en distintas formas. Por ejemplo, la gran multiparidad se asocia con un mayor riesgo de situación transversal, hemo-

tragia posparto e hipertensión. A su vez, un daño está asociado con muchos factores de riesgo. Por ejemplo la hipertensión se asocia con multiparidad pero también con pobreza, con nutrición, con primiparidad, etc. (figura 9)

**Figura 9: Relaciones entre factores de riesgo y daños a la salud**

Un factor puede asociarse con	.....	tres daños
Analfabetismo	.....	Mortalidad perinatal Mortalidad materna Mortalidad infantil
Tres factores pueden asociarse con	.....	un daño
Hipertensión Edad materna avanzada Bajo peso al nacer		Mortalidad Perinatal

Para un daño específico, la elección inicial del factor o factores de riesgo a ser investigados (que deben incluirse en las hipótesis específicas) se basa, frecuentemente, en experiencias previas, en resultados de estudios pasados o en variables sospechosas de ser factores de riesgo así como en el contexto sociopolítico donde se desarrolla el estudio. La lista puede tomarse larga por lo que debe hacerse una selección de los factores más relacionados con el daño a la salud que quiere estudiarse.

La primera pregunta a formularse se relaciona con la posibilidad de medir el factor. ¿Es apropiada la metodología que permite identificar la presencia del factor en el contexto en el que se realiza la investigación? Un factor de riesgo que sólo puede identificarse con una prueba complicada de laboratorio no será de mucha utilidad en situaciones en que la mayoría de los partos es atendida por personal auxiliar. En algunas ocasiones se encuentra una medida

preliminar más práctica del factor de riesgo. La anemia, diagnosticada de ordinario por hemoglobímetro, puede ser sospechada también por la palidez, aunque este es un medio mucho más subjetivo. El valor predictivo del diagnóstico de anemia basado en la palidez será inferior al del diagnóstico que se apoya en el laboratorio. *El sistema de medición usado en una investigación sobre riesgo debe ser el mismo que el que se empleará en los programas de acción o control que se diseñan sobre la base de los resultados de la investigación.*

La segunda pregunta es si la información deseada puede obtenerse. Si se realiza una encuesta retrospectiva de nacimientos\*, es probable que no se disponga de datos relacionados con muchos acontecimientos prenatales. Es posible imaginarse que en Ficticia, donde sólo el

\* Es decir, incluyendo nacimientos que han ocurrido en el pasado.

40% de las mujeres campesinas recibió atención prenatal, no se podrá obtener información sobre presión sanguínea durante la gestación. El diseño de la investigación probablemente excluya una cantidad de variables que hubiera sido deseable usar. En algunas culturas las mujeres desconocen su edad exacta; en países en los que el aborto inducido es ilegal, es probable que las mujeres nieguen haberlo experimentado, etc.

La tercera pregunta que debe hacerse a esta altura es si se incluirán factores de riesgo para los que no se dispone aún de una *estrategia de intervención* o de un tratamiento. Es probable que deba eliminarse los factores de riesgo que no pueden ser modificados o compensados mediante acciones especiales de salud en el país en el que se está actuando; no obstante, aquellos factores que son reconocidos como importantes, pobreza o desnutrición, por ejemplo, deben siempre incluirse o medirse en el estudio, debido a su carácter de determinantes de salud, o por su naturaleza de posible variable distractora en la exploración de una asociación del daño con otros factores de riesgo.

Hay que *definir* con claridad todas las variables. Por ejemplo, ¿qué se entiende por paridad?: ¿número total de gestaciones previas?, ¿nacidos vivos?, ¿nacidos vivos y mortinatos? Hay que escoger la *definición*. La definición adecuada de las variables es un elemento importante para:

- mantener la calidad de la información reunida;
- reducir la variación del observador o sesgo; y
- permitir la replicación del estudio en otros lugares, en épocas diferentes y por investigadores diferentes.

Hay que determinar la *escala de medición* que se va a usar. Algunas variables son claramente discretas (por ejemplo, sexo); otras son ordinales, en el sentido que definen la situación del sujeto dentro de una serie (por ejemplo, educación: 1-5 años, 6-9 años, etc.; o primaria, secundaria, etc.). Finalmente otras son continuas (por ejemplo, la talla o el peso). Las varia-

bles continuas pueden dividirse en clases o categorías. Debe haber un número suficiente de categorías. La edad puede medirse en años y meses; años solamente; períodos de 5 a 10 años; o por ejemplo, menores o mayores de 40 años. Hay que analizar detenidamente esas categorías. Si se sospecha que la edad materna es un factor de riesgo, la unidad debería ser el año, porque no se conoce a qué edad aumenta el riesgo. En un estudio se emplearon agrupaciones de cinco años: 10-14, 15-19, 20-24 años, etc. En el análisis de los resultados, el grupo de investigadores no pudo establecer si el riesgo aumentaba en las madres que se encontraban por debajo de 19, 18, 17, 16 ó 15 años de edad. Si los datos se obtienen de grupos ya establecidos (edad 10-14), no se puede conocer posteriormente la edad exacta en la que se producen los cambios de riesgo. Si la información se recaba con su valor real, siempre puede ser agrupada posteriormente; este es el método más recomendable. Hay que indicar por escrito, en el protocolo del estudio, la definición y escala de medición que se adopta.

El proceso de definición de variables es un paso importante para determinar los métodos de recolección de datos. Puede presentar problemas al hacerlo detalladamente y, con frecuencia, provoca desacuerdos y falta de unanimidad entre los diferentes investigadores. Pero es un paso que hay que cumplir para evitar problemas insuperables en etapas posteriores del estudio.

## 5.6.2 Fuentes de datos y métodos de recolección de la información

Los métodos y procedimientos usados en la recolección de datos dependerán de las variables a estudiar, de las fuentes de información disponibles para cada una de las variables y del diseño del estudio. Las principales fuentes y métodos de recolección de información en el estudio epidemiológico con enfoque de riesgo en la atención maternoinfantil son:

- fuentes documentales (fichas de registro individual,

gráficos de crecimiento, registros de hospitales);

- entrevistas (con mujeres embarazadas o madres);
- observación (exámenes clínicos a mujeres embarazadas);
- exámenes de laboratorio.

Las mismas variables pueden obtenerse por diferentes procedimientos y un solo procedimiento puede ofrecer información sobre muchas variables de la misma persona. Comúnmente es provechoso combinar varios procedimientos. Por ejemplo, si se realiza un estudio de mujeres embarazadas, se seleccionará, examinará y entrevistará a una muestra de mujeres. Las fichas de registros de los centros de salud pueden ofrecer información complementaria sobre las mismas mujeres y los registros de hospitales pueden proporcionar datos más detallados, especialmente sobre daños a la salud como, por ejemplo, las muertes perinatales.

Por medio de encuestas a mujeres embarazadas y de las historias clínicas maternoinfantiles puede analizarse, por ejemplo, la diferencia entre aquellas mujeres que reciben control prenatal y aquellas que no asistieron a la consulta prenatal.

#### A) Fuentes documentales

Los registros de salud maternoinfantil, tanto de centros de salud como de hospitales, son fuentes adecuadas de datos. Pueden usarse siempre que exista buena cobertura de la población estudiada, que las fichas sean completas, que la información posea buena calidad y que ofrezcan la información necesaria en forma conveniente.

Por lo común los registros no ofrecen toda la información necesaria con suficiente validez y precisión. Una anotación de anemia en una historia clínica no es muy útil, porque es poco específica. Con frecuencia se comprueba que las historias clínicas son incompletas, y que cuando se registra la información, la caligrafía es tal que no se la puede descifrar.

Por lo tanto es fundamental examinar el sistema de registros de salud que se usa en el presente para determinar si puede emplearse en la investigación: ¿qué porcentaje de la población estudiada está registrado? ¿están completos los registros? ¿contienen toda la información necesaria? ¿corresponden las variables registradas a las que se necesitan para el estudio?

Si no se puede usar el sistema de información existente, podrían plantearse algunos cambios que lleven a:

- examinar la historia clínica maternoinfantil;
- dar nuevas instrucciones sobre el modo de completarla;
- adiestrar al personal relacionado con ella; y
- revisar el sistema de archivo y obtención de las historias clínicas.

En países que tienen un buen sistema de registro civil, los certificados de nacimiento y defunción pueden servir como fuentes de datos, especialmente si es posible ligar ambos certificados. Es probable que se registren datos maternos en el certificado de nacimiento y no en el de defunción del niño. Por lo tanto, si se comparan los certificados de nacimiento y defunción es posible estudiar los factores de riesgo presentes durante el embarazo que se asocian con la mortalidad infantil.

#### B) Entrevistas

Si no puede obtenerse la información necesaria a partir del sistema de registro actual a causa de alguna de las razones antes mencionadas, debe llevarse a cabo un estudio especial basado en entrevistas. Las variables seleccionadas y el plan de estudio determinarán el procedimiento exacto a seguir.

El formato de las entrevistas varía. Las preguntas pueden estar definidas con claridad y presentarse en una secuencia rigurosa, o la entrevista puede tener carácter de discusión del entrevistado con el entrevistador, quien actúa

como moderador. La entrevista menos estructurada es, en general, útil para encuestas rápidas, para obtener impresiones acerca de actitudes o creencias. Sin embargo, cuando se requiere información más exacta sobre acontecimientos pasados y hechos presentes, como en el estudio que nos interesa, la entrevista debe estar más estructurada. La técnica de la entrevista entraña muchas dificultades. El modo de preguntar, el tipo de pregunta y la relación establecida entre el entrevistador y el entrevistado producen resultados diferentes. Debe probarse cuidadosamente el cuestionario empleado, adiestrarse adecuadamente al entrevistador y elaborarse un manual con instrucciones muy claras para su llenado.

### 5.6.3 El cuestionario

El instrumento de recolección de datos de las entrevistas es un cuestionario o formulario que sirve para registrar toda la información sobre un individuo o sobre dos individuos relacionados (por ejemplo, madre/hijo). Habitualmente se completa durante una entrevista y/o con información complementaria de historias clínicas, por ejemplo.

El diseño del cuestionario es una tarea calificada. Si el cuestionario no es claro o las preguntas son ambiguas, los resultados del estudio pueden ser inservibles. Por ello es muy importante que se dedique mucha atención a su planificación y que antes de emplearlo se pruebe en condiciones prácticas (estudio piloto) similares a las que se espera encontrar durante la investigación. Luego del estudio piloto debe prepararse la versión final del cuestionario.

Los cuestionarios pueden ser precodificados, en cuyo caso los tipos de información, como estado civil y número de nacidos vivos, se especifican en el formulario y las posibilidades de respuestas se limitan a ciertas categorías predefinidas. Por lo común la codificación plantea el problema de la agrupación de los datos. Es importante recordar que cuando se registra los datos en forma agrupada, no pueden obten-

erse nuevamente en su valor exacto. En caso de duda es mejor conservar los datos en forma detallada y no agrupada.

La codificación puede realizarse al llenar el cuestionario o en una etapa posterior, pero su planificación no es exclusiva del experto en computación. Todos los miembros de los grupos de estudio deben convenir en los detalles del plan de codificación. Puede ser una tarea prolongada, pero siempre se obtiene un beneficio de ella.

Independientemente del momento en que se haga la codificación de los cuestionarios (precodificación o codificación posterior), la información terminará siendo codificada para su procesamiento. Siempre debe elaborarse un manual de instrucciones claro, con lenguaje sencillo, sobre la forma de llenar el cuestionario o de codificarlo posteriormente.

Si además de cuestionarios precodificados se usan otros instrumentos de investigación, debe elaborarse formularios uniformes de registro y probarseles previamente. Tales formularios de registro deben ser completos, claros, breves y coherentes. En el manual de instrucciones se uniforma el criterio para completar los formularios.

#### a) Números de identificación

Cada serie de datos (toda la información relacionada con una mujer o con la pareja madre/hijo) necesita un número que permita remitirse a la fuente original de los datos (la tarjeta de salud maternoinfantil, la persona, el certificado de nacimiento) para la verificación de la información.

Si se usan datos de fuentes diferentes téngase en cuenta un punto fundamental: *los datos de fuentes diferentes deben vincularse*. En el estudio perinatal, hay que relacionar la información sobre la mujer embarazada y el niño que resulte del embarazo, para que sea posible estudiar las consecuencias de las características de la madre en el infante. Debe elaborarse un sistema seguro de numeración que permita la comparación de las dos series de datos.

**b) Adiestramiento de los entrevistadores**

La capacitación garantiza la recolección uniforme de datos y ayuda a minimizar el sesgo del investigador. Una buena técnica de adiestramiento es el método de psicodrama. Se llevan a cabo entrevistas delante de todos los miembros del equipo de investigación, quienes anotan la entrevista en el formulario uniforme. Posteriormente se examina los formularios y se compara los datos registrados.

**c) Selección de los entrevistadores**

Se determina la clase de personal que hay que seleccionar, enumerando:

1. *Las características de la población estudiada:*  
edad, sexo, grupo étnico, aspectos culturales, etc.
2. *El tipo de información que debe recolectarse:* médica, personal, etc.
3. *La complejidad de la información que debe recolectarse* (necesidad de aptitud para la interpretación de las respuestas;  
determinación del nivel de educación necesaria, sin sobreestimar ni subestimar la tarea).
4. *La duración del estudio:* (necesidad de contratar personal capaz de mantener interés y motivación en tareas repetitivas y potencialmente aburridas a largo plazo).

Para el adiestramiento de entrevistadores así como para la evaluación de su rendimiento, se requiere completar en una etapa previa una documentación concreta sobre los procedimientos de recolección de datos de la investigación. Hay que simular las tareas, ensayándolas en una situación real, a fin de identificar adecuadamente áreas problemáticas. Una buena técnica es la de disponer que todos los investigadores

realicen cada fase de la recolección y ordenación de datos, inclusive la redacción y codificación.

**d) Elaboración de los planes para las entrevistas y los cuestionarios**

Los investigadores deben consultar las referencias clásicas acerca de diseño y prueba de cuestionarios y sobre planes de recolección de información.

Los elementos del diseño y contenido de los planes de recolección de información y de los cuestionarios que comúnmente tienen especial aplicabilidad e importancia para la recolección de datos son:

- la protección de la confidencialidad, conservando alguna forma de identificación individual;
- el tipo de preguntas para la clasificación de los entrevistados;
- la secuencia de las preguntas;
- la redacción de las preguntas;
- la eliminación de frases y términos confusos y de metáforas;
- la formulación de preguntas específicas, no generales;
- la elección de la forma abierta o cerrada para las preguntas; y
- la precodificación de las preguntas.

Esos y otros factores deben estudiarse en detalle, idealmente con la ayuda de un texto especializado sobre planificación de cuestionarios. En principio, la tarea conviene realizarla con una persona competente y experta en diseño y prueba de cuestionarios.

**5.6.4 Exactitud, validez, confiabilidad y sesgos en la recolección de datos**

En la planificación de un estudio epidemiológico el investigador ha de considerar la calidad de la información que se obtiene con los procedimientos propuestos. A continuación se discuten los principales criterios metodológicos em-

pleados para evaluar la calidad de la información obtenida.

#### A) Exactitud

Debe seleccionarse procedimientos de recolección de información que permitan la obtención de datos exactos. La finalidad es obtener información que se acerque lo más posible a la realidad. La exactitud es el indicador que da respuesta a preguntas tales como:

- ¿en qué medida, en un cuestionario contestado por las personas que se quiere estudiar, obtiene un valor de la edad igual o cercanamente aproximado a la edad promedio de una muestra de los mismos individuos?;
- ¿en qué medida, un análisis de sangre determina el número verdadero de mujeres anémicas (según una definición preestablecida) durante el primer trimestre de embarazo?

La exactitud depende, a su vez, de otras dos características de los datos: validez y confiabilidad.

#### B) Validez

La validez de una técnica de medición (ej. análisis de sangre, cuestionario, etc.) es su capacidad de medir lo que el investigador quiere estudiar. La validez de las técnicas de medición puede definirse con relativa facilidad, por ejemplo:

- el peso al nacer es una medida válida de nacimiento prematuro;
- la auscultación de sonidos de Korotkov mediante un esfigmo-manómetro y un estetoscopio es una medida válida de la presión sanguínea.

Estos dos enunciados pueden tener una aceptación general, pero es posible cuestionar su *validez* y el grado en que reflejan la realidad comparados con otra forma fiel de medición. Es probable que se quiera saber, en los ejemplos mencionados, cuán buenas fueron esas medidas:

- cuando se comparan con una evaluación pediátrica del neonato, o
- con un registro de presión sanguínea intra-arterial.

Si se pretende medir actitudes o estilos de vida (dieta, hábito de fumar, alcoholismo) puede resultar más difícil encontrar técnicas válidas de medición. En los tipos de estudios considerados (relación entre factores de riesgo sociales y médicos, y desenlace del embarazo), el investigador debe considerar también la validez de las medidas en términos de cuán *importante* es una característica determinada para la predicción de un daño a la salud. Los textos de epidemiología y de métodos de encuestas en ciencias sociales ofrecen material de lectura adicional. En ellos suelen emplearse distintas definiciones de validez, entre las que se cuentan con más frecuencia las dos que se mencionan en seguida.

##### a) Validez nominal

Es el término que se emplea para describir la relación que existe entre una técnica de medición y la variable a ser medida. Por ejemplo, se puede afirmar (basándose en el conocimiento de los procedimientos fisiológicos y bioquímicos) que un nivel de glucosa más elevado en la sangre a la mitad del embarazo señala la presencia de diabetes gestacional. Un análisis de glucosa en orina tendrá menor validez nominal por razones clínicas y fisiológicas; una búsqueda de síntomas de diabetes, tendrá menor validez aún. Es evidente que el investigador debe considerar la validez nominal, recurriendo a fuentes actualizadas, informes publicados e intercambios con expertos en el tema.

##### b) Validez de criterio

Representa el proceso de relacionar la prueba o procedimiento de medición con otras medidas de la característica estudiada. En el caso de un problema médico que requiera el uso de pruebas para el diagnóstico diferencial (por ejemplo tuberculosis y radiografía del tórax), la validez del criterio puede determinarse con la evalua-

ción de los resultados de las pruebas usando otros procedimientos (por ejemplo, examen clínico, cultivo de esputo, o autopsia). Puede resultar difícil o imposible aplicar ese enfoque a algunos tipos de mediciones de actitudes y comportamiento. Pero le corresponde al investigador considerar los métodos para la validación de cada procedimiento de medida. Por ejemplo, para la información relacionada con actitudes y comportamiento, se puede optar por entrevistas adicionales que proporcionen un método de verificación de datos alternativo e independiente. En la sección dedicada a la validez de las pruebas de tamizado o de diagnóstico diferencial se analiza con más amplitud la validez de criterio.

### C) Confiabilidad o repetitividad

Estos son términos empleados para describir la capacidad de la prueba de dar resultados similares cuando se la aplica al mismo individuo bajo las mismas condiciones en ocasiones diferentes. La confiabilidad puede verse afectada por los dos tipos de fenómenos que se describen a continuación.

#### a) Confiabilidad en los instrumentos de medición

Para que la técnica de medición sea provechosa en la investigación, la *confiabilidad* debe alcanzar un nivel aceptable. Por ejemplo, la confiabilidad o repetitividad es evidentemente baja si se emplea el instrumento unificado de recolección de datos y se descubre que los individuos se clasifican en una ocasión en un grupo y en la segunda oportunidad en el otro. Por otro lado la confiabilidad de una técnica es alta cuando en mediciones independientes un mismo individuo es clasificado consistentemente en el mismo grupo.

#### b) Confiabilidad en el observador

En los estudios epidemiológicos las observaciones necesarias requieren la participación de un gran número de indivi-

duos. Es decir, más de un observador participa en la recolección de datos. En el curso del estudio, los investigadores deben identificar las variaciones que tenga el observador al usar la técnica o instrumentos de medición y las diferencias que existen en las mediciones entre los observadores. Los tipos de variaciones introducidos por el observador son:

- 1) variación o sesgos en un mismo observador (intra)
- 2) variaciones entre observadores (inter)

Comúnmente el sesgo en un mismo observador es de carácter aleatorio y no sigue ninguna pauta particular. Puede indicar falta de pericia, de capacitación o de claridad en las instrucciones sobre el uso de los instrumentos de medición. Por lo tanto, al mejorar todos esos aspectos se puede reducir su efecto.

La *variación o sesgo entre observadores* es importante cuando en el estudio participa más de un investigador. La fuente de esa variación o sesgo incluye la originada por el observador y por el medio ambiente y las técnicas e instrumentos empleados. Si diferentes observadores examinan diferentes subgrupos (por ejemplo, casos y controles), un elemento de amplia variación entre observadores será sumamente grave. Este problema puede reducirse asegurando una asignación al azar de los individuos de diferentes subgrupos de estudio a los distintos observadores juntamente con un entrenamiento adecuado de los observadores.

### D) Tipos de sesgos

Al momento de realizar una observación, puede producirse tres tipos de sesgos: sesgo aleatorio, sesgo sistemático y sesgo en el sujeto. Los dos primeros dependen del observador y/o sus instrumentos. El último depende del observado.

#### a) Variación o sesgo aleatorio

Mediante la repetición de mediciones en el mismo individuo (por ejemplo, presión sanguínea) puede minimizarse el efecto de la *variación aleatoria*. Desafortunadamente, de ordinario se cuenta con una sola oportunidad de obtener los datos en el estudio. El sesgo aleatorio puede reducirse también con explicaciones claras a las personas que serán medidas, con información adecuada de los observadores y con la uniformidad de los métodos. Si se amplía el tamaño de la muestra y se repiten las medidas en los mismos individuos es poco probable que la *variación aleatoria* tenga mucha importancia en los estudios epidemiológicos.

#### b) Sesgo sistemático

En comparación con los otros observadores, uno de ellos puede sobreestimar o subestimar, constantemente, el nivel de la variable bajo estudio. Este *sesgo sistemático* puede relacionarse con alguna característica peculiar del observador o con las condiciones (ambientales, temperatura, nivel de ruido, privacidad, etc.) en las que se realizan las observaciones, o con los instrumentos utilizados. Hay que identificar y, si es posible, eliminar el sesgo sistemático. Nótese que éste puede variar con el tiempo, por lo que es necesario un control periódico y regular durante el estudio.

#### c) Sesgo en el sujeto

Las variaciones en las mediciones pueden proceder de cambios en la condición del sujeto. Por ejemplo, los lípidos sanguíneos o la tensión arterial fluctúan durante el día; pueden ocurrir diferencias estacionales importantes en las variables y síntomas fisiológicos. Si se da esta particularidad debe tenerse en cuenta la hora en que los observadores realizan las medidas para el estudio.

Los problemas de salud pueden manifestarse de diferentes maneras. Es probable que los síntomas y signos no siempre se relacionen adecuadamente con la gravedad patológica. La *percepción individual* de un problema de salud puede ser muy diferente entre los participantes del estudio. Las variaciones o sesgos en la percepción individual pueden reflejarse tanto en la repetitividad como en la validez de una prueba. El investigador debe anticiparse a la presencia de *variación o sesgo subjetivo* y procurar establecer métodos de recolección de información que impidan posibles distorsiones por él originadas.

### E) Porcentaje de respuestas

¿Se obtuvo información sobre *todos* los individuos encuestados? ¿Se obtuvo *toda* la información sobre cada uno de esos individuos? Hay que eliminar a las personas no incluidas en la población estudiada: por ejemplo, mujeres de fuera del área del estudio, que han dado a luz en instituciones de salud que forman parte de la investigación. Particularmente en los estudios basados en poblaciones, debe excluirse sistemáticamente a aquellas personas que pertenecen a otras regiones.

Es probable que no aparezcan algunos individuos que deberían haber sido incluidos, por diversos motivos, por ejemplo:

- negativa a responder las preguntas;
- ausencia del sujeto durante la visita del entrevistador;
- el entrevistador no pudo contactar al sujeto por causa de la distancia; o
- pérdida de la ficha de registro.

Los individuos no incluidos podrían presentar características muy diferentes de aquellos incluidos (las mujeres que usan el sistema de atención de salud materno-infantil son, en general muy diferentes de las que no lo usan; pueden ser más alfabetizadas, con mayores ingresos, vivir cerca del centro de salud, tener menos hijos, etc.). Por lo tanto, debe tomarse alguna

medida para solucionar la situación de no respuesta. El primer paso consiste en determinar el grado de no respuesta y las razones para no responder. Puede ser de importancia comparar los que respondieron con aquellos que no lo hicieron en términos de algunos variables disponibles (edad, sexo, paridad, etc.). El siguiente paso, consiste en procurar decididamente tomar contacto en una segunda fase con todos aquellos que no fueron incluidos en la primera. Si esta segunda fase resulta costosa o impráctica puede intentarse la elaboración de una muestra de aquellos que no respondieron. Sólo como recurso final debe sustituirse a quienes no pudieron ser entrevistados.

### 5.6.5 Examen preliminar del instrumento para la recolección de la información

Todos los métodos y procedimientos de recolección de datos deben someterse a prueba en el ambiente de estudio. Cuando se preparan los documentos y los planes de estudio, debe asignarse tiempo y recursos suficientes para su examen preliminar. Debe hacerse referencia a los diagramas y decisiones originales; se requiere programar estudios operativos para determinar si todos los aspectos logísticos de la investigación son viables, aceptables, eficientes y eficaces. En esta etapa hay que prestar atención especial a la *variación o sesgo del observador*.

El examen preliminar del instrumento debe responder a las siguientes preguntas: ¿Se entendió bien el cuestionario? ¿Es práctico? ¿Se obtiene la información que se requiere? ¿Es correcto el esquema de codificación?

Debe evaluarse los resultados del examen preliminar de los instrumentos y propiciarse modificaciones. Puede ser necesario, incluso, una prueba ulterior. La fase de prueba puede incluir verificaciones *ad hoc*, limitadas a algunos aspectos de los métodos e instrumentos de recolección de la información.

Para ellos es muy útil el contacto con representantes de la comunidad y del sistema de atención (por ejemplo, maestros, líderes, cen-

tros de salud, hospitales, administradores locales, personal de salud).

Los procedimientos deben explicarse y discutirse con diferentes grupos, lo que puede dar una mayor practicabilidad y aceptabilidad al estudio; lo mismo se aplica a los representantes de la comunidad.

### 5.6.6 Manejo, procesamiento y control de los datos

#### A) Consideraciones generales

El registro de datos de cada individuo estudiado es el resultado de diversas etapas de la investigación. Por ejemplo, selección de las preguntas, definición de variables que deben recolectarse, diseño de plan de tabulación, elección de métodos de recolección de datos. El registro puede ofrecer información procedente de una o varias fuentes de datos. Estos pueden ser de una clase (por ejemplo, médicos) o de varios tipos (médicos, sociales, demográficos).

A continuación se presentan algunos aspectos para tener en cuenta en la identificación de registros, codificación y procesamiento de datos y el análisis preliminar de los resultados.

#### B) Identificación de registros

Es muy importante en cualquier estudio, la identificación de individuos mediante sistemas fidedignos. Esos sistemas son necesarios para asegurar que los registros estén: i) *disponibles para identificar individuos y grupos concretos*; ii) *clasificados* en una forma predeterminada, y iii) *enlazados*, a fin de que los diferentes registros del mismo individuo puedan unificarse.

Antes de describir algunas etapas importantes del manejo de los datos recolectados debe insistirse en que es absolutamente imprescindible tener un método apropiado para archivar y recuperar las hojas de recolección de datos durante el estudio.

##### a) Revisión

La revisión de los datos recolectados puede realizarse en varias etapas del estudio, por ejemplo:

- *sobre el terreno*, durante etapas claves de la recolección de datos;
- *en la sede del estudio*, a medida que se reciben y cotejan las hojas de datos; y
- *en las fases finales* de la preparación de datos.

La revisión y la codificación de datos pueden realizarse simultáneamente, pero son tareas diferentes. Muchas tareas importantes de revisión deben llevarse a cabo antes de las etapas finales de preparación y ordenación de los datos. Aunque la disponibilidad de computación permita realizar algunas tareas de revisión en la etapa preliminar del análisis, debe hacerse todo el esfuerzo de editar la información durante la etapa de recolección. Esto permitirá volver a la fuente de los datos inmediatamente si esto fuera necesario y completar los ítem faltantes.

En la planificación del estudio, los investigadores deben indicar con claridad las etapas en las que se procederá a la revisión, por ejemplo:

- al final de la sesión de recolección de datos;
- al final del día de la recolección de datos;
- al concluir una fase del estudio.

En los estudios de pequeña escala, la revisión puede limitarse a una sola etapa de la recolección y preparación de datos.

Las tareas de revisión incluyen las verificaciones de:

#### 1. Exactitud:

Se refiere a errores tales como anotación de la edad (por ejemplo, 60 años para edad materna, o años en lugar de meses para la edad del lactante), ingreso declarado que no corresponde a la ocupación, tipo de vivienda o nivel de vida. Las verificaciones de esa clase pueden resultar difíciles porque en la realidad los datos

aparentemente incorrectos pueden ser correctos.

#### 2. Legibilidad:

Las anotaciones en las hojas de recolección de datos deben ser legibles. Los codificadores no deben confundirse por las anotaciones en los formularios ni verse obligados a adivinar la información correcta.

#### 3. Integridad:

En general, el diseño de las hojas de recolección de datos debe hacerse de tal forma que exista un tipo de anotación para cada rubro. Los asistentes de investigación pueden remediar deficiencias de la información si realizan verificaciones de integridad poco después de la recolección de datos. Por supuesto es probable que el punto omitido no se aplique al individuo entrevistado, pero esa posibilidad deberá tenerse en cuenta en el momento de la preparación de la hoja de datos.

#### 4. Consistencia interna:

Se trata de descubrir anotaciones contradictorias o incompatibles en términos científicos. En algunos casos, el tipo de verificaciones se asemejará a la realizada para determinar la exactitud. Las verificaciones de consistencia pueden predeterminarse en las primeras fases de la elaboración del estudio. Suponen un examen sistemático de cada rubro de datos y su relación con otros rubros importantes de la información. Por ejemplo, una sección de la hoja de recolección de datos puede describir el número de embarazos previos; la suma de los abortos, niños nacidos vivos y mortinatos no puede ser superior al número indicado de embarazos.

La computadora puede ser programada para realizar la revisión. El programador debe asegurar que el código apropiado se encuentre en la columna correspondiente.

Por ejemplo, si los únicos códigos disponibles para la columna de identificación de sexo son 1 y 2 (para hombre y mujer) la computadora rechazará la entrada de 3, 4, 5, 6, etc. Si se especifica el alcance de un resultado de laboratorio, la computadora rechazará los valores que se hallan fuera de ese rango.

### C) Codificación

La codificación es la transformación de la información contenida en el cuestionario en una forma que puede usarse en la ordenación de datos. En la codificación se asigna un número a cada respuesta posible, de manera que sea posible transferirla a registros de ordenación de datos. La codificación puede realizarse en el cuestionario mismo o en una hoja de codificación separada. De ahí puede transferirse la información a las tarjetas de ordenación de datos, perforándose en las partes correspondientes, o pueden ingresarse directamente los datos en la computadora. A continuación se describe un proceso de codificación para introducir datos a la computadora, aunque tales principios también son aplicables a otras formas de procesamiento.

El plan de codificación debe elaborarse y ultimarse en las primeras fases del estudio; pero es probable que se requieran modificaciones, e inclusive una recodificación fundamental, en etapas ulteriores.

#### a) Códigos simples

Los códigos son números asignados a rubros individuales y a grupos de datos. Los mismos códigos se aplican a cada individuo. Puede analizarse el método de codificación empleando tarjetas de 80 columnas. Cada casilla de codificación lleva una numeración según la columna correspondiente. Las casillas se dividen en grupos, llamados *campos de datos*. El campo de datos para nacimiento incluye seis columnas:

#### DDMMAA

El campo de datos de duración del embarazo, en semanas, incluye dos columnas

#### SS

Comúnmente se procura evitar el uso del 0 como código; los códigos 1 a 9 son de uso corriente, reservándose el 8 para "no aplicable" y el 9 para "desconocido". Los números de códigos se anotan comenzando por la derecha, no por la casilla de la mano izquierda. Si hay casillas vacías en el campo de datos se las llena con ceros. Por ejemplo, en la toma de presión sanguínea es posible tener un campo de datos de tres cifras, pero una anotación típica de la presión arterial diastólica puede ser:

#### 080

Si se desea, por ejemplo, describir la gran variedad en el uso del alcohol, se puede escoger un número para cada tipo de combinación de alcohol (por ejemplo, cerveza: 1, cerveza y vino: 2, vino y aguardiente: 3, etc.).

#### b) Códigos aditivos

Cuando existe la posibilidad de combinaciones múltiples de rubros, el uso de códigos aditivos es una forma económica en términos de cantidad de columnas usadas, y de la identificación de las combinaciones de varios rubros en un individuo. Por ejemplo: cerveza = 1; vino = 2; vino fortificado = 4; aguardiente = 8.

Ese código geométrico permite que todas las combinaciones se identifiquen en la forma de un puntaje, que ocupará sólo dos columnas. Por ejemplo, el código para el bebedor de cerveza y vino será 3, y para el de vino y aguardiente, 10. Algunos investigadores prefieren evitar los códigos aditivos porque pueden prestarse a error en las etapas de recolección o codificación de datos.

c) *Precodificación*

Las hojas de recolección de datos pueden *precodificarse*. En tal caso se especifica en el formulario el tipo de datos. Es probable que se necesite una hoja con instrucciones suplementarias para la codificación de referencia. Alternativamente se produce una clave de codificación, separada de la hoja de recolección de datos, para uso de los codificadores. La preparación de instrucciones de codificación es relativamente simple cuando se trata de datos concretos. Es tarea difícil cuando se trata de *preguntas de interpretación abierta*: exige vasta experiencia en el uso de métodos de investigación y requiere extensas notas e instrucciones. En general, la codificación plantea el problema de la agrupación de datos. Puede ser necesario y conveniente en el caso de la agrupación por edad, por ejemplo. Sin embargo, la agrupación significa que las hojas precodificadas de recolección de datos sólo tienen una serie de secciones arbitrarias y que los datos no podrán desagregarse posteriormente. En algunas variables importantes es preferible conservar los datos en forma detallada para la codificación. Esto será de suma importancia cuando sea necesario separar variables continuas por dicotomía, conservando al mismo tiempo la posibilidad de escoger diferentes límites de inclusión para una variable en particular.

d) *Formato de las hojas de datos*

En el lado izquierdo de la hoja precodificada de recolección de datos se ofrece una descripción de los datos a reunirse juntamente con las instrucciones para la codificación. En el lado derecho están las casillas de codificación con sus números de columnas.

Según se indicó, el formato de la hoja de datos debe corresponder al *plan* de tarjetas de 80 columnas. El número de

*campos de datos* no debe sobrepasar 80 columnas. En realidad, puede haber varias columnas desocupadas en cada hoja, que pueden servir más adelante para datos complementarios. Si el número de casillas de codificación en la hoja de datos excede 80 columnas, debe usarse dos o más tarjetas. Esa solución es perfectamente legítima pero debe tenerse en cuenta en la preparación del formulario y del método de codificación. Nótese que cada hoja debe llevar casillas de codificación para el número de estudio de *cada tarjeta*. Además, cada *tarjeta* debe *identificarse* con un número por tipo de tarjeta.

e) *Estructuras de codificación*

Es un término que designa el *método de codificación* usado para una determinada clase de datos. Comúnmente el método de codificación usado se aplica sólo a una sección particular de la hoja de datos. Por ejemplo, en la hoja de datos el marco de codificación para *estado civil* es necesariamente diferente del que corresponde a la *historia obstétrica*. En general, el marco de codificación se deducirá del tipo de datos a ser codificados.

D) *Procesamiento—Ordenación*

Son varias las técnicas disponibles para el manejo de datos. La elección de una de ellas depende de la cantidad de datos a obtenerse, de los fondos asignados para el análisis, y del tipo de medio de procesamiento y ordenación de datos.

a) *Tipos de tarjetas*

1. *Tarjetas de bordes perforados*: Este sistema permite asignar una tarjeta a un individuo de la encuesta. La tarjeta tiene una serie de agujeros en los bordes. Se los puede recortar con un *perforador manual* abriéndolos hacia la parte exterior de la fi

cha. Si se usan varias combinaciones de agujeros recortados y no recortados, el investigador puede registrar diferentes categorías de datos en la tarjeta. Esa clase de método manual tiene ventajas y desventajas potenciales. Puede ser muy provechoso si se lo usa cuando:

- el equipo de investigación no posee experiencia previa en el uso de máquina de ordenación automática de datos;
- no se dispone de personal competente para el manejo computarizado de datos;
- no se dispone de computadora;
- se cuenta con escasos recursos;
- el estudio se encuentra en sus etapas exploratorias.

2. **Tarjetas de 80 columnas:** Este sistema de manejo de datos tiene la ventaja de que se puede analizar las tarjetas mediante clasificadores mecánicos, es decir que se las puede usar para ingresar datos a la computadora. Actualmente en muchos centros se usan métodos de entrada directa de información en cintas magnéticas o "diskettes" y cada vez resulta más difícil obtener y mantener los equipos de perforación y análisis; éstos pueden resultar más costosos que los nuevos sistemas de microcomputadora.

#### b) *Perforación*

La columna puede perforarse en las diez posiciones marcadas (09) y en dos posiciones complementarias en la parte superior que no están marcadas. Además de las perforaciones de agujeros en las columnas, el equipo perforador de teclado imprime también la información en el borde superior de la ficha. Ese es un elemento útil, conocido también como inter-

pretación de la tarjeta, porque permite tomar con facilidad, manualmente, las fichas individuales. No todas las máquinas perforadoras de teclado disponen de esa capacidad de impresión.

#### c) *Verificación*

Se requiere un método para verificar la perforación de los agujeros apropiados correspondientes a la información codificada en las hojas de datos. Los errores de perforación serán relativamente insignificantes si hay operadores competentes de perforadoras de teclado (por ejemplo 2% de las columnas perforadas) pero pueden ser mucho más importantes en el caso de un operador inexperto.

#### d) *Clasificación*

Las tarjetas de 80 columnas pueden clasificarse mecánicamente para producir tabulaciones y tabulaciones combinadas. Algunos clasificadores de tarjetas son relativamente simples y sólo producen montones de tarjetas clasificadas sobre la base de una entrada en las columnas especificadas. Otros clasificadores cuentan el número de fichas con un agujero perforado en una columna especificada, y otros, aun más complejos, cuentan e imprimen números, obtienen totales y realizan tabulaciones combinadas.

#### e) *Computadoras y paquetes estadísticos*

Si se proyecta usar computadoras para el procesamiento de los datos puede ser necesario preparar un programa estadístico especial. En la actualidad, afortunadamente, existen programas generales ya preparados que permiten realizar los análisis básicos y muchos de los específicos sin necesidad de preparar un programa especial. Ejemplos de paquetes estadísticos disponibles en el mercado son los conocidos bajo las siglas de SPSS, SAS, MINI-

TAB. Para la utilización de estos paquetes estadísticos es necesario contar con computadoras de gran capacidad. Debe recordarse que pueden hacerse análisis básicos (pero importantes) en microcomputadoras o con la ayuda de calculadoras de escritorio. No es necesario tener una gran computadora para hacer análisis epidemiológicos de relevancia en salud pública.

### Ejercicio 6

Usted decidió realizar un estudio de corte transversal mediante entrevistas a un grupo de madres.

- (i) Formule preguntas para incluir en el cuestionario del estudio que le permitan obtener información de la relación entre la mortalidad perinatal y las siguientes cinco características de la madre:
  - a) lugar de residencia;
  - b) quién la asistió en el parto;
  - c) si tuvo toxemia o preeclampsia;
  - d) edad; y
  - e) número de hijos vivos.
- (ii) ¿Precodificará usted el cuestionario? En caso afirmativo, ¿cuáles variables y por qué? Indique los códigos que usaría en aquellas variables para las que precodificaría las preguntas.
- (iii) Bosqueje el plan de codificación que usará para las respuestas abiertas.
- (iv) Prepare una verificación de la consistencia interna para la variable "número de hijos vivos".
- (v) ¿Qué método aplicará para verificar la validez de sus datos?
- (vi) ¿Qué instrucciones impartirá a los verificadores de la información acerca de la variable "número de hijos vivos", en relación con la consistencia interna, la integridad de la información, el rango de los valores y la exactitud aproximada de los datos?

## Sección 5.7 Interpretación de datos

### Resumen

La interpretación de datos debe permitir la evaluación de los tres tipos posibles de asociación entre un factor de riesgo y un daño a la salud. El primer paso es utilizar una prueba estadística para evaluar la posibilidad de que la asociación encontrada se deba al azar. Si se puede demostrar una asociación estadística, el segundo paso consiste en determinar si la asociación podría explicarse por un sesgo o error en la selección de la población estudiada o de la información de ella obtenida. También debe explorarse cuidadosamente la existencia de posibles variables distractoras que puedan explicar la asociación hallada. Si se acepta que la asociación no es artificial la etapa final consiste en determinar si se encuentran presentes los criterios para establecer causalidad. Los criterios de uso corriente para evaluar la causalidad son la intensidad de la asociación, la especificidad, la consistencia, la relación temporal y la plausibilidad biológica.

### 5.7.1 Introducción

Después de haber identificado los principales daños a la salud, se puede emplear el enfoque de riesgo como una guía para la elección de las estrategias más eficaces para solucionar esos problemas. Algunas de esas estrategias procurarán identificar individuos o poblaciones con mayor exposición a los factores de riesgo a fin de que se les pueda ofrecer atención especial o compensatoria. Para esas estrategias se requiere el uso de factores de riesgo que tengan una alta capacidad predictiva, es decir, que permitan inferir el probable desenlace de aquellas personas que tienen el factor de riesgo comparado con el de aquellas que no lo tienen. Si se trata de comunidades, habrá también interés en conocer la prevalencia de los factores de riesgo. Con esa información se tendrá una idea de la medida en que la estrategia escogida puede reducir el problema de salud. Hay que distinguir entre las

acciones compensatorias de los factores de riesgo y las que modifican esos factores. Para que ese último tipo de estrategia sea eficaz debe existir una relación de causalidad. Algunos factores de riesgo son simples indicadores y la atención dentro del sistema de servicios de salud sólo puede compensar el aumento de riesgo producido por el factor.

En la determinación de asociaciones causales entre un factor de riesgo y un daño a la salud debe procurarse establecer tres consideraciones:

- A. Reconocer si el factor de riesgo y el daño a la salud que está siendo estudiado se asocia con una frecuencia mayor que la esperada por el azar, es decir, si existe una *asociación estadística*.
- B. Reconocer si la asociación estadística se debe a un sesgo o error de la información y, por lo tanto, es artificial o espuria.
- C. Determinar si los criterios usados se aplicaron científicamente para precisar si la *asociación es causal*.

### 5.7.2 Asociación estadística

La asociación estadística indica el grado de probabilidad que tiene un valor observado, o una serie de valores, o las diferencias entre dos valores de ser producto del azar. Para determinar esta asociación se emplean las *pruebas de significancia*.

La asociación estadística indica el grado de probabilidad que tiene un valor observado, o una serie de valores, o las diferencias entre dos valores de ser producto del azar. Para determinar esta asociación se emplean las *pruebas de significancia*.

Cada tipo de medición estadística tiene una prueba de significancia apropiada. Una de las más simples consiste en comparar los valores observados con sus errores estándar, como se hizo con la diferencia de las dos tasas de prevalencia.

Una prueba útil para el tipo de análisis necesario en la investigación sobre riesgo es la llama

mada prueba del "Chi cuadrado". Esa prueba permite rechazar o admitir la hipótesis nula, es decir establecer la falta de asociación entre los factores de riesgo y los daños a la salud. Si se analiza, por ejemplo, mujeres con o sin anemia, la prueba permite comparar el número de mujeres en cada categoría posible de ausencia o presencia del factor de riesgo y del daño a la salud, con el número esperado de mujeres si no hubiera asociación. La prueba indica la probabilidad de que los resultados se deban al azar, es decir, el grado de confianza con el que se puede rechazar o aceptar la hipótesis nula.

Nótese que esa asociación no indica nada respecto a la *causalidad*, sino, simplemente, enuncia que un determinado daño a la salud habitualmente está acompañado por una cierta característica (factor de riesgo). Esto puede resultar de utilidad para predecir un daño a la salud en un individuo cuando estudian. Si se analiza, por ejemplo, mujeres con o sin anemia, la prueba permite comparar el número de mujeres en cada categoría posible de ausencia o presencia del factor de riesgo y del daño a la salud, con el número esperado de mujeres si no hubiera asociación. No se requiere conocer el mecanismo por el cual el factor de riesgo produce el efecto, ni asumir forma alguna de relación causal.

Por ejemplo, en una tabla hipotética de 2 x 2, se tienen cuatro celdas, marcadas con las letras a, b, c y d, las cuales contienen el número de individuos que poseen dos características al mismo tiempo. La celda a, que es la de mayor interés para el estudio epidemiológico, expresa el número de individuos con el factor de riesgo que experimentaron el daño a la salud que se está estudiando (muerte perinatal).

Factor De Riesgo	Daño a La Salud		Total
	Presente	Ausente	
Presente (+)	a	b	a+b
Ausente (-)	c	d	c+d
Total	a+c	b+d	N*

$$*N = a + b + c + d$$

Se obtiene para cada celda de la tabla el valor esperado y luego se comparan ambos valores de la siguiente forma:

$$\frac{(\text{Observado} - \text{Esperado})^2}{\text{Esperado}}$$

En otras palabras, se calcula la diferencia entre el valor observado y el valor esperado, se eleva al cuadrado esta diferencia y se divide por el valor esperado. Para todas las celdas se hace lo mismo y se suma el resultado de cada una de las operaciones efectuadas:

$$\text{Chi cuadrado} = \text{la suma del } \frac{(O - E)^2}{E}$$

para cada celda.

El valor esperado se obtiene de la tabla de  $2 \times 2$ , es decir, el valor que se encontraría si los daños a la salud se distribuyeran proporcionalmente entre aquellos sujetos con el factor de riesgo presente y aquellos sin el factor de riesgo. El valor esperado para cada celda se obtiene de la siguiente manera:

$$E = \frac{(\text{total de las columnas}) \times (\text{total de las hileras})}{\text{total de casos (N)}}$$

En el caso de la celda "a" el valor esperado será:

$$\frac{(a+c) \times (a+b)}{N}$$

Después de obtener el valor del Chi cuadrado éste debe interpretarse usando la tabla que se ofrece más adelante, en la que se presentan algunos valores de la distribución del Chi cuadrado. En otros términos, usando el ejemplo de factores de riesgo y de daños, la tabla ofrece la distribución de valores entre factores de riesgo y daños, que se puede esperar hallar debido solamente al azar cuando en la realidad no existe ninguna asociación. Nuevamente, volviendo a la sección 5.1 sobre la presentación de las hipótesis, la distribución de valores del Chi cuadrado en la tabla es la que se espera encontrar si la hipótesis nula o de no asociación, no es falsa.

En este caso, la probabilidad de hallar el valor Chi cuadrado de 2,71 o más, es de aproximadamente 10% y la de hallar el valor de 3,84 o más es 5%. En otras palabras, la posibilidad de rechazar incorrectamente la hipótesis nula, de no asociación, cuando el Chi cuadrado es mayor que 3,84, es sólo 5%.

La selección del nivel de significancia depende de varios factores, además de los meramente matemáticos. Por ejemplo, la importancia médica puede determinar la selección de un nivel de significancia alto o bajo. Si se analiza las posibles consecuencias de un tratamiento clínico donde la muerte es probable, puede considerarse un valor de por ejemplo 5% de nivel de significancia; por otra parte, en la compara-

ción de un tratamiento nuevo con otro tradicional que tiene buenos resultados, es probable que se requiera alcanzar un nivel más bajo (por ejemplo 1%) antes de rechazar el tratamiento tradicional.

*Cuando el Chi cuadrado con un grado de libertad es mayor o igual que:*

1,64	0,20
2,71	0,10
3,84	0,05
6,63	0,01
10,83	0,001

Valores del Chi cuadrado para ser usados en una Tabla de 2 x 2:

**Ejemplo**

**Daño a la salud**

		Presente	Ausente	Total
Factor de Riesgo	Presente	150 (100)*	350 (400)	500
	Ausente	50 (100)	450 (400)	500
Total		200	800	1.000

\*Los valores esperados aparecen entre paréntesis ( ).

$$X^2 = \frac{(O - E)^2}{E} \text{ Sumando para las cuatros celdas}$$

$$X^2 = \frac{(150 - 100)^2}{100} + \frac{(350 - 400)^2}{400} + \frac{(50 - 100)^2}{100} + \frac{(450 - 400)^2}{400}$$

$$X^2 = \frac{(50)^2}{100} + \frac{(-50)^2}{400} + \frac{(-50)^2}{100} + \frac{(50)^2}{400}$$

$$X^2 = \frac{2.500}{100} + \frac{2.500}{400} + \frac{2.500}{100} + \frac{2.500}{400}$$

$$X^2 = 25,0 + 6,25 + 25,0 + 6,25$$

$$X^2 = 62,5$$

O = Observado    E = Esperado    X<sup>2</sup> = Chi cuadrado

En la tabla anterior se indica que este valor de Chi cuadrado es superior a 10,83 (el valor para el nivel 0,001 ó 0,1%); por consiguiente, se puede afirmar que la probabilidad de rechazar incorrectamente la hipótesis nula es de menos de 0,1% o 1 en 1.000.

La anotación usada para registrar ese evento es:

$$\text{Valor de "p" menos de } 0,001 \\ \text{p} < 0,001$$

El método anterior para llevar a cabo la prueba del Chi cuadrado en la tabla de  $2 \times 2$  resulta complicado. Lo que se ofrece a continuación es una simplificación, usando una fórmula en la que se sustituyen los valores a, b, c y d. A continuación se presenta dicho procedimiento simplificado usando el mismo ejemplo anterior.

mula en la que se sustituyen los valores a, b, c y d. A continuación se presenta dicho procedimiento simplificado usando el mismo ejemplo anterior.

Factor de Riesgo	Resultado		
	Negativo	Positivo	Total
Presente	150 (a)	350 (b)	500 (a + b)
Ausente	50 (c)	450 (d)	500 (c + d)
Total	200 (a + c)	800 (b + d)	1.000 (N)

La fórmula simplificada del Chi cuadrado es:

$$\text{Chi cuadrado} = \frac{(ad - bc)^2 N}{(a + b)(c + d)(a + c)(b + d)}$$

En el ejemplo usado:

$$\text{Chi cuadrado} = \frac{(150 \times 450 - 350 \times 50)^2 1.000}{500 \times 500 \times 200 \times 800} = 62,5$$

Se compara nuevamente el valor del Chi cuadrado con los que aparecen en la tabla de valores de  $\chi^2$  presentada previamente.

Repetiendo, el valor (62,5) es superior al nivel de significancia 0,001. Por consiguiente, se puede decir que es significativo en el nivel de 0,1%. Es decir, la posibilidad de rechazar incorrectamente la hipótesis nula es inferior a 1 por mil.

*Nótese que la prueba del Chi cuadrado no ofrece indicación sobre causalidad o sobre la magnitud de la asociación.*

### 5.7.3 Asociación artificial o espuria

Se tiene ahora una lista de factores relacionados estadísticamente con el daño estudiado. El paso siguiente consiste en verificar cada factor de riesgo en la lista a fin de determinar si la

asociación hallada se debe a sesgos en la elaboración o interpretación y, por consecuencia, si es artificial o espuria.

#### A) La asociación hallada es falsa o debida a sesgos

El equipo de investigación debe hacer preguntas acerca de una asociación observada y los datos en los que se basa a fin de determinar si dicha asociación es producto de algún sesgo, y sea en el modo en que se escogió la población estudiada, o en la forma en que se realizaron las mediciones u otros tipos de problemas de recolección de datos. Por ejemplo, si fueran varios los entrevistadores se podría sospechar un sesgo sistemático debido a que el modo de entrevistar a las personas o la forma de registrar las respuestas pudo originar una asociación *aparente* pero no causal.

Por supuesto, también podría estarse eliminando una asociación que en realidad sí existe. Para evitar esto, debe asegurarse que el diseño del estudio, los métodos de muestreo o los instrumentos y procedimientos de recolección de los datos se elaboren y usen en tal forma que reduzcan el riesgo de introducir un sesgo indeseado.

La búsqueda de errores o sesgos que expliquen un fenómeno se realiza, fundamentalmente, mediante el intercambio de opiniones entre los miembros del grupo acerca de las etapas en que pudo haberse cometido algún error. A continuación se exponen dos ejemplos en los que puede producirse un sesgo en la asociación.

### **Ejemplo 1**

Puede ocurrir un sesgo cuando la población estudiada se selecciona a sí misma, por ejemplo, si hay mayor probabilidad de que las mujeres enfermas se pongan en contacto con los servicios prenatales, mientras que las sanas no se preocupan por visitarlos. Si hay mayor probabilidad de que las mujeres enfermas tengan un daño a la salud, los resultados indicarán una asociación entre atención prenatal temprana y complicaciones del embarazo. Obviamente, el indicador "atención prenatal temprana" no puede usarse como factor de riesgo.

### **Ejemplo 2**

Hay una tendencia entre las mujeres de condición socioeconómica más elevada a no practicar la lactancia materna. En los registros se advierte que sus hijos tienen menos enfermedades diarreicas. Si se estudia la asociación de diarrea con lactancia materna, sin considerar la condición socioeconómica, podría llegar a encontrarse una asociación positiva entre lactancia materna y enfermedades diarreicas.

Obviamente aquellos factores de riesgo que tengan una asociación estadística producto de algunos de los sesgos anteriormente referidos, deben ser eliminados de la lista.

Un caso especial de sesgos es el que ocurre cuando se selecciona pacientes de los hospitales. Dos problemas pueden surgir con el uso de pacientes hospitalarios: a) su capacidad de recordar puede estar afectada (aumentada o sesgada) por la causa de su internación, b) aun en ausencia de una asociación entre el factor de riesgo y la enfermedad, los patrones de referencia de los distintos subgrupos de la población hacia el hospital pueden afectar o anular la asociación.

### **B) La asociación hallada se debe a variables distractoras**

Una variable bajo estudio puede estar aparentemente asociada con un daño a la salud, sin ser ésta una asociación causal. En tal caso la variable o factor de riesgo y el daño a la salud de interés están asociados con un segundo factor de riesgo que es la causa de la enfermedad. Esta variable o factores de riesgos que tienen una asociación real pero no causal con el daño, se llaman variables distractoras, la variable causal y el daño están relacionados entre sí, el uso de ambas características como factores de riesgo no añadirá mucho a la predictibilidad del daño a la salud que interesa.

### **Ejemplo 3**

Las cifras del cuadro 12 representan una tabulación cruzada de edad, orden de nacimiento y mortalidad perinatal.

Cuadro 12

Tasa de mortalidad perinatal (TMP) específica por edad de la madre y orden de nacimiento

Edad de la madre (años)	Orden de nacimiento			Total
	1	2-4	5 ó más	
17 o menos	110	110	-	110
18-25	48	47	48	47
26-34	50	50	50	50
35 o más	90	90	90	90
<b>Total</b>	<b>62</b>	<b>52</b>	<b>67</b>	<b>40</b>

Un análisis de las TMP totales para las diferentes edades de las madres y las TMP totales para los distintos grupos de orden de nacimiento indica que existe una relación aparente (asociación) entre las TMP y las edades extremas, así como los extremos en el orden de nacimiento. Pero cuando se observa las TMP a diferentes niveles de edad materna, se puede ver que no hay diferencia por orden de nacimiento en la mortalidad perinatal. Por el contrario cuando se varía la edad materna a distintos niveles en el orden de nacimiento se observa un efecto de la edad materna en la TMP. De ahí se puede deducir que la diferencia en la mortalidad perinatal observada (por el orden de nacimiento), es mediada por la edad materna y que no se obtendrá una predictibilidad adicional conociendo ambas variables. La elección de la variable a utilizarse dependerá también de la facilidad con que puede obtenerse la información.

#### 5.7.4 Asociación causal

Algunos factores de riesgo son sólo indicadores de problemas. En algunos casos la modificación del factor no producirá ningún efecto sobre el daño a la salud. Otros factores pueden contribuir directamente a ser la causa del problema y

esos son los que principalmente deben ser considerados.

El conocimiento del grado de causalidad ayudará a determinar cuáles son los factores de riesgo que se intentará reducir o eliminar prioritariamente mediante una intervención de salud u otros tipos de estrategias.

No existe un método matemático para establecer el grado de causalidad. La causalidad no puede probarse, sino sólo inferirse. Muchos son los elementos que pueden emplearse para inferir el grado de causalidad, y los más importantes de ellos son:

- A) Fuerza de la asociación.
- B) Especificidad de la asociación.
- C) Asociación temporal.
- D) Consistencia de la asociación.
- E) Credibilidad biológica (coherencia con el conocimiento existente).

#### A) Fuerza de la asociación

El riesgo relativo (que posteriormente se examinará) cuantifica el grado o fuerza de una

asociación. Además, se podría esperar una relación del tipo de dosis—respuesta, en la cual al aumentar el nivel de la exposición al factor se encontrarían mayores niveles de daños a la salud. En otras palabras, cuanto más fuman las madres, más niños de bajo peso se esperarían encontrar.

El riesgo relativo indica también la importancia de un factor o grupo de factores de riesgo en la causalidad de la enfermedad. Por ejemplo, un riesgo relativo de 8.0 probablemente indicará que el factor de riesgo estudiado es de suma importancia como causa de un problema de salud y, probablemente, su eliminación reducirá marcadamente el problema. Pero esa marcada disminución es posible sólo si la prevalencia del factor de riesgo es relativamente alta. La reducción de la mitad de un factor de riesgo muy raro influye muy poco en la situación global del problema, aunque es muy importante para los contados individuos vulnerables. Por ejemplo, las malformaciones congénitas aumentan el riesgo de bajo peso al nacer ( $RR = 5$ ). No obstante, dado que son relativamente raras, su eliminación va a tener un efecto pequeño en la incidencia total del bajo peso en una población. Por el contrario, fumar durante el embarazo también aumenta el riesgo de bajo peso ( $RR = 2$ ). Con todo, si el 20% de las mujeres fuma, la eliminación del hábito de fumar baja más la incidencia total de bajo peso que la eliminación de las malformaciones congénitas.

#### B) Especificidad de la asociación

Si se descubre que la asociación observada entre un factor de riesgo y un daño a la salud es específica para ese factor y para ese daño, se puede concluir que, muy probablemente, existe una relación causal. Si el factor de riesgo examinado se relaciona, por el contrario, con una gran cantidad de problemas de salud ello disminuye, aunque no necesariamente elimina, la posibilidad de una asociación causal. Algunos factores de riesgo como el fumar y la pobreza están asociados *causalmente* con una gran variedad de enfermedades.

#### C) Asociación temporal

Se requiere tener evidencia de que el factor de riesgo precede al daño a la salud que se estudia, para considerar la posibilidad de una asociación causal.

#### D) Consistencia de la asociación

En ésta y otras secciones del documento se consideran diversas maneras en que algunos errores o sesgos pueden presentarse en los estudios epidemiológicos y distorsionar los resultados. Por lo tanto, se requiere evidencias en el sentido de que las asociaciones observadas han sido o serán encontradas en otros estudios (valiéndose de los mismos métodos y criterios), en otras poblaciones, en otras áreas geográficas y que han sido realizadas por otros grupos de investigadores.

#### E) Credibilidad biológica

Es importante poder explicar los mecanismos de una asociación observada en términos de los conocimientos científicos actuales, es decir, la fisiología o patología del proceso o los factores psicológicos o sociológicos que están asociados. Debe aceptarse un margen de lagunas en el conocimiento, pero si se requiere formular nuevas hipótesis para explicar la asociación observada, puede disminuir la credibilidad de que la asociación sea causal. *No obstante, en salud pública no es imprescindible conocer el mecanismo de acción exacto para recomendar una intervención que está fundamentada en investigaciones epidemiológicas.*

### 5.7.5 Un método para la planificación de estrategias

Las estrategias de riesgo pueden dividirse en dos categorías generales: las orientadas al individuo y las dirigidas a la comunidad. Para las de la primera categoría, el trabajador de salud deberá evaluar el riesgo que corre un individuo de experimentar un daño a la salud cuando se lo compara con otros individuos de la población

que no estuvieron expuestos al factor de riesgo. Hay que determinar las características relacionadas con el individuo y el daño a la salud bajo estudio, a fin de predecir quién tiene mayor riesgo de experimentar dicho daño a la salud o enfermedad. Se supone, aunque no es necesario que así sea, que tales características están relacionadas con la causa. Los factores causales representan en realidad problemas concretos y se podrán intentar soluciones conocidas previamente. Otros serán sólo factores predictivos. El objetivo es identificar el mayor número posible de individuos que presentarán el daño a la salud (enfermedad) con el menor número de factores de riesgo.

Para la descripción de individuos con muchas características o para la comparación de un individuo con otros, se requiere alguna forma de

expresión que valore el efecto de un factor de riesgo sobre el daño o la enfermedad. Aquí se usará una medición que es el *riesgo relativo*.

Para las estrategias dirigidas a la comunidad, en oposición a las dirigidas al individuo, es necesario determinar las causas de un daño a la salud e identificar la repercusión de un factor de riesgo específico en la salud general de la comunidad. Esto se mide con el *riesgo atribuible en la población* mediante el cual se puede determinar qué estrategia de intervención tendrá más repercusión en la reducción del daño a la salud en la comunidad.

En las secciones 5.8 y 5.9 se exponen los métodos propuestos para el uso del riesgo relativo y del riesgo atribuible en la población, para la planificación de las estrategias individual y comunitaria.

### Ejercicio 7

La tabla siguiente muestra el número de nacidos vivos y de muertes perinatales de niños nacidos de madres que presentaron alguna(s) de las cinco características estudiadas.

En la última columna se da el valor del "Chi cuadrado" para algunas de las características, excepto la de "3 o más hijos vivos."

Característica	Número de nacimientos	Muerte perinatal		Valor del "Chi cuadrado"
		Sí	No	
Región montañosa	4.339	307	4.032	21
Personal auxiliar	22.280	1.448	20.832	69
Toxemia	1.636	388	1.248	1.067
Edad hasta 19 años*	9.291	778	8.513	174
3 ó más hijos vivos**	4.286	297	3.989	—
Total***	51.110	2.832	48.278	

\* Edad materna no registrada en 208 casos, incluyendo 10 casos de muerte perinatal.

\*\* "Número de hijos vivos" no registrados en 19.424 casos incluyendo 1.045 casos de muerte perinatal.

\*\*\* Población total

**Ejercicio 7 (continuación)**

- (i) Utilizando la información del ejercicio 5 complete las tablas de  $2 \times 2$ , como la que se presenta a continuación, para cada una de las cinco variables. Calcule el valor del Chi cuadrado para la variable de "3 ó más hijos vivos."

Edad	Muerte perinatal		Total
	Sí	No	
Hasta 19 años	778	8.513	9.291
Más de 19 años	2.044	39.567	41.611
Total	2.822	48.080	50.902

- (ii) Usando la fórmula simplificada presentada para la prueba del  $X^2$  (chi-cuadrado) determine la asociación estadística entre el número de hijos vivos ("2 hijos vivos o más" contra "menos de dos hijos vivos") y la mortalidad perinatal usando la siguiente información:

Número de hijos vivos		Muerte perinatal		Total
		sí	no	
2 o más	2 o más	586	9.878	10.464
	Menos de 2	1.201	20.021	21.222
	Total	1.787	29.899	31.686

Compare el resultado con el obtenido en el ejercicio (i) Comente las diferencias.

- (iii) Valiéndose de los datos adicionales recolectados (cuadros 23, 23A y 24 de Ficticia); ¿qué otras variables usaría para describir el problema perinatal en Ficticia?
- (iv) ¿Cuáles de esas variables describen mejor las características de las madres cuyos niños tienen un riesgo alto de morir en el periodo perinatal?

## Sección 5.8 La cuantificación del riesgo en individuos y la elaboración de estrategias de intervención

### Resumen

El riesgo relativo es una medida que cuantifica la fuerza de la asociación entre un daño a la salud y un factor de riesgo. Se le puede usar para determinar si los factores de riesgo son interdependientes; el riesgo relativo ajustado puede servir para describir con mayor precisión la fuerza de la asociación. Las técnicas de la prueba de tamizado dependen de: 1) *por qué* se realizan 2) *quién* las lleva a cabo 3) *cuándo* se efectúan 4) *dónde* se realizan y 5) *qué* tecnología se usa para el análisis. Hay varios métodos para elaborar un sistema de análisis. Varían según los cálculos matemáticos usados. La sensibilidad, especificidad y poder predictivo de una prueba positiva y negativa dependen de los cinco factores mencionados y del modelo matemático empleado.

### 5.8.1 Introducción

En la sección anterior se identificaron las características de los individuos que experimentaron un daño a la salud. El paso siguiente consiste en valerse de ese conocimiento para formular procedimientos que permitan identificar a otros individuos en condiciones similares. Posteriormente, basándose en el conocimiento de sus características se puede calcular la probabilidad de que ellos experimenten un daño a la salud en el futuro, y de esta forma poder tomar las medidas preventivas necesarias. Cuando se habla de individuos en riesgo se quiere decir con mayor riesgo porque en cierta medida toda la población está en riesgo. Esa noción de la distribución del riesgo en la población constituye la base de la discusión acerca de las pruebas de tamizado y de los límites de inclusión correspondientes.

Hasta ahora se trató cada factor de riesgo aisladamente. Sin embargo un mismo individuo

puede tener diversos factores de riesgo y éstos pueden ser sometidos a diferentes intervenciones.

Por lo tanto, la tarea consiste en vincular en forma práctica la información clínica, epidemiológica y de recursos al nivel del conocimiento, competencia y medios disponibles del investigador. Hay diversos métodos de enfocar ese problema; y algunos de ellos son complejos y altamente matemáticos. No obstante, existen métodos más sencillos que son suficientemente adecuados para la mayoría de las aplicaciones. Seguidamente se exponen los pasos básicos en el proceso de cálculo del riesgo relativo así como la derivación de un sistema de puntaje para la reasignación de recursos a los grupos de riesgo elevado.

### 5.8.2 El concepto de riesgo relativo—principios generales

Debe introducirse aquí el concepto de *riesgo relativo* a fin de ofrecer los antecedentes necesarios para las secciones siguientes. El riesgo relativo es una medida de la probabilidad que experimenten un daño a la salud (enfermedad) aquellos individuos con un determinado factor de riesgo en comparación con los que no tienen ese factor. No refleja el riesgo individual en la población que experimenta el evento—ese aspecto está representado en la tasa de incidencia; no indica la medida del cambio del riesgo de la población de experimentar el daño en caso de la eliminación del factor de riesgo—ese aspecto se deriva del riesgo atribuible en la población y se expone en otra sección.

#### A) Cálculo del riesgo relativo en estudios de cohortes

Si se entiende que:

$p_1$  = tasa de incidencia del daño a la salud de los individuos que tienen el factor de riesgo (expuestos).

$p_2$  = tasa de incidencia del daño a la salud de los individuos que no tienen factor de riesgo (no expuestos).

RR = Riesgo Relativo,

Entonces el cálculo consiste en una comparación directa de las tasas de incidencia en aquellos con el factor de riesgo y en los que no lo tienen:

$$RR = p_1/p_2$$

**B) Cálculo de riesgo relativo en casos y controles**

Puesto que la incidencia es desconocida en estudios retrospectivos, el método de *estimación* del riesgo relativo con datos provenientes de estudios de casos y controles es diferente. Simplemente, al tratarse de una tabla de  $2 \times 2$  se usa la razón de los productos cruzados de las cuatro celdas.

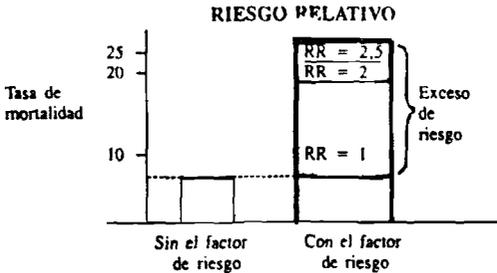
Estudio de casos y controles daño a la salud

		Presente (casos)	Ausente (controles)	Total
Factor de riesgo	Presente (Expuesto)	a	b	a + b
	Ausente (No expuesto)	c	d	c + d
TOTAL		a + c	b + d	a+b+c+d

La *estimación (e)* del riesgo relativo es la razón de los productos cruzados de las cuadro celdas (comúnmente conocido como razón de productos cruzados, o también llamado razón de disparidad), es decir:

$$RR (e) = \frac{a \times d}{b \times c}$$

*Nota importante:* Esta proporción es una aproximación al riesgo relativo, pero es muy cercana al valor verdadero cuando la frecuencia del daño a la salud (enfermedad) es baja, como en muchos estudios epidemiológicos.



## Ejemplo

Cálculo de riesgo relativo para estudios de cohortes y de casos y controles

## a) Estudio de cohorte

		Daño a la salud		TOTAL
		Presente (casos)	Ausente (controles)	
Factor de riesgo	Presente (Expuestos)	60	2.000	2.060
	Ausente (No-Expuestos)	40	8.000	8.040
	TOTAL	100	10.000	10.100

$$p_1 = 60/2.060 = 0,0291$$

$$p_2 = 40/8.040 = 0,00497$$

$$RR = 0,0291/0,00497 = 5,85$$

## b) Estudio de casos y controles

		Daño a la salud		TOTAL
		Presente (casos)	Ausente (controles)	
Factor de riesgo	Presente (expuestos)	60	20	80
	Ausente (no-expuestos)	40	80	120
	TOTAL	100	100	200

$$RR (c) = \frac{60 \times 80}{40 \times 20} = \frac{4.800}{800} = 6,0$$

Nota: La proporción de casos que estuvieron expuestos, es la misma en ambos estudios. El cálculo sin sesgo del riesgo relativo en el estudio de cohorte es:  $RR = 5,85$  y en el de casos y controles es:  $RR = 6,0$  que es bastante aproximado. No obstante el número de sujetos (tamaño de la muestra) necesarios es mucho más reducido para el estudio de casos y controles (10.100 vs. 200).

### C) Introducción a la prueba de hipótesis utilizando el riesgo relativo

Es importante remarcar que el riesgo relativo obtenido por un muestreo es un cálculo estimado del riesgo relativo verdadero de la totalidad de la población de la cual se escogió la muestra. Le corresponde al investigador preguntarse qué grado de confianza tiene en que el

cálculo de por ejemplo  $RR = 3.0$  obtenido en la población en estudio (la incluida en la muestra), sea bastante aproximado al riesgo relativo de la población en estudio. El nivel de confianza aceptado por el investigador depende en parte del tamaño de la muestra. Por ejemplo, si el tamaño de la muestra es de 40 mujeres en total, la tabla  $2 \times 2$  observada puede parecerse a la siguiente:

		Daño a la salud		TOTAL
		Presente	Ausente	
Factor de riesgo	Presente	6	14	20
	Ausente	2	18	20
TOTAL		8	32	40

$$RR = \frac{p^1}{p^2} = 3.0$$

$$X^2 = 2.50$$

(*p*es mayor que) 0,10

El valor del riesgo relativo es 3.0, pero evidentemente el tamaño reducido de la muestra va a disminuir la confianza en el cálculo estimado. De hecho, puede considerarse que aún si el factor de riesgo *tiene un efecto verdadero* sobre

el daño, se pueden observar resultados como el ejemplo anterior. Pero la no detección del efecto sería menos probable si el análisis se basara en 1.000 casos, como se muestra en el siguiente cuadro:

		Daño a la salud		Total
		Presente	Ausente	
Factor de riesgo	Presente	150	350	500
	Ausente	50	450	500
TOTAL		200	800	1.000
RR $\frac{p^1}{p^2} = 3$			$X^2 = 62,5$	<i>p</i> es menor que) 0,001

Aquí las proporciones y el RR se mantienen iguales, mientras que el valor del  $X^2$  aumentó de 2.5 a 62.5 como consecuencia del incremento en el tamaño de la muestra estudiada.

A continuación se discute un método estadístico para (i) verificar si el riesgo relativo es significativamente diferente de 1 (la unidad) y, (ii) crear un intervalo dentro del cual se tenga

confianza de que el riesgo relativo estaría incluido si se realizara un estudio similar que abarcara a toda la población.

#### D) Prueba para determinar si un riesgo relativo es diferente de la unidad

En la presente sección se pondrá a prueba la hipótesis nula de no diferencia indicada por  $H_0$  y usando  $H_1$  para denotar la hipótesis alternativa:

$H_0: RR = 1$  es equivalente a  $H_0: P_1 = P_2$

$H_1: RR \neq 1$  es equivalente a  $H_1: P_1 \neq P_2$

Nota: La hipótesis alternativa es que  $p_1$  no es igual a  $p_2$

porque si  $p_1 = p_2$ , entonces  $RR = P_1 / P_2 = 1$

Y si  $p_1$  no es igual a  $p_2$  entonces  $RR$  no es igual a 1.

#### Ejemplo

Si  $p_1 = 0,05$  y  $p_2 = 0,05$  entonces  $p_1 = p_2$  y  $RR = 0,05 / 0,05 = 1,0$  Por consiguiente,  $H_0: p_1 = p_2$  y  $RR = 1,0$

$H_0: RR = 1$

Pero si  $p_1 = 0,10$  y  $p_2 = 0,05$  entonces  $p_1$  no es igual a  $p_2$  y

$RR = 0,10 / 0,05 = 2$  que no es igual a 1. Por consiguiente,  $H_1: p_1 \neq p_2$

no es igual a  $p_2$  y

$H_1: RR \neq 1$ .

Cuando se prueban las siguientes hipótesis:

$H_0: RR = 1$  ó  $H_0: p_1 = p_2$

$H_1: p_1$  no es igual a  $p_2$  ó  $RR$  no es igual a 1.

Se le conoce como prueba de dos colas. La explicación es que puesto que la hipótesis alternativa es una relación "no igual a", entonces o bien  $p_1$  es superior a  $p_2$ , o  $p_1$  es inferior a  $p_2$  ( $RR$  es mayor que 1 ó  $RR$  es menor que 1), satisfaciendo el criterio de "no igual a". De ahí que el error de tipo I pueda hallarse en cualquier extremo de la curva de distribución de probabilidad, es decir, a ambos lados, en las dos colas.

Cuando únicamente se prueban las siguientes hipótesis:

$H_0: p_1 = p_2$  o  $RR = 1$

$H_1: p_1$  mayor que  $p_2$  o  $RR$  mayor que 1.

se le designa como prueba de una cola. Ahí el tipo de error I puede darse sólo en un extremo de la curva de probabilidad, por eso se refiere como de una cola. Lo mismo sucede con  $H_1: p_1$  menor que  $p_2$  o  $RR$  menor que 1.

A continuación se trabajará con las pruebas de dos colas usadas comúnmente. Las pruebas de una cola deben emplearse en ocasiones en que se conoce la direccionalidad del efecto del factor de riesgo bajo estudio.

#### E) Intervalos de confianza para el riesgo relativo

Aquí se examina un método de evaluación de la importancia estadística del riesgo relativo mediante la creación de un intervalo de confianza, alrededor del riesgo relativo observado. que: (a) resume qué tan cercano se encuentra el cálculo del riesgo observado (en el muestreo) al riesgo verdadero (en la población objetivo); y (b) permite determinar si el valor observado de  $RR$  es marcadamente diferente de la unidad (1.0). Así los intervalos de confianza indican los límites probables dentro de los que se encuentra el valor verdadero del riesgo relativo de la población objetivo. Igualmente, si el límite inferior del intervalo de confianza incluye el valor uno o sea la "Hipótesis Nula" ( $H_0$ ), se interpreta que el riesgo relativo encontrado no es estadísticamente distinto al  $RR$  de 1, y por tanto no existe exceso de riesgo debido al factor expuesto. Para calcular los intervalos de confianza se procede de la forma siguiente: Si se entiende que

$RR$  = riesgo relativo obtenido.

$X$  = la raíz cuadrada del  $X^2$  obtenida de la tabla de  $2 \times 2$ ;

$X = X^2$  (la raíz cuadrada de  $X^2$ )

$z$  = valor del desvío normal correspondiente al nivel de confianza deseado, por ejemplo  $z = 1,64$  para límites de 90% y  $z = 1,96$  para límites de 95%.



		Daño a la salud		Total	TASA DE INCIDENCIA
		Presente	Ausente		
Factor de riesgo 1	Presente	150	250	400	0.375
	Ausente	50	550	600	0.0833
	Total	200	800	1.000	

El riesgo relativo se calcula así:

La probabilidad del daño a la salud dada

$$RR = \frac{\text{la presencia del factor de riesgo}}{\text{la ausencia del factor de riesgo.}}$$

$$= \frac{150/400}{50/600} = \frac{0,375}{0,0833} = 4,5$$

Comúnmente los investigadores querrán examinar varios factores de riesgo potencialmente importantes y calcular el riesgo relativo relacionado con sus efectos (multiplicativos) sobre el daño a la salud. Considérese el caso de dos factores de riesgo en los siguientes cuadros.

Factor 1	Daño a la salud			Factor 2	Daño a la salud		
	+	-	Total		+	-	Total
Presente	150	250	400	Presente	130	300	430
Ausente	50	550	600	Ausente	70	500	570
Total	200	800	1.000	Total	200	800	1.000
Riesgo Relativo = $\frac{150/400}{50/600} = 4,5$				Riesgo Relativo = $\frac{130/430}{70/570} = 2,46$			

Si los dos factores de riesgo seleccionados son independientes y ambos contribuyen a un daño específico, cuando ambos se hallan presentes en el mismo individuo, el riesgo relativo total es el producto de los dos riesgos relativos individuales.

$$4,5 \times 2,46 = 11,1$$

Esto se llama *RR conjunto multiplicado*. El total de riesgo relativo de tres o más factores se obtiene multiplicando los valores de riesgo rela-

tivos por cada factor siempre que contribuyan independientemente al riesgo de un daño a la salud determinado.

No hay necesidad de comparar cada factor con el resto de factores. Basándose en intercambios de ideas y en discusiones con los miembros del grupo de estudio y consultores externos (clínicos), enfermeras obstétricas y otros, se crea una lista de factores que pueden ser interdependientes.

Si se sospecha que dos factores se relacionan (no son independientes) como por ejemplo, edad materna y estado civil, o vivienda y nivel de ingreso, se puede examinar esa relación mediante tabulaciones cruzadas simples. Un método consiste en *controlar* un factor (mantenerlo constante) y analizar los efectos del otro inde-

pendiente de los posibles efectos del primero.

Tomando por ejemplo el caso de los dos factores de riesgo presentados previamente (Factor 1, RR = 4,5; Factor 2, RR = 2,46).

En el ejemplo señalado, pueden examinarse los efectos del factor 2 con la presencia o ausencia del factor 1.

Factor 1 ausente				Factor 1 presente			
Daño a la salud				Daño a la salud			
Factor 2	Presente	Ausente	Total	Factor 2	Presente	Ausente	Total
Presente	10	100	110	Presente	120	200	320
Ausente	40	450	490	Ausente	10	50	60
TOTAL	50	550	600	TOTAL	150	250	400

Riesgo Relativo =  $\frac{10/110}{40/490} = 1,1$

Riesgo Relativo =  $\frac{120/320}{10/60} = 1,0$

Nótese primero que la presencia o ausencia del factor de riesgo 1 determina la distribución de individuos entre los dos cuadros. La frecuencia del factor de riesgo 1 en la población estudiada es de 400/1.000 ó 40%. Segundo, la razón de la población con el daño a la salud cuando el factor de riesgo 1 está ausente es 50/600 ó 8,3% y cuando está presente es 150/400 ó 37,5%.

Tercero y último, en los cuadros se indica que dentro de esos subgrupos (presencia o ausencia del factor de riesgo 1) la presencia o ausencia del factor de riesgo 2 no ejerce influencia sobre la proporción del daño a la salud. El riesgo relativo relacionado con el factor de riesgo 2 es (o se aproxima a) 1,0 en los grupos con o sin factor de riesgo 1. Evidentemente, el factor de riesgo 2 no contribuye como factor independiente al riesgo del individuo a experimentar el daño a la salud.

Al seguir interesados en determinar el riesgo relativo verdadero de un daño a la salud en la población que tenga esas características, para obtener el valor verdadero del riesgo relativo debe considerarse la posible influencia de

ambos factores de riesgo y luego calcularlos como RR conjunto crudo:

$$RR \text{ conjunto crudo} = \frac{\text{Probabilidad del daño a la salud suponiendo la presencia de ambos factores de riesgo}}{\text{Probabilidad del daño a la salud suponiendo la ausencia de ambos factores de riesgo}}$$

Basados en las dos últimas tablas de 2 x 2 de la página anterior se calcula el RR conjunto crudo de la siguiente manera:

$$RR \text{ conjunto crudo} = \frac{120/320}{40/490} = \frac{0,375}{0,0813} = 4,60$$

Según se calculó anteriormente, si los factores 1 y 2 son independientes, el RR conjunto multiplicado sería 11,1. El RR conjunto crudo obtenido, 4,60, es más de 60% inferior e indica nuevamente que los dos factores no son independientes.

Se analiza la independencia porque si los dos factores de riesgo contribuyen independientemente al riesgo de un daño concreto, se los puede usar conjuntamente para evaluar la situación de riesgo. Si existe una relación estrecha, sólo se requiere el uso de uno de ellos para esa evaluación. Una razón importante para usar las técnicas estadísticas de variables múltiples es que ellas automáticamente consideran o controlan los efectos de las variables distractoras.

Sin embargo, esas técnicas de variables múltiples no se trataron en este documento de trabajo por dos motivos: 1) son métodos complejos y es casi imposible enseñarlos en el tiempo de duración del taller; y 2) existen instrumental y programación de calculadoras o computadoras que puede no estar disponible en algunas investigaciones. Por ese motivo, se procuró elaborar un método sencillo y directo de determinación de las contribuciones independientes de cada variable al riesgo de un daño a la salud.

El siguiente aspecto a discutir aquí es una regla de decisión por la cual se declare que un factor de riesgo es un contribuyente independiente a un daño a la salud. Cualquier regla elaborada puede siempre ser cuestionada estadísticamente porque no tiene en cuenta todas las posibles influencias sobre un daño a la salud.

Para obtener el riesgo relativo conjunto ajustado se usan dos riesgos relativos que se identificarán como RR1 y RR2. Escoja el menor, al que se llamará RR1, y luego calcule el riesgo relativo conjunto ajustado.

$$\text{El RR conjunto ajustado} = (0,5(RR1 - 1) + 1)RR2.$$

La interpretación de esta medida es que si el riesgo relativo conjunto ajustado es superior al riesgo relativo conjunto ajustado entonces los dos factores de riesgo son independientes; si es menor son considerados no independientes.

### Ejemplo

Riesgo relativo conjunto ajustado. En un estudio sobre mortalidad neonatal, después de examinar varios posibles factores de riesgo, se

determinó que cuatro eran estadísticamente significativos a un nivel de 5%. Esos factores y sus correspondientes Riesgos Relativos fueron:

1. Edad materna de 19 años o menos RR = 2,5
2. Complicaciones relacionadas con el embarazo RR = 3,0
3. Estado civil soltera RR = 2,0
4. Embarazos múltiples RR = 4,0

Se supuso que no todos los factores de riesgo eran contribuyentes independientes, en especial las dos series siguientes de factores de riesgo:

1. "Edad materna de 19 años o menos" y "soltera".
2. "Complicaciones relacionadas con el embarazo" y "Embarazos múltiples".

Primeramente se determinó que el Riesgo Relativo conjunto crudo de "Edad Materna de 19 años o menos" y "Soltera" era igual a 3,1.

El Riesgo Relativo conjunto ajustado =  $(0,5(2,0 - 1) + 1)2,5 = 3,75$

Como 3,1 es menor que 3,75 se concluyó que "edad materna de 19 años o menos" y "soltera" no son factores independientes.

Debido a que la primera edad se relaciona más estrechamente con riesgo, y por razones de credibilidad biológica, se decidió seleccionarla como predictor de riesgo.

Seguidamente se determinó que el Riesgo Relativo conjunto crudo de "complicaciones relacionadas con gestación" y "embarazos múltiples" era 11,3.

Riesgo Relativo conjunto ajustado =  $(0,5(3,0 - 1) + 1)4,0 = 8,0$

Como 11,3 es mayor que 8,0, se concluyó que "Complicaciones relacionadas con el embarazo" y "embarazos múltiples" son contribuyentes independientes y que ambas deben formar parte de la ecuación de predicción de riesgo.

De los cuatro factores de riesgo que se tenían al inicio, quedan tres para incluirse en la ecuación de predicción de riesgo:

1. Edad de madre de 19 años o menos.
2. Complicaciones relacionadas con el embarazo.
3. Embarazos múltiples.

El siguiente es un resumen de los pasos a cumplir en la identificación de factores de riesgo:

1. No incluya posteriormente aquellas variables que sólo tienen poco o ningún efecto sobre el daño a la salud bajo estudio.
2. Use la prueba  $\chi^2$  (Chi cuadrado) para determinar aquellas variables que son estadísticamente significativas en relación con el daño a la salud bajo estudio.
3. Analice los factores de riesgo importantes para determinar si son contribuyentes independientes.
4. Bajo criterios técnicos seleccione uno de los factores de riesgo interdependientes.

#### 5.8.4 Elaboración de un programa de tamizaje

##### A) El plan global

Un enfoque práctico de la planificación y aplicación de un programa de tamizaje se puede resumir de la siguiente manera:

- a) Determinar *por qué* se realiza el examen de tamizaje, *quién* lo llevará a cabo, *cuándo* y *dónde* se realizará, y *qué* tecnología se empleará.

Tómese el caso de examinar mujeres embarazadas para prevenir un daño a la salud perinatal (por qué), en el cual se dispone básicamente de tres períodos (cuándo): prenatal, intraparto y postparto. Un auxiliar de enfermería (quién) que realiza el examen en la casa (dónde) y determina el equipo necesario (qué). Ese sistema difiere del de la enfermera obstétrica (quién) que examina en el centro de salud (dónde) para prevenir lo mismo, usando diferente equipo (qué).

- b) Calcular el riesgo relativo de los factores que se usarán.

- c) Analizar el efecto de la selección de diferentes límites de inclusión para las variables y el puntaje total del riesgo en cuanto a:

1. sensibilidad y especificidad<sup>1</sup>;
2. poder predictivo; y
3. número total de individuos que son asignados al grupo de riesgo.

- d) Recorrer los puntos (a), (b) y (c) en varias oportunidades y describir cómo la aplicación del puntaje de riesgo afectará el uso de los recursos de salud.

En general, el investigador debe estar preparado para encontrar una relación inversa entre sensibilidad y especificidad y entre el poder predictivo de las pruebas positivas y negativas. *La elección final de la prueba de tamizaje dependerá fundamentalmente de las prioridades y de los recursos disponibles.*

##### B) Análisis y puntaje de riesgo

Algunos individuos no tienen factores conocidos de riesgo respecto de determinada enfermedad, otros tienen dos, y otros, muchos. ¿Cómo combinar los factores de riesgo que cada persona tiene para encontrar una expresión numérica, un puntaje que pueda cuantificar su riesgo de tal forma que corresponda al nivel de atención requerida?

Diferentes factores de riesgo afectan en diferentes medidas al daño; el riesgo relativo varía entre factores. Si son varios los factores presentes, el daño depende del efecto de cada uno de ellos y de su interacción. Diferentes combinaciones de factores de riesgo afectan en formas diferentes a los daños. Se puede definir el riesgo relativo de cualquier factor o cualquier combinación de factores y combinarlos en una ecuación de puntaje de riesgo. Si no se conocen

<sup>1</sup> sensibilidad: proporción de resultados positivos verdaderos detectados por la prueba de tamizaje.

especificidad: proporción de resultados negativos verdaderos detectados por la prueba de tamizaje.

todos los riesgos relativos o no se dispone de una computadora, se cuenta con métodos más simples para la combinación de factores de riesgo. Los más sencillos son:

a) *El método arbitrario*

En ese método, las opiniones de personas con experiencia y conocimientos clínicos y de salud pública forman la base para el sistema de puntaje.

En su versión más simple, el sistema de puntaje se reduce a una lista de los indicadores más importantes de riesgo. Si se descubre la presencia de uno o más factores, se dice que la persona tiene mayor riesgo.

En la versión más refinada, se asigna un puntaje (un número) a cada factor de riesgo, por ejemplo de 1 a 5. Los números se suman o multiplican. Si la persona tiene un puntaje total, dígase, de 10 o más, se dice que tiene un mayor riesgo.

La persona con menos de 10 puntos se clasificará entre los del grupo de bajo riesgo. En este ejemplo, el puntaje 10 se usa para separar los sujetos de alto de los de bajo riesgo. Este nivel se denomina "límite de inclusión". El método arbitrario ha sido ampliamente utilizado en la literatura.

b) *Métodos matemáticas simples*

Los puntajes pueden derivarse también de una transformación relativamente simple de los riesgos relativos, obtenidos por tabulaciones cruzadas del tipo expuesto en 5.8.1. A continuación se ofrece un ejemplo de un sistema sencillo de puntaje derivado de la evaluación de tres factores de riesgo que, para los fines del presente ejercicio, se supone que son independientes. A cada factor se le atribuye una importancia proporcional a su riesgo relativo.

Ejemplo

Factor de riesgo	Factor 1			Factor 2		
	Daño a la salud		Total	Daño a la salud		Total
	Presente	Ausente		Presente	Ausente	
Presente	150	250	400	130	300	430
Ausente	50	550	600	70	500	570
TOTAL	200	800	1.000	200	800	1.000
Riesgo Relativo	$\frac{150/400}{50/600}$			$\frac{130/430}{70/570}$		
=	0,375 0,083			= 0,302 0,123		
	= 4,5			= 2,5		

Factor de riesgo	Factor 3 Daño a la salud		Total
	Presente	Ausente	
Presente	150	350	500
Ausente	50	450	500
<b>Total</b>	<b>200</b>	<b>800</b>	<b>1,000</b>

$$\text{Riesgo Relativo } \frac{150/500}{50/500}$$

$$= \frac{0.30}{0.10}$$

$$= 3.0$$

Debe adoptarse ahora un método para combinar los tres RR. Por ejemplo, se podría multiplicarlos, pero ese sistema no es práctico para los que trabajan en el campo y podría ocasionar errores. Una alternativa a este método consiste en asignar a cada factor un peso proporcional a su logaritmo. En el cuadro siguiente se indican los pesos asignados usando el puntaje de 1 a 10.

Conversión de riesgo relativo a puntajes aditivos

Riesgo Relativo	Log10	Puntaje
1.0	0	0
1.2	0.08	1
1.4	0.15	2
1.6	0.20	2
1.8	0.26	3
2.0	0.30	3
2.2	0.34	3
2.5	0.40	4
3.0	0.48	5
4.0	0.60	6
5.0	0.70	7
6.0	0.78	8
7.0	0.85	9
8.0	0.90	9
9.0	0.95	10
10.0	1.00	10

Usando esta tabla para los tres factores se obtiene:

	Factor 1	Factor 2	Factor 3
Riesgo Relativo	$\frac{4.5}{7}$	$\frac{2.5}{4}$	$\frac{3.0}{5}$
Puntaje			

Según ese esquema, una mujer con factores de riesgo 2 y 3 tendrá un puntaje de 9, y otra con factores 1 y 3 un puntaje de 12. Este esquema es fácil de usar, y representa una aproximación muy buena del riesgo asociado por la combinación de factores.

### c) Técnicas más complejas

Valiéndose de computadoras y programas, se puede elaborar un sistema de puntaje más refinado, basado en algunos de los diversos tipos de análisis con variables múltiples. Tienen la ventaja de que contemplan simultáneamente las variables importantes (por eso se denomina "con variables múltiples") y, por consiguiente, pueden examinar y compensar la falta de independencia de los factores.

Un método alternativo consiste en analizar primero los efectos de las variables interdependientes mediante un simple análisis de dos variables. Luego, valiéndose de esas variables, se establece una función discriminatoria para identificar las variables que podrán diferenciar mejor los

grupos. Por ejemplo, diferenciar una mujer embarazada cuyo hijo murió en el período neonatal de otra cuyo neonato no murió. El análisis llega a ese resultado ponderando y combinando linealmente las variables independientes. Se aprecia así la contribución especial que tiene la variable de diferenciar entre grupos, después de controlar simultáneamente los efectos de todas las otras variables independientes. Para que se incluya una variable en la función discriminatoria su contribución especial debe ser estadísticamente significativa. Las variables deben ser objeto de una dicotomía o sea la división de cada variable en dos grupos para facilitar el uso la función. Aunque se pierde parte de la información, la facilidad de uso compensa abundantemente esa pérdida.

### 5.8.5 Distribución del riesgo en la población

#### A) Evaluación del sistema de tamizaje

Se ha asignado ya un peso a cada factor que se incluirá en el sistema de puntaje y a cada mujer de la población se le ha asignado un puntaje que es la suma de los pesos de los factores de riesgo que posee. Por ejemplo, 35 de las 1.000 mujeres estudiadas tienen un puntaje

de 10 o más, y 320 un puntaje de 0. El resto tiene valores intermedios. La distribución completa de sus puntajes es esta

#### Ejemplo

*Distribución de puntajes de riesgo del grupo de mujeres embarazadas*

Puntaje de riesgo	Número de mujeres
0	320
1	136
2	105
3	94
4	79
5	63
6	57
7	48
8	37
9	26
10+	35
TOTAL	1.000

Ahora hay que decidir a partir de qué puntaje se considerará que la persona está "a alto riesgo". La finalidad de ese paso es la de identificar el mayor número posible de casos de daños a la salud incluyendo el menor número de personas (para un costo y molestia mínimos). Si se añaden dos columnas adicionales a la tabla de puntajes se pueden realizar los cálculos necesarios.

#### Ejemplo

Puntaje	Daño a la salud				
	Total nacidos		N. muertos		N. vivos
	Número	Acumulativo (%)	Número	Acumulativo (%)	Número
10+	35	4	30	15	5
9	26	6	20	25	6
8	37	10	23	37	14
7	48	15	18	46	30
6	57	20	22	57	35
5	63	27	20	67	43
4	79	35	23	78	56
3	94	44	20	88	74
2	105	54	11	94	94
1	136	68	8	97	128
0	320	100	5	100	315
TOTAL	1.000		200		800

Por razón de conveniencia se colocaron los puntajes de riesgo en orden de magnitud descendente en la tabla. En ella se indica que si la regla de decisión dispuso que sólo los individuos con un puntaje de riesgo de por lo menos 10 se incluyeran en el grupo de riesgo elevado, el sistema de puntaje identificó correctamente a sólo 30 (15%) de los 200 con daño a la salud, incluyendo 35 (4%) de todos los nacimientos. Por otra parte, sólo incluiría a una proporción insignificante (5/800 = 0.6%) de los que *no* tuvieron el daño a la salud.

### B) Evaluación del sistema de puntaje de riesgo

Como consecuencia de la prueba de tamizaje un individuo puede ser considerado de alto o bajo riesgo; el investigador desea conocer cuál es la probabilidad de que ese individuo desarrolle o no el daño a la salud. A continuación se discuten indicadores para evaluar esta probabilidad en la población bajo estudio.

Si se presentan los resultados en una tabla  $2 \times 2$ , se puede describirlos en la población estudiada. Nótese que  $(a + c)/(a + b + c + d)$ , en el estudio de tamizaje representa la proporción de los daños a la salud en la población general es decir, que el investigador *no* fija la proporción de daños a la salud como en el estudio de casos y controles. Entonces pueden usarse los siguientes métodos para la predicción del daño a la salud basándose en la prueba de tamizaje.

RIESGO	Daño a la salud		Total
	Presente	Ausente	
ALTO	a	b	a + b
BAJO	c	d	c + d
TOTAL	a + c	b + d	N

$$N = a + b + c + d$$

Se llama *sensibilidad* de la prueba de tamizaje a la aptitud para identificar correctamente a aquellos que tuvieron el daño a la salud, y *especificidad* a la capacidad de identificar co-

rectamente a aquellos que no tuvieron el daño a la salud (sanos). Se puede expresar a esas variables como *proporciones*:

$$\text{Proporción positiva verdadera} = \frac{a}{a + c} \\ (\text{sensibilidad})$$

$$\text{Proporción negativa verdadera} = \frac{d}{b + d} \\ (\text{especificidad})$$

El denominador de la proporción positiva verdadera es el número total de casos confirmados, más aquellos que pese a tener el daño no lo mostraron en el examen de tamizado conocidos también como "negativos falsos" (celda "c").

Igualmente, la proporción negativa verdadera tiene un denominador que incluye todos los casos negativos confirmados (sanos), incluyendo los identificados como verdaderos negativos (d) más los que son falsos positivos (b).

Las proporciones positiva-verdadera (sensibilidad) y negativa-verdadera (especificidad) en las pruebas de tamizaje son características permanentes de la prueba, las cuales son más o menos constantes con cambios en la prevalencia del daño y se les cita ampliamente como índices de eficiencia. Pero ellas *sólo* ofrecen un índice del modo en que la prueba de tamizaje funcionará dentro de esos subgrupos confirmados de individuos con o sin el daño. *No* le dan al usuario una indicación de cómo funcionará la prueba de tamizaje cuando se la emplee para distinguir personas normales de anormales en la población general. La sensibilidad y especificidad sirven para examinar las proporciones positiva y negativa en grupos concretos, aquellos con o sin el daño. Ese procedimiento es diferente del que se emplea cuando el investigador desea *predecir* la presencia de enfermedad en la población. En la sección siguiente se discutirá por qué las pruebas de tamizaje con proporciones elevadas de positivo-verdaderos (especificidad) pueden ocasionar algunos problemas cuando se les aplica en las estrategias de intervención en la población general.

La tabla siguiente puede servir para resumir lo antes expuesto.

		Daño a la salud	
		Presente	Ausente
Factor de riesgo 1	Positivo verdadero	a	b
	Positivo falso		
Prueba de tamizaje 0	Negativo falso	c	d
	Negativo verdadero		

$$\text{Poder Pred. Positivo (PPP)} = \frac{a}{a + b}$$

$$\text{Falso positivo} = \frac{b}{a + b}$$

$$\text{Poder Pred. Negativo (PPN)} = \frac{d}{c + d}$$

$$\text{Falso negativo} = \frac{c}{c + d}$$

$$\text{Sensibilidad} = \frac{a}{a + c} \quad \text{Especificidad} = \frac{d}{b + d}$$

El poder predictivo de una prueba positiva es reducido cuando se la emplea en una población con una frecuencia relativamente baja del daño. La siguiente tabla ilustra lo expuesto.

Poder Predictivo Positivo (PPP) y Negativo (PPN) como función de la frecuencia del daño con sensibilidad y especificidad constantes, (N = 1.000)

Frecuencia del daño a la salud	Sensibilidad	Especificidad	PPP	PPN
			0,30	0,90
0,20	0,90	0,90	0,69	0,97
0,10	0,90	0,90	0,59	0,99
0,05	0,90	0,90	0,32	0,99

Nótese que el PPP disminuye proporcionalmente con la reducción de la frecuencia de daño a la salud aunque la sensibilidad y la especificidad permanecen constantes. Por otra parte, el PPN aumenta con la disminución de la frecuencia del daño a la salud. Eso significa que cada vez que se estudia un daño a la salud de baja frecuencia habrá que esperar un número alto de falsos positivos para un nivel de sensibilidad y especificidad dado.

### C) Límites de inclusión

Usando la tabla de la página 123 que muestra las distribuciones de riesgo en una población, se puede comprobar, por ejemplo, que si se decide considerar en riesgo a todas las mujeres con puntajes de riesgo de 7 o más, se obtiene un total de 146 (48 + 37 + 26 + 35) de las cuales en 91 (18 + 23 + 20 + 30) se encontró daño a la salud y en 55 no se encontró daño a la salud es decir, 91 "positivos verdaderos" y 55 "positivos falsos". Un límite de inclusión de 5 o más dará resultados diferentes. El siguiente cuadro ofrece un resumen derivado del cuadro de la página 123 para todos los límites de inclusión superiores a 3.

Límite de inclusión para un puntaje de:

	4+	5+	6+	7+	8+	9+	10+
Positivos verdaderos	156	133	113	91	73	50	30
Positivos falsos	189	133	90	55	25	11	5
Negativos verdaderos	611	667	710	745	775	789	795
Negativos falsos	44	67	87	109	127	150	170

Es evidente que la fijación del límite de inclusión implica un intercambio entre las cuatro categorías. Si se decide pasar de un límite de 7+ a otro 6+ se identifican otros 22 con el daño a la salud (de 91 a 113), a un costo de 35 falsos positivos (de 55 a 90) pero todavía quedan sin ser incluidos 87 con el daño a la salud (los negativos falsos).

Un método útil de resumir la eficacia de una prueba consiste en calcular sus poderes predictivos. El poder predictivo de una prueba positiva es el número de positivos verdaderos dividido por el número total de los identificados como positivos.

Empleando los datos anteriores se tiene:

### Ejemplo

Poder predictivo de un límite de inclusión de:

	4+	5+	6+	7+	8+	9+	10+
Poder predictivo positivo	45%	50%	56%	62%	74%	82%	86%
Poder predictivo negativo	93%	91%	89%	87%	86%	84%	82%

En la práctica, la fijación del límite de inclusión depende de varios factores, ninguno de los cuales puede deducirse exclusivamente de los datos epidemiológicos obtenidos en el estudio analizado en el capítulo presente. A corto plazo, la gravedad del problema y los recursos disponibles (recursos humanos, instalaciones de salud, cobertura a nivel primario) determinarán la cantidad de individuos que se incluirá en la categoría de riesgo.

A largo plazo, la estrategia de riesgo procurará la reasignación de recursos y tareas de tal forma que puedan compararse entre sí todas las consideraciones importantes, incluyendo no sólo la probabilidad de no detectar un daño a la salud sino también la inconveniencia y el peligro de los falsos positivos, ponderándolos el uno contra el otro. En esta etapa puede ser necesario reconsiderar los límites de inclusión previamente determinados.

(ii) Usando el riesgo relativo seleccione de la lista siguiente las 12 variables que se relacionan más estrechamente con la mortalidad perinatal. Nótese que la selección de las variables se hace en base a RR y no al  $X^2$  debido a que el RR es el indicador que mide la fuerza de la asociación.

### Ejercicio 8

Hay que determinar ahora la fuerza de asociación entre las variables y el daño a la salud.

(i) Calcular el riesgo relativo de las cinco variables de las hipótesis del ejercicio 2.

Los siguientes son datos derivados de Ficticia (cuadros 23, 23A y 24)

Factor	TMP con Factor de riesgo	Valor X <sup>2</sup>	Riesgo relativo
Estatura/145	145	480,7	2,84
Estatura/155	94,5	884,0	2,6
Agua (contaminada)	86	390	2,0
Personal auxiliar	67	69	--
Atención prenatal (ninguna)	67	388,46	2,13
(ninguna + sólo 3er trimestre)	68,6	308,01	2,18
Soltera	101	477,04	2,26
Grupo étnico B	91	962,98	3,35
Paridad (.../3)	65,1	36,7	1,26
Paridad (.../4)	83	126,05	1,65
Región (montañosa)	71	21,33	--
Sector (rural)	58,8	9,03	1,13
Sexo (masculino)	615	41,01	1,27
Complicación médica durante el parto	182,7	1.846,0	4,50
Distancia (.../5 Km)	65	56,76	1,31
Complicación obstétrica durante el parto	130,0	1.704,9	4,06
Ocupación (obrero, agri- cultor, minero, pescador)	62,3	176,35	2,12
Educac. Mat. (analfabeta)	74	255,0	1,74
Número hijos vivos (.../2)	56,6	0,04	1,0
Número hijos vivos (.../3)	69,3	15,3	--
Complicación durante el presente embarazo	122,4	1.859,68	4,53
Toxemia	237,0	1.067,0	--
Complicación previa al embarazo	88,4	158,0	1,73
Tipo de parto (pélvico)	242,0	1.064,0	4,87
Peso al nacer (.../2.500)	262,0	6.629,0	12,67
Aborto (presente)	74,0	55,57	1,46
Parto prematuro (presente)	95,0	94,30	1,80
Edad materna (.../19)	83,7	173,9	--
Edad materna (.../17)	83,7	60,6	1,57
Número de visitas prenatales (.../3)	55,0	0,10	0,98
Madre que trabaja fuera de casa	55,0	0,17	0,98
Fumar durante el embarazo	56,0	0,0024	1,02

(iii) Se desea eliminar las "variables distractoras" a fin de reducir la cantidad de información necesaria. En su opinión, ¿qué combinaciones de factores se confunden?; ¿cómo las identificaría usted y

cómo determinará que Valiéndose del cuadro 11 que si la estatura de la ción de alfabetismo s

ESTATURA	Persona alfabetizada		Analfabeta	
	Muertes perinatales	Total de nacimientos	Muertes perinatales	Total de nacimientos
Menos 145 cms	133	1.077	201	1.223
= 6 mayor 145 cms	1.031	27.346	1.467	21.464
Total	1.164	28.423	1.668	22.687

- (iv) En un estudio sobre mortalidad neonatal, después de analizar muchos factores de riesgo potenciales se descubrió que cuatro eran estadísticamente importantes a nivel de  $p$  menos de 0,01. Esos factores de riesgo y sus correspondientes RR fueron:
- Número de embarazos anteriores RR = 2,5  
 Toxemia RR = 2,8  
 Aborto anterior RR = 3,1  
 Embarazos múltiples RR = 4,0
- a) Existía una sospecha de que "embarazo anterior" y "aborto anterior" no eran factores independientes. Su RR conjunto crudo = 4,1. ¿Son o no independientes?
- b) Asimismo, se supuso que no eran independientes toxemia y "embarazos múltiples". Su RR conjunto crudo = 10,1. ¿Son o no independientes?
- c) ¿Cuáles son las variables que formarán la ecuación de predicción de riesgo?
- v) Los investigadores decidieron que en el primer nivel de atención sería práctico utilizar las siguientes variables para seleccionar las embarazadas de alto riesgo:
- estatura de las mujeres;
  - si tiene agua contaminada;
  - mes aproximado en que comenzó la atención prenatal;
  - número de niños vivos;
  - si sabe leer;
  - su edad;
  - a que grupo étnico pertenece;
  - si está casada;
- la ocupación de su marido;
- historia previa de partos prematuros o abortos;
- si vive en área rural;
- distancia de la casa donde vive la mujer con respecto al centro de salud;
- qué lugar ocupa el presente niño entre los nacidos vivos.
- Posteriormente escogieron dos métodos diferentes para describir la distribución de riesgo en Ficticia. Al comienzo decidieron realizar un simple recuento del número de mujeres con algunos de los factores, de la lista anterior, presentes durante: 1) el embarazo; 2) el trabajo de parto y el parto; y 3) el postparto. ¿Qué importancia le atribuye a cada variable esa decisión? Use los cuadros 26, 27 y 28 de Ficticia para discutir las respuestas.
- (vi) Luego los investigadores, decidieron usar el riesgo relativo para asignar pesos a las variables. ¿En qué forma afecta al puntaje esa medida? Use el cuadro 29 de Ficticia para discutir la respuesta.
- (vii) Escogiendo un puntaje de ocho o más como definición de alto riesgo para el período preparto compare los dos sistemas de puntaje (cuadros 26 y 29 de Ficticia) en cuanto a:
- a) número de mujeres supuestamente con alto riesgo
  - b) número de mujeres pronosticadas correctamente
- (viii) Usando el cuadro 26 de Ficticia elabore un diagrama indicando cada uno de los límites de inclusión.

Límite de inclusión para un puntaje.

de  
6+      7+      8+      9+

- Los positivos verdaderos
- Los positivos falsos
- Los negativos verdaderos
- Los negativos falsos

Los investigadores escogieron el límite de 8+. Analice las consecuencias de esa decisión.

- (ix) Hay cuatro métodos para medir la eficacia del sistema de puntaje, es decir, la proporción falso-positivo y falso-negativo, y el poder predictivo positivo y negativo. ¿Cuál de ellos escogería usted en Ficticia y por qué?

## Sección 5.9. La cuantificación del riesgo en la comunidad

### Resumen

Las estrategias seleccionadas para las comunidades deben insistir en la necesidad de identificar los determinantes principales de los problemas de salud. Puede usarse el riesgo atribuible como medio para determinar las intervenciones prioritarias dentro de la comunidad. El riesgo atribuible en la población es tanto una función del riesgo relativo como de la prevalencia del factor de riesgo en la comunidad. Expresa el porcentaje de la incidencia del daño a la salud que podría reducirse si el factor de riesgo fuese eliminado o su efecto neutralizado.

### 5.9.1. Uso de factores de riesgo para formular estrategias de salud para una comunidad

En la determinación de la estrategia de salud para una comunidad se requiere identificar el aporte de un factor de riesgo particular al daño(s) de salud prioritarios. En otros términos, en qué proporción la mortalidad o morbilidad (daño) se reducirá si el factor(es) de riesgo se elimina de la comunidad. Esta medida permite al planificador de salud calcular hasta dónde un daño particular es producto de un factor de

riesgo específico y predecir en qué medida un programa de salud orientado al factor de riesgo reducirá la proporción de daños a la salud.

En la sección 5.8. se expuso la cuantificación del riesgo (riesgo relativo) a nivel individual se realizó el estudio epidemiológico para determinar la magnitud del problema y se identificaron los factores de riesgo que son predictivos relacionados causalmente, o ambos, con el problema. En la elaboración de las estrategias para la comunidad se debe usar primordialmente los factores de riesgo relacionados causalmente con el daño a la salud. La prevalencia del factor de riesgo se relaciona con su contribución relativa al daño bajo estudio. Los factores relacionados pueden agruparse de tal forma que se formen conjuntos a ser considerados en intervenciones.

### 5.9.2 Identificación de factores causales

En la sección 5.7. se expusieron los criterios de asociación causal por lo que aquí sólo se hace breve mención de ellos. Corresponde al investigador determinar los factores de riesgo que relacionan causalmente con los problemas usando los criterios de la fuerza, especificidad, asociación temporal, consistencia y plausibilidad biológica. Habrá que dividir en dos grupos los factores causales: aquellos que pueden ocurrirse durante el presente embarazo y aquellos que deben prevenirse en embarazos futuros. Por ejemplo si se trata de una mujer embarazada a los 15 años de edad, será difícil superar la contribución de la edad al embarazo (se podría alterar el efecto mediante un nivel superior de atención). Una estrategia más amplia podría ser la ejecución de un programa comunitario para prevenir embarazos en esa edad y/o prevenir futuros embarazos en las ya embarazadas.

### 5.9.3 Uso del riesgo atribuible

El riesgo atribuible en la población total es una medida de asociación influenciada por la prevalencia del factor de riesgo en la población.

total. De ahí que sea un método que pueda servir para evaluar la contribución de un factor causal al problema global. El riesgo atribuible es el *margen de exceso de mortalidad* que puede atribuirse a la presencia de un factor de riesgo particular. En otras palabras, el riesgo atribuible(%) en la población mide el descenso porcentual en el número de muertes que habría si ese factor causal fuera eliminado (o neutralizado) totalmente.

### Ejemplo

#### Mortalidad Perinatal

Controles prenatales	Mortalidad Perinatal		Total
	Presente	Ausente	
0	2.056	25.891	27.947
1 o más	776	22.387	23.163
<b>TOTAL</b>	<b>2.832</b>	<b>48.278</b>	<b>51.110</b>

Riesgo Relativo = 2,2

$$RA \text{ pobl (\%)} = \frac{\text{tasa de incidencia del daño en la población total} - \text{tasa de incidencia del daño en los no expuestos al factor}}{\text{tasa de incidencia del daño en la población total}} \times 100$$

Una fórmula alternativa es:

$$RA \text{ pobl (\%)} = \frac{F(RR-1)}{1 + F(RR-1)} \times 100$$

Donde F es la proporción de la población expuesta al factor de riesgo y RR es el riesgo relativo. Esta segunda fórmula es más importante porque permite percibir con claridad cómo cambia el impacto de una intervención con la prevalencia de un factor de riesgo. Además, en estudios de casos y controles, en los cuales no

se formula la pregunta: ¿cuál sería el número de muertes si todas las mujeres recibieran por lo menos un control prenatal?

La tasa de MPN entre las mujeres que tuvieron por lo menos un control prenatal es:

TMPN control =  $776/23.163 = 0.0335$   
y por lo tanto el número de muertes sería, respondiendo a la pregunta anterior, igual a  $0.0335 \times 51.110 = 1.712$ . El descenso porcentual es entonces:

$$\frac{2.832 - 1.712}{2.832} \times 100 = 39,5\%$$

y este es el riesgo atribuible (%) en la población para el factor de riesgo "sin control prenatal".

El riesgo atribuible (%) en la población se calcula de la siguiente manera:

se puede obtener las tasas reales de incidencia del daño, uno puede usarla tomando la razón de productos cruzados como una aproximación para el riesgo relativo. La proporción de la población expuesta al factor de riesgo se puede obtener de otras fuentes, de un estudio piloto o de la proporción encontrada entre los controles, cuando ésta es relativamente baja.

## Ejemplo

Con los datos del cuadro anterior (atención prenatal), la proporción de la población expuesta al factor de riesgo se calcula:

$(a + b)/N$  ó  $27.947/51.110 = 0,55$  (o 55 %) y el riesgo relativo es 2,2.

$$RA \text{ pobl. (\%)} = \frac{0,55(2,2-1) \times 100}{1 + 0,55(2,2-1)} = \frac{0,55 \times 1,2 \times 100}{1 + 0,55 \cdot 1,2}$$

$$RA \text{ pobl. (\%)} = \frac{0,66 \times 100}{1 + 0,66} = \frac{66}{1,66} \approx 40\%$$

La razón de productos cruzados en este ejemplo es  $(2.056 \times 22.387) / (25.891 \times 776) = 2,3$  muy próximo al valor exacto del  $RR = 2,2$ . Si se hubiera usado como aproximación de  $RR$  se obtendría  $RA \text{ pobl. (\%)} = 42\%$ .

El cuadro 13 permite tener una idea de los valores de  $RA \text{ pobl. (\%)}$  para diferentes proporciones de la población expuestas al factor de riesgo y algunos valores selectos de  $RR$ .

**Cuadro 13: Riesgos atribuibles en función de valores selectos de riesgo relativo y de la proporción de la población con el factor de riesgo<sup>1</sup>**

F = Proporción de la población con el factor de riesgo.	RR = Riesgo Relativo			
	2	4	10	12
%				
10	0,09	0,23	0,47	0,52
30	0,23	0,47	0,73	0,77
50	0,33	0,60	0,82	0,84
70	0,41	0,67	0,86	0,89
90	0,47	0,73	0,89	0,91
95	0,49	0,74	0,90	0,92

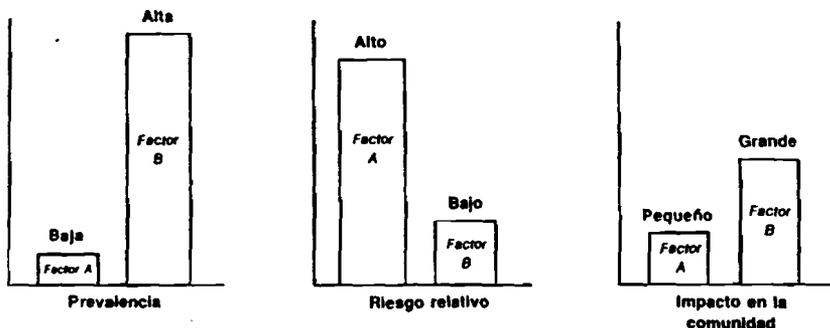
Cuando la frecuencia de una característica o factor es baja, pero el riesgo relativo es elevado, una proporción considerable del riesgo atribuible puede relacionarse con ese factor. Si la frecuencia de un factor de riesgo es alta pero el riesgo relativo es reducido, el riesgo atribuible va a ser más elevado. De los dos factores que contribuyen al RA, la *prevalencia* del factor de riesgo o característica corresponde al que hace

la mayor contribución al valor del RA%. La figura 10 siguiente ejemplifica este punto.

Para determinar prioridades en la comunidad se sugieren tres etapas. La primera consiste en identificar y asignar prioridades, usando el

<sup>1</sup> Adaptado de Lilienfeld, A.M. *Foundations of Epidemiology*, Oxford University Press (1982)

**Figura 9: Efecto de la remoción de dos factores de riesgo de distinta prevalencia**



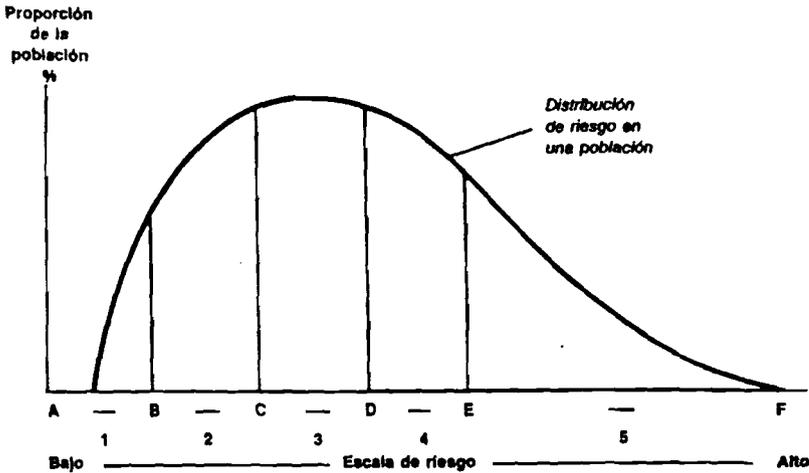
riesgo atribuible, a la contribución de los problemas parciales (toxemia) al problema total (mortalidad perinatal) (ver figura 8). Una vez establecidas las prioridades de los problemas parciales, la siguiente etapa consiste en cuantificar el riesgo atribuible de los factores de riesgo (nulípara) relacionados con los problemas parciales (toxemia). La tercera fase es la de comparar el riesgo atribuible de los factores de riesgo nulípara relacionados con los problemas parciales y determinar cuál de los factores de riesgo es el que más contribuye al problema total (mortalidad perinatal).

#### 5.9.4 Determinación de los límites de inclusión

La determinación de límites de inclusión

afectará significativamente a la estrategia. La determinación de límites de inclusión que permitan que la proporción de la población con el (los) factor(es) de riesgo sea baja aunque el riesgo relativo sea alto, reduce el riesgo atribuible en la población. Sin embargo esto puede hacer que el proceso sea más eficiente. La figura 11 muestra una distribución hipotética de la escala de riesgo usando los RR de 1 a RR 5 y los distintos niveles de inclusión (A,B,C,D,E,F). Puede observarse que dependiendo de la selección del límite de inclusión se podrá reconocer una proporción distinta de los sujetos a riesgo. Será responsabilidad del administrador de salud determinar los límites de inclusión más eficaces, cuando se diseñe la estrategia.

Figura 10



### Ejercicio 9

- (i) Comente la importancia relativa de cada factor listado en el ejercicio 7 desde el punto de vista del riesgo para:
  - a) una mujer en particular
  - b) la comunidad
- (ii) De la lista de factores del ejercicio 8 seleccione aquellos que, en su opinión, se relacionan causalmente con la mortalidad perinatal.
- (iii) Fundamente su respuesta en (ii).
- (iv) De estos factores, ¿cuáles pueden modificarse?
- (v) ¿Cuál es el riesgo atribuible de cada uno de los factores que usted identificó como relacionados causalmente?

## Sección 5.10 Presentación de datos y preparación de informes

### 5.10.1 Consideraciones generales

La preparación de informes y documentos para publicación u otras formas de divulgación de las conclusiones del estudio, es una tarea importante que obliga al investigador a cuestionar nuevamente los datos y a reconsiderar el enfoque empleado en el análisis y la interpretación de resultados. La posibilidad de que observadores independientes examinen los datos produce normalmente un efecto poderoso y beneficioso en el comportamiento del investigador.

En la presentación del informe deben relacio-

narse claramente las preguntas originales e hipótesis de la investigación con los resultados. Hay que resumir los hallazgos en una forma tal que contenga los datos esenciales para indicar el resultado del estudio, con algunas informaciones adicionales que comprendan contribuciones complementarias a las conclusiones. Esa tarea es sumamente difícil para los investigadores que en general consideran indispensable cualquier información derivada de sus estudios.

*Hay que practicar la aptitud de escribir informes breves de las investigaciones.* Debe ser posible escribir un resumen completo en una cuartilla. En la preparación del informe debe usarse como modelo los encabezamientos de secciones del documento original. Además, hay que considerar si es necesario añadir o suprimir encabezamientos.

En algunos casos, hay que preparar en reportes separados información obtenida en distintas etapas. Como parte del proceso de evaluación y divulgación, se podrán preparar informes sobre las diferentes metodologías y los resultados de su aplicación, por ejemplo:

- muestreo;
- métodos de recolección de datos;
- seguimiento de individuos desaparecidos, migrantes o que no responden; y
- técnicas de análisis.

En muchos informes médicos y sociales hay que depurar y simplificar los hallazgos de tal forma que puedan ser fácilmente leídos y asimilados por un grupo heterogéneo.

### 5.10.2 Directrices específicas

El investigador puede adaptar y adoptar diversas directrices o modelos. Una ampliamente citada, es el Informe Estadístico de las Naciones Unidas (1964) titulado "Recommendations for the Preparation of Sample Survey Reports". En ese documento se indica que deben prepararse tres informes.

#### A) Informe preliminar

Puede ser estimulante e importante presentar los resultados preliminares. Se recalca la necesidad de sujetarse a la disciplina de realizar un análisis preliminar, quizás de una submuestra representativa, y de presentar los resultados para comentarios. Esto puede revelar problemas del estudio que necesiten corregirse y que de otra manera permanecerían ocultos. Los informes preliminares pueden ocasionar intercambios de opiniones que constituyen un estímulo eficaz para el equipo de investigación. Es posible mantener y mejorar la motivación demostrando que el hastío de la recolección rutinaria de datos tiene un producto final definido y meritorio.

#### B) Informe general

Ese informe emplea todos los encabezamientos del documento original y le comunica al lector una visión panorámica del estudio, incluyendo los resultados y sus interpretaciones, sin exposición detallada de los aspectos técnicos y estadísticos.

#### C) Informes técnicos

Comúnmente son documentos complementarios y ofrecen datos sobre los aspectos administrativos e institucionales; selección y formación de personal; aspectos estadísticos de muestreo, exactitud, repetibilidad y análisis; eficacia de los métodos de recolección de datos; costos del estudio.

### Ejercicio 10

- (i) Al terminar el estudio epidemiológico los investigadores principales decidieron preparar un resumen de los resultados. Prepare un sumario, que no exceda una página, de los resultados que usted hubiera presentado.

## REFERENCIAS

- Bolt, W.J., J Day, N.E. (1979)  
Synergism & Interaction : Are  
They Equivalent?  
*AM J Epidemiology*. 110(1) : 99-100
- Boy D, A.V. (1979)  
"Testing for association of  
diseases".  
*J. Chronic Dis*. 32 : 667-672
- Fleiss, J.L.: *Statistical Methods for Rates and  
Proportions*. 2nd Edn. New York, Wiley. 1981  
(Intervalos de Confianza para razón de Produc-  
tos Cruzados)
- Kahn H.A : *An Introduccion to Epidemiologic  
Methods*  
Oxford University Press New York 1983.
- Katz, D., Baptista, J., Azen, S.P.,  
Pike, M.C. "Obtaining Confidence Intervals  
for the Risk Ratio In Cohort Studies": *Bio-  
metrics*, 34(3), 469-474, 1978
- Kleinbaum DG, Kuppen L L ,  
Morgenstern H:;  
*Epidemiologic Research*  
Belmont, Calif.  
Lifetime Learning Publications 1982
- Leung, H. M. Kupper, L.L. 1981  
Comparisons of Confidence Intervals For  
Attributable Risk.  
*Biometrics* 37(2) : 293-302
- Mantel, N. (1963) "Chi-Square Tests"  
with One Degree of Freedom: Extension  
of the Mantel-Haenszel Procedure.  
*J. Am. Stat. Assoc.* 58 : 690-700
- Massey, F.J. (1970): "Computers in  
epidemiology" en : *Data Handling in  
epidemiology*  
Ed. W.W. Holland. Oxford  
University Press, London.
- Rothman, K. J., Boice J.D. (1979)  
"Epidemiologic Analysis with a  
Programmable Calculator".  
N I H Publ. N° 79-1649  
Washington D C : Government Printing  
Office.
- Schlesselman, J.J.  
*Case-Control Studies*.  
New York 1982
- Walter SD. (1976) "The Estimation  
and Interpretation of Attributable Risk"  
*Biometrics* 32 : 829-899

## Capítulo VIII

### Evaluación de las intervenciones basadas en el enfoque de riesgo

Sección 8.1	Conceptos básicos sobre evaluación	Sección 8.6	La prueba de los instrumentos para el estudio de evaluación
Sección 8.2	Desarrollo de los criterios de evaluación	Sección 8.7	Instrumentación simultánea de la nueva estrategia de atención maternoinfantil y del proceso de evaluación
Sección 8.3	Diseño del estudio evaluador de la nueva estrategia de riesgo	Sección 8.8	Supervisión
Sección 8.4	Población de las áreas de estudio y de comparación	Sección 8.9	Análisis de los datos y conclusiones
Sección 8.5	Preparación de los planes para el análisis de los datos		

#### Objetivos de aprendizaje

Al finalizar este capítulo, los participantes deberán ser capaces de:

1. Describir los pasos principales del proceso de la evaluación y del componente de supervisión de la estrategia maternoinfantil basada en el enfoque de riesgo.
2. Preparar indicadores que puedan ser usados durante la evaluación de la estrategia maternoinfantil basada en el enfoque de riesgo.
3. Describir los posibles diseños de investigación utilizados en el proceso de evaluación de la estrategia maternoinfantil basada en el enfoque de riesgo.

#### Sección 8.1 Conceptos básicos sobre evaluación

##### Resumen

La evaluación es uno de los aspectos más importantes de la ejecución de la estrategia maternoinfantil basada en el enfoque de riesgo, ya que permite reconocer el impacto de una nueva estrategia en la efectividad, eficiencia, equidad y aceptabilidad de los servicios de salud para resolver las distintas necesidades de salud maternoinfantil de la comunidad. Una de las primeras actividades del proceso de evaluación es la selección de los conceptos, indicadores y fuentes de información a utilizarse. Los indicadores de evaluación deben responder a las hipótesis de trabajo y a los objetivos

de la evaluación. Asimismo, los indicadores seleccionados deben permitir la cuantificación del cambio ocurrido en los niveles de salud de la población. En general existen tres tipos de indicadores de evaluación: a) los indicadores de estructura, b) los indicadores de proceso, y c) los indicadores de impacto o resultado.

El proceso de evaluación<sup>1</sup> de las intervenciones o estrategias de salud basadas en el enfoque de riesgo constituye la base fundamental para reconocer su impacto sobre los niveles de salud de la comunidad.

El objetivo fundamental del proceso de evaluación es identificar en qué medida la estrategia basada en el enfoque de riesgo es una alternativa mejor que las intervenciones tradicionales. Por lo tanto, los tres posibles resultados de la evaluación son:

- A) Que *no* exista cambio en los niveles de salud de la comunidad, según lo reflejado por los indicadores seleccionados para la evaluación.
- B) Que exista un cambio *negativo* en los niveles de salud de la comunidad, es decir, que exista un deterioro de la salud manifestado a través de los indicadores seleccionados.
- C) Que exista un cambio *favorable* en los niveles de salud de la comunidad, el cual se refleje, a su vez, en los indicadores de salud seleccionados para la evaluación.

Un aspecto fundamental de la evaluación es la posibilidad que ofrece de reconocer si el impacto positivo o negativo, o la ausencia de impacto, es consecuencia de la estrategia basada en el enfoque de riesgo y no de otros factores. Este problema metodológico puede controlarse únicamente a través de un cuidadoso diseño del estudio de evaluación y de una selección apropiada de los indicadores.

En términos generales la evaluación puede realizarse a través de tres tipos de indicadores:

- i) indicadores de estructura
- ii) indicadores de proceso
- iii) indicadores de impacto o resultado

Los *indicadores de estructura* son aquellos que reflejan información cuantitativa sobre la infraestructura de los servicios de salud, su nivel de organización, los tipos de servicios disponibles para la población, y los recursos con que se cuenta. El tipo de evaluación que puede efectuarse con estos indicadores presupone que la calidad de la atención a la salud y los niveles de salud dependen de los recursos disponibles.

Los *indicadores de proceso* incluyen información sobre los pasos seguidos en el diagnóstico, el tipo de tratamiento, las condiciones y las formas de acceso al sistema de atención a la salud, los patrones de referencia y el tipo de servicios ofrecidos en los distintos niveles de atención.

Los *indicadores de impacto o resultado* son los utilizados para evaluar el nivel de salud alcanzado como consecuencia de la intervención. Buscan identificar las variaciones y cambios en los niveles de enfermedad, invalidez o muerte en la comunidad, producidos por la nueva estrategia basada en el enfoque de riesgo. La elaboración de este tipo de indicadores requiere en muchas ocasiones períodos largos, ya que los cambios en las tasas de mortalidad o morbilidad se expresan con mayor claridad al mediano plazo. No obstante la inconveniencia del tiempo, estos indicadores son directos y efectivos para evaluar el impacto de las intervenciones en los niveles de salud de la población materno-infantil.

## Sección 8.2 Desarrollo de los criterios de evaluación

### Resumen

Una vez decidida la nueva estrategia de salud materno-infantil basada en el enfoque de riesgo, debe ser evaluada para determinar su efectividad, comparada con la estra-

<sup>1</sup> Para mayor información sobre evaluación consulte: *La evaluación de la atención sanitaria en Planificación Familiar*, Serie No. 569, Organización Mundial de la Salud, Ginebra, 1975.

tegia existente anteriormente. Esta sección trata de la selección de indicadores para tal evaluación. Al seleccionar los indicadores debe considerarse fundamentalmente:

- a) los objetivos e hipótesis de la evaluación (en términos cuantitativos);
- b) el tipo de evaluación;
- c) las características de los individuos con factores de riesgo;
- d) las áreas de estudio y control;
- e) los niveles de atención;
- f) el tiempo;
- g) los costos; y
- h) las actitudes de la población y de los trabajadores de la salud.

### 8.2.1 Introducción

El grupo de investigación se propone medir si la efectividad de la estrategia basada en el enfoque de riesgo es superior, igual o inferior a la de la estrategia actual. Hay varios criterios que se podría utilizar para medir la efectividad, por ejemplo:

- A) ¿Se está proporcionando la atención adecuada a las madres con riesgo elevado?
- B) ¿Está haciéndose un uso más eficiente de los diferentes niveles de atención para las madres con diferentes características de riesgo?
- C) ¿Es más efectiva la nueva estrategia para la solución de problemas maternoinfantiles?
- D) ¿Es más aceptable la estrategia basada en el enfoque de riesgo entre la población que hace uso de los servicios de atención en salud?
- E) ¿Cuenta con la aceptación de parte de los trabajadores de la salud y de los círculos políticos?
- F) ¿Es eficiente en términos financieros y sociales?

El diseño del estudio de evaluación se selecciona de acuerdo con los objetivos de la evaluación y la(s) hipótesis de cambio a estudiarse. Las hipótesis de cambio son cuantificables ya que es posible verificarlas en términos cuantitativos.

Una hipótesis cuantificable debe ser redactada de manera que se evite ambigüedades o imprecisiones tales como:

"La estrategia de riesgo mejoró la cobertura de la atención maternoinfantil". En su lugar debe preguntarse: "¿Qué porcentaje de madres que recibieron al menos una visita prenatal de una enfermera durante el segundo o tercer trimestre del embarazo recibió el paquete de atención descrito en la estrategia de riesgo?" Este tipo de hipótesis cuantificables es de gran utilidad ya que además de orientar sobre el tipo de diseño de evaluación, permite la selección de indicadores y fuentes de datos a ser utilizados. Para cualquier diseño de evaluación se requirirá una combinación de diversos indicadores. La selección de indicadores es una tarea multidisciplinaria. Algunos investigadores estarán preocupados con las medidas de eficiencia y efectividad. Los administradores en el campo de la salud, así como los políticos, insistirán en tener evidencia de la cobertura y de la satisfacción de la comunidad pero también estarán preocupados por la eficacia en función de los costos. El desarrollo de indicadores apropiados debe basarse entre otras cosas, en tareas secuenciales, aspecto examinado al revisar la planificación de la estrategia de intervención, descrita en el capítulo VII.

Gran parte de los datos necesarios para una evaluación surge de los registros de rutina de los servicios de atención personal en salud y de los procedimientos de intervención que constituyen la estrategia basada en el enfoque de riesgo. Sin embargo, en cada etapa del desarrollo de la estrategia de riesgo debe considerarse qué información adicional se necesita para la evaluación, que no sea requerida para la atención individual y que deberá recogerse a través de procedimientos especiales. Es posible que haya diversos tipos y fuentes de información utilizables para escoger los indicadores para la evaluación. Por ejemplo, un indicador que puede utilizarse en la evaluación de los programas de control prenatal es la reducción de la incidencia de bajo peso al nacer; por otra parte, el tiempo necesario para observar una modificación significativa en la

mortalidad infantil, así como el tamaño de la población que deba seguirse para medir ese cambio, pueden resultar imprácticos. Los indicadores para medir el proceso e impacto de la intervención también son útiles en otros aspectos del proceso de atención. Las medidas de reducción de los factores de riesgo en la población también deben ser consideradas como indicadores de la evaluación.

### 8.2.2 **Importancia de una línea básica**

Se entiende por línea básica el conjunto de datos fundamentales reunidos con anterioridad a la intervención, a manera de información mínima, que refleje las condiciones de salud en la población materno-infantil antes de la puesta en práctica de la estrategia de riesgo y que permita medir los cambios y efectos producidos por la estrategia basada en el enfoque de riesgo.

Los indicadores de la evaluación deben utilizar la información de la línea básica con el fin de cotejar los cambios que han ocurrido en el área de estudio. Los criterios utilizados en la evaluación pueden considerar las siguientes opciones:

- A) la situación existente en el área de estudio antes de la aplicación de la estrategia, o
- B) la situación en un área comparable (distinta a la población en estudio) en relación con la situación prevalente en el área donde se aplicó la estrategia de riesgo.

La selección de indicadores guarda una relación estrecha con el diseño del estudio para la evaluación. A su vez el diseño del estudio reflejará la existencia o no existencia de grupos control y el tipo de información básica que será utilizada.

### 8.2.3 **Tipos y fuentes de indicadores para la evaluación**

La información necesaria para el empleo de los indicadores de evaluación puede provenir de diversas fuentes de datos incluyendo los contac-

tos entre el personal de salud y la población. Los procedimientos de evaluación que finalmente se adopten dependerán, hasta cierto grado, de la habilidad del personal que realiza la evaluación, de sus destrezas, así como el tiempo y recursos disponibles.

Los procedimientos usados para la recolección de datos deben ajustarse a las posibilidades y limitaciones del caso (recursos, sistema de registros y de informe, etc.) Los datos necesarios para la evaluación pueden obtenerse de:

- a) registros sanitarios
- b) estadísticas rutinarias
- c) encuestas de corte transversal, especialmente asociadas con el diseño antes/después
- d) evaluación continua de grupos de madres a través de cuestionarios o entrevistas
- e) historias clínicas y estadísticas hospitalarias
- f) combinación de las anteriores.

Recuérdese que los indicadores seleccionados deben ser cuantificables, por ejemplo: proporción de reducción de la tasa de bajo peso al nacer, de reducción de la tasa de morbilidad por gastroenteritis, etc.

Las actitudes del personal y de la comunidad deben evaluarse a través del uso de instrumentos de investigación adecuados. Los estudios de tiempos y movimientos y otros métodos de análisis de tareas, pueden utilizarse a fin de medir cambios en la actividad del personal. Toda evaluación debería incluir un análisis de costos comparando antes y después de la intervención, así como la relación costo-efecto de las dos estrategias de atención.

### **Ejercicio 1**

Basado en el conocimiento de Ficticia y en la nueva estrategia materno-infantil desarrollada en el capítulo VII:

- (i) Formule una hipótesis cuantificada del cambio que resultará de la aplicación de la nueva estrategia de riesgo dentro de un período dado.

- (ii) Mencione tres tipos de indicadores que usted puede utilizar a fin de evaluar la nueva estrategia.
- (iii) Para cada uno de los tres tipos de indicadores, desarrolle los indicadores específicos de evaluación que usted usaría.
- (iv) Para cada indicador específico, mencione las limitaciones que usted usará para obtener los datos necesarios para su construcción.
- (v) Para cada indicador específico, mencione las limitaciones posibles tanto para la obtención de tales datos como para su uso en el procedimiento de evaluación.

## Sección 8.3 Diseño del estudio evaluador de la nueva estrategia de riesgo

### Resumen

La evaluación debe ser simultánea a la ejecución de la nueva estrategia materno-infantil. Dado que el diseño experimental no siempre puede ser utilizado en la evaluación, debe darse especial atención a la línea básica y a la selección de áreas de comparación al diseñar el estudio de evaluación. Pueden emplearse diseños "antes y después", con o sin área o grupo de comparación, y sus diversas combinaciones. Es importante reconocer que el diseño del estudio de evaluación tome en cuenta las limitaciones existentes de orden ético, financiero, temporales y de factibilidad.

### 8.3.1 Diseño del estudio evaluador de la nueva estrategia de riesgo

En la literatura sobre metodología de evaluación de los servicios de salud se menciona gran cantidad de diseños posibles para el estudio de evaluación de nuevas intervenciones. Sin embargo, el primer punto a señalar es que el enfoque metodológico para evaluar la estrategia de intervención debe reflejar muy de cerca la metodología adoptada para la aplicación de la propia estrategia, aunque dentro del esquema de tra-

bajo puede haber cierto grado de flexibilidad. Por ejemplo a los efectos de la evaluación, puede realizarse encuestas *ad hoc* que no sean parte de la estrategia misma de intervención.

De acuerdo con la estrategia de intervención, el diseño del estudio de evaluación puede basarse en:

- A) Un simple diseño *antes/después*, sin área (s) de comparación. Este diseño se basa en encuestas de corte *transversal*, la primera realizada antes y la segunda realizada *después* de la intervención.
- B) Un diseño *antes/después*, con áreas de comparación, donde toda la información se basa en estadísticas existentes, recolectadas rutinariamente. (este diseño es apropiado cuando las estadísticas de salud rutinarias son relativamente completas y de calidad aceptable).
- C) Un estudio *prospectivo* en el área de intervención (información especialmente recolectada), con comparaciones *antes/después*, utilizando indicadores reunidos rutinariamente en el área de comparación.
- D) Un diseño *antes/después* en el área de intervención con estudios de corte *transversal* y usando los datos nacionales como área de comparación.

Para la evaluación de una nueva estrategia de intervención, la situación ideal sería tener una o más áreas de comparación donde no se introduzcan cambios en el sistema de atención. Tanto en las áreas de estudio como en las de comparación, debería evaluarse los niveles de daños a la salud en los grupos de riesgo y de no riesgo y comunicarse las diferencias encontradas.

Si hay diferencias estadísticamente significativas en los niveles de daños a la salud en los grupos de riesgo y no riesgo entre las áreas de estudio y comparación, podría decirse que la nueva estrategia introducida modificó la situación de salud en el área de estudio, y si tal cambio fué positivo, que es más efectiva que la estrategia anterior a la intervención.

### 8.3.2 Factores que afectan el impacto de la intervención

Las personas y los sistemas de atención a la salud no pueden controlarse como animales de laboratorio. Inevitablemente surgirán cambios en el área de estudio y en las áreas de control, independientemente de la nueva estrategia de riesgo. En ocasiones se construyen nuevas carreteras, se cambia al personal sanitario, se presentan epidemias, las actitudes y percepciones de las personas varían. Generalmente ninguno de estos factores puede ser controlado por los investigadores y, sin embargo, pueden afectar las áreas de estudio y comparación en grados diferentes y pueden modificar el sistema de salud y los problemas sanitarios independientemente de la intervención de la nueva estrategia. Las áreas de comparación o de control en el sentido clásico de *diseño experimental*, pueden, por lo tanto, no resultar factibles.

A continuación se presentan ejemplos de los factores que pueden afectar los resultados y que deben ser considerados en el estudio de la evaluación.

#### A) El efecto de Hawthorne

El solo hecho de que se muestre interés en las personas y en sus actividades con frecuencia mejora la moral y la actuación en el trabajo, sin que ello guarde relación con el proceso de trabajo, la organización o los recursos. Inevitablemente, las áreas de estudio recibirán más atención que las áreas de control. Esto, por sí solo, puede cambiar la actuación en el trabajo. El llamado "efecto de Hawthorne" puede compensarse brindando algún adiestramiento al personal en las áreas de control, o a través de otras formas de atención o de apoyo, similares a las de la nueva estrategia aplicada en el área de estudio, por ejemplo un sistema de supervisión más frecuente.

#### B) El efecto de divulgación

Cuando se aplica una nueva estrategia en un área de estudio, con frecuencia resulta difícil

evitar que la comunidad y el personal sanitario en el área de control conozcan el hecho. Esto puede estimular al sistema de atención a la salud en las áreas de control a introducir cambios difíciles de medir, y que modificarán el área en tal medida, que sería mejor ya no considerarla como de comparación, es decir como un área donde no se introducen cambios en las estrategias. Para evitar el efecto de divulgación, a veces se selecciona las áreas de comparación a distancia de las áreas de estudio aunque, si es posible, dentro de la misma área administrativa. Otra forma de resolver este problema es el uso del comportamiento promedio de un país completo como área de comparación. Esto tiene sus propios problemas de costo y disponibilidad de datos.

### 8.3.3 Consideraciones éticas

No es ético ~~retacear~~ los servicios de salud a las personas con alto riesgo, con el pretexto de que el diseño de investigación ha colocado a esa persona en un área de control. Esto hace imposible dividir a la población en el grupo de control en grupos con riesgo y sin riesgo y no tomar acciones especiales para las personas con riesgo. Quizá esto podría hacerse sobre la base de un diseño retrospectivo, pero sólo en casos donde se reúna la información rutinariamente. Si hay que recoger datos adicionales, sería poco ético no usar aquella información para aplicar la tecnología existente para madres y niños que se prevé que tengan alto riesgo. Sin embargo, las consideraciones éticas no deberían impedir la evaluación de una estrategia particular, cuyos beneficios aún no han sido probados. Existe una analogía interesante sobre la evaluación del manejo de hipertensión leve. En los Estados Unidos, el Programa de Seguimiento de Detección de Hipertensión (HDFP) midió la morbilidad y mortalidad de un grupo de hipertensos leves que estaban expuestos a una estrategia basada en riesgo con una programación cuidadosa de atención escalonada. El grupo de control recibió la atención usual disponible de un especialista o de los servicios de atención primaria.

## Sección 8.4 Población de las áreas de estudio y comparación

### Resumen

Es necesario determinar si los planes seleccionados para el estudio de la estrategia de riesgo y su evaluación:

- A) pueden instrumentarse en la población seleccionada para el estudio;
- B) son congruentes con los datos disponibles para su comparación; y
- C) son congruentes con los métodos disponibles de recolección de datos.

Una vez que los indicadores de evaluación, el diseño y los métodos del estudio hayan sido seleccionados, debe procederse a escoger las poblaciones de estudio, incluyendo, si se usara, el área de comparación. Aquí se revisa lo referente a la población de estudio, no desde el punto de vista de ejecución de la estrategia sino en términos de la necesidad de recolectar información para propósitos de evaluación. Es importante considerar cómo se articula el plan propuesto para la evaluación con la ejecución de la estrategia del enfoque de riesgo.

#### 8.4.1 Población de estudio

Para propósitos de evaluación, la población bajo estudio incluye tanto las áreas de estudio (intervención) como las de comparación (si se usa). Las áreas de estudio son aquellas donde se aplica la nueva estrategia. En las áreas de comparación no se introduce ningún cambio.

Generalmente se define las áreas de estudio y de comparación dentro de las mismas áreas administrativas de atención sanitaria: sectores, distritos, o cualquier otra división administrativa del sistema de atención a la salud. Según el diseño de investigación que se haya seleccionando, se incluyen o no áreas de comparación. El tipo y el grado de detalle de los datos recolectados en tales áreas pueden ser distintos. Por ejemplo, algunas veces sólo los datos que se recogen en forma rutinaria serán utilizados en las áreas de comparación.

Si se utiliza en la evaluación los datos recolectados rutinariamente, es necesario que las

áreas de estudio y comparación sean parte de las mismas áreas administrativas sanitarias.

El tamaño mínimo de la población a estudiarse dependerá de los siguientes criterios:

- el grado de cambio previsto en los principales indicadores de evaluación; y
- el grado de certeza estadística requerido para llegar a conclusiones en cuanto a los efectos de la intervención.

Los métodos utilizados para seleccionar el tamaño de las poblaciones de estudio y comparación fueron descritos en el capítulo V.

Las consideraciones estadísticas para determinar el tamaño mínimo de la población, la cantidad de datos a recolectarse, el tiempo disponible y el grado de certeza estadística requerida, deberán revisarse a la luz de las hipótesis que van a someterse a prueba.

Para algunos indicadores de evaluación, sólo puede ser necesaria una muestra de la población bajo estudio; este puede ser el caso para indicadores donde se prevén grandes cambios en un período corto.

Por otra parte se puede seleccionar y utilizar el mismo muestreo para todos los indicadores de evaluación a fin de usar un procedimiento uniforme. No existe una regla para esto y deberá considerarse las ventajas y desventajas de los diferentes enfoques.

Algunas de las principales características que se utilizan para seleccionar las áreas de control y las de estudio en una situación particular son las siguientes:

- niveles de problemas de salud
- edad/sexo/grupos étnicos/"status" socioeconómico
- desarrollo especial de proyectos en marcha o planificados
- tipo de cobertura de atención en salud
- comunicaciones
- servicios generales.

La situación ideal consiste en reunir una población de estudio y de comparación lo más similar posible en cuanto a las características más importantes, exceptuando, desde luego, el

procedimiento de *intervención* que se está examinando. Por lo tanto la diferencia de los resultados debería estar dada por el impacto de la instrumentación de la intervención. No debe olvidarse que esta es una situación ideal y que siempre se encuentra diferencia entre las poblaciones que debe considerarse en el análisis.

Un problema frecuente en los estudios comunitarios de evaluación es que las intervenciones se realicen en áreas de estudios universitarios y áreas de adiestramiento que suelen ser muy diferentes de las otras áreas de salud en términos de tipo de personal y de otros recursos. Pueden no ser diferentes a las del resto del país en términos de recursos, pero sí hallarse mejor administradas, y es posible que el personal y la población estén acostumbrados a ser sujetos de estudios. A menudo difieren del resto de las áreas de salud en cuanto a características que son difíciles de describir y cuantificar, pero que tienen un profundo impacto en los resultados del estudio y en su aplicabilidad en el resto del país. Sin embargo, en vista de los aspectos prácticos de las áreas de influencia de los centros de salud universitarios los estudios de evaluación con frecuencia se llevan a cabo allí. Con las reservas del caso, los hallazgos de tales estudios en ocasiones han resultado útiles a pesar de las dudas que existen sobre ellos.

#### 8.4.2 Población de comparación

Deberá seleccionarse las áreas de comparación similares a las áreas de estudio en cuanto a sus características más relevantes. Es importante que tanto las áreas de estudio como las de comparación correspondan a las áreas administrativas para que las estadísticas recolectadas rutinariamente puedan utilizarse, pues de otra manera deberá prepararse nuevas agrupaciones estadísticas causando problemas y confusión.

Finalmente, es importante conocer los planes más significativos de desarrollo existentes que afecten a las áreas de estudio y comparación. Asimismo es necesario observar los planes de agrupamiento de habitantes o de migración, ma-

yores recursos de atención en salud, nuevas industrias u otros proyectos de desarrollo. Estas circunstancias deben ponerse de manifiesto y tomarse en consideración al interpretar los hallazgos de los estudios. No sería ético tratar de aplazar o impedir el desarrollo de un sector específico en las áreas de control o de estudio. Así, si se prevén cambios importantes únicamente en el área de estudio o en el área de control, deberá considerarse la selección de otras áreas u otro diseño para la evaluación de la estrategia de riesgo.

Sin duda alguna, los cambios en la salud y en la atención de salud resultantes de modificaciones en el desarrollo general en el área pueden ser mayores que los provenientes de las nuevas estrategias de riesgo. Si esto fuera así los indicadores de evaluación deben ser reconsiderados y modificados para que reflejen en mejor forma los cambios específicos esperados en la población materno-infantil con la introducción de la estrategia de riesgo.

#### Ejercicio 2

- (i) Con el fin de evaluar la nueva estrategia de intervención materno-infantil, use su conocimiento general de Ficticia y sus respuestas al ejercicio 1 para describir las áreas de estudio y de comparación que usted eligiría para esta evaluación.
- (ii) Señale como evaluaría el proceso de identificación de las mujeres embarazadas de alto riesgo en un área de la región montañosa de Ficticia.
- (iii) Calcule el tamaño de la muestra requerida para el área de estudio y control para evaluar el impacto de una intervención que espera reducir la tasa de mortalidad perinatal del 55 por mil al 40 por mil— (use  $\alpha = 0.05$ , poder = 0.90).

5344 EMBD-92-2

## Sección 8.5 Preparación de los planes para el análisis de los datos

### Resumen

En esta etapa deberá tenerse el plan completo para el análisis de la información, y contarse con una lista de los datos a obtenerse. También las fuentes, los métodos y la frecuencia de la recolección. Se debe preparar los esqueletos de las tablas y cuadros a utilizar. Con base en estos esqueletos se ajusta el diseño de la evaluación que se esquematóz previamente.

### 8.5.1 Esquema general para la construcción de esqueletos de las tablas utilizadas en la evaluación

Los datos deben ser transferidos, codificados, agrupados y presentados adecuadamente. Es importante que estas actividades se realicen en forma eficiente. Se debe planificar antes de empezar la recolección de los datos ya que la planificación puede tener implicaciones en el tipo de procedimientos, los métodos de verificación de datos, la codificación, la revisión de la consistencia, etc., aspectos a los que se ha hecho referencia en el capítulo V. Los mismos principios se aplican al manejo y control de datos en el estudio de evaluación.

Los planes para el análisis de los datos deben ser realizados previamente, incluyendo la preparación de los esqueletos de los cuadros requeridos. La mayoría de los cuadros en esta parte del estudio, será de dos entradas, dando la información relacionada por cada indicador antes y después y/o en áreas de estudio y de comparación.

La preparación de estos esqueletos de los cuadros o tablas de  $2 \times 2$  sin contar con los datos reales puede parecer una tarea tediosa, pero evita la necesidad de llamar a los estadísticos para que digan cómo se puede utilizar los datos recolectados, que suele surgir al final de los estudios.

### Ejemplo 1: Diseño de evaluación antes/ después sin grupo de comparación

Una de las hipótesis generales que pudo formularse en el ejercicio 1 (i) es: "Como consecuencia de la intervención de la estrategia de riesgo se espera que en un año, en las áreas de estudio se reduzca la tasa de bajo peso al nacer del 25% al 15% de todos los nacimientos". El esqueleto del cuadro en que se presentaría esta información podría ser el siguiente:

#### Diseño antes/después sin grupo de comparación

	Año 1 (antes) No. %	Año 2 (después) No. %
Nacidos + 2.500 gms	75	85
Nacidos - 2.500 gms	25	15
<b>Total</b>	<b>100%</b>	<b>100%</b>

El grupo de trabajo habrá decidido el total de nacimientos ("n") que será recolectado a fin de obtener una diferencia significativa entre antes y después. Esto representa el tamaño de la muestra, que se calcula siguiendo los procedimientos enunciados en el capítulo V. Puede utilizarse diversas pruebas de significancia estadística para comparar el porcentaje de bajo peso al nacer existente antes de la intervención (25%) con el encontrado posteriormente a la intervención (15%) y establecer si la diferencia encontrada es estadísticamente significativa.

### Ejemplo 2: Diseño de antes/después usando una población de comparación.

La hipótesis de este tipo de diseño podría enunciarse así: "Como consecuencia de la intervención de la estrategia de riesgo se espera que el porcentaje de bajo peso al nacer disminuya de 25% a 15% entre todos los nacidos vivos en el área de estudio en un año y que no haya cambio en el área de control".

A partir del siguiente diagrama, puede observarse que existe un cambio (disminución) del porcentaje de recién nacidos con bajo peso en el área de estudio donde se introdujo la nueva intervención, y sin embargo no se observa modificación en el porcentaje de niños con bajo peso al nacer en el área de comparación donde no se introdujo la nueva intervención. El siguiente paso es aplicar una prueba de significancia estadística para comparar las dos áreas y los cambios observados entre los años 1 y 2. El cuadro correspondiente a este diseño es el siguiente:



**Diseño antes y después usando población control**

	Año 1 (antes)	Año 2 (después)
	Nacimientos con menos 2.500 gms (%)	Nacimientos con menos 2.500 gms (%)
AREA A (con intervención)	25	15
AREA B (Sin intervención)	25	25

### Ejercicio 3

- (i) Prepare el esqueleto de un cuadro que incluya información para evaluar cada uno de los tipos de indicadores presentados en el ejercicio 1 (ii) indicadores de estructura, proceso y resultado o impacto).
- (ii) Para cada uno de los esqueletos preparados sugiera qué tipo de comparaciones y pruebas estadísticas aplicaría.
- (iii) Discuta qué otras variables pueden estar asociadas con el resultado encontrado.

## Sección 8.6 La prueba de los instrumentos para el estudio de evaluación<sup>1</sup>

### Resumen

Los formularios, cuestionarios, hojas de resumen, e instrucciones de codificación deben prepararse y probarse en el terreno para hacer frente a la necesidad de evaluación del estudio. La metodología ha sido discutida en el capítulo V, sección 5.6. Es importante establecer y mantener un enlace activo entre la comunidad y el personal de salud. La participación del personal de salud es vital, por lo cual deberá ser informado de todos los aspectos relevantes de la evaluación.

Los principios que se señalaron en el capítulo V (sección 5.6) deben revisarse en esta etapa de la evaluación. La tarea consiste en vincular los instrumentos de recolección de datos usados en la evaluación con los registros rutinarios de identificación, referencia y manejo clínico de los casos de riesgo. Asimismo deberá contemplarse el tipo de adiestramiento que el personal de salud recibirá para asegurar el cumplimiento de la estrategia de riesgo.

### 8.6.1 Enlace entre la comunidad y el personal

Antes de la aplicación de la nueva estrategia, debe consultarse e informarse nuevamente a los trabajadores sanitarios y los miembros de la comunidad relacionados con salud, así a como la comunidad en general. Los investigadores deberán discutir y explicar el proceso y propósito de todos los procedimientos y registros. A través de estas consultas surgen con frecuencia ideas útiles sobre la forma de aplicar la estrategia en la práctica.

Las actividades de información y consulta están estrechamente relacionadas con la prepa-

<sup>1</sup> Para mayor información sobre cómo desarrollar un manual de atención primaria véase:

The Primary Health Care Worker  
A Working guide. Guidelines for Training  
Guidelines for adaptation  
World Health Organization, Geneva, 1980

ración de manuales y guías para el adiestramiento del personal y para la prueba de los nuevos procedimientos.

### 8.6.2 Preparación y prueba de los instrumentos técnicos para el personal

La nueva estrategia materno-infantil a probarse puede sugerir cambios en cuanto a la estructura, organización, procedimientos técnicos y manejo del sistema de salud. Más adelante deberá documentarse y desarrollarse en detalle todos los cambios. La participación del personal de salud en esta etapa *es vital* por dos razones:

- A) para asegurar que las instrucciones preparadas sean factibles y aceptables, y
- B) para asegurar la colaboración del personal en la aplicación de los nuevos procedimientos.

Con frecuencia las instrucciones existentes están dispersas en una masa de diferentes cartas, circulares, manuales, etc. Esta es una buena ocasión para revisar todo ese material, retener lo que es útil e insertar nuevo material. Todo el material puede entonces reunirse en manuales relacionados con diferentes niveles de atención a la salud y para diferentes tipos de trabajadores de salud, asegurándose la consistencia en los distintos manuales. La elaboración de los manuales es difícil y sólo a través de pruebas de campo y de una revisión constante se puede lograr que sean realmente efectivos.

Asimismo, deberá prepararse los lineamientos generales para la preparación de manuales. Cuando las diferentes partes de los manuales estén terminadas, se discutirán y probarán con el personal en la forma práctica, proceso que puede tomar de uno a dos meses; pero vale la pena tomar este tiempo ya que asegurará el logro de una estrategia realista y práctica.

Pequeños grupos pueden hacer borradores de las diferentes partes del manual. Un comité nombrado al efecto, así como los miembros del equipo de salud que llevará a cabo el estudio,

revisarán el borrador del manual y cuando se haya llegado a un acuerdo general, se puede probar la efectividad del procedimiento. Para propósitos de evaluación, interesa particularmente que la información registrada, que haya sido generada por cada procedimiento, sea completa y confiable.

Las personas que elaboraron el borrador del manual deben explicar o demostrar los procedimientos a los trabajadores de salud y luego proseguir con una aplicación práctica. De esta forma los registros sanitarios son probados y modificados en una situación real antes de aplicarse en el área de estudio.

Una vez aprobados y documentados los diferentes aspectos de la instrumentación de la nueva estrategia, se elaboran los materiales de enseñanza que se empleará para adiestrar al personal en uso del manual, organizándolo para entrenar a *todo* el personal de salud en el área de estudio.

La secuencia utilizada en la elaboración de manuales y el desarrollo de las actividades de adiestramiento podría ser la siguiente:

- a) Borrador de los nuevos procedimientos basados en la nueva estrategia.
- b) Revisión y aprobación del primer borrador por parte del comité encargado de su elaboración y por personal sanitario seleccionado.
- c) Prueba de los nuevos procedimientos a pequeña escala.
- d) Modificaciones basadas en la experiencia práctica.
- e) Revisión de la consistencia de los borradores de los manuales.
- f) Preparación de material de enseñanza basado en lo anterior.

### 8.6.3 Adiestramiento del personal

El adiestramiento deberá cubrir todos los componentes de la nueva estrategia y también habrá de reforzar los elementos importantes de las estrategias existentes. El tipo y tamaño de las sesiones de adiestramiento dependerán de su

contenido, el nivel y número de personas a ser adiestradas, la disponibilidad de instructores, las facilidades de adiestramiento y otros recursos. Ha de formularse objetivos educacionales específicos para hacer frente a las necesidades de los distintos tipos de personal de salud. Si el personal involucrado es poco, la forma más apropiada es el adiestramiento en servicio, que puede variar desde un día hasta una semana en cada una de las clínicas, trabajando con cada uno de los trabajadores sanitarios. En otras situaciones puede organizarse grupos de adiestramiento más formales.

## Sección 8.7 Ejecución simultánea de la nueva estrategia de atención maternoinfantil y del proceso de evaluación

### Resumen

Una vez que estén listos los borradores de los manuales, se haya iniciado el adiestramiento del personal y la comunidad esté debidamente informada, la nueva estrategia debe ser aplicada en la(s) área(s) de estudio. Este proceso se define y discute en el capítulo VII. Aquí preocupa la relación de los procedimientos de evaluación con este proceso. Esta sección recuerda al lector que introducir modificaciones en los sistemas de salud es una labor difícil, que requiere mucha paciencia, atención y tiempo. Esto mismo es válido en cuanto a la evaluación.

### 8.7.1 Ejecución de la nueva estrategia en las áreas de estudio

*Resolución de problemas en los procedimientos de la evaluación*

Esta etapa se refiere a la ejecución de la estrategia de atención maternoinfantil que va a ser incorporada en el área de estudio.

Si se planea realizar un número considerable de cambios en los enfoques y procedimientos, la puesta en práctica plena de la nueva estrategia

puede requerir varios meses. Se espera tener un período de ajuste en el cual se identifique los problemas resultantes de la aplicación de la estrategia. Las dificultades deben ser solucionadas inmediatamente.

Durante este período es importante la disponibilidad del grupo de investigación para guiar y ayudar al personal de campo, no limitada sólo a visitas cortas, ya que puede ser necesaria la participación total del grupo de investigación durante varios días o semanas.

La recolección de datos para la programación de la evaluación se inicia en este período. Debe supervisarse cuidadosamente los datos recolectados, ya que es una experiencia común que el nuevo sistema sólo se estabilice después de un período de dos o tres meses, o aún más largo. Los datos obtenidos durante este período con frecuencia pueden descartarse en la evaluación final ya que no reflejan el funcionamiento normal del nuevo sistema. Sin embargo, esos datos son muy importantes para propósitos de supervisión y para determinar si hay una fecha límite después de la cual los datos deberán incluirse en la evaluación final.

### 8.7.2 Recolección de datos

La metodología de recolección de datos ha sido descrita en el capítulo V. Esta sección tiene por objeto únicamente insistir en que el control de calidad en la recolección de datos es tan importante para los propósitos de evaluación como para la atención médica. Debe prepararse procedimientos cuidadosos y detallados para la recolección de los datos, que sean supervisados cuidadosamente.

## Sección 8.8 Supervisión

### Resumen

La evaluación continua y la supervisión del funcionamiento de la atención maternoinfantil permitirán detectar los problemas básicos encontrados en la ejecución de la nueva estrategia de riesgo. Asimismo proveerán la información necesaria para tomar las acciones correctivas requeridas. Gran

parte de la información usada en la evaluación y la supervisión puede utilizarse en la evaluación final.

- b) ¿Son de buena calidad y completos?
- c) ¿El manejo de los datos se está realizando de acuerdo con lo planeado?

### 8.8.1 El proceso de supervisión

La lista de actividades a ser supervisadas es casi interminable, por lo que deberá hacerse una selección de las más importantes o críticas, a fin de evitar la reunión de información excesiva, que luego no podrá ser analizada. El proceso de supervisión debe cubrir al menos dos áreas problemáticas:

#### A) Relación con el sistema de salud ya establecido

En esta área se buscará dar respuesta al siguiente tipo de preguntas:

- a) ¿El trabajador de salud está aplicando la nueva estrategia de la salud según lo planeado?
- b) ¿Las actividades de apoyo están cumpliéndose según lo planeado (adiestramiento, cambios de procedimientos, cambios en el suministro de drogas, etc.)?
- c) ¿Cuál es la actitud del personal de salud hacia los cambios?
- d) ¿Cómo reacciona la comunidad ante la nueva estrategia?
- e) ¿La mejoría en la atención maternoinfantil trae como consecuencia un descuido de otras actividades prioritarias?
- f) ¿Es exitoso el adiestramiento del personal en términos de nuevo conocimiento? ¿En términos de cambios de prácticas?
- g) ¿Están siguiendo los individuos (pacientes) la nueva estrategia? ¿Hasta qué grado los casos referidos acuden al próximo nivel de atención?

#### B) Relación con el estudio de riesgo y los procedimientos de evaluación

De la misma manera, debe contestarse el siguiente tipo de preguntas:

- a) ¿Los datos están siendo recolectados a tiempo?

El listado de funciones y actividades podría ser mucho más extenso. Es importante que el grupo de investigación considere todos los aspectos de la ejecución de la estrategia y esté conciente de que puede equivocarse en la selección de los indicadores apropiados, la determinación del método y la frecuencia de la recolección de datos. No obstante en esta etapa no puede proponerse un esquema fijo. La preparación cuidadosa del sistema de supervisión trae como resultado un mejor control del estudio y menos problemas en las etapas finales de la evaluación.

Durante la aplicación de la estrategia surgirán nuevas ideas de supervisión que, de ser posible, deben ser incluidas. Las actividades de supervisión son útiles tanto para sugerir acciones correctivas como para dar idea de los problemas que necesitan mayor clarificación a través de estudios cortos y como parte de la evaluación final.

### 8.8.2 Algunos ejemplos de supervisión

A) El proceso de referencia de los casos con alto riesgo por el personal técnico auxiliar hacia el médico debe ser cuidadosamente supervisado. Este proceso de supervisión ha de establecerse en todos los niveles de la cadena de referencia.

El grupo de investigación decide utilizar como indicador de supervisión el número de casos referidos para ser atendidos por el médico y que en efecto llegaron a ese nivel de atención.

El número de casos referidos así como el número de aquellos casos que habiendo sido originalmente referidos llegaron a ser atendidos por el médico, puede documentarse una vez cada trimestre, preparándose un cuadro para presentar esta información a los equipos de atención a la salud. A continuación se muestra como ejemplos, dos esqueletos modelos que podrían ser utilizados en la supervisión.

Cuadro 1.\* Evaluación del sistema de referencia de pacientes por el personal auxiliar al médico por trimestre. Ficticia 1980.

	Año				Total
	Enero Marzo	Abril Junio	Julio Set.	Oct. Dic.	
Nivel personal auxiliar					
1. No. casos referidos al médico por el auxiliar.	20				
2. Total de casos que contactaron al auxiliar	100				
Nivel médico					
3. Total de casos atendidos por el médico	30				
4. No. de casos referidos y atendidos	15				
5. No. de casos que originalmente acudieron al médico	15				

\*En este cuadro se incluye información del trimestre enero-marzo para ejemplificar los indicadores de evaluación usados en el cuadro 2.

Cuadro 2. Evaluación del sistema de referencia de pacientes por el personal auxiliar al médico por trimestre. Ficticia 1980.

	Año 1980				Total
	Enero Marzo	Abril Junio	Julio Set.	Oct. Dic.	
A. % de casos referidos que fueron atendidos por el médico. (4/1)*	15/20				
B. % de todos los casos atendidos por el médico (3/2 + 5)	30/100 + 15				

\* Los números entre paréntesis se refieren al cuadro número 1

Sobre la base de los resultados obtenidos en estos cuadros puede tomarse acciones correctivas en caso de que el porcentaje de casos referidos que fueron atendidos por el médico sea bajo, o en caso que el porcentaje total de pacientes atendidos por el médico sea alto o tenga una tendencia a aumentar.

B) El personal auxiliar técnico será adiestrado para establecer un puntaje simplificado de riesgo en mujeres embarazadas. Los médicos supervisores verificarán los puntajes de riesgo

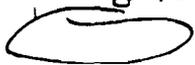
obtenidos por el personal auxiliar técnico, examinando los registros de salud de un número determinado de primeras visitas prenatales durante cada trimestre del estudio. Puede prepararse un cuadro similar al que se presenta para evaluar la capacidad de cada funcionario auxiliar, para seleccionar correctamente las pacientes de riesgo. De estas tablas de  $2 \times 2$  puede calcularse el porcentaje de acuerdo existente entre el personal auxiliar y el supervisor con respecto al puntaje de riesgo.

**Cuadro 3. Verificación por el médico del puntaje de riesgo determinado por el personal auxiliar técnico (número).**

		Puntaje determinado por el médico		Total
		Alto riesgo	Bajo riesgo	
Puntaje determinado por el personal auxiliar técnico	Alto riesgo	A	B	A + B
	Bajo riesgo	C	D	C + D
Total		A + C	B + D	Total

Medición del grado de acuerdo  $(A + D) / \text{Total} \times 100 = \% \text{ total de acuerdo}$

*ojo ver*



C) Puede utilizarse el peso al nacer como el indicador principal al evaluar el impacto de la intervención. Por lo tanto son importantes la calidad de la información y su uniformidad en los distintos hospitales o centros de salud. Considérense dos componentes de este ejemplo de supervisión: primero, la estandarización de la toma del peso del recién nacido (para esto debe

entrenarse a los participantes y los instrumentos deben ser estandarizados y calibrados previamente); segundo, el momento de la toma del peso puede ser un factor que influya en la uniformidad y validez de la información. Un cuadro similar al que se presenta a continuación puede ser de utilidad para supervisar esta variable.

**Cuadro 4. Tiempo transcurrido entre el nacimiento y la toma del peso neonatal.**

Tiempo transcurrido en horas	Centro de salud A		Centro de salud B		Total	
	No. de Niños	%	No. de Niños	%	No.	%
	0-12					
13-24						
25 y más						
Total						

### Ejercicio 4

- (i) Basado en sus conocimientos de Ficticia y la nueva estrategia de atención maternoinfantil seleccione un indicador que le permita supervisar uno de los componentes que se listaron en la parte B de la sección 8.8.1
  - Para ese indicador establezca
    - la posible fuente de los datos
    - la forma de recolección
    - las posibles limitaciones que puede encontrar.
  - (ii) Prepare un cuadro para presentar al equipo de salud los resultados obtenidos.
  - (iii) ¿Puede ser usado este indicador en la evaluación final de la estrategia? ¿Cómo?

## Sección 8.9 Análisis de datos y conclusiones

### Resumen

Debe documentarse cuidadosamente todas las diferencias y conclusiones que pueda obtenerse de los datos. El análisis de los datos permitirá derivar conclusiones acerca del impacto alcanzado por la nueva estrategia de riesgo.

### 8.9.1 Las conclusiones del estudio de evaluación

Ha llegado ahora el momento de tratar de establecer conclusiones a partir de los datos producidos por la evaluación. Una pregunta importante a formularse es la siguiente: ¿Es la nueva estrategia maternoinfantil basada en el enfoque de riesgo, más efectiva para resolver las necesidades de salud de la población y maternoinfantil que la establecida anteriormente en términos de los criterios seleccionados en la sección 8.2.?

Hay que tener gran precaución en la interpretación de los datos. Lo que a primera vista puede parecer un juicio simple sobre la diferencia entre dos series de datos, es en realidad un proceso complicado. Diferentes personas pueden presentar distintas conclusiones después de haber estudiado series de datos similares. Sus conocimientos de hechos más allá de los datos, la relativa importancia dada a los diferentes indicadores (basados en la experiencia personal) y los esfuerzos más o menos concientes de resaltar o suprimir ciertas conclusiones, son factores que constituyen elementos críticos en el proceso de evaluación. En la evaluación de la atención a la salud es a veces difícil producir datos que sean fácilmente interpretados. Los métodos estadísticos usados en el análisis deben ser trabajados con sumo cuidado. En particular,

puede ser necesario estandarizar los datos antes de realizar comparaciones para tomar en cuenta la distribución desigual de algunas variables entre los grupos de estudio y de comparación. La utilización de los indicadores de impacto puede ser difícil. Asimismo con frecuencia no son muy específicos porque pueden influirlos muchos factores ajenos a la estrategia de riesgo.

La evaluación de los servicios de la atención a la salud y de las estrategias tiene algunos problemas metodológicos; sin embargo existen varias formas de minimizar estos problemas. Los indicadores deberán ser seleccionados desde una etapa temprana del proceso. La participación en el proceso de evaluación deberá ser amplia e incluir al grupo de estudio, al personal de salud, y de administración, al personal con poder en las decisiones políticas, y a los representantes de la comunidad. Cada uno de estos grupos tiene intereses propios que influyen en la interpretación de los datos. Cuando se realiza una evaluación como un proceso de grupo, es más difícil sesgar los resultados. Por otra parte, la participación amplia ayudará a facilitar la aplicación de los resultados. El estudio no debe perder su valor científico en general y no se ocultarán datos; no obstante las interpretaciones del estudio deberían ser aceptadas como parte de un proceso socio-político tendiente al desarrollo de los niveles de salud.

### Ejercicio 5

(i) Revise los indicadores de evaluación y los criterios que usted seleccionó en el ejercicio 1.

Discuta la importancia relativa que usted le daría a cada uno de estos indicadores si usted fuera:

- El ministro de salud
- El líder del partido de oposición
- El oficial médico a cargo del área de estudio
- Un hombre del pueblo
- Una enfermera/partera
- Una mujer embarazada.