PD-ACD-062
98452

EPI INFO TRAINING OF TRAINERS
RACHA PROJECT

23 March-10 April 1998

Kingdom of Cambodia

Patrick Kelly

BASICS Technical Directive 000-CB-01-016
USAID Contract No HRN-C-00-93-00031-00

A

TABLE OF CONTENTS

ACKNOWLEDGMENTS

EXECUTIVE SUMMARY 1

PURPOSE 2

ACTIVITIES 2

RESULTS 4
- Training 4
- Training Evaluation 4
- Field Trips 6

PROBLEMS AND RECOMMENDATIONS 7
- Problems 7
- Recommendations 7

APPENDIXES

A DATA ENTRY QUESTIONNAIRE FOR EVALUATION
B LIST OF TRAINEES
C TRAINING MANUAL (separate document)

ACKNOWLEDGMENTS

We had a great time, worked hard, and the enthusiastic trainees took the first steps on the journey of becoming competent and future teachers of the world of Epi Info

As is often the case on an assignment, the consultant found that everyone was extremely busy No problem! The support from the RACHA staff and the training facilities, and the collaboration provided by Dr Ouk Chhieng and his staff at the Royal Phnom Penh University was excellent

Special thanks are in order for Dr Sieng Nam, my co-facilitator and trainer, who would not let me go until he understood every line of code in my many programs

EXECUTIVE SUMMARY

Purpose

The purpose of the mission was to support the RACHA project in survey capacity building, specifically, questionnaire design, data processing, and analysis using the statistical package Epi Info

Activities

The three-week assignment was divided into two phases computer training during the first two weeks, followed by visits to two of the four participating provinces during the third week

Results

The training was conducted at two levels simultaneously—that of the beginners and that of the co-facilitators Co-facilitators is in the "plural" because not only Dr Sieng Nam, but also Neon Vongsa, the new RACHA MIS coordinator, and Taing Peng Sreng from the university assumed the dual roles of trainee/trainer in the time-honored method of "see one, do one, teach one "

An evaluation of the training was conducted that showed that the beginners know a lot more now about Epi Info than they did at the start of training They are eager to practice what they have learned and want to learn more

Problems And Recommendations

The two categories addressed are training and data entry for the ongoing health facility surveys

Training

The major constraints were as follows large class size (21), low levels of English comprehension, lack of basic computer and DOS skills, and only one week for the Epi Info training Khmer translation of training materials has been initiated, but was not ready by the time of the training

Data Entry

Review of data already collected in the provinces of Pursat and Kampot revealed specific areas where improvement can be achieved in questionnaire design, interviewer training, and day-to-day supervision during data collection

PURPOSE

The purpose of the assignment was to support the RACHA project in survey capacity building specifically, questionnaire design, data processing, and analysis using the statistical package Epi Info

The ambitious objectives called for—

1 The design of a 1-week course to teach 10-12 participants in Epi Info so that they could train others and provide local technical support within their own provinces and institutions

2 Help initiate the process of translation of key course materials

3 Work with co-facilitators to conduct the training, using an actual data base if possible

4 Obtain evaluation data from the participants

ACTIVITIES

The length of the mission was three weeks, apportioned as follows

- **Week 1**

The computer training was conducted in the Faculty of Science of the Royal Phnom Penh University The first week, 23-27 March, Professor Hang Vannak introduced the trainees to Windows 95, the use of WORD, and basic DOS commands During this time, the consultant observed and facilitated the training and worked closely with his co-facilitator, Dr Sieng Nam, to introduce him to the materials—manuals and diskettes—in preparation for the Epi Info training

- **Week 2**

The Epi Info training, per se, was conducted from 30 March-3 April for 21 trainees, each with a computer The agenda was as follows

Day 1

1 Participants asked about their objectives for the week
2 Brief introduction to the world of RACHA, surveys, computers and the Internet
3 Review of basic computer and DOS skills
4 Begin installation of the programs on the computers (including Epi Info 6 04b, update to 6 04c, F-PROT for virus protection, and PKZIP to decompress files)

Day 2

1 Review and objectives of the day
2 Finish installation of the programs
3 Use of EPED
   a For questionnaire writing QUEST0 TXT
   b For data entry QUEST0 QES

Day 3

1 Review and objectives of the day
2 Use of ENTER
   a To create REC files QUEST0 REC
   b To enter data
3 Use of CHECK to minimize data entry errors QUEST0 CHK
4 Use of ANALYSIS
   a Data processing and descriptive analysis
   b Introduction to writing PGM files

Day 4

1 Review and objectives of the day
2 Introduction to menu driven programs EPIGLUE
3 Introduction to EPINUT for nutrition analysis
4 Use of REC2QES to recover an untimely QES file deletion

Day 5

1 Overall review and objectives of the day
2 Ad lib practice
3 Copying programs on diskettes for each province/institution
4 Course evaluation
   a Self evaluation
   b Skill evaluation by co-facilitators

- **Week 3**

During the third week of the assignment, a team comprised of the consultant, the RACHA MIS coordinator, the training co-facilitator, and the trainee from the university, spent a day in each of two provinces—Pursat and Kampot—to assist in the installation of the programs in the computers and begin to attempt to do data entry for the health facility survey recently conducted The last day of the mission was spent at the RACHA office for the final debriefing, installation of programs, and discussion of problems, solutions, and next steps

RESULTS

Training

The training was conducted at two levels simultaneously—that of the **beginners** and that of the **co-facilitators** The beginners know a lot more now than they did two weeks ago about Epi Info, Murphy's Law, and GIGO, and they are eager to practice what they have learned and to learn more They will require frequent follow up, supervision, refresher training, and participation in surveys for any of this new knowledge to "stick "

Co-facilitators is in the "plural" because not only Dr Sieng Nam, but also Neon Vongsa, the new RACHA MIS coordinator, and Taing Peng Sreng from the university have an good understanding of the contents of the programs and accompanying manual, and they assisted in the training They will improve their skills as they adapt the files to the needs of the RACHA surveys and provide ongoing support to the beginners

Training Evaluation

An evaluation of the training was conducted on the last day and consisted of two parts

1 Self evaluation questionnaires filled out by the trainees (the computerized version of the questionnaire is presented in Appendix A)

2 Assessment of the trainee skill levels by Dr Sieng Nam and Prof Hang Vannak

The trainees were asked to evaluate their perceived levels of understanding of the seven major components of the Epi Info training EPED, ENTER, CHECK, ANALYSIS, EPIGLUE, EPINUT, and BOOLEAN expressions Responses were coded from 1-5

1 = Not at all
2 = A little
3 = Good
4 = Very Good
5 = Excellent

Out of a possible total of 35, overall scores ranged from 9 to 17, with a median of 12 Summary results are presented by group in Table 1

TABLE 1
Self Evaluation by Group

Group	N	Median Total Score
Kampot	4	12 0
NMCH	2	16 5
NPH	2	17 5
Pursat	4	12 0
Siem Reap	5	10 0
Stung Treng	3	12 0
RPP University	1	19 0
Total	21	12 0

Siem Reap has the lowest score at 10, and the other three provinces—Kampot, Pursat, and Stung Treng—all are at the overall median of 12

The skill levels of the trainees were also independently evaluated by Dr Sieng Nam and Prof Hang Vannak The possible scores were defined from 1-5, as shown for Table 1 The two evaluations never varied by more than one point for any trainee, and when this did occur, the mean of the two scores was used Table 2 shows the results by group

TABLE 2
Skill Evaluation by Co-Facilitators

Group	N	Median Total Score
Kampot	4	3 0
NMCH	2	3 5
NPH	2	3 5
Pursat	4	3 0
Siem Reap	5	3 0
Stung Treng	3	3 0
Total	20	3 0

The range was from 2-4, with an overall median of 3, and there was little variation among the groups It is encouraging to note that each participating province/institution had one person judged to have above average skill (score > 3) at the end of the course These individuals are identified in Table 3

TABLE 3
Trainees Achieving Above Average Skills

Group	Trainee
Kampot	Sol Swath
NMCH	Sau Sokun Mealiny
NPH	Sam Sophan
Pursat	Chengli Bunty
Siem Reap	Sor Sothy
Stung Treng	Soeur Kea Eng

The self-evaluation instrument has several open ended questions

- What did you like/dislike about the training?
- What are your recommendations for improvement?
- What further training/experience would you like?

The responses were similar for all the trainees and can be summed up as follows

- The training was too short
- The presentation of the material was too fast
- We want more practice and further training
- Translation of the training documents into Khmer would be very useful

The evaluation results presented here have been put into a menu driven program entitled CAMEVAL that was written at the end of the mission This program, written using EPIGLUE, contains the data entry forms, the data, the program files for the analysis, and an EPIAID presentation of the results

Field Trips

The trips to the provinces of Pursat and Kampot were very useful to encourage the trainees and to initiate the training for data entry of the health facility survey questionnaires Numeric codes

were created for the health facilities in the HCENTERP and HCENTERK files Revisions in the QES and CHK files for QUEST1 (child observation and exit interview) and QUEST4 (health facility resources) were made based on actual data entry practice The revised files for the health facility survey are organized in a menu driven program entitled RACHA

PROBLEMS AND RECOMMENDATIONS

Problems

The main problems encountered are outlined as follows

- **Training**

  - Large class size - 21
  - Low levels of English comprehension
  - Lack of basic computer and DOS skills
  - Only one week for the Epi Info training
  - Translation of the manual into Khmer not done yet

- **Data Entry for the Health Facility Survey**

  - **Pursat**

    - No questionnaires for antenatal and birth spacing visits
    - The questionnaires for the sick child and facility resources differ from the revised version for which Epi Info data entry files had been prepared

  - **Kampot**

    - There are no client names or identification numbers on the antenatal and birth spacing questionnaires This prevents the matching of clinical observation and exit interviews
    - In QUEST4, the name of the drug dispenser rather than the title is noted

Recommendations

- **Training**

As previously indicated, the trainees need to be followed, encouraged, and aided for any of this brief experience to stick They all appeared to be motivated and expressed interest in improvement of their skills To be a teacher is to see, do, and then teach They have seen, but

have not yet done enough to be independently competent or to transmit their knowledge to others

Even though people do not acquire computer skills by reading manuals, Khmer translation of at least part of the manual will be useful It contains many examples that have been useful in real surveys and can be easily adapted to local needs

- **Data Entry for the Health Facility Survey**

Evaluation of the questionnaires in Pursat and Kampot indicate that improvement is indicated for survey implementation in the following areas

1. Questionnaire design
   a. Question numbering would aid the data entry people
   b. Making sure required variables are present, identification numbers, titles, etc
   c. Verifying that the Khmer version corresponds exactly to the English version

2. Improved interviewer training

3. Improved day-to-day supervision during data collection

APPENDIXES

APPENDIX A

DATA ENTRY QUESTIONNAIRE FOR EVALUATION

APPENDIX A

EPI INFO TRAINING EVALUATION
ROYAL UNIVERSITY OF PHNOM PENH
RACHA PROJECT
23 MARCH - 3 APRIL 1998

```
 NAME    <A                      >         {PROV}INCE     #
-------------------------------------------------------------------------------
          1=NOT AT ALL, 2=A LITTLE, 3=GOOD, 4=VERY GOOD, 5=EXCELLENT)

 {EPED}                           #          {ENTER}         #
 {CHECK}                          #          {ANAL}YSIS      #
 {EPIG}LUE                        #          {EPIN}UT        #
 {BOOL}EAN EXPRESSIONS            #          {TOT}AL SCORE  ##
-------------------------------------------------------------------------------
 What did you LIKE about training?
 {L1}                                      <A                    >
 {L2}                                      <A                    >
 {L3}                                      <A                    >

 What did you {NOT L}IKE about training?
 {NL1}                                     <A                    >
 {NL2}                                     <A                    >
 {NL3}                                     <A                    >
 Recommendations for improvement?
 {R1}                                      <A                    >
 {R2}                                      <A                    >
 {R3}                                      <A                    >

What more training and experience liked?
 {TE1}                                     <A                    >
 {TE2}                                     <A                    >
 {TE3}                                     <A                    >

Any other comments?
 {OC1}                                     <A                    >
 {OC2}                                     <A                    >
 {OC3}                                     <A                    >
-------------------------------------------------------------------------------
SKILLS EVALUATION
 {SHV}                  #
 {SSN}                  #
 {AVE}RAGE SCORE      # ##
-------------------------------------------------------------------------------
```

APPENDIX B

LIST OF TRAINEES

APPENDIX B
LIST OF PARTICIPANTS

TRAINER	PATRICK KELLY
CO-FACILITATORS	DR SIENG NAM (PURSAT) PR HANG VANNAK (ROYAL PHNOM PENH UNIVERSITY)
TRAINEES	
KAMPOT	KHUY RATHANA MINNOUN SRAPCHANTHA SOL SOWATH SOR SAMEDY
NMCH	KEO SONY SAU SOKUN MEALINY
NPH	RETH CHANN SAM SOPHAN
PURSAT	CHENGLI BUNTY CHHUN MUYKECH MAO SOVANNA NGOR SIEKKHEANG
SIEM REAP	CHHENG HAL CHHOY CHAN HUN CHEANG SOR SOTHY TAN KANG
STUNG TRENG	PHKAY SAMBATH SAY PROLEUNG SOEUR KEA ENG
RPP UNIVERSITY	TAING PENG SRENG

APPENDIX C

TRAINING MANUAL (SEPARATE DOCUMENT)

Kelly, P Epi Info Training of Trainers Cambodia--RACHA Program. Phnom Penh: RACHA/BASICS/USAID, April 1998

EPI INFO TRAINING OF TRAINERS

CAMBODIA - RACHA PROGRAM

BASICS

PATRICK KELLY
TECHNICAL OFFICER
BASICS

MARCH 1998

TECHNICAL DIRECTIVE # 000-CB-01-016

Table of Contents

		Page
	How to Use this Manual and Accompanying Programs	ii
1	The Internet	1
2	The Importance of Measurement	2
3	Review of basic computer and DOS skills	3
4	Installation of Epi Info 6 and other programs on your computer	5
5	Goals and objectives of the training workshop	12
6	Getting Started in Epi Info	13
7	The Epidemiologist's Editor EPED	15
8	Creating a questionnaire with EPED	23
9	Entering data with ENTER	32
10	Improving data entry with CHECK	39
11	Data preparation with ANALYSIS	52
12	Analysis with ANALYSIS	64
13	Processing Summary Records with ANALYSIS	73
14	Evaluation of Nutritional Status with EPINUT	76
15	Linking files with MERGE and RELATE	81
16	Putting it all together with EPIGLUE	84
17	The Beginning	96

How to Use this Manual and Accompanying Programs

Computer skills are learned mostly at the keyboard supplemented by learning aids that include both Computer Assisted Instruction and written materials Epi Info comes with a variety of such tools including Help Messages in context tutorials for self-instruction and a comprehensive manual (online and in book format) The quantity of information is sometimes overwhelming to the beginner, and may be confusing when definitions and examples of certain processes and commands are spread among several chapters

The RACHA manual attempts to synthesize and simplify the material for the use of Epi Info 6 0b or 6 0c into a more user friendly format for both training and reference It consists of essential excerpts from the official Epi Info manual complemented by examples and suggestions for working in real survey conditions It is meant to be used together with two menu driven programs (CAMTOT and CAMHFS) that organize the Epi Info family of programs and present examples of actual files that have been successfully used in Health Facility Surveys

The manual can be used at any level from beginner to advanced users and trainers While not coming close to exhausting the possibilities of Epi Info it does provide the basis for serious and competent use of the program in real survey conditions The accompanying files will facilitate the data processing of surveys by the user in that it is much easier to modify something that has already worked rather than creating a system from scratch

The manual starts at **Chapter 1** by talking about the Internet, the information highway for the 21st century, that is already here It was the major source of information and materials for the preparation of this training course A simulation of this net surfing" has been included in CAMTOT to provide information that will be useful for future reference and downloading of software upgrades

Chapter 2 is a reminder of the importance of measurement in international public health programs and the role that the BASICS project has undertaken to contribute to the strengthening of the capacities of health professionals in developing countries to do child survival program monitoring and evaluation

Chapter 3 includes check lists of basic computer and DOS skills that are essential for successful use of Epi Info For those who know them a quick review will suffice Others will do well to invest time to master them before proceeding Epi Info requires a lot of file manipulation and precise commands

Epi Info and other programs must be successfully installed on the user s computer **Chapter 4** provides the step by step details of the process using the accompanying diskettes

- Green (5) Epi Info and PKZIP
- Red (5) Epimap 2 (OPTIONAL)
- Orange (3) Doepi (OPTIONAL)
- Blue (1) CAMTOT and CAMHFS (Menu driven programs)
- Yellow (2) F-Prot (Anti Virus Protection)

Chapter 5 states the goals of the training workshop and reminds us of the steps in preparing and implementing a scientific survey Although strictly speaking, the issues of data processing addressed by this manual only start at step 4 it is advised that the computer person(s) be involved from the beginning of the planning process The time and resources budgeted for data processing are often underestimated in a survey thus the sooner one starts the better

At **Chapter 6** we actually start using Epi Info This is best done by running and becoming familiar with the contents of three menu driven programs EPI6, CAMTOT and CAMHFS

The use of EPED as a general word processor is presented in **Chapter 7** This is a program that comes with Epi Info for the writing and editing of text files Those that are needed include ones with the following extensions, QES, CHK, and PGM To the extent that this tool is mastered, the process of using Epi Info will be greatly facilitated

Chapter 8 describes the creating of computerized questionnaires with EPED The three essential components of fields (variables) are presented and explained They include field name, field type and field length Files containing excerpts of a questionnaire (Clinical Observation of the Sick Child) in Word Perfect format (QUEST1 W51) and a computerized version (QUEST1 QES) have been installed at C \CAMHFS\QUEST1\ Loading, modifying and actually typing and saving these files to other names will help the user achieve the required EPED skills

Chapter 9 covers data entry with the program ENTER After the questionnaire file has been written, ENTER is used in a two step process

- Creation of the data (REC) file from the questionnaire (QES) file
- Entry of the data into records with editing as necessary

Theoretically ENTER can be run by itself without the use of an accompanying CHECK (CHK) file It is, however, strongly recommended that a check file be written to facilitate data entry and automatically minimize errors both at entry and in the calculation of indicators

The improvement of data entry with CHECK as described in **Chapter 10** is particularly useful for people with minimal skills Skilled people also when tired or working fast make their fair share of errors which can be caught and corrected by well written check files The check file (QUEST1 CHK) is included in the chapter Spending time actually typing and using it together with QUEST1 REC will provide the user practice involving the essential CHECK commands including an Boolean expressions An introduction to writing Help files for CHECK is also provided

And now we are ready for the analysis Well not quite **Chapter 11** introduces the program ANALYSIS and describes how it can be used for further data preparation verification and cleaning and new variable definitions and coding The writing of optional programs (PGM) is introduced along with an actual file (EDIT1 PGM) that can be used with QUEST1 REC to sort records and allow data updating Program files can be extremely useful to reduce manual errors by means of automatic functions and to provide a "memory" of the set of commands that actually produced a given result

Chapter 12 provides the basic introduction to descriptive analyses and presentation of the results with charts and graphs Two program files (STATS1 PGM and BARHIST PGM) provide more programming examples Discussion of the theory needed to understand and interpret the many statistical indicators generated by Epi Info is not attempted within the scope of this manual

Chapter 13 describes how to do analyses when summary data rather than individual observations (cases) are contained in each record This may be the case, for example, when the data come from monthly or quarterly reports The commands also allow the analyst to do further processing based on tables rather than the "raw data" This option is useful in surveys with very large sample size

With the growing recognition of the importance of nutritional status in child health it is likely that most surveys concerned with child survival will include anthropometric data -- age, sex, weight and height **Chapter 14** shows how to use the excellent programs that Epi Info provides to do the calculations based on international growth reference curves recommended by WHO Two accompanying program files are included

- NMISS PGM (A modification of AMISS1 PGM provided by Epi Info)
- ANTHRO PGM (How to create a separate data base when memory is otherwise insufficient)

There are times when data files may need to be linked together either top to bottom or side to side **Chapter 15** describes how this is done using the MERGE and ANALYSIS programs

Chapter 16 introduces and explains EPIGLUE an optional feature of Epi Info that can be used to create menu driven programs that automatically organize Epi Info programs and files It is extremely useful in survey situations when time is short the pressure is on and computer people with minimal skills are being used The program CAMHFS a menu driven program using EPIGLUE is presented as an example of how a Health Facility Survey can be organized The steps are then described, along with accompanying programs to enable the user to create his or her own program for another survey by modifying the CAMHFS files

The manual ends at **Chapter 17** with encouragement to the user to begin the Epi Info journey and invest the time necessary for success

1 The Internet

The Internet also known as the World Wide Web is the information highway 'now' two years before the arrival of the 21st century It was my major source of information and materials for the preparation of this training course

The starting point in the search was the home page of the Centers for Disease Control and Prevention in Atlanta Georgia Two clicks of the mouse later I arrived at Epi Info's Homepage which gave me the following information

Epidemiology Program Office

Epi Info is a series of microcomputer programs for word processing, data management and epidemiologic analysis designed for public health professionals Epi Info is easy to use, but also offers programming languages for both data input and analysis so that permanent health information systems can be developed The options for continuing included

- Download Software
  - Epi Info Epimap DoEpi and SSS1
- Technical Support
  - Instructions for contacting the Epi Info hotline by phone fax or email
- General Information About Epi Info
  - History Manual Tutorials and Translations
- Frequently Asked Questions
  - Explanations of error messages, progress on Epi Info 2000 the Windows version
- Epi Info Discussion Group
  - How to join the email group
- Links outside of CDC
  - Worldwide links
  - Add a link
  - Vendors (Brixton Books ToucanEd Publications and USD Incorporated

You too will be able to take a simulation of this trip in cyberspace once you have installed Epi Info and CAMTOT on your computers Run CAMTOT place the bar cursor on Internet and press <F1> The three homepages visited include

1 The CDC (Epi Info)
2 Epiconcept (Epi Info in French)
3 F-Prot Anti-Virus (An antiviral program)

2 The Importance of Measurement

Before we plunge into the waters of Epi Info let's ask ourselves why in a world of increasing needs and decreasing budgets we should spend so much time effort and money to conduct studies surveys, monitoring and evaluation

The urge to measure has existed for a long time Stephen Hales an English scientist born in 1677 made the following claim

> "Since we are assured that the all-wise Creator has observed the most exact proportions of number weight and measure in the making of all things, the most likely way, therefore, to get any insight into the nature of those parts of the Creation which comes within our observation must in all reason be to number weigh, and measure "
>
> Source The Columbia Dictionary of Quotations

Now in the latter half of the 90's aided by the widespread use of microcomputers the mandate to measure has intensified Look at the goals of three major agencies involved in international health

A **WHO** Progress Towards Health for All by 2000
  1 Reduction of 1990 levels of IMR by 1/3 or 50/1000 live births
  2 Reduction of under fives mortality by 1/3 or to 70/1000 live birth
  3 Reduction of maternal mortality by 50%
B **UNICEF** Mid-decade and Decade Goals for Child Health
C **USAID** Stabilizing World Population and Protecting Human Health
  1 Strategic Objective 3 Reduction of Infant and Child Mortality
  2 Results Packages The Bottom Line
    a) There is a need for evaluation of impact and evaluation of trends
    b) Indicators of sustainability will serve as a litmus test that will determine whether projects get or do not get funded

What has the BASICS project with its focus on child survival been doing in the realm of measurement? It also is committed to "Managing for Results" Together with the CDC and USAID it has developed a conceptual framework, the Pathway to Survival, to assist in the development monitoring and evaluation of integrated case management programs The mission statement of the Evaluation and Monitoring Division of BASICS is To increase the capacity of developing countries to plan implement and sustain effective public health programs The End of Project objective is to have developed tested and implemented indicators and instruments for child survival program monitoring and evaluation

One of the types of quantitative surveys that BASICS has undertaken in many countries has been a Health Facility Survey that evaluation access demand for care and the quality of health services offered in health clinics This method is easily adaptable to a specific country program allows rapid data collection analysis and presentation of results and is relatively inexpensive And of course an important tool that has helped achieve this result is Epi Info

3 Review of basic computer and DOS skills

The current version of Epi Info works in the DOS and not the Windows environment To install the programs on your computer and run them effectively and efficiently requires a familiarity and hopefully mastery of basic computer and DOS skills Thus before going farther let's be sure we understand the following

- **COMPUTER**

A Computer vocabulary
- 1 Hardware
  - a) Hard drive and floppy capacities in bytes (1 4 MB) and bytes free
  - b) Memory (RAM) capacities in bytes and bytes free
- 2 Software
  - a)Programs and floppy disks

B Booting the computer
- 1 Cold boot (Turn on the computer)
- 2 Warm boot Ctrl + Alt + Del
- 3 AUTOEXEC BAT and CONFIG SYS files
- 4 Special keys on the computer keyboard
- 5 Esc Function Keys (F1-F10), Scroll Lock
- 6 Tab Insert Home Page Up, Page Down End Num Lock
- 7 Caps Lock Shift Ctrl Alt Delete Enter Backspace Arrow keys (4)

C Mathematical symbols
- 1 + plus - minus divide * multiply
- 2 > greater than >= greater than or equal < less than <= less than or equal
- 3 <> not equal ^ exponent

D Pay attention to detail when typing
- 1 Letter "l" <> number "1", letter "O" <> number "0," / <> \
- 2 There are 3 kinds of parentheses ( ), { }, [ ]

ESSENTIAL DOS COMMANDS

A Where am I?
- 1 Disk addresses C A B
- 2 The Prompt and Cursor

B Disk Management
- 1 FORMAT
  - a) FORMAT A
  - b) FORMAT A /S
  - c) DISKCOPY A A
- 2 Directories and Subdirectories
  - a) Root Directory C \ A \
  - b) Change Directory CD XXX, CD\, CD
  - c) Make Directory MD XXX
  - d) Remove Directory RD
  - e) Tree

C File Management
- 1 File Names Prefix (Up to 8 letters) +" " + Extension (0-3 letters)
  - a) Examples XXX EXE, XXX BAT, XXX WPG
- 2 Look at file names DIR, DIR/P
- 3 Copy files
  - a) Copy from Source to Destination Copy A XXX TXT C
  - b) * = wildcard ? = unknown letter
- 4 Rename files REN Oldname Newname
- 5 Delete files DEL Filename

D Text files ASCII file, DOS text file
- 1 Plain text files that can be easily created and edited
  - a) Examples Config sys * bat, * txt * qes * chk * pgm Read me
- 2 Reading text files TYPE <Filename> TYPE <Filename|More>
- 3 Creating and editing text files EDIT <Filename>
- 4 Programmed DOS commands in a BATCH (* BAT) file

E Typing ASCII codes Numlock On + ALT + 3 digit code (1 - 255)

F Repeat last command
- 1 Complete Command - F3
- 2 Character by Character - F1 or Right Arrow

G Set date and time
- 1 Date Allows you to enter New Date (mm-dd-yy)
- 2 Time Allows you to enter New time (hh mm ss)

H BASIC Programming
- 1 QBASIC comes with DOS
  - a)We may use it to practice Boolean Algebra Expressions

4 Installation of Epi Info and other programs on your computer

And now without further ado let s install Epi Info on your computers All of the necessary files for Epi Info 6 04b downloaded from the Internet have been saved as self-expanding compressed files on five green diskettes

1 EPI604_1 EXE (File Size = 1 367 649 bytes)
2 EPI604_2 EXE (File Size = 1 341 995 bytes)
3 EPI604_3 EXE (File Size = 1,360 925 bytes)
4 4BUPDATE EXE (File Size = 1 274 457 bytes) (Updates to Version 6 04b to 6 04c) (Update for year 2000 date capability)
5 Manual (Microsoft Word 6 Format) - EPI6MAN EXE (Optional) (Download File Size=963 165 bytes) (Expanded File Size=3,996 160)

[Note File sizes have been made to fit on 1 4 MB floppy diskettes]

To get a complete working system on your computer requires a lot of disk and file manipulation The files must be copied to a directory on the hard disk decompressed and then installed To facilitate and accelerate the process I have written some BATCH files that will automatically do most of the work

1	START BAT	Creates the temporary subdirectory C \CAMTOT\EPI and copies 4 batch files into it (EPI6 BAT, MANUAL BAT, UPDATE BAT and PK BAT
2	EPI6 BAT	Copies the Epi6 files from Disks 1-5 to C \CAMTOT\EPI
3	MANUAL BAT	Copies EPI6MAN EXE to C \EPI6 after Epi Info is installed
3	UPDATE BAT	Copies 4BUPDATE EXE To C \EPI6 after Epi Info is installed
4	PK BAT	Copies PKZ204G EXE to C \DOS

A Now we are ready to begin [Note Keys or commands to be typed are indicated by "<>"]

1 Make sure you have the green Epi Info disks 1-5
2 Place Disk 1 in Drive A (On some computers the floppy disk drive may be at B)
3 Go to drive A
4 Type <START>
5 The cursor is now at C \CAMTOT\EPI>
6 Type <EPI6> and change disks when prompted to do so When this is done, you no longer need the 5 diskettes

B Now decompress and install the Epi Info files

1 Verify that the cursor is still at C \CAMTOT\EPI>
2 You may verify that the necessary files are there by typing <DIR> or <DIR/P>

3 EPI INFO

a Decompression of the 3 Epi Info files

1 Type <Epi604_1> At the prompt Continue Extraction? Type (Y) plus <Enter>

2 Do likewise for Epi604_2 and Epi604_3

b Installation of Epi Info into the C \EPI6 subdirectory Type INSTALL and follow the prompts on the screen There are many questions so to speed up the process we will note them and indicate the appropriate response

Prompt	**Response**
Source Drive	
C (Must be a letter)	<Enter>
Destination Drive	
_ (Must be a letter)	<C>
Does this drive have floppy disks (Y/N) _	<N>
For a normal installation press <Enter>	<Enter>
Would you like to install the system for use? _	<I>
Select Video Drivers <F4>-Continue	<F4)
Select Print Drivers	
<F8>-Select All	<F8>
<F4>-Continue	<F4>
Select Groups to Install <F4>-Continue	<F4>
Press <F4> to install or <Enter> to change choices	<F4>
Is it alright to create or edit your AUTO EXE and CONFIG SYS files?	
Type Y for Yes and N for No _	<Y>
Press <Enter> to leave the installation program	<Enter>

c **Congratulations!** With any luck you have been successful In order to activate any modifications in the AUTO EXE and CONFIG SYS files you should now **REBOOT** the computer

C Upgrade Epi Info 6 04b to 6 04c and add the manual

Return to C \CAMTOT\EPI> and run 2 more batch files that copy files to C \EPI6

1 Type <UPDATE>

2 Type <MANUAL>

3 Now to install the Epi Info 6 04b to 6 04c upgrade

a Go to C \EPI6>

b Type <4BUPDATE> At the prompt Continue Extraction Type <Y>

1 The Epi Info files will be updated from Version 6 04b to 6 04c (Note ANALYSIS version 6 04b works better than 6 04c in running some of the commands in menu driven programs (CAMHFS for example)

c To free disk space type <Del 4BUPDATE EXE>

4 Installation of the Epi Info Version 6 manual

a Still at C \EPI6>

b Type <EPI6MAN> At the prompt Continue Extraction Type <Y>

1 The file EPI6MAN EXE will be decompressed to EPI603 WD6 which is the full manual in Word 6 format

2 To free disk space type <Del EPI6MAN EXE>

5 To run Epi Info just type <EPI6> "et voila" the Epi Info 6 menu appears

D Installation of other programs

a Installation of **PKZIP** files The compressed files we have been using by have created by a shareware program called PKZIP You too may wish at times to combine many files into one and compress them For example, this is useful for backup copies of important files fitting information containing more than 1 4 MB onto a floppy diskette and attaching files to email Since PKZIP files are utilities we will install them in the DOS subdirectory

1 Return to C \CAMTOT\EPI> and type <PK> to run the final batch file

2 Go to C \DOS> and type <PKZ204G> This will decompress the files into functional ones along with accompanying documentation

3 To free disk space type <Del PKZ204G EXE>

b Now free up disk space by removing transfer files and temporary subdirectories

1 Go to C \CAMTOT\EPI >

2 Type <Del * *>

3 At the prompt All files will be deleted! Are you sure (Y/N)? Type <Y>

4 Go to C \CAMTOT> Type <CD >

5 Remove the EPI subdirectory Type <RD EPI>

c Other programs we mav install on our computers include DOEPI and EPIMAP2 DOEPI comes on three orange diskettes and EPIMAP2 on five red ones The installations require a series of steps analogous to those we used for Epi Info that is 1) Copying the compressed files to a temporarv file on the hard drive 2) Decompressing the files 3) Running the INSTALL file 4) Answering the screen prompts appropriatelv and 5) finallv deleting the transfer files and temporarv directory If all went well DOEPI will be installed in C \DOEPI and EPIMAP2 will be in C \EPIMAP2 You can run the programs as follows

1 At C \DOEPI> Type <DOEPI> 2 At C \EPIMAP2> Type <EPIMAP>

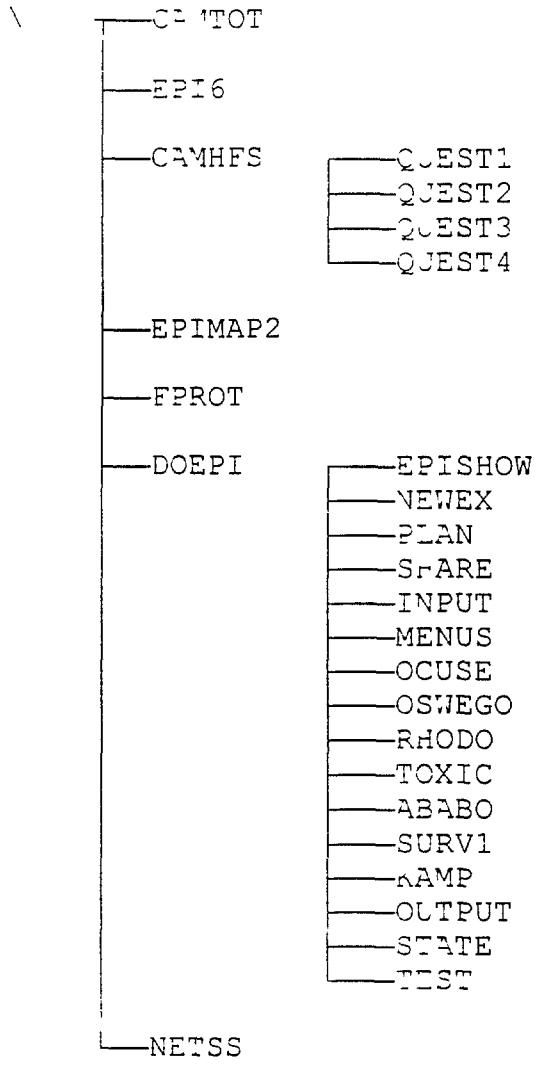
d One of the powerful features of Epi Info is the ability to create menu driven programs of your own design complete with pull-down menus and popup dialog boxes with a program called EPIGLUE These programs are the "glue" that brings together the manv directories and programs vou have installed on vour computer and allows them to be run as a single unified svstem Two such programs (**CAMTOT** and **CAMHFS**) will provide the basis of the CAI (Computer Assisted Instruction) for Epi Info They have been compressed into the file CAM ZIP on the blue diskette labeled Epi Info CAI along with a batch file (CAM BAT) that will copy the files in the proper places on the hard drive To install the programs insert the blue disk into the computer and at A \> Type <CAM>

e And now that you thought the installation process was finallv over, there is just one more item **Murphy's law** states that bad things often happen at the most inopportune times There are, indeed two kinds of researchers 1) Those who have lost important files and data due to disk crashes mistaken deletions or computer viruses and 2) Those who will do so in the future Anti-virus protection is essential preventive medicine for the latter problem F-PROT is an excellent shareware program that gives such protection The latest version (V 228 - Januarv 1998) can be downloaded for free via the Internet It has been saved in a compressed file (FP-228~1 ZIP) on the yellow Anti-Virus Protection Disk 1 The batch file (FPROT BAT) on Disk 1 will copy the program to C \FPROT> and decompress it The yellow Disk 2 is a F-Prot Boot Disk that can be used in an emergencv to detect and eliminate viruses when a computer does become infected To proceed with the installation, place the yellow Disk 1 in the computer go to A and Tvpe <FPROT> To scan disks for viruses, all you have to do is type <F-PROT> and follow the instructions in the menu

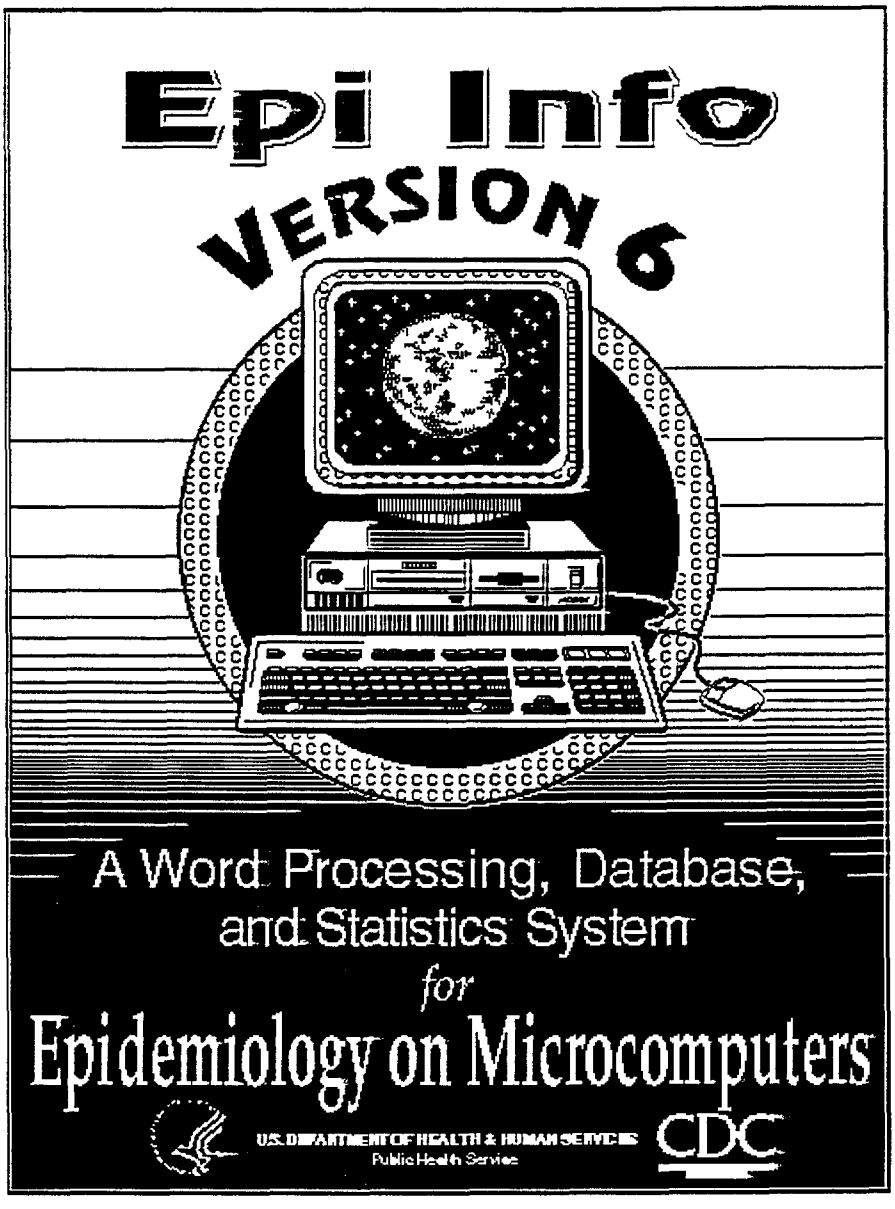
D Summary of the installations

Can you remember all of the programs vou installed and where we they are on your computer? Directories and subdirectories are to a disk as chapters are to a book- they help organize the material so that we can find it The "tree" on the next page shows the directories that we have created We also placed the PKZIP program into the DOS directory, which existed already

DIRECTORY TREE FOR C DRIVE



Hundreds of files exist in these directories so let's get started and see how many we can learn to use The following two pages show the cover pages of Epi Info 6


Epi Info
VERSION 6
A Word Processing, Database,
and Statistics System
for
Epidemiology on Microcomputers
U.S. DEPARTMENT OF HEALTH & HUMAN SERVICES
Public Health Service
CDC

Epi Info, Version 6

A Word-Processing, Database, and Statistics Program for Public Health on IBM-compatible Microcomputers

Program design by Andrew G Dean Jeffrey A Dean
Denis Coulombier, Anthony H Burton Karl A Brendel Donald C Smith, Richard C Dicker Kevin M Sullivan and Robert F Fagan

Programming by
Jeffrey A Dean Denis Coulombier Donald C Smith,
Karl A Brendel Thomas G Arner, and Andrew G Dean

Manual by Andrew G Dean

Revised for Version 6.03, January 1996

Produced by
The Division of Surveillance and Epidemiology
Epidemiology Program Office
Centers for Disease Control and Prevention (CDC)
Atlanta, Georgia 30333
in collaboration with
The Global Programme on AIDS
World Health Organization (WHO)
Geneva, Switzerland

Suggested citation Dean AG, Dean JA, Coulombier D, Brendel KA, SmithDC Burton AH, Dicker RC, Sullivan K, Fagan RF, Arner, TG Epi Info, Version 6 a word processing, database, and statistics program for public health on IBM-compatible microcomputers Centers for Disease Control and Prevention, Atlanta Georgia, U S A , 1996

Epi Info Hotline for Technical Assistance
(404) 728-0545
FAX (404) 315-6440
EpiInfo@CDC1 CDC.GOV

The good news is that this manual and the accompanying programs are in the public domain and you are encouraged to freely copy translate and distribute them to your friends and colleagues The bad news is that there are over 600 pages **Furthermore, "No warranty is made or implied for use of the software for any particular purpose" In other words-- Caveat Emptor, buyer beware**

5 Goals and Objectives of the Training Workshop

The goal of this training workshop is to train people to use Epi Info who will in turn be able to train others to perform data collection and analysis activities aimed at the strengthening of the implementation of the RACHA program The implementation of a scientific survey is a complicated process involving the following steps

1 Plan the survey
    a Selection of the principal investigator
    b Identification of the survey objectives
    c Selection of the sample design
    d Preparation of the budget
    e Acquisition of adequate resources
        1 Personnel recruitment Supervisors interviewers, data entry people
        2 Véhicles chauffeurs
        3 Creation adaptation and translation of survey instruments
2 Prepare to conduct the survey
    a Administrative arrangements
    b Pretesting of survey procedures and instruments
    c Training of the participants
3 Conduct and supervise the survey
4 Data Processing
    a Coding the questionnaires EPED
    b Entering the data ENTER CHECK
    c Verification and data cleaning ANALYSIS * PGM FILES
5 Analysis ANALYSIS * PGM FILES
    a Preliminary - Descriptive (May include key indicators)
        1 Frequencies proportions central tendency indicators
    b Secondary - Analytic (Cross tables statistical comparisons etc)
6 Interpretation
7 Recommendations
8 Rapid Feedback
9 Formal report

This workshop is concerned essentially with steps 4 and 5a The Epi Info manual defines three skill levels

I Basic Create questionnaires enter data and do analysis without programming using EPED ENTER and ANALYSIS
II Intermediate Do more refined data entry and analysis by using built-in functions in the CHECK program and writing basic program files for ANALYSIS IMPORT and EXPORT data from and to other formats MERGE files, and create menu driven programs using EPIGLUE
III Advanced Do more advanced programming for CHECK and ANALYSIS Use EPINUT for nutritional analysis, link files for relational analysis, write help files for menu driven programs and * CHK files and write tutorials with EPIAID

The objective of this workshop is to arrive at the intermediate level and be exposed to other features of the range of possibilities of Epi Info Computer skills are learned mostly by "hands-on" time at the computer A lot of Computer Assisted Instruction (CAI) comes with Epi Info in the form of tutorials to which I have added some menu driven programs of my own Not only will these programs facilitate the learning process but they will serve as examples that you can use for your future needs It is much easier to adapt than to create a program from scratch Thus, for example, an EPIGLUE program that has been written for a survey in one country, can be adapted to another in a matter of minutes We may look at examples used in Health Facility Surveys conducted by BASICS in Madagascar Mali Niger, and Senegal

6. Getting Started in Epi Info

Go to C \EPI6> and type <EPI6> On the screen you see the main Epi Info menu that "glues" together most of the components of the program The titles across the top line are the choices of pop-up menus

- PROGRAMS--The main programs of the Epi Info system
- TUTORIALS--Interactive tutorials to introduce the main concepts of Epi Info
- EXAMPLES--Sample files that illustrate particular aspects
- MANUAL--The entire manual for Epi Info in easily accessible form
- FILE--Functions to open create view or print text files from the EPI6 menu
- EDIT--Functions to edit copy cut and paste text files opened with the FILE menu
- SETUP--Functions for setting up the menu and the making of questionnaires in ENTER

To select an item from the menu

1 If you do not have a mouse move the highlighted bar with the arrow keys When the correct item is highlighted press the <Enter> key to run the program If the cursor leaves the menu bar press the <Alt> key to make it reappear

2 If you have a mouse you may move the cursor with the mouse and press the left mouse button to select an item

3 You may also type the highlighted (red) letter of the selection for example **P** for **P**rograms or **T** for **T**utorials and then press <Enter>

Pressing the <F10> key will exit from the menu or any other Epi Info program Leaving the menu will bring back the DOS prompt unless your computer has a **customized menu system**

Command-Line Parameters

Any of the Epi Info programs may be run directly from DOS rather than from the Epi menu Some users find this faster than using the menu it also saves the small amount of memory space (about 12 K) that is used by the menu

Each of the programs has a set of command-line parameters" that gives information normally obtained through interactive screens in the program For example, to run EPED the word processor, and work on the file called QUEST1 QES you would type

C:\CAMHFS\QUEST1>**EPED QUEST1 QES**

from the DOS prompt EPED will then automatically display the file on the screen ready for editing or additions

From the Epi6 menu you can also enter command-line parameters by pressing <Ctrl-Enter> while the bar cursor is on the desired program choice in the menu A dialog box will appear so that you can enter the parameters for example CAMHFS MNU to edit a file by that name if the bar cursor is on EPED in the PROGRAM menu

Viewing and Editing Files from the Menu

The EPI6 menu has a built-in editor that offers basic editing of text files To create a new file, choose NEW from the FILE menu To open an existing file choose OPEN from the same menu

We do indeed have a customized menu and information system that ties together all of the components of this training course including Epi6 which we just saw To access it exit Epi 6 and type <CAMTOT> Now we see another choice of pop-up menus

- INTERNET--A simulation of surfing the Internet for information and programs
- PROGRAMS--Programs within and related to Epi Info
- NETSS--A telecommunication system for epidemiological surveillance
- SURVEY--A customized menu for a Health Facility Survey (CAMHFS) – New Survey
- TOOLS--A Calendar calculator and table of ASCII codes
- SETUP--Parameters for the hard drive and source and destination for copying files
- INSTALL--Copying all the CAMTOT files to another disk(ette)
- FILE--Functions to open, create or view text files from the CAMTOT menu
- EDIT--Functions to edit text files including creating large letters

You will notice that there is a clock to the far right of the top line Several sources of HELP are incorporated into this program When the bar cursor is on a pop-up menu name pressing <F1> will go to the World Wide Web menu When the cursor is on the name of a specific program <F1> will give an "in context" message And finally, sometimes a message will appear on the bottom line of the screen

We will examine the contents of this menu together and you are encouraged to do more exploration on your own Would you believe that as complicated as this menu program may seem it is, in fact, quite easy to write Now it's time to go work Let's get to know the Epidemiologist s Editor EPED

7. The Epidemiologist's Editor, EPED

EPED is a general word processor of text files If you don t have a more sophisticated program such as Word or WordPerfect you would be able to use it to type letters memos reports etc We will use it to create and edit questionnaires and other text files (* CHK * PGM * SCR, * MNU * HPT, * TXT, * EPA) This chapter covers the basic essentials of chapter 6 of the Epi Info manual To run EPED goto C \EPI6> and Type <EPED>

The EPED Screen

Across the top of the screen is a command prompt line listing the 10 function keys that perform major functions in EPED Each one either performs a command directly or causes a menu to pop up in the middle of the screen Pressing <F2> for example, shows the menu of commands affecting files Holding down the control (<Ctrl>) key displays a second menu of commands that are activated by holding down the <Ctrl> key while pressing a function key Many of these commonly used functions also have entries in the pull-down menu, but appear on the <Ctrl-Function> menu for easier access

At the bottom of the screen is the STATUS LINE From left to right it shows

I The number of the currently active window if more than one file is open (Up to 4 files may be open at one time with easy switching from one to another)
II The name and directory of the current file if any
III The amount of memory (RAM) available for holding the document
IV The function you are carrying out EPED for general word processing, EPIAID for programmed assistance with word processing
V The line (L) and column (C) numbers occupied by the cursor
VI A series of letters that indicate the status of important settings WW, TXT, or QES for the current mode--Word Wrap for word processing, TXT for line-oriented documents such as computer programs and QES to make an Epi Info questionnaire After this will appear INS for Insert mode ON OVR for Overwrite mode ON, IND for AutoIndent mode ON JST for Right/Left Justification ON

Tutorials

If you would like to proceed directly to the tutorials type <F3> to take you to the EPIAID MAIN MENU which contains a Word Processing tutorial consisting of five lessons that provide a good review and hands-on practice for many of the subjects in this chapter The lessons include

1 The Basics of Word Processing
2 Creating a Document with EPED
3 Retrieving and Editing a Document
4 Find/Replace & Cursor Movement Commands
5 Block Commands

Creating a Document

Creating a document in EPED is easy Simply type the document on the screen as you would like it to appear At any point you may press the "save" or <F9> key A window will appear, asking for a file name and after this is given the file will be saved on the current or the designated disk At the end of the document, press <F10> to leave EPED The program will remind you if any portion of the file has not been saved and then exit to the EPI menu or to DOS The completed document will be on the disk waiting to be called up for viewing, editing, or printing

Setting Word Wrap, Text, and Questionnaire Modes

EPED offers a choice of three *modes*--Word Wrap(WW) Text (TXT) and Questionnaire (QES) The abbreviation for the current mode is shown in the lower right corner of the screen To change modes press the <F6> key to show the SET menu and use the space bar to change the current setting <Ctrl-F9> or <Ctrl O-W> will toggle between TXT and WW modes rapidly

Each mode comes with a built-in series of settings on the <F6> menu WW mode is initially set with a right margin of 65 for example, and TXT and QES with margins of 80 (full screen) For any of the modes, you can change the settings while in that mode The settings are saved in a file called EPED CFG when you choose SAVE SETTINGS THIS MODE in the <F6> menu When you exit from EPED the current settings and file name are saved If RELOAD is set to ON in the SET menu the same file (if any) and mode will be reloaded the next time EPED is run

Page Markers

When PAGINATION (in the <F6> SET menu) is ON, markers appear on the first line of every page on both sides of the screen as solid rectangles This feature is very useful when you are making questionnaires

Editing a Document

To edit a document use the "**Open file this window**" command in the <F2> (File) menu When the file name window appears give the name of the file name you would like to edit If you have spelled it correctly and the disk letter and directory are correct the file will appear on the screen ready for editing

If you are not sure of the spelling or location of the file, use a "wildcard" file name, like "* LET" for all file names ending in " LET" or 'jones *" for all files beginning with JONES or "* *" for all files If you wish to see files on another disk the name might be "A * *" or "D \ DATA\ * LET" for example The names of files matching the description will be displayed Move the cursor bar to the one desired and press <Enter> to bring the file into EPED for editing

Now use the cursor keys to move to a point where you would like to insert or delete text If the letters INS are not at the lower right corner of the screen, press the <Ins> key to turn INSERT on Inserting is easy, just type the material for insertion and it will be inserted at the cursor location Deleting a few letters is done with the backspace or delete keys An entire line can be deleted by pressing <Ctrl-Y> when the cursor is on that line

You can turn off the insert feature so that characters typed "write over" the ones on the screen rather than pushing them to the right Pressing the <Ins> or Insert key will turn off the Insert mode and turn on Overwrite mode pressing it again will reverse the process

When you have finished editing press the <F9> (Save) key and then <F10> (Quit) to return to the menu You may prefer to use <F10> alone and be prompted to save any new material

Entering Text and Moving the Cursor Around on the Screen

To see the effect of the keys that move the cursor, type a few paragraphs of text on the screen This is done exactly as on a typewriter except that with Word Wrap ON (WW displayed at the bottom of the screen), you should **avoid pressing the <Enter> key** except at the end of a **paragraph** Do **not** press <Enter> after every line unless you want the line to be considered a separate paragraph

The cursor may be moved to any part of the text with the following commands

I The **arrow** keys move the cursor right, left, up, or down
II The <PgUp> and <PgDn> move up and down one entire screen (22 or 23 lines)
III The <Home> key moves to the beginning of the current line
IV The <End> key moves to the end of the current line

Note that the cursor will not move past the end of the document unless you insert more blank lines with the <Enter> key

Pop-up Menus, the <Esc> Key, and Giving Commands

Pressing a function key indicated on the command prompt line at the top of the screen will cause a menu to pop up in the middle of the screen To choose one of the commands on the menu, move the cursor bar to the choice with the down or up arrows and press <Enter> Alternatively you can enter the highlighted letter in the command to execute the command directly

To close a menu on the screen press the <Esc> (Escape) key This key will remove any of the menus or help windows and is a good one to press experimentally if you have trouble deciding how to move around the menu system

Note as you move the cursor bar up and down that a series of command keys is shown in the top left corner of the screen These keystrokes are alternative commands for the current menu choice Pressing these keys will call the command when you are in the middle of text processing without using the pop-up menu To center a line for example you can choose **Center Current Line** from the <F4> TXT menu with the cursor bar press C to choose this command from the menu or press <Ctrl-O C> without popping up the menu At first you will probably use the menus but later you may find the <Ctrl> key sequences faster for some often-repeated operations

Handling Files

The documents in EPED are stored as files on the computer's floppy or hard disk The File menu that pops up when you press <F2> provides commands for opening new or existing files and for saving material on the screen in files It also gives access to file directories, information on the current file and a command for moving from window to window when more than one file is open

When you open a file or save one that has not been named, you will be asked to type the file name in a window that pops up on the screen If you try to leave EPED without saving the current file you will be reminded to save it if desired before the program exits to the EPI menu or to DOS

Do you remember **Murphy's Law**? Here is some important advice on how to minimize losses

> Since the document under construction is held in Random Access Memory (RAM), recent changes can be lost if there is a power failure or if someone turns off your computer To minimize the frustration from such events, you should save the current document approximately every 20 minutes as you work so that only bearable amounts of text will be lost in the event of such disasters
>
> If you work on a hard disk it is important to copy ("back up") newly created or edited files to a floppy disk periodically, since hard disks, given enough time, can be counted upon to "crash" and make files unavailable, with or without viruses Backup files on a floppy disk can be made using the "Save File to " command in the <F2> File menu, or can be copied with the copy command in DOS
>
> Text that has been "saved" is safely on the disk, but subsequent changes are only in RAM and will be removed by an electrical power failure During a thunderstorm or when someone is working on the electrical system, it is usually best to save your file turn off the computer, and wait until later If you do continue working, frequent saves (<F9>) are in order!

Finding and Replacing Text

Word processors are good at finding a particular word or phrase and replacing it with another combination of characters Suppose for example that you have a document in which the word "color" appears many times You wish to change the spelling to "colour" for a European colleague

Press the <Ctrl> key and note the two functions FIND and REPL on the menu at the top of the screen To find the word "color" without changing it, you would use <Ctrl-F6> FIND, but in this case <Ctrl-F7> REPL is needed to replace the word with another A dialog box appears asking for the characters to be found You type "color," and respond "colour" to REPLACE WITH A series of options appears

C=Case Sens	Match upper/lower case of your entry and ignore those that fail to match Normal default is to match regardless of case
B=Block	Find "color" only within a block of text that has been selected with the block commands (see next section)
U=Upward	Look upwards in the file from the current cursor position
G=Global	Find all instances of "color," not just the first one
W=Word	Find "color" only if surrounded by spaces Do not find "colorize," for example
N=No ask	Do not ask for confirmation before changing "color" to "colour" **(Be careful, this is a very dangerous command!)**

You may combine options For example the options GN will change all instances of "color," regardless of case, to "colour " without asking for confirmation for each one Entering no options will change only the first instance of "color " "Color," etc regardless of case pausing for confirmation before doing so

Block Commands

So far we have described elementary methods for dealing with files characters, and lines EPED also offers commands for marking, copying, moving, and deleting **blocks** of text A block is a body of text that has been selected for further manipulation with the block commands The common commands for block operations--BEGIN, END, COPY, MOVE, DELETE--are shown on the menu that appears when the <Ctrl> key is held down These commands are also in the Blk or Block menu reached by pressing <F8>

To use the block commands you must first mark the beginning and end of the block To mark the beginning of a block, place the cursor at the place desired and then choose BEGIN by pressing <Ctrl-F1> Alternatively, you may choose BEGIN BLOCK from the <F8> Blk menu No visible results occur, but you have marked the block nevertheless

To mark the end of the block move the cursor to the end and choose END by pressing <Ctrl-F2> The block will now be highlighted Any subsequent block commands will act on this block as long as it remains selected To copy the block to another location leaving the original marked text intact first place the cursor at the point where the copied text should be inserted then select COPY with <Ctrl-F3> or Copy Block to Here from the <F8> Blk menu A copy of the selected block will be inserted at the new location leaving the original block where it was

To move the block to another location ("cut and paste") put the cursor at the point where the block should be inserted and choose the MOVE command <Ctrl-F4> or Move Block to Here from <F8> The block will appear at the cursor location and disappear from the old location

To delete a block use DELETE <Ctrl-F5> or Delete Block from <F8> The text will disappear To remove the grey highlighting of the block use Hide clock display from <F8>

Reading of WordPerfect Files

Version 6 of EPED reads WordPerfect, Version 5 xx files directly They are converted automatically to as close an approximation of the original as possible within EPED, although embedded graphics and WordPerfect tables will not be converted When you save a file that was originally in WordPerfect format it will be saved as an EPED file in the ASCII or plain DOS format You should therefore give it another name before saving This version of EPED does not read WordPerfect, Version 6 files

Boxes

Boxes are useful for emphasizing text making professional-looking menus on the screen or for organizational diagrams

EPED allows drawing single or double horizontal and vertical lines or boxes on the screen with the Arrow (Cursor) Keys Boxes and shading are produced with the arrow keys after adjusting the type of character with the <Home> and <End> keys The following instructions will appear at the top of the screen after you press the <Scroll Lock> key

Arrows=SINGLE box, Home SCRL LOCK,=Double, End=Erase/Shade

To draw a box on the screen using single lines press the <Scroll Lock> key, and then simply draw the box with the up, down right, and left arrow keys The corners will be taken care of automatically If your last character is on a corner, be sure to "turn" the corner with another arrow key so the program knows that you want to complete the box

To draw boxes with double lines press the <Home> key The information line at the top of the screen will change to "DOUBLE Any lines drawn in this mode will be double Single and double lines may be mixed in the same box and the program will take care of putting the correct joining characters at the intersections To erase a box press the <End> key and use the arrow keys to trace the part you want to erase To start drawing boxes again press the <Home> key

Non-English-Language Characters

This is a very useful function For example, to write a French character, hold down the <Alt> key and then use the following keys to produce vowels and other characters with special marks

Use <Alt> with	a	e	i	o	u	c
to produce	aa	eeêe	îi	ô	û	ç

The first time the key is pressed the first of the special characters appears To produce the others continue to hold down the <Alt> key and press the same letter key repeatedly until the correct character is obtained

You can also write mathematical symbols

Scroll Lock on

Type	2	4
To obtain	½	¼

Scroll Lock on--Shift key pressed

Type	+	<	>
To obtain	±	≤	≥

8 Creating a questionnaire with EPED

There are several tutorials that you may want to use

- Go to C \EPI6>
  1 Type <EPI6> (This is the most basic one)
  From the **Tutorials** menu choose **Making a Questionnaire** and follow the instructions on the screen
  2 Type <EPED> Press <F3> to see the EPIAID menu and choose **Make Epi Info Questionnaire** Follow the instructions in the tutorial carefully and make a questionnaire If the save the questionnaire at the end the file name should end in QES
- Go to C \DOEPI>
  1 Type <DOEPI>
  Choose **Exercises** from the main menu followed by **Programming Data Input** With the down arrow go to Instructions and press <Enter> to begin and proceed to **Making a Questionnaire**

Overview

Before beginning to enter data, you must provide Epi Info with a questionnaire that describes the data to be entered The ENTER program then uses the questionnaire to automatically construct a data file called a REC file

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, 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

Setting Up EPED To Make a Questionnaire

At C \EPI6> Type <EPED> or run it from the EPI6 menu Press <F6> to see the SETUP menu The first choice is WW/TXT/QES mode Press the space bar one or more times until QES mode is selected, and then press <Esc> This establishes page size and other settings with convenient values to make a questionnaire

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,** what to name them and what kind of data to accept at these locations The field formats are described below and on the back cover of the Epi Info manual

One way to easily insert fields into your questionnaire is to press <F4> and then <Enter> at the **QUESTIONS** ^Q^Q prompt A menu of field types will appear on the screen The ^Q ^Q reminds you that you have gotten the same menu by pressing <CTRL Q> together followed by <Q> To insert a field type into 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 characters for the field into your questionnaire without using the command The field types are shown below, with information about their meaning

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 including the upper case "A"

<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 Yes/No fields are always one character in length

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

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 In creating a questionnaire however, be sure to type the field format exactly as they appear above

<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>)

<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

The number of digits is the number of characters between the brackets including the letters "IDNUM " and must be 5 or more 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 The first IDNUM in a file can also be set by using commands in a CHK file

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 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 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 ADDRESSand 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

A Sample Questionnaire

Questionnaires may be up to 500 lines Here is the beginning of one that we might use for a Health Facility Survey You can find it at C \CAMHFS\QUEST1\ QUEST1 W51 The format is WordPerfect 5 1

KINGDOM OF CAMBODIA
MINISTRY OF PUBLIC HEALTH/BASICS/USAID
Health Facility Survey
for Integrated Management of the Sick Child

1 Clinical Observation of the Sick Child

District _____ Date __/__/__
Interviewer _____
Health Center Name _____
Type of HC ___ (1=Hosp, 2=Clinic, 3=Health Post)
Caretaker _____ (Mother, Grandmother, Father, Etc)
Health Agent ___ (1=Doc, 2=Nurse, 3=MW, 4=Dispenser)
Child's Identification Number ___
Date of Birth ___/___/___ AGE ____(Months)
Sex ___ (M or F)
Weight _____ (KG) Height ____(CM)

Time at start of observation ___H___

PURPOSE OF VISIT

1 What are the reasons given by the caretaker for the visit?
(Check all correct responses)

Fever ___
Diarrhea ___
Cough ___
Other ___
Specify, if other _____

HISTORY OF THE ILLNESS

2 Did the health agent ask (Check correct answers)	Yes	No
Duration of the illness	___	___
If the child can drink or breastfeed	___	___
If the child has vomited everything	___	___
If the child has had convulsions	___	___
If there was blood in the stool	___	___
If the child had difficulty breathing	___	___

CLINICAL EXAMINATION

	Yes	No
3 Did the health agent evaluate (Check correct answers)		
General status of the child (Normal, irritable, lethargic, etc.)	___	___
The child's temperature (By hand or thermometer)	___	___
The presence of thirst (Offered the child liquid)	___	___
The presence of dehydration (Abdominal skin pinch)	___	___
Respiratory frequency (Breaths/minute while child is calm)	___	___
Chest in-drawing (Sub-costal)	___	___
The presence of a stiff neck	___	___
The signs of measles (Rash, runny nose or red eyes)	___	___
Time at start of observation	___H___	
Duration of observation	___	

You can load this questionnaire directly into EPED by choosing <F2> on the command line, followed by **Open File this window** and typing the path and file name

C \CAMHFS\QUEST1\QUEST1 W51

You may now press <F9> to save the file give it the name QUEST QES and press <F10> to quit EPED If you now go to ENTER and create a new data file using this file, you will see that Epi Info has automatically created field names and types Unfortunately the results are far from being user friendly For example dates are broken up into 3 fields and many field names, with up to 10 letters, are rather confusing In other words a lot of editing will be necessary to create an appropriate data entry file such as the one (QUEST1 QES) that follows

KINGDOM OF CAMBODIA
MINISTRY OF PUBLIC HEALTH/BASICS/USAID
Health Facility Survey
for Integrated Management of the Sick Child

- Clinical Observation of the Sick Child {CASEID} <idnum>

{DIS} District # (1=Kampot, 2=Kandal, 3=Takev {Date} <dd/mm/yyyy> {MT} __
{INTN} Interviewer Number ## {Month ##
{CARE} Caretaker # (F9 {OTHS} If Other, Specify <A >
{NUM} Number ## (F9 Health Center {HCEN} Name _____
{TYP} Type # (1=Hosp, 2=Clinic, 3=Health Post)
{HAG} Health Agent # (1=Doc, 2=Nurse, 3=Midwife, 4=Dispenser)
{IDN} Child's Identification Number ##
{DOB} Date of Birth <dd/mm/yyyy> {AGE} ## ## (Months)
{SEX} Sex # (1=Male, 2=Female)
{WEIGHT} Weight ## ### (KG) {HEIGHT} Height ## # (CM)

Time at start of observation {H1} ## {M1} ##

PURPOSE OF VISIT

1 What are the reasons given by the caretaker for the visit?
(Type [1] for all correct responses)

{FVR} Fever #
{DIAR} Diarrhea #
{CGH} Cough #
{OTH2} Other #
{SOR} Specify <A >

HISTORY OF THE ILLNESS
2 Did the health agent ask (1=Yes, 2=No)

{DURI} Duration of the illness #
{DRNK} If the child can drink or breastfeed #
{VMT} If the child has vomited everything #
CONV} If the child has had convulsions #
BLD} If there was blood in the stool #
{DIFF} If the child had difficulty breathing #

CLINICAL EXAMINATION

3 Did the health agent evaluate (1=Yes, 2=No)

{CGS} General status of the child # (Normal, irritable, lethargic, etc)
{TMP} The child's temperature # (By hand or thermometer)
{THST} The presence of thirst # (Offered the child liquid
{DEH} The presence of dehydration # (Abdominal skin pinch)
{RESF} Respiratory frequency # (Breaths per minute while child calm)
{CHIN} Chest indrawing #
{STFN} The presence of a stiff neck #
{SMSL} The signs of measles # (Rash, runny nose or red eyes)

KEY INDICATOR

DCEV: Diarrhea correctly evaluated = (1=Yes, 2=No

Definition

(If DIAR=1 and DURI=1 and BLD=1 and at least 2 or CGS=1 or THST=1 or DEH=1, then DCEV=1, else DCEV=2)

Time at end of observation (H2) ## (M2) ##

{LOOM}	Length of Observation (Minutes)	###
{LOOH}	Length of Observation (Hours)	# ##
{TEND}	The end #	

9 Entering data with ENTER

Tutorials

- Go to C \EPI6>
  1 Type <EPI6>
  From the **Tutorials** menu choose **Entering Data** and follow the instructions on the screen This will work if you have created FIRST QES in the Making a Questionnaire tutorial

- Go to C \DOEPI>
  1 Type <DOEPI>
  Choose **Exercises** from the main menu followed by **Programming Data Input** With the down arrow, go to Instructions and press <Enter> to begin and proceed to **Making a Data (REC) File and Entering Data**

Overview

The ENTER program will create a data file from a questionnaire Once the file is created, ENTER guides the data entry process so that any number of data records ("copies" of the questionnaire) may be added

To use ENTER, first create a questionnaire as described in the previous chapter using EPED in the QES mode You can also create the questionnaire in another word processor such as Word Perfect if you save it as a plain text (DOS) file and give it a name ending in QES The ENTER program will create an Epi Info database (REC) file using the questionnaire to provide information on the file structure Once created the file may be loaded into ENTER for adding more records or editing those already entered If the questionnaire is revised ENTER can be instructed to revise the REC file accordingly

Epi Info 6 also contains a new program called ENTERX that uses Random Access Memory (RAM) above 640 K It allows more complex questionnaires, larger check files and more related files at one time than the standard ENTER If you are having difficulty with RAM capacity try running ENTERX rather than ENTER

Running ENTER

Go to C \EPI6 Type <ENTER> to go directly to the program If you wish to use the EPI6 menu program, type <EPI6> to go to the main EPI6 menu At the Programs command press <Enter> move the cursor bar to the ENTER data choice and press <Enter> or simply type <N>

In either case, a screen will appear asking for the name of the data file and displaying the following menu choices

1 Enter or Edit data
2 Create new data file from QES file
---
3 Revise structure of data file using revised QES
4 Reenter and verify records in existing data file
5 Rebuild index files(s) specified in CHK file

On the bottom line of the screen you see

F2-Setup F9-List Files F10-Done

F2 is useful for defining the default data path, F9 lists REC or QES files and F10 allows you to exit **ENTER**

Editing an Existing Data File (Menu Choice 1)

Epi Info database files have names ending in " REC " because they contain data RECords To enter new data or edit existing data in an existing data (REC) file enter the name of the data file or press <F9> to see those available Choose the file from the menu by moving the bar cursor with the arrow keys and then press <Enter> to choose the file Enter a "1" for the menu choice press <Enter> and the file will appear on the screen ready for editing or addition of new records More details on entering and editing data are given in subsequent sections of this chapter

Making a New Data (REC) File (Menu Choice 2)

ENTER constructs a new data file by reading the questionnaire (QES) file and using the information to set up a new data (REC) file To make a new data file, first make a QES file as described in the previous chapter In the ENTER program give the name of the data file to be created (the REC file), press <Enter>, and then enter "2" for the menu choice ENTER will now ask for the name of the QES file Use the same file name for the data and questionnaire files, calling them QUEST1 REC and QUEST1 QES, for example There is no need to type the suffixes REC and QES because they are automatically added by the program

When the cursor is in either the "Data file" or "Questionnaire file" blanks you may press <F9> 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 <F9> Moving the cursor bar around with the arrow keys and then pressing <Enter> will choose a file from the list

When you have specified both the REC and the QES 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

Murphy's Law Strikes Again

Once you have created a * REC file from your * QES file, you have protection against the fateful day when you will accidently delete the only copy of your * QES file There is no need to panic for Epi Info provides the cure to this malady with the program REC2QES EXE If you have just deleted QUEST1 QES just type <REC2QES QUEST1> and it will be reconstructed The results will not be perfect and will require some editing but this is surely a lot better than having to start again from scratch

Entering Data

Entering data means typing the appropriate responses in the blanks on the screen The cursor will move 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 and <Esc> returns the original command line

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 the cursor one character to the 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

54

<End>	Moves to the last entry field in the questionnaire
<Ins>ert	Turns the insert mode on or off 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> in a blank field without entering data enters a missing value
- When a field is full, the cursor moves automatically to the next field unless this feature has been turned off
- 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 another opportunity to 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 EPI6 menu or to DOS

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 <F7 > <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 <F3> for "Find 1st" You might type SMITH in a "Lastname" field, for example, and then press <F3> All the records with LASTNAME = SMITH will be listed on the screen

Use the up and down arrow keys to move to the record desired and select it by pressing <Enter> The record will then appear on the screen in questionnaire format

ENTER will search in this way using **any combination of fields that you enter after pressing <Ctrl-F> for the FIND command**--a multi-field 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

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 return to 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 <F6> will mark the current record as deleted The computer will beep and 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 <F6> again This key thus alternates (toggles) between deleting and undeleting

Revising the Structure of a Data File (Menu Choice 3)

If ENTER detects that the date/time of a REC file is older than that of the QES file with the same name you will be notified that the questionnaire file is newer that the data file and asked if you would like to update the REC file from the QES

At any time, 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 question naire file Then return to the ENTER program Enter the name of the REC file under "Data file " and press <Enter> Enter "3" for the menu choice and then give the name of the revised QES file for the questionnaire

When you 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 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

Reentering and Verifying Records in a Data File (Menu Choice 4)

After records have been entered in a data file ENTER allows another operator to enter them again and verify that they have been entered correctly To do this, run ENTER and give the name of the existing data file choosing "Reenter and verify records in existing file" from the menu **In reality, you will probably never have an occasion to use this option, since one is usually hard pressed to get the data entered once**

Rebuilding Index Files (Menu Choice 5)

Index files specified in CHK files (see later chapters) are normally maintained automatically Occasionally an index file may become damaged This can happen if a power failure occurs while an index is being updated or a copying error destroys an index file In such cases, the Choice 5 on the ENTER menu is available to correct the problem After specifying a data file, Choice 5 will rebuild the index files for the database

Using ENTER When a Check File is Present

CHK files are 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 They perform checks only on data entered while the CHK file is present not on previously entered records

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 that will be accepted If no CHK file is active, this may be "All entries allowed," or "Integers allowed "

10. Improving data entry with CHECK

At this point in the Epi6 manual the next chapter discusses the ANALYSIS program While it is true that CHECK is not absolutely essential for data entry, it is extremely useful to facilitate the process and minimize entry errors It can t be emphasized enough that a survey is only as good as the quality of it's data We have all heard of GIGO -- Garbage In Garbage Out There is no such thing a survey without data errors Those you discover may be correctable, the others will constitute "noise" in the data Knowing and using the basic essentials of CHECK will go a long ways to reduce a variety of error types such as

- **Range** Minimum and Maximum values
- **Transposition of numbers** 13 instead of 31
- **Coding errors** Placing a health center in district 2, when it belongs in district 1
- **Logical errors** Asking a man his pregnancy history
- **Calculation errors** A child's age, length of the interview
- **Missing values** for key indicators can be reduced by forcing data entry

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 Even skilled people however, will also make their fair share of mistakes

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

Running CHECK

Before running CHECK a REC file must already exist for the questionnaire you wish to enhance 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**	**Shft-F6-Display**	**Ctrl-F6-Delete**
F3-Repeat	**F9-Edit field**	**Jump**	**F7-Add**	**Shft-F7 Display**	**Ctrl-F7-Delete**
F4-MustEnter	**F10-Quit**	**Codes**	**F8-Add**	**Shft-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 Any of the following commands may also be entered manually using <F9>

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, for example

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 MustEnter is specified, missing (blank) values are not allowed This feature is useful to ensure that fields essential for the calculation of key indicators are not missing To specify Must Enter for a field, place the cursor in the proper field and press <F4> Pressing <F4> again removes the MustEnter condition

Repeat <F3>

During entry, the value of the field will be set automatically to the value in the last record entered This is useful when a series of questionnaires all contain the same date and health center, for example The value of the field 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 the child only presented with fever, for example, the cursor should jump over the section on diarrhea Also if certain fields are calculated automatically, the cursor should jump over them to ensure that the data entry person does not enter another value

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 (**AutoJump**) , press <F7> while the field is blank Such unconditional jumps are useful for changing the order in which fields are entered on a questionnaire A jump may be removed with <Ctrl-F7>

Linking of Fields and Automatic Coding <F5> and <F8>

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 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 The keys <Shift-F8> will display all codes for a field

The Edit Field Command <F9>

The Edit Field function allows you to see and edit the commands that are inserted in the file by CHECK It confines the view to commands for the current field--the one containing the cursor You will often make errors when editing the file directly Sometimes CHECK is nice and informs you that at a certain line number there is an unknown command that cannot be processed Sometimes it rectifies the error for example, adding an END that you forgot At other times however, it kicks you out of the program and you lose anything you have typed since the last SAVE Thus to minimize frustration **write data to disk often**

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
      1
      2
   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
      1 Female
      2 Male
   END
END
```

Now the ENTER program will display the codes in a pop-up window if the <F9> 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

Other CHECK Features

CHECK can be programmed to do many interesting and complex operations beyond the automatic functions built into the menu These include mathematical and logical operations with more than one field, pop-up help windows and linking to other files for codes A check file QUEST1 CHK) has been written to accompany QUEST1 REC It contains many of the CHECK commands and examples of Boolean[1] mathematical formulas that give power and elegance to your data entry program The commands are as follows

Autojump	Autosearch
Clear	**Comment Legal**
Codefield	**Codes**
Goto	Goto Writerec
Help	Include
If Then Endif	**Jumps**
Key	**Legal**
Let	**MustEnter**
Range	**Repeat**

[1] An introduction to the history, definition and use of Boolean expressions has been prepared as an EPIAID tutorial in the text file BOOLE EPA You can access it directly at C \EPI6> by typing <EPED /BOOLE> It is also included in the CAMTOT program At C \Epi6> type <CAMTOT> In the **Programs pop-up** menu choose **Boolean Algebra** with the bar cursor, and you're on your way

The **bold** commands have already been discussed above The contents of QUEST1 CHK are printed below By actually producing this file you will begin to understand how to use CHECK And you will refer back to it often to remind yourself how to program certain functions

```
BEFORE RECORD
  HELP QUEST1 HLP 1 1 1 CONTEXT
  HIDE MT
  * MT WILL NOT APPEAR ON THE SCREEN
  * FILE NAME  QUEST1 CHK
  * A LINE STARTNG WITH AN ASTERISK, IS A COMMENT IGNORED BY THE COMPUTER)
  * THIS COMMAND LOADS A HELP FILE THAT WAS WRITTEN AS QUEST1 TXT
  * AND COMPILED TO QUEST1 HLP BY TYPING <HELPPREP QUEST1 TXT>
  * IT CAN BE VISUALIZED BY TYPING <BINDER QUEST1>
  * HELP IS ACTIVATED ANYWHERE IN THE QUESTIONNAIRE WITH <F1>
END

DIS
  HELP QUEST1 HLP 2 1 -1 CONTEXT
  REPEAT
  MUSTENTER
  LEGAL
    1
    2
    3
  END
  * <SHIFT F1> ACTIVATES THE HELP FILE FOR THIS FIELD
  * REPEATS THE VALUE IN THE PREVIOUS RECORD
  * YOU MUST ENTER A LEGAL VALUE TO PROCEED
  * JUST PRESSING <ENTER> WILL NOT WORK
  * YOU COULD HAVE ALSO WRITTEN  RANGE 1 3
END

DATE
  REPEAT
  MUSTENTER
  LET MT = DATE [4,2]
  * LET IS OPTIONAL
  * DATE IS A STRING OR TEXT VARIABLE
  * SUBSTRINGS CAN BE EXTRACTED FROM TEXT VARIABLES
  * MT CONTAINS 2 CHARACTERS FOR MONTHS, STARTING AT POSITION 4 OF DATE
  * MT IS A TEXT (NOT NUMERIC  FIELD
END
```

```
INTN
  REPEAT
  AutoJump CARE
  COMMENT LEGAL
    INCLUDE INTERV T T
  END
  * THE CODES ARE IN THE FILE INTERV TXT RATHER THAN BEING WRITTEN HERE
  * VISUALIZE THEM BY PRESSING F9
  MONTH = (MT="01")+(MT="02")*2+(MT="03")*3+(MT="04")*4+(MT="05")*5
  MONTH = MONTH+(MT="06")*6+(MT="07")*7+(MT="08")*8+(MT="09")*9
  MONTH = MONTH+(MT="10")*10+(MT="11")*11+(MT="12")*12
 * MONTH IS CALCULATED TO BE THE NUMERIC VALUE OF MT
END

MONTH
  NOENTER
  * THE COMPUTER ENTERS THE VALUE  NOT YOU
END

CARE
  COMMENT LEGAL
    1 MOTHER
    2 GRANDMOTHER
    3 FATHER
    4 OTHER
  END
  Jumps
    1 NUM
    2 NUM
    3 NUM
  END
  * SKIP OTHS IF THE VALUE IS 1-3, OTHERWISE GO TO OTHS
  * ERASE OTHS
END

OTHS
  IF OTHS <> "" AND CARE <> 4
  THEN
    GOTO CARE
  ENDIF
  * IF CARE IS NOT EQUAL TO 4, THEN ANYTHING WRITTEN
  * IN OTHS IS NOT LOGICALLY APPROPRIATE
END
```

```
UM
  KEY 1
  AutoJump FLAG
  CODEFIELD HCEN
  CODES HCENTER REC
  * KEY CREATES - INDEX FILE * I  FOR THIS FIELD
  * THIS WILL SPEED UP THE SEARC- FOR RECORDS CONTAINING A PARTICULAR VALUE
  * GO DIRECTLY TO FLAG  BECAUSE HCEN AND T P WILL BE FILLED IN AUTOMATICALLY
  * LINKS NUM TO HCEN
  * CODES ARE CONTAINED I  2 FILES  HCENTER REC AND HCENTER CHK)
  IF NUM=1
  THEN
    TYP = 1
  ENDIF
  * IF NUM = A CERTAIN CENTER,  E KNOW IT IS A HOSPITAL
END

HCEN
  KEY 2
  CODEFIELD NUM
END

TYP
  RANGE 1 3
  * 1 IS THE MINI UM AND 3 IS T-E MAXIMUM ACCEPTED VALUE
END

FLAG
  RANGE 1 4
END

DOB
  CLEAR AGE
  IF DOB > DATE
  THEN
    HELP " THE C-ILD HASN'T BEEN BORN YET" +0 +0
    CLEAR DOB
    CLEAR DATE
    CLEAR MONTH
    GOTO DATE
  ENDIF
```

```
  IF DOB <> ""
  THEN
    AGE = (DATE-DOB)/30.44
    GOTO SEX
  * SKIP AGE SINCE IT WILL BE FILLED IN AUTOMATICALLY
  * AGE IS CALCULATED AUTOMATICALLY
  ENDIF
END

AGE
  RANGE 0 59
 * 0 TO 59 MONTHS
END

SEX
  RANGE 1 2
END

H1
  RANGE 7 17
  * WORK HOURS ARE 7 00 AM TO 5 00 PM
END

M1
  RANGE 0 59
  * 0 TO 59 MINUTES
  IF M1 =
  THEN
    M1 = 0
  ENDIF
  * CHANGE A MISSING VALUE TO 0 FOR THE DURATION CALCULATIONS
END

OTH2
 CLEAR SOR
 IF OTH2 <> 1
 THEN
  GOTO DURI
 ENDIF
END
```

```
SOR
 CODES
    VOMIT 'OMITI G
    /OMITS /OMITI'G
    ' VOMITING
    * AUTOMATIC CODING OF A SWERS TO OPEN ENDED QUESTIONS
 END
 IF SOR <> '" AND OTH2 <> 1
 THEN
   GOTO OTH2
 ENDIF
END

SMSL
  AutoJump H2
  DCEV = (DIAR=1)*(DURI=1 *'BLD=_ *'((CGS=1)+(THST=1)+(DEH=1))>1)
  * CALCULATE T-E VALUE FOR DIARRHEA CORRECTLY EVALUATED
  DCEV =  DCEV=_)+(DCEV=0)*2
  * 1 = CORRECT, 2 = NOT CORRECT
END

H2
  RANGE ~ _7
END

M2
  RANGE 0 59
  AutoJump TEND
  IF M2 =
  THEN
    M2 =
  ENDIF
  LOOM =  H2-H1 *60 + (M2-M_
  * DURATION OF OBSERVATION I  MINUTES
  LOOH =  H2-H1) + (M2-M1)/60
  * DURATION OF OBSERVATION IN HOURS
END
```

```
TEND
    * WRITEREC ASKS IF YOU WANT TO WRITE DATA TO DISK  SAVE T-E RECORD
    * THIS HELP STATEMENT IS WRITTEN DIRECTLY IN THE * CHK FILE
  IF TEND = 1
  THEN
    GOTO WRITEREC
  ELSE
    HELP ""YPE 1 TO SAVE RECORD\nPress ESC" 21 21
    GOTO TEND
  ENDIF
END
```

The four files used by this check file contain the following commands

HCENTER QES

```
CODE ##
HCENTER  <A                >
```

HCENTER CHK

```
CODE
  KEY _
  AUTOSEARCH CODE
END
```

INTERV TXT

```
_ "Interviewer #1"
2 "Interviewer #2"
3 "Et Cetera"
```

QUEST1 TXT

```
  IDTF 80
 TOPIC 1 OVERVIEW OF HELP
This is general help for the overall questionnaire  It is displayed when <F1>
is pressed, because QUEST1 CHK contains
  BEFORE FILE

    HELP QUEST1 HLP 1 1 1 CONTEXT
  END
The key word CONTEXT specifies that this help is displayed in response to
pressing the help key  The declaration's location within the BEFORE FILE block
determines that the help key is <F1> and that the help can
be displayed from anywhere within this check file system

'TOPIC 2 DIS
This is help for the DIS field  It is displayed when the cursor
is in the DIS field, and either <Shift-F1> or <?> is pressed
<?> activates help ONLY IF the cursor is in the first column of the field,
in any other location, <?> behaves as a normal character key
This help is displayed because QUEST1 CHK contains

  DIS
    HELP QUEST1 HLP 2 1 +- CONTEXT
  END
```

All four files are ASCII (DOS text) files that you can write using EPED Data entry files (HCENTER REC) for HCENTER QES must be created by ENTER and then one health center code and name are entered per record The advantage of this system as opposed to typing the codes into the original check file is that you could have hundreds of codes without worrying about memory limitations Furthermore if such a data base already existed in the Ministry of Health it could be downloaded directly into the Epi Info format without having to be retyped The help file (QUEST1 TXT) must be compiled to QUEST1 HLP using HELPPREP

Many of the commands in QUEST1 CHK are advanced and don t need to be learned by a beginner However if you practice creating and using QUEST1 CHK and its accompanying files you are well on your way to becoming a master of check files

11. Data Preparation with ANALYSIS

Overview

ANALYSIS produces lists frequencies tables statistics and graphs from Epi Info files Simple commands will cause ANALYSIS to select records using specified criteria sort or list records do frequencies or cross tabulations and do logical or mathematical operations on a variable (VARIABLE is a synonym for FIELD) put the results in a new variable and direct the results to the screen, a printer, or a disk file The commands constitute a programming language and they may be entered one by one from the keyboard or placed in a program file (* PGM) that is then "RUN" from ANALYSIS

Getting Acquainted

To run ANALYSIS, go to C \EPI6 and Type <ANALYSIS> or you may run it from the EPI6 menu You will see now see two windows on the screen, a lower one for entering commands and a larger upper one where the results of the commands will appear

At the bottom of the screen some key commands are indicated

F1-Help F2-Commands F3-Variables F4-Browse F5-Printer on F9-DOS F10-Quit

and information about data files and available memory is displayed at the top

Whenever the cursor is at the EPI6> 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

Pressing the function keys shown at the bottom of the screen allow selection of help topics <F1>, commands <F2>, and variable names <F3> from lists that appear on the screen After pressing one of the function keys 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

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 You may also type commands from the keyboard

Whenever a command has been entered--READ for example--pressing Help <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 EPI6> 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 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

EPI6> **READ**

read and press <Enter> A directory of files will appear in a window Move the cursor bar with the arrow keys and choose a file by pressing <Enter> Or you may type the appropriate path and file name such as

C \CAMHFS\QUEST1\QUEST1

ANALYSIS will now use this file for all subsequent operations until another READ is performed It has become the "active data set

Examining and Editing Records--BROWSE and UPDATE

There are several methods of reviewing and editing data The ENTER program may be used to examine and perhaps change each record In ANALYSIS, frequencies can be performed on each field with FREQ *, and outlying data points can be examined for errors ANALYSIS offers another method, the BROWSE and UPDATE commands, in which records are displayed in tabular fashion like a spreadsheet

The difference between the two commands is that UPDATE allows the columns of data in the file to be changed permanently while BROWSE only allows you to look Thus before you use UPDATE, protect yourself by making a backup copy of the REC file **Mistakes do happen!**

To use BROWSE press <F4> The records in the current file appear on the screen in spreadsheet format By moving the cursor to the right or down you can see all the fields and all records If you prefer the original questionnaire format press <F4> as indicated at the bottom of the screen for "Full Screen " The questionnaire will appear as if in ENTER and additional records can be viewed with the <F8> and <F7> keys Home/End is a toggle that will switch you back and forth from the first to the last record Either <Esc> or <F10> will bring back the EPI6> prompt

UPDATE works in exactly the same way as BROWSE, but you type <UPDATE> at the EPI6> prompt rather than pressing a function key The same spreadsheet format appears, but the current field is presented at the top of the page, ready to receive an entry

If you make an entry and then press the <Down arrow> key rather than <Enter> the cursor will subsequently move down the column rather than across the page when you press <Enter> greatly simplifying the editing of a column of data BROWSE and UPDATE can display selected subsets of the data file SELECTing is described below SORTed data appear in the sorted order, a useful feature for eliminating duplicate records in a data file Note that a record can be deleted just as in ENTER by pressing the <F6> key, and that the entire questionnaire can be displayed as in ENTER if desired

When you have completed the changes on a record you will be asked if you want to save the changes If you respond with "Y " the record will be permanently changed in the data file You will find UPDATE a convenient way to examine a column of entries and correct typing or coding mistakes made on data entry

Data Verification and Cleaning

As previously mentioned real-life surveys produce "dirty" data, and often the most time-consuming part of the analysis is the "cleaning" of the data set Methods for viewing and editing data in ANALYSIS include sorting records into numeric or alphabetical order, selectively including or excluding records from analysis, defining and assigning values to new variables, grouping or otherwise recoding variables performing conditional operations (IF THEN), managing dates and time and setting a number of options in ANALYSIS that can affect results

These operations are not individually complex but a given data set may require many such operations before all fields are in satisfactory condition for analysis It is useful to incorporate the commands for these operations into a program file so that the program can be run many times as small additions and changes are made We will therefore describe how to make program files before we proceed

Writing Programs

Programs are text files created with a word processor that contain the same ANALYSIS commands you have been using They can be created in EPED, EDIT or most other word processors in text-file format (Word WordPerfect etc)

If you create a program called EDIT1 PGM the RUN command in ANALYSIS is used to perform all the commands that are contained in the program

EPI6> RUN C \CAMHFS\QUEST1\EDIT1

EDIT1 PGM was written to assist in the data preparation of QUEST1 REC It can also be run from the CAMHFS menu by going to **Edit** on the command line placing the bar cursor on **1 Clinical Observation** and pressing <Ctrl-Enter> as indicted by the happy face symbol ☻

```
* FILE NAME = EDIT1 PGM
READ \CAMHFS\QUEST1\QUEST1 REC
ET PROCESS=UNDELETED
SORT NUM IDN
ROUTE TEMP$$$ REC
WRITE RECFILE
COPY TEMP$$$ REC \CAMHFS\QUEST1\QUEST1 REC
ERASE TEMP$$$ REC
READ \CAMHFS\QUEST1\QUEST_ REC
UPDATE
QUIT
```

It reads QUEST1 REC into ANALYSIS sorts the records on Health Center Number (NUM) and Child Identification Number (IDN), copies the sorted records back to QUEST1 REC via a temporary file (TEMP$$$ REC) deletes the temporary file a presents the data in UPDATE mode, allowing you to make correct errors that you may see

Not all program files are this long They may be as short as one line They are most useful when the operations are complex enough to be a typing burden or when they will be performed repeatedly

Programs are also useful for "debugging" statements with complicated logic (lots of IF's, AND's and OR's) These can be typed once to form a program and then revised as the "bugs" are worked out of the statements by repeated RUNning It is helpful to set up the READ, SELECT, RECODE IF and LET statements for processing a REC file, and then to RUN the program before using the keyboard to enter commands for tables, lists frequencies, or means Commands that have been entered from the keyboard may be saved to form programs using the SAVE command as described in the next section

Recalling, Saving, and Loading Analysis Commands

As commands are entered ANALYSIS saves each line for possible reuse When the cursor is at the EPI6> 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

One way to make a program file is to enter commands interactively (at the keyboard) at the ANALYSIS Epi6> prompt If you wish to save the commands already entered (up to the limit of the "Command Stack", usually 20) use the SAVE command and the name of a file, for example

SAVE FIRST PGM

All the commands in the current Command Stack will be placed in the file, which can then be edited in EPED to remove unwanted commands or add new ones SAVE is useful if you have entered a number of commands and decide you would like to make a permanent program file Use the SAVE command to make the file and then EPED to edit and perhaps refine the program so that it can be RUN repeatedly in ANALYSIS If you have SAVEd a command stack and would like to recall it from the file this can be done using the LOAD command LOAD <File name> will place the commands on the Command Stack where they are available for editing, but, in contrast to RUN commands will not be executed until you recall them and press <Enter>

The number of commands that can be SAVEd can be changed by typing <SET COMMAND STACK = 30> or any number up through 100

Access to DOS Commands Directly from ANALYSIS

Any DOS command, and even other programs can be run directly from ANALYSIS by pressing the <F9> DOS key as indicated at the bottom of the screen The DOS prompt appears and you may give any DOS command or run other programs Typing "EXIT" at the DOS prompt brings back ANALYSIS with all previous processing intact This procedure is useful, for example for making a backup copy of a data file (Copy QUEST1 REC QUEST1B REC)

Putting Records in Order -- The SORT Command

We have seen the SORT command in the above EDIT1 PMG file Sometimes it is important to have the records in a file arranged in a particular order -- alphabetically by NAME, for example, or numerically by AGE This is particularly useful for the LIST and WRITE commands The SORT command can do this for several fields at once temporarily arranging the records into the new order The order resulting from a SORT disappears with the next READ or on exiting from ANALYSIS unless a permanent file of the sorted records is made by using the ROUTE and WRITE RECFILE commands as seen in EDIT1 PGM It is not necessary to SORT records to do tables or frequencies, since tabulations are kept in order automatically by the program

Selecting Records for Processing

ANALYSIS normally processes all records in a file You can however SELECT records for processing based on specific criteria The following commands would select only males with ages 20 through 49

SELECT SEX = "M"
SELECT AGE < 50
SELECT AGE > 19

Successive SELECT statements accumulate as though AND had been inserted, so that the same commands could be combined as

SELECT (SEX = "M") AND (AGE > 19) AND (AGE < 50)

Parentheses are needed to separate conditions in some cases and it is safest to use them routinely as they do no harm SELECT will also accept the shortened form for multiple conditions

SELECT AGE <50 and >19

Be careful with the logic the statements

SELECT GROUP = 1
SELECT GROUP = 2

will not select any records for processing, since no record has the variable called GROUP set to 1 and simultaneously to 2! Probably SELECT (GROUP = 1) OR (GROUP = 2) is what was intended This will select all the records that are in either group 1 or group 2

SELECT can be used to omit records also If you want to eliminate records with GROUP equal to 9, they can be marked as deleted in the ENTER program, or SELECT can be used to omit them from processing with the command

SELECT GROUP <> 9

SELECT conditions remain in effect until you READ another file, leave ANALYSIS, or cancel them by issuing the command

SELECT

without any criteria A permanent file containing only the new records can be made by using the ROUTE and WRITE RECFILE commands as described above

Defining New Variables

For many data "cleaning" operations such as recoding variables new variables must be created The results of recoding or other operations are placed in the new variables temporarily while screen or printed output is obtained These **defined variables** disappear on exit from ANALYSIS, but can be preserved in a permanent file

The commands that create the variables can be preserved in program files so that they can be created again each time the program is run

To define a new variable in ANALYSIS the same special characters used in the questionnaire can be used to describe the type and length of the variable Thus

```
Define NAME _____
DEFINE AGE ###
DEFINE ID ##########
DEFINE PRICE ### ##
DEFINE ONSET <mm/dd/yy>
DEFINE ONSET <dd/mm/yy>
DEFINE NAME <A        >
```

are all allowed

When Commands Are Processed

In writing programs and using defined variables it is important to understand when commands are processed Certain commands such as LIST, FREQ, TABLES MEANS, and the graphing commands cause ANALYSIS to read every record in the current data file and send the results to the screen or other output device Only records currently SELECTed are processed

Assigning a Single Value: the LET Command

The LET command like the RECODE and IF commands, is performed each time a record is processed LET assigns a value to a defined variable possibly as a result of an arithmetic calculation

The simplest operation is to assign the value of an existing variable to a newly defined variable If your questionnaire contains the variable AGE, the following commands define a new variable called AGE2 and copy the value of AGE to AGE2 for each record

```
DEFINE AGE2 ##
LET AGE2 = AGE
```

The use of the word "LET" is optional and the second command may be written as

AGE2 = AGE

Missing Values

Missing values are represented in ANALYSIS programs by a period Thus, to set AGE to the missing value, the command is

AGE =

If a new file is written from ANALYSIS missing values appear in the records as blanks, just as though entered from a questionnaire The " " convention is for ANALYSIS programming only The " " symbol does not have special meaning in ENTER, where missing values are entered as blanks by pressing <Enter> in an empty field

Recoding Data

The data items entered are not always in the form needed for processing Frequently ages are entered to the nearest year, and tables are desired by 5-year age groups County names or codes may be given but tables are needed by district or region, or you may want to substitute descriptive words for numeric codes The RECODE command allows you to assign new values to existing data and to place the results in defined variables

To use RECODE, it is usually best to DEFINE a new variable of the right type to receive the data

DEFINE AGEGROUP _____
RECODE AGE TO AGEGROUP BY 10

will place ages into 10-year age groups so that you can use AGEGROUP for tables or frequencies

You can also specify the exact codes desired for each value in a RECODE statement

```
DEFINE CTYNAME _____
RECODE CTYCODE TO CTYNAME 1=ROME 2=AUGUSTA
```

A frequency of CTYNAME will then list the number of records for "ROME" and "AUGUSTA" rather than "1" and "2"

IF Statements

Suppose that food histories have been collected from ill and not ill attendees at a church supper, as in the sample file OSWEGO REC You would like to know how many of the ill and well people ate any dairy product at the dinner The dairy products are MILK CHOCOLATE, and VANILLA (ice cream) First, define a variable called DAIRY and choose codes (1=Yes, 2=No)

```
DEFINE DAIRY #
```

and then use an IF statement to give it values

```
IF (VANILLA= "Y" ) OR (MILK = "Y") OR (CHOCOLATE = "Y")\
  THEN DAIRY = 1 ELSE DAIRY = 2
```

Note that the statement is too long for one line and that the "\ " (backslash) has been used to tell ANALYSIS that the next line is to be considered part of the same statement in a program Statements can be up to 254 characters long, with lines subsequent to the first one indented from the left margin Also notice that the command did not end in ENDIF as required in CHECK files

In writing conditional statements in the IF and SELECT commands it is never *in*correct to surround each complete condition with parentheses and to separate parts of a statement with spaces Values of text (String) variables should be enclosed in quotation marks "Nested" IF statements, in which one IF results in another IF statement, are not allowed The ELSE clause at the end of the IF statement is important to specify what should happen if the conditions are not met Of course these commands could also be written using Boolean expressions The following formulas would give the same result

```
DAIRY = ( (VANILLA= "Y" )+(MILK = "Y")+(CHOCOLATE = "Y"))>0
DAIRY = (DAIRY = 1)+(DAIRY = 0)*2
```

Dates

Dates have been a nuisance both before and since Pope Gregory XIII produced the modern calendar in 1582 One must cope with leap years European and American date order varying numbers of days in a month and the fact that leap years are omitted every hundred years unless the year is divisible by 400 Scholars say that the event which marked the beginning of the *Anno Domini* count actually occurred in 4 to 6 B C Against this background Epi Info does its best to make date handling easy

Date variables are created on the original questionnaire or as defined variables during analysis The standard variable types for dates within questionnaires are

<mm/dd/yyyy>	U S Format
<dd/mm/yyyy>	European and U S Military Format

To sort records using dates use the SORT command Epi Info will take care of sorting so that the year is placed first and the sort will be in correct order Arithmetic operations may be performed on date variables, with the results being in days Remember how we calculated a child's age in months the QUEST1 CHK where DATE was the date of the survey and DOB was the Date of Birth

AGE = (DATE - DOB)/30 44

Dividing by 365 25 instead of 30 44 would have given AGE in years These same commands could have been executed in ANALYSIS

We have also already seen when we calculated the child's MONTH of birth in QUEST1 CHK that dates are actually recorded in the file as text variables and must be surrounded by quotation marks if used as literal parts of commands For example, how many days before 1 January 2000 was the survey conducted?

LET DAYS = "01/01/2000" - DATE

12 Analysis with ANALYSIS

For this chapter we will assume that you have chosen the file OSWEGO.REC supplied with the Epi Info system by giving the following the command after running ANALYSIS

EPI6>**READ OSWEGO**

OSWEGO.REC often used in teaching exercises contains interview information from a foodborne outbreak. The file contains the results of interviews with attendees at a church supper that was followed by a gastroenteritis outbreak.

Producing a Line Listing

The first step in data analysis is to scan the data visually to gain an overall impression and see what further analysis might be appropriate. A "line listing" is helpful for this purpose. To produce a listing of the records in the file, type

EPI6>**LIST**

The command LIST will display only as many variables as will fit across the current screen width. If you would like to list all variables, use

EPI6>**LIST** *

The "*" is a wildcard for "all variables". LIST followed by one or more variable names lists only those variables.

LIST * NOT NAME ADDRESS will list all fields *except* NAME and ADDRESS.

Frequencies

The frequencies command (FREQ) will count each category for a specified variable and give the absolute and relative frequencies for each category.

EPI6>FREQ SEX produces this result

```
SEX    |  Freq  Percent    Cum
-------+----------------------
F      |    44   58 7%     58 7%
M      |    31   41 3%    100 0%
-------+----------------------
Total  |    75  100 0%
```

The number (frequency) in each category is given first, followed by the percentage of the total and the cumulative percentage If statistics is SET to ON (SET STATISTICS = ON) and the field types are numeric the sum mean and standard deviation are also printed (**NOTE** If the numbers represent codes for nominal or ordinal data, then the statistics have no meaning An example of nominal data represented by numbers would be Male and Female coded 1 and 2)

The command FREQ * will produce frequencies for all the variables in your questionnaire, a convenient way to begin the analysis of a new data set STATS1 PGM performs this function for QUEST1 REC of the CAMHFS survey

```
* FILE NAME  \CAMHFS\QUEST1\STATS1 PGM
SET NOECHO=ON
SET STATISTICS=OFF
READ \CAMHFS\QUEST1\QUEST1 REC
ERASE \CAMHFS\QUEST1\STATS1 OUT
ROUTE \CAMHFS\QUEST1\STATS1 OUT
TITLE 1 "\C FREQUENCIES FOR QUESTIONNAIRE 1"
TITLE 2 "\CDate  @SYSTEMDATE
FREQ *
ROUTE SCREEN
SET NOECHO=OFF
CLS
?Press <Enter> to continue  <F5> to print  <F10> to quit  ?
CLS
TYPE \CAMHFS\QUEST1\STATS1 OUT
?Press <Enter> to terminate and return to menu  ?
```

The data are read into ANALYSIS, any existing output file is erased, the path and name for the output file (STATS1 OUT) are defined two titles are written and frequencies are run on all variables STATS1 OUT is a text file, that can be loaded into any word processor

FREQ with a series of variable names following separated by spaces will do separate frequencies for each of the variables listed

Cross Tabulations the TABLES Command

Now back to ANALYSIS and OSWEGO The TABLES command will count the records in which the values fulfill criteria for two fields at the same time Thus the command

EPI6>**TABLES ILL VANILLA**

will result in

```
                      VANILLA
ILL          |      +      - |  Total
-------------+---------------+-------
           + |     43      3 |     46
           + |     11     18 |     29
-------------+---------------+-------
       Total |     54     21 |     75
```

followed by a many statistical indicators including three variations of Chi square
Note that the values for ILL and VANILLA are presented as +/- instead of **Y/N**

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 example, tables of VANILLA consumption by ILLness stratified by SEX will be generated by the command

EPI6>**TABLES VANILLA ILL SEX**

This will result in separate tables for male and female cases, followed by a multitude of statistical indicators, unless you SET STATISTICS = OFF

The MEANS Command

The TABLES command is for discrete 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 statistical analysis

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"

Missing Values

Missing values in Epi Info are entered as blanks in the actual records During data entry, pressing <Enter> in a field rather than entering data will result in a missing value In the MEANS, TABLES and FREQ procedures missing values will be ignored if SET IGNORE is ON (the default condition) 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 these titles 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 Commands consist of mandatory and optional components shown surrounded by curly brackets { } Up to three titles may be added to the graph using the TITLE 1 <Text> etc

- **BAR**

The BAR command will make a bar chart of the variable named Bar charts are useful for displaying the counts for values of both ordered an d non ordered variables Normally, vertical bars represent counts of records with different values of the variable The syntax for the commands is as follows

BAR <Variable name> {/Y=<min>-<max>} {/LABEL=<Text>}

Bar charts differ from histograms in having the bars separated by spaces and in omitting counts for values in which the count is zero The maximum vertical axis value can be increased by adding the parameter 'Y=min-max in which min is zero and max can be any number larger than the number automatically assigned by the program This adjustment is useful if you are making a series of graphs and want the vertical axis values to be comparable There should be a space before the "/", but no spaces in 'Y=min-max"

The default label on the vertical axis is "Count " If the /LABEL=<Text> parameter is given the label used is the one given as Text

Examples

BAR ROLLS

BAR ROLLS /Y=0-75

BAR ROLLS /LABEL=Cases

- **HISTOGRAM**

The HISTOGRAM command will make a histogram of the variable named and vertical bars represent counts of records with different values of the variable The commands and syntax are the same as for BAR

HISTOGRAM <Variable name> {/Y=<min>-<max>}
{/LABEL=<Text>} (All on one line or with continuation symbol \)

Histograms are useful for displaying the numbers of cases occurring over a period of time, and this type of graph is often called an epidemic curve The dates may be single or grouped, but if grouped, should be in groups containing the same numbers of years days etc Histograms differ from bar graphs in having the bars contiguous--not separated by spaces-- and in displaying counts even for values in which the count is zero

Examples
HISTOGRAM ONSETDATE
HISTOGRAM ONSETDATE /Y=0-75

- **LINE**

The LINE command will make a line chart of the variable named **first** Line charts are useful for displaying the counts of ordered data such as age, blood pressure or a series of dates Grouping the data often makes the line smoother and easier to understand Points on the line represent counts of records with different values of the variable

LINE <Variable name> {<Variable name>}
{/LABEL=<Text>}

More than one line may be graphed by giving the name of an additional stratifying variable The command LINE AGE ILL will produce one line for the ages of the ILL and one line for the ages of those not ILL if ILL is a Yes/No variable

Examples
LINE AGEGROUP
LINE AGE ILL

- **PIE**

The PIE command will make a PIE chart of the variable named Normally, segments of the pie will represent counts of records with different values of the variable Here is the syntax

PIE <Variable name>

Pie charts should be used for non ordered data such as race or sex (The sum is 100%), but are usually not appropriate for ordered data such as age or blood pressure where a bar or line chart would be better

Example
PIE RACE

- **SCATTER**

SCATTER will make a scatter plot with Variable1 on the horizontal axis and Variable2 on the vertical axis Scatter plots are most useful when both fields contain continuous numeric data like age or blood pressure rather than coded or yes/no information The syntax is

SCATTER <Variable1> <Variable2> {/R}{/X=<min>-<max>
/Y=<min>-<max>}

If "/R" is included after the field names a least-squares regression line will be drawn through the data points The maximum and minimum numbers on the X (horizontal) and Y (vertical) axes can be made more extreme by specifying new numbers with the X and Y parameters No spaces are allowed in the expressions Axis=<min>-<max>, but a space must be present before each slash ("/") The minimum cannot be increased nor the maximum decreased from the values chosen by the program

Examples
SCATTER SYSBP BODYMASS
SCATTER SYSBP BODYMASS /R
SCATTER AGE SYSBP /X=0-100 /Y=0-250

■ The vertical axis on BAR and HISTOGRAM Graphs

When you run a BAR chart or HISTOGRAM you see that in the default or "normal setting the Y (vertical) axis represents the counts of records with different values of the variable Often you are more interested in presenting the percentage than the number of cases Happily there is a way to do this as shown in following program file (BARHIST PGM) that runs in ANALYSIS and READs OSWEGO This file also introduces the command PICKLIST a feature that causes a list of items to be displayed on the screen, allowing the user to move the bar cursor to the desired item and choosing it by pressing <Enter>

```
* FILE NAME  BARHIST PGM
* BAR GRAPH OF ILL BY VANILLA
* NUMBER OF PEOPLE OR PERCENTAGE ON THE Y AXIS
* RACHA - CAMBODIA
* MARCH 1998
* BASICS
CLS
ECHO OFF
READ C \EPI6\OSWEGO REC
TITLE 1 PEOPLE SICK FROM EATING VANILLA ICE CREAM
DEFINE ILL_P  ## #### GLOB
DEFINE NREC      #### GLOB
DEFINE CT        #### CUMULATIVE
CT = CT + 1
LIST  NOECHO
IMMEDIATE NREC = CT
ILL_P = 100/NREC
CLS
PICKLIST 20 10
"BAR - NUMBER OF PEOPLE" BAR ILL VANILLA /LABEL="NUMBER OF PEOPLE"
"BAR - PERCENTAGE OF PEOPLE" BAR ILL VANILLA /SUM=ILL_P /LABEL=PERCENTAGE
END
PICKLIST 20 10
"RUN PROGRAM AGAIN' RUN BARHIST
"QUIT PROGRAM"  QUIT
END
```

The function of the command SUM will be explained in the following chapter on processing summary records

Sending Results to the Printer or to a File

In general it is easier to route results to a file that can later be printed by loading it into a word processor rather than by going directly to a printer

The ROUTE command

EPI6>**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 All files are closed by leaving ANALYSIS with <F10> When you wish to save a graph to a file, you must first indicate the format, followed by the file name

EPI6> SET GRAPH=SPCX
EPI6> ROUTE SEX PCX
EPI6> PIE SEX

You can visualize the file by exiting to DOS and using the PICEM EXE program

C \EPI6>PICEM SEX
C \EPI6>PICEM (will show all the * PCX files and you can choose with the bar cursor)

13 Processing Summary Records with ANALYSIS

Up until now we have been dealing with data where each record contains only one observation Many situations require the use of summary records to represent more than one person or case For example the district health center may process outpatient registers (pre-natal visits child immunizations etc) and send the cumulative results on a monthly or quarterly basis to the district These reports are then sent to the sent to the regional level, where results from many districts are combined and sent on the the national level SUMFREQ and SUMTABLES rather than FREQ and TABLES are used for processing these records The option {/SUM=<Field Name>} may also be added to the BAR, HISTOGRAM, LINE and PIE commands to produce graphs You may remember that SUM was used in the program file (BARHIST PGM) on page 71 to produce charts with percentages on the vertical axis

It is also sometimes useful to be able to use results from an Epi Info table and conduct further processing on the table itself rather than having to return to the "raw" data which might contain thousands of records The OUTPUT command allows us to do this

- **OUTPUT**

The OUTPUT command when used with the TABLES or FREQ commands makes an Epi Info data record from each cell in the table When combined with the ROUTE command this can create an Epi Info file of records summarizing the cells in the table Thus if you run ANALYSIS and READ OSWEGO the commands

```
EPI6> ROUTE VANILLA REC
EPI6> OUTPUT TABLES ILL VANILLA SEX
```

will create a file called VANILLA REC that will contain records summarizing the number of cases of ILL by VANILLA stratified by SEX The file will be a standard Epi Info REC file and will have the following variables

```
EPI6> ILL VANILLA SEX COUNT
```

The variable COUNT was added and for each record will contain the contents of one cell in the table produced by TABLES ILL VANILLA SEX in OSWEGO

- **SUMFREQ and SUMTABLES**

These commands are used for processing summary records They produce tables similar to those from the FREQ and TABLES commands but with cell values representing sums rather than counts Let s READ VANILLA into ANALYSIS to clarify what we have done

First you will notice that there are only 7 records compared to 75 in OSWEGO

EPI6> **LIST** shows their contents

REC	SEX	VANILLA	ILL	COUNT
1	F	+	+	27
2	F	+	-	4
3	F	-	+	3
4	F	-	-	10
5	M	+	+	16
6	M	+	-	7
7	M	-	+	8

EPI6> **SUMTABLES COUNT ILL VANILLA** produces

	VANILLA		
ILL	+	-	Total
+	43	3	46
+	11	18	29
Total	54	21	75

This is the identical result that we got with this command for the OSWEGO data set on page 66

EPI6>**TABLES ILL VANILLA**

To see the outcome stratified by sex type as follows to produce two tables

EPI6>**SUMTABLES COUNT ILL VANILLA SEX**

```
            SEX =F
                          VANILLA
ILL           |      +      - | Total
--------------+-------------+------
          +   |  27 0    3 0 |  30 0
          -   |   4 0   10 0 |  14 0
--------------+-------------+------
      Total   |  31 0   13 0 |  44 0


            SEX =M
                          VANILLA
ILL           |      +      - | Total
--------------+-------------+------
          +   |  16 0    0 0 |  16 0
          -   |   7 0    8 0 |  15 0
--------------+-------------+------
      Total      23 0    8 0 |  31 0
```

accompanied by single and two-table statistical indicators

Graphing Summary Data

These commands will work for the graphic presentation you desire

EPI6> **BAR ILL /SUM=COUNT**
EPI6> **HISTOGRAM ILL /SUM=COUNT**
EPI6> **PIE ILL /SUM=COUNT**

14. Evaluation of Nutritional Status with EPINUT

Child malnutrition remains an important public health challenge at the end of the 20th century While in the past nutrition intervention programs have targeted only severally malnourished children it is now recognized that even mild and moderate malnutrition are significant risk factors for both child morbidity and mortality

The good news is that Epi Info provides the tools for measuring and classifying the three major indicators for child malnutrition **Weight for Age, Height for Age** and **Weight for Height** The bad news is that Epi Info considers this to be an advanced feature of the program in other words, inaccessable to the beginning analyst Let's see if we can rectify this situation

Overview of Growth Reference Curves

The anthropometric calculations described in this chapter are based on the growth reference curves developed by the **National Center for Health Statistics** (NCHS) and **CDC** using data from the Fels Research Institute and US Health Examination Surveys These growth curves are recommended by the **World Health Organization** (WHO) for international use

To calculate the anthropometric indices information is needed on each individual's **sex, age, weight**, and **height** From these data it is possible to form different indices including those that relate to height-for-age (**HA**), weight-for-age (**WA**), and weight-for-height (**WH**) These indices can be expressed in terms of Z-scores percentiles and percent of median relative to the international growth reference population mentioned above The following abbreviations are used

HAP Height-for-Age Percentile
HAZ Height-for-Age Z-score
HAM Height-for-Age percent of Median

WAP Weight-for-Age Percentile
WAZ Weight-for-Age Z-score
WAM Weight-for-Age percent of Median

WHP Weight-for-Height Percentile
WHZ Weight-for-Height Z-score
WHM Weight-for-Height percent of Median

Interpretation and Uses of Anthropometry

Anthropometry can be used to assess nutritional status at both the individual and the population level. Ideally, individuals should have several weight and height measurements over time so that growth velocity can be assessed. A decline in an individual's anthropometric index from one point in time to another could be an indication of illness and/or nutritional deficiency that may result in serious health outcomes. In some situations, a single set of measurements may be used for screening populations or individuals to identify abnormal nutritional status and priority for treatment.

The two preferred anthropometric indices for determining nutritional status are WH and HA, as these discriminate between different physiological and biological processes. Low WH is considered an indicator of **wasting** (i.e. "thinness") and is generally associated with failure to gain weight or a loss of weight. Low HA is considered an indicator of **stunting** (i.e. "shortness"), which is frequently associated with poor overall economic conditions and/or repeated exposure to adverse conditions. The third index, **WA**, is primarily a composite of **WH** and **HA** and fails to distinguish tall, thin children from short, well-proportioned children. However, if height hasn't been measured it will be the only index available.

The distribution of the indices can be expressed in terms of Z-scores, percentiles, and percent of median. Z-scores, also referred to as standard deviation (SD) units, are frequently used. The Z-score in the reference population has a normal distribution with a mean of zero and standard deviation of 1. Z-score cutoff points are recommended by WHO to classify nutritional status. For example, if you are evaluating WAZ, the following categories are used:

No malnutrition	**WAZ ≥ -2**
Mild malnutrition	**WAZ < -2 and ≥ -3**
Severe malnutrition	**WAZ < -3**

Calculating these indices with Epi Info

The program EPINUT EXE can add anthropometric indices to an Epi Info file that already contains data in fields called AGE SEX WEIGHT and HEIGHT You may remember that we did indeed include them in QUEST1 REC

Here are the rules for definitions of these variables

AGE should represent biologic age in months
SEX can be coded as a numeric 1 "m", or "M" for boy and numeric 2 , "f", or "F" for girls
WEIGHT and **HEIGHT** should be numeric in metric units

When all indices are calculated 10 new variables are added to the file **HAZ HAM HAP WAZ WAM, WAP, WHZ WHM WHP** and **FLAG** Any existing variable with one of these names will be overwritten and replaced by the newly calculated value If HEIGHT was missing then only WAZ, WAM and WAP FLAG will be calculated

What is this strange variable called **FLAG**? It is a signal (a flag being waved) that some of the data are too extreme to be biologically plausible and thus represent errors at the time of data collection or data entry Flag has eight possible values from 0 through 7 Zero means that none of the indices were flagged However this does not necessarily mean the information is correct Either sex, age weight or height could be incorrect but not extreme enough to be flagged

Luckily we do not have to know what the errors coded 1-7 mean for Epi Info provides us the program (AMISS1 PGM) that runs with ANALYSIS It identifies any missing or extreme values (outliers) and codes the corresponding indices to " " (Remember that "period" is the code for missing value)

I have modified and saved it to the name NMISS PGM to save any changes to the ANTHRO REC file described below in **Calculating the Nutritional Indices** and to define the malnutrition category codes (1=None 2=Mild 3=Severe)

```
*************************************************************************
* NMISS PGM Modification by BASICS of AMISS1 PGM, MARCH 1998
*************************************************************************
*AMISS1 PGM  Version 2 0  18/OCT/1993                                    *
*An Epi Info program to code missing variables for anthropometry values*
*  See Epi Info documentation for more information,                      *
*  by Kevin Sullivan  Division of Epidemiology, Emory University        *
*************************************************************************
READ \CAMHFS\QUEST1\ANTHRO REC
*Declare missing variables based on Record Flag                         *
*************************************************************************
If (flag=1) or (flag=3) or (flag=5) or (flag=7) then HAZ =
If (flag=1) or (flag=3) or (flag=5) or (flag=7) then HAP =
If (flag=1) or (flag=3) or (flag=5) or (flag=7) then HAM =
If (flag=2) or (flag=3) or (flag=6) or (flag=7) then WHZ =
If (flag=2) or (flag=3) or (flag=6) or (flag=7) then WHP =
If (flag=2) or (flag=3) or (flag=6) or (flag=7) then WHM =
If (flag>3) and (flag<8) then WAZ =
If (flag>3) and (flag<8) then WAP =
If (flag>3) and (flag<8) then WAM =
*************************************************************************
DEFINE HA #
HA = (HAZ>=-2)+((HAZ>=-3)*(HAZ<-2))*2+(HAZ<-3)*3
DEFINE WH #
WH = (WHZ>=-2)+((WHZ>=-3)*(WHZ<-2))*2+(WHZ<-3)*3
DEFINE WA #
WA =  WAZ>=-2)+((WAZ>=-3)*(WAZ<-2 ))*2+(WAZ<-3)*3
ROUTE TEMP$$$ REC
WRITE RECFILE CASEID HA HAZ WH WHZ WA WAZ FLAG
COPY TEMP$$$ REC \CAMHFS\QUEST1\ANTHRO REC
ERASE TEMP$$$ REC
QUIT
*************************************************************************
```

Calculating the Nutritional Indices

From what we have said it should be easy to proceed Once all of the data have been entered verified and cleaned for QUEST1 REC all we have to do is run EPINUT and NMISS PGM, and we are done

Unfortunately the process is not always that easy! Questionnaires often are quite long, and I have found that it is sometimes impossible to run EPINUT on the entire data set because of **Insufficient Memory**

What we can do is create a separate data base containing only the necessary variables This can be done in ANALYSIS with the program file ANTHRO PGM to produce ANTHRO REC by typing <ANTHRO> at the EPI6> prompt in ANALYSIS

```
* FILE NAME  C \CAMHFS\QUEST1\ANTHRO PGM
SET NOECHO=ON
READ \CAMHFS\QUEST1\QUEST1 REC
ERASE \CAMHFS\QUEST1\ANTHRO REC
ROUTE \CAMHFS\QUEST1\ANTHRO REC
WRITE RECFILE CASEID AGE SEX WEIGHT HEIGHT
QUIT
```

ANTHRO REC contains only five variables so there should be no memory problems when we run EPINUT and NMISS with the following commands

At the C 'EPI6> in DOS type <EPINUT ANTHRO REC>
At the EPI6> prompt in ANALYSIS type <NMISS>

When doing the nutritional analysis on ANTHRO, we are interested in the correlation of a child's nutritional status with other variables such as the presenting illness, its duration, etc To do this we can link ANTHRO REC back to QUEST1 REC This can be done by the MERGE program in DOS and the RELATE command in ANALYSIS which we will explain in the next chapter

15 Linking files with MERGE and RELATE

■ MERGE

Overview

MERGE can be used for combining Epi Info files in several different ways or for updating records in one file, using data in another file It operates in "batch" mode making a permanent file containing the results of merging two existing Epi Info files Many of the functions of MERGE can be accomplished in a dynamic way with the relational features of RELATE in ANALYSIS

Running MERGE

Running MERGE from the DOS EPI6> prompt or by choosing MERGE from the EPI6 menu will display the screen shown below If your computer has a mouse you can move the cursor and then select items by pressing the left mouse button Without a mouse move the cursor from field to field with the <TAB> key (<Shift-Tab> to go backward>), and select with <Enter> Help is available at any time by pressing the <F1> key

Merge

File 1
File 2
Output file

Merge Options

( ) Concatenate **( ) Update**
( ) Join **( ) Revise**

OK **Cancel**

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 Output File is the REC file to be created by MERGE that will contain the results of the merge

CONCATENATE (Concatenating Similar Files Top to Bottom)

Two files that have the same format (came from the same questionnaire) can be concatenated with one appended to the end of the other Suppose for example that your survey has been conducted in three districts and the data for QUEST1 REC has been entered using the same questionnaire in three different computers The three files can be joined **end to end** by using MERGE as follows

File 1	**Name of one file**
File 2	**Name of the other file**
Output file	**Name of a new file to be created**
Merge Option	**Concatenate**

After you have answered the questions on the screen choose OK or press <Enter> twice The program will convert the file and then return to the merge screen Press <F10> to return to the EPI6 menu or to DOS

Since you want to join three files MERGE will have to be executed twice and there will be some file renaming to do, to have the final combined file retain the original name of QUEST1 REC

Merging All Fields of Dissimilar Files (Side to Side) -- JOIN

This is a procedure that we could use to join our files QUEST1 REC and ANTHRO REC
You can now see why the variable CASEID was included in ANTHRO REC This is because each file must contain a common identifier variable in order to be joined together

Now run MERGE as follows

File 1	**QUEST1 REC**
File 2	**ANTHRO REC**
Output file	**Name of a new file**
Merge Option	**JOIN**

When you choose OK or press <Enter> twice, a dialog box appears with the names of fields present in both files Using the mouse button or arrow keys and <Spacebar> select a field or fields that uniquely identify each record in File 1 (CASEID) Pressing <Enter> twice will begin the merging process

- **Using ANALYSIS to link files**

Joining End to End

Suppose the QUEST1 REC files from the districts have been renamed QUEST1A REC, QUEST1B REC and QUEST1C REC Read all the files together as follows

EPI6> READ QUEST1A QUEST1B QUEST1C

Now you can save them in one file joined top to bottom, with the commands

EPI6> **ROUTE QUEST1 QES**
EPI6> **WRITE RECFILE**

Joining Side to Side

Joining QUEST1 REC and ANTHRO REC requires two commands

EPI6> READ QUEST1
EPI6> RELATE CASEID ANTHRO

Here again CASEID is the common identifier that allows the records to be matched and linked Don t forget to save the newly constituted file with the commands

EPI6> **ROUTE QUEST1J QES**
EPI6> **WRITE RECFILE**

You can now quit ANALYSIS DELete QUEST1 REC (Only after saving a backup copy, in case something goes wrong) and REName QUEST1J REC QUEST1 REC

16 Putting it all together with EPIGLUE

We have seen that working with Epi Info requires working with a bewildering number of directories subdirectories files and commands Do you think you can remember them? Do you have trouble keeping DOS and Epi Info commands distinct in your mind? What about in the heat of the battle of an actual survey when questionnaires are being revised up until the last minute new indicators are being proposed one of your star data entry people cannot participate and you re trying to save backup copies of all the survey files to a diskette EPIGLUE to the rescue! This is the program with which I wrote two menu driven programs 1) CAMTOT to organize and provide CAI for training, and 2) CAMHFS as an actual example of how to organize the files for a Health Facility Survey Getting to know these programs will facilitate your learning of EPIGLUE for it is much easier to modify programs that exist than to start from scratch You may also run EPIGLUE as it comes with EPI INFO by typing the following command

C \EPI6> EPIGLUE

Don t forget to use Help <F1> for in context tutorials

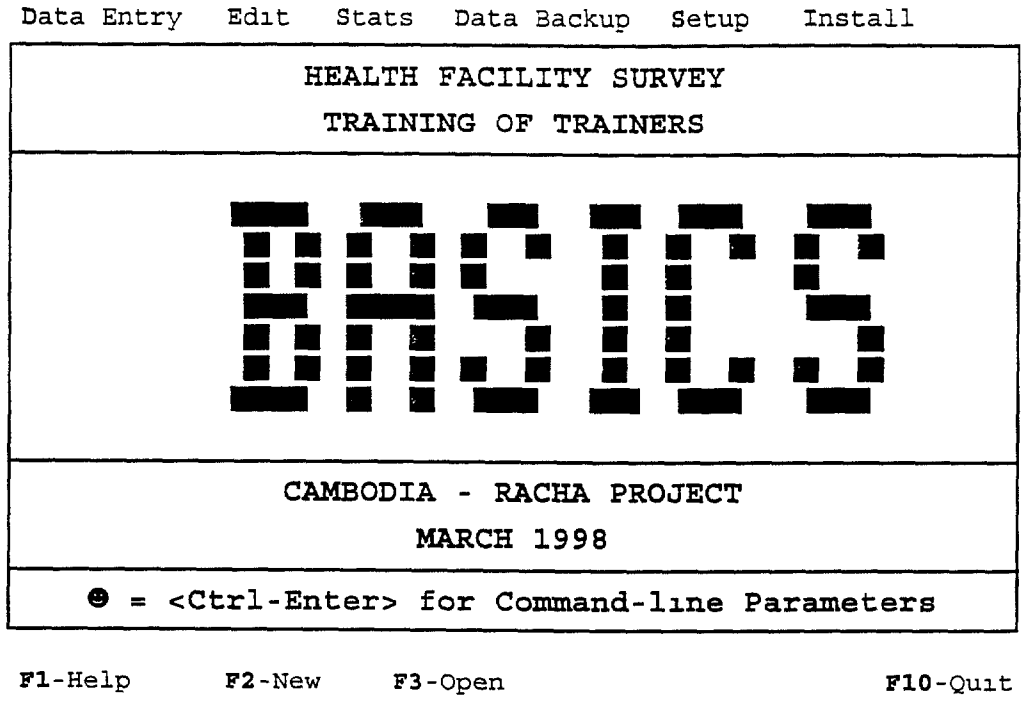
Overview

EPIGLUE provides features for creating menus of your own design, complete with pulldown menus and popup dialog boxes simply by editing a text file (* MNU) of menu names and commands It also contains editing and display features that provide quick access to text files The menu functions with or without a mouse and a decorative background can be inserted from another text file (* SCR) of your choice

EPIGLUE provides facilities to display text reports or tables either as text or as "hypertext " and for navigating and invoking actions by clicking on highlighted key words Programs such as Harvard Graphics or Epi Map can be called from EPIGLUE to display graphs or maps As illustrated in the EPIGLUE tutorial you can also add editing and file management functions or a calendar calculator or ASCII table to the menu by adding one-line commands to the * MNU file controlling the menu

Using a program for a Health Facility Survey -- CAMHFS

Now let's get ready for a survey When at C \EPI6> you type <CAMHFS> you see the following screen



Across the top line you see commands for pop-up menus They may be activated by moving the green bar cursor to your choice with horizontal arrow keys or a mouse or typing the RED letter in the command On the bottom line we see commands activated by function keys **Note that for this exercise only files for QUEST1 have been prepared** Procedures for QUEST2-4 are exactly analogous to those for QUEST1

- **Data Entry**--Data can be entered into questionnaires 1-4 by pressing <Ctrl-Enter> at the same time Nutritional indicators are calculated for QUEST1 REC in options 5-7 by EPINUT
- **Edit**--The records are sorted by health center number (NUM) and the child identification number (IDN) and listed by the update command in ANALYSIS which allows making changes to the data
- **Stats**--For QUEST1 REC frequencies are run on all variables with the results outputted to the text file STATS1 OUT

- **Data Backup**--A good habit to develop to save a backup copy of data after each days data entry Place a diskette in drive A and press <Ctrl Enter> When asked for the DESTINATION DRIVE type <A > and QUEST1 QES and QUEST1 REC will be saved to the diskette
- **Setup**--For CAMHFS to function on the hard disk and be installed to a diskette the default parameters must be setup as shown Advance the cursor from Labels to Parameters by pressing <Tab> If you have changed the parameters they do not become effective until you have left the program and restarted it This can be done quickly by typing <F10>, <F3> and <Enter> If you are installing CAMHFS from a diskette to a hard drive then the parameters for source and destination are reversed (Source = A and Destination = C)

Labels	Code	Parameters
Hard Disk	%P1	C
Source	%P2	C
Destination	%P3	A

- **Install**--With just the keystroke <Enter> all of the necessary directories and files are copied from hard disk to diskette or vice versa
- **F1-Help**--In a real survey you probably won't have the luxury of having enough time to prepare help messages in a text file CAMHFS HPT that is compiled by HYPER to CAMHFS HLP
- **F2-New**--This is a built in editor that allows you to create text files, a * PGM file, for example
- **F3-Open**--You may open and edit a text file that already exists
- **F10-Quit**--You leave CAMHFS and return to the DOS environment

Creating your own EPIGLUE program for a Health Facility Survey

An EPIGLUE program consists of several files with the same prefix and different extensions Thus when you run CAMHFS you are actually using the following files

CAMHFS EXE	This is EPIGLUE EXE supplied by Epi Info renamed for our program
CAMHFS MNU	This is a text file that we write containing the commands to run our program
CAMHFS CFG	This text file is created automatically to store the parameters that we have set up
CAMHFS SCR	We write this text file to create the background screen
CAMHFS HPT	This is the HELP program text file that we may write
CAMHFS HLP	If CAMHFS HPT exists then it must be compiled to this file by the program HYPER
CAMHFS LST	If CAMHFS exists this text file is created automatically to keep track of the Help **TOPIC**s

The first three files in the list are necessary to run the program All the others are optional but the presentation is more attractive if you add the ,SCR file Suppose that you have the responsibility of writing a menu driven program for a new Health Facility Survey that uses the CAMHFS methodology You can start by naming the survey FIRST, since it is your first attempt at writing an EPIGLUE program Here are the steps that have been programmed into CAMTOT to make this a simple task

- **Create the background file (FIRST SCR)**

Run **CAMTOT** go to Surveys and New Survey Here you see three options

1 Make Lletters
2 Make * SCR
3 Make New Files

The screen that comes up when you run CAMTOT has three colored stripes four title lines and the word BASICS in large letters **Make Lletters** tells you that you can make your own banner by going to **File** and opening a **New File** Now press <**ALT E**> and choose **Large Letters** or go directly to **Large Letters** by typing <**ALT F9**> Type <**FIRST**> using the **Big Font** and press <**Enter**> Save the file to the name **FIRST TXT** and quit the editor

Now choosing **Make * SCR** runs the program **SCREEN EXE** (Written in Power BASIC) to create **FIRST SCR** Here are some sample responses for the prompts

LL FILE NAME?	<FIRST)
OUTPUT FILE NAME?	<FIRST>
8 COLOR CODES GIVEN	
FIRST COLOR?	<3>
SECOND COLOR?	<4>
THIRD COLOR?	<1> These codes will produce red white and blue from top to bottom
TITLE 1?	<Health Facility Survey>
TITLE 2?	<Training of Trainers>
TITLE 3?	<RACHA Project - Cambodia>
TITLE 4?	<March 1998>

To check the results type **FIRST** and you should be dazzled by the red white and blue

- Create the necessary files and subdirectories copy files to them Remembering how the directories are organized for CAMHFS we want an analogous tree for FIRST

\ **CAMHFS** -------	**QUEST1**	\ **FIRST** ------	**QUEST1**
	QUEST2		**QUEST2**
	QUEST3		**QUEST3**
	QUEST4		**QUEST4**

The seven EPIGLUE files (FIRST *) must be placed in the EPI6> directory This is easily done by the following command

C \EPI6>COPY CAMHFS * FIRST *

Now files need to be copied to the QUEST directories (* QES * CHK * PGM, etc) The program files must be modified to reflect the FIRST tree rather than the CAMHFS one

If you had to do all of these files manipulations manually there is a good chance that you would make a mistake or forget something The above **new survey** option **Make New Files** makes life easy by doing all this automatically with a program called SURVEY EXE (Also written in Power BASIC)

■ Modifying the menu file CAMHFS MNU

It is possible that the methodology for FIRST is identical to CAMHFS This could be the case, for example in baseline and End Of Project surveys In such an instance the only modification necessary, replacing every occurrence of the word CAMHFS by FIRST, has already been done above Of course, a survey is never perfect and there are always changes to be made This can be done by loading CAMHFS MNU which is a text file into EPED Here we will examine excerpts related to QUEST1

```
*------------------------ MENU TITLE ---------------------------
CAMHFS MENU
*---------------- START OF MENU LAYOUT BLOCK --------------------
BEGIN
  POPUP "&Data Entry"
  BEGIN
    MENUITEM "1  Clinical Observation^#2"  DoEnter1
    MENUITEM SEPARATOR
    MENUITEM "5  Create ANTHRO REC^#2", DoANTHRO
    MENUITEM "6  Run EPINUT^#2", DoEPINUT
    MENUITEM "7  Clean with NMISS^#2"  DoNMISS
  END
  POPUP "&Edit"
  BEGIN
    MENUITEM "1  Clinical Observation^#2", DoEditQ1
  END
  POPUP "&Stats"
  BEGIN
    MENUITEM "1  Clinical Observation^#2", DoStatsQ1
    MENUITEM SEPARATOR
    MENUITEM "5  Indicators Q1^#2"  DoIndQ1
  END
```

```
  POPUP "Data &Backup"
  BEGIN
    MENUITEM "1  Clinical Observation^#2"  DoBackQ1
  END
   ENUITEM "Se&tup"  IDSETUP
  POPUP "&Install"
  BEGIN
    MENUITEM "&Copy CAMHFS files' DoCopy
  END
END
*-------------------- END OF MENU LAYOUT BLOCK ------------------------

*--------------------- START OF COMMAND BLOCKS ------------------------
DoEnter1
BEGIN
  -P1
  FILEPARAM "\CAMHFS\QUEST1\QUEST1 REC"
  Enter %1
END

DoANTHRO
BEGIN
  %P1
  FILEPARAM "\CAMHFS\QUEST1\ANTHRO PGM"
  Analysis %1
END

DoEPINUT
BEGIN
  %P1
  FILEPARAM "\CAMHFS\QUEST1\ANTHRO REC"
  EPINUT %1
END

DoNMISS
BEGIN
  %P1
  FILEPARAM "\CAMHFS\QUEST1\NMISS PGM"
  ANALYSIS %1
END
```

```
DoEditQ1
BEGIN
  %P1
  FILEPARAM "\CAMHFS\QUEST1\EDIT1 PGM"
  IF NOT EXIST %EIDIR%\EPI6\ANALYSIS EXE GOTO NOFILE
  Analysis %1
END

DoStatsQ1
BEGIN
  %P1
  FILEPARAM "\CAMHFS\QUEST1\STATS1 PGM"
  IF NOT EXIST %EIDIR%\EPI6\ANALYSIS EXE GOTO NOFILE
  Analysis %1
END

DoIndQ1
BEGIN
  %P1
  FILEPARAM "\CAMHFS\QUEST1\IND1 PGM"
  IF NOT EXIST %EIDIR%\EPI6\ANALYSIS EXE GOTO NOFILE
  Analysis %1
END

DoBackQ1
BEGIN
  %P1
  PARAMETER "  DESTINATION DRIVE" '<A >"
  If %1'==' GOTO INFO
  ECHO BACKING UP DATA
  COPY \CAMHFS\QUEST1\QUEST1 REC %1
  COPY \CAMHFS\QUEST1\QUEST1 QES %1
  GOTO DONE
   INFO
  ECHO         For BACK-UP all you need to do is type
  ECHO         the destination drive, as shown below
  ECHO         ---------------------------------------
  ECHO                              A
  ECHO         QUEST1 QES and QUEST1 REC will be copied to Drive  A
  ECHO         -----------------------------------------------
  PAUSE
```

```
   DONE
  CLS
END

DoCopy
BEGIN
  %P2
  ECHO OFF
  CLS
  ECHO Installing
  MD %P3\EPI6
  MD %P3\CAMHFS
  MD %P3\CAMHFS\QUEST1
  MD %P3\CAMHFS\QUEST2
  MD %P3\CAMHFS\QUEST3
  MD %P3\CAMHFS\QUEST4
  COPY \EPI6\CAMHFS *                %P3\EPI6
  COPY \CAMHFS\* W51                 %P3\CAMHFS
  COPY \CAMHFS\QUEST1\QUEST1 *       %P3\CAMHFS\QUEST1
  COPY \CAMHFS\QUEST1\HCENTER *      %P3\CAMHFS\QUEST1
  COPY \CAMHFS\QUEST1\* PGM          %P3\CAMHFS\QUEST1
  COPY \CAMHFS\QUEST1\INTERV TXT     %P3\CAMHFS\QUEST1
  COPY \CAMHFS\QUEST2\QUEST2 *       %P3\CAMHFS\QUEST2
  COPY \CAMHFS\QUEST2\* PGM          %P3\CAMHFS\QUEST2
  COPY \CAMHFS\QUEST3\QUEST3 *       %P3\CAMHFS\QUEST3
  COPY \CAMHFS\QUEST3\* PGM          %P3\CAMHFS\QUEST3
  COPY \CAMHFS\QUEST4\QUEST4 *       %P3\CAMHFS\QUEST4
  COPY \CAMHFS\QUEST4\* PGM          %P3\CAMHFS\QUEST4
  ECHO                        done
END
*--------------------------- END OF COMMAND BLOCKS -------------------------

*---------------------- START OF STRINGTABLE BLOCK -------------------------
STRINGTABLE
BEGIN
  DoEnter1      "« DATA ENTRY FOR CASE MANAGEMENT OBSERVATION - Q1"
END
*---------------------- END OF STRINGTABLE BLOCK -------------------------
```

Although this file may seem long and complicated at first the logic is quite clear and there are surprisingly few commands to know An asterisk (*) as the first character on a line can be used to indicate comments anywhere in a menu file for anything on the same line following an asterisk will be ignored by the program The file is organized into four sections

- The menu title
- The menu layout
- The command blocks
- The string table

Describing the Contents of the Menu

Menu Title

A menu starts with the name of the menu followed by a space and the reserved word MENU In this case it is **CAMHFS MENU**

Menu Layout

The menu layout contains sets of instructions starting with BEGIN and ending with END In this block you will find three types of instructions MENUITEM, MENUITEM SEPARATOR, and POPUP

- **MENUITEM**

  Syntax MENUITEM "Text to display { ^<Symbol>}",<Command Block Name>
  Example MENUITEM "1 Clinical Observation^#2", DoEnter1

MENUITEM represents one option of the menu The string of characters between quotes that follows is what appears in the menu The optional ^<Symbol> is used to add a symbol to the right side of the item name in the pull down menu <Symbol> can be a character or, if preceded by a number sign (#), a numeric ASCII code The musical notes in the EPI6 PROGRAM menu, signify a program with command line parameters that may be entered after pressing <Ctrl-Enter> I have changed this to a 'happy face 'in CAMHFS MNU (ASCII Code 2)

<Command Block Name> after the comma names a block of commands to be performed when the item is selected There are several built-in commands that refer to predefined Command Blocks with special meaning IDSETUP is the command that displays a dialog box for setting-up the default directories

Here are two others used in CAMTOT

Command	Function
IDCALC	Displays a popup calculator
IDASCII	Displays a table of ASCII characters and their codes

- **MENUITEM SEPARATOR**

  Syntax MENUITEM SEPARATOR

MENUITEM SEPARATOR displays a horizontal line in the submenu It is useful for separating groups of functions in the menu

- **POPUP**

  Syntax
  ```
  POPUP '&Text to appear"
  BEGIN
    MENUITEM  ( as above  1 or more)
  END
  ```

  Example
  ```
  POPUP '&Data Entry '
  ```

POPUP defines a popup menu that will contain one or more menu items A POPUP menu may be an item on the main menu or on a lower-level menu In either case, a new menu pops up when the defined item is chosen POPUP requires that the contents of the submenu it represents be defined by instructions between the words BEGIN and END

Command Blocks

They define the commands to be executed for each of the menu options The block of instructions between BEGIN and END is similar to a DOS batch file Any DOS instruction appearing there will be executed after shelling (exiting) to DOS An external program such as ANALYSIS can be called and almost any DOS batch file command can be used FILEPARAM often appears in the command blocks

- **FILEPARAM**

Syntax FILEPARAM "<File path>","
Example FILEPARAM \CAMHFS\QUEST1|QUEST1 REC"

FILEPARAM is a function which is invoked only if <Ctrl-Enter> key If the <Enter> key only is used the FILEPARAM function is ignored, and the **%1** parameter is reset to an empty string

- **%Px Parameters**

Remember in **Setup** we defined **%P1** as the address for the Hard Drive, and **%P2** and **%P3** as the Source and Destination addresses for copying files

- **DOS Commands**

Many DOS commands are used in the **Backup** and **Install** command blocks including COPY CLS ECHO and MD

String Table

```
Syntax      STRINGTABLE
              BEGIN
                <Command Block Name>  "<Help phrase>"
              END
Example     STRINGTABLE
              BEGIN
                DoEnter1    "« DATA ENTRY FOR  QUEST1"
              END
```

The STRINGTABLE is an optional list of phrases or "strings' corresponding to the various menu options surrounded by a BEGIN END block, and immediately following a STRINGTABLE statement The strings appear at the bottom of the menu screen when the cursor rests on that item It may be used to provide short help phrases that further instruct the user about the nature of a menu choice If a menu item has no help phrase the status line at the bottom of the screen is blank

The phrase **/CLOCK NOSEC** displays a digital clock at the top right corner of the screen without displaying the seconds It was used in CAMTOT

17 The Beginning

You have learned the rules, now it's time to play the game This manual and accompanying files do not come close to exhausting the possibilities of Epi Info use but do provide the basis, and more for serious and competent use of the program There are three keys to success

Practice, Practice and more Practice

Remember you are not alone Don't be afraid to experiment Share your experiences and questions with your colleagues and the international community of Epi Info users

If you are having a problem that you can't solve take a break Fatigue and errors go hand in hand I have often solved a problem after fifteen minutes of jogging with which I had earlier struggled unsuccessfully at the keyboard

And don't forget to have fun!