

第十五章 Epi Info

§15.1 简介

Epi Info 是1988年美国疾病控制中心(CDC)流行病学室编制的用于流行病学调查分析的软件,运行于IBM PC及其兼容机。1990年该室与WHO全球爱滋病控制小组合作,推出了功能更齐全的第5版。它同时具有文字处理、数据库管理与统计分析这三种功能。在此基础上的汉化版Epi Info在菜单提示帮助方面全为中文,且可使用中文变量名(字段)和文件名,便于不熟悉英文者使用和在国内推广。除了在建立调查表时要输入中文外,以后在使用中文时必须输入中文的时候不多而可由文件菜单及变量名菜单中直接调用汉字文件名及字段名,因而使用汉化版也十分方便。EPI INFO新版本可在ftp.cdc.gov得到,在功能上有更多的改进。EPI INFO是流行病学研究设计、数据录入和分析的最佳选择之一。以dBASE III为例,编写修改屏幕格式文件是一件费力的工作,而且由于它至多只能处理128个变量,极易在变量数较多时造成数据库间关系的混乱,Epi Info则很好解决了这一问题。

Epi Info主菜单中有主控程序EPI.EXE连结主菜单中的主要程序。这些主要程序的名称及功能如下:

1. EPED 文本编辑程序。可用于建立调查表或编写各种程序。其中的EPIAID程序可帮助准备文稿及设计流行病学调查。EPED类似Wordstar的编辑功能。计算机所编辑的表格样式,可以直接作为dBASE那样的屏幕格式文件进行数据录入。每数据项后加上相应类型和长度的格式描述符后进入录模块,即可进行录入。该模块的EPAID可用于流行病学调查辅助设计和报告撰写。
2. ENTER 数据录入程序。在由EPED或由其它文字处理器建立调查表的基础上,形成Epi Info数据文件。可用于输入、查找以及修改数据文件格式。开始时,一个以.QES为扩展名的调查表文件是必要的,.REC一旦建立,可被ENTER反复直接调用。

ENTER可以进行光标的任意移动,并提供有特定记录的定位方法。每个记录录入结束后,软件提问是否将录入的数据记入磁盘,也可以取消提示。

3. ANALYSIS 数据分析程序。可由Epi Info或dBASE数据文件中显示记录及频数;处理交叉表。进行流行病学中常用的统计计算:如比数比、相对危险度、可信限,Fisher精确检验及卡方检验,Mantel Haenszel分层分析,单向方差分析,线性回归及配比病例对照研究的分析。此外,可对记录作选择、排序;有"IF"语句及数学和逻辑运算、绘图、编制复合报表及产生新数据集的功能。还有一个程序语言可供编程计算。
4. CHECK 数据核对。可用于设置变量值范围、合法数值、自动编码及数据输入(ENTER)的跳越模式。可进行字段间数学和逻辑运算、复合跳跃、同一过程中访问几个文件。并可支持用户提供的自编程序。
5. STATCALC 统计计算。可对由键盘输入表格的数值进行统计。包括 $2 \times n$ 表单一的或分层的 2×2 表、样本大小估计及单一或分层的趋势分析。
6. CONVERT 数据文件输出。把Epi Info的数据文件转换到12种数据库或统计软件的格式,大多数计算机程序可以接受至少一种格式。所支持的文件格式有:

1.EAS、2.SAS3、3.定界的、4.BASIC、5.SPSS-X、6.EpiStat、7.dBase II、8.dBase III、9.Lotus 1-2-3、10.SPSS/PC+、11.Statpac、12.固定长度的格式数据。

7. IMPORT 数据文件输入。从其它系统输入数据供Epi Info 使用。IMPORT 可将其它系统产生的文件调入Epi Info 作处理, 所支持的文件格式有: 1. 固定长度的格式记录; 2. 逗号定界的; 3. Lotus 1-2-3; 4. dBASE II/III。对其中的第1 和第2 项, 首先必须产生一个描述数据的调查表并用ENTER 产生一个空的.REC 文件, 对第3 和第4 则不需要。
8. MERGE 数据文件连接。将相同或不同格式的调查表产生的文件合并。可用以合并各计算机输入的数据文件并可用新输入的数据对过去的记录进行更新。
9. VALIDATE 比较由两个操作员录入的Epi Info 文件并报告其差别。如果能指定一个包含特定的标识的关键字段则两个文件中的记录不要求顺序相同。

除以上主要程序外, Epi Info 还有一些其它文件。

1. Help 文件。该文件包括了指导手册中大部分内容。在Epi Info 程序中可用;F1;键查到。
2. Sample 程序。提供了两个完整的样本监测系统; 两个流行病学调查文件及一个营养调查的人体测量学系统以供使用。汉化版提供了一个上海市肿瘤报告的例子。
3. Tutorials。一个完整的介绍EPED 和ANALYSIS 特性的人机对话课程。

在文字处理方面, 除一般文字处理外, 能以全屏幕编辑方式建立调查表, 同时也建立了数据结构。根据调查表输入数据即可建立相应的数据库。

软件的统计功能限于一些在流行病学方面常用的统计分析。但由于它能同包括SAS, SPSS, dBASE 及BASIC 等12 种数据类型进行互相转换, 因而所建立的数据很容易由其它统计软件进行复杂的统计处理。

§15.2 汉化版的运行环境和安装

§15.2.1 硬件配置要求

Epi Info 对硬件的要求不高, 如:

IBM\PC机或其兼容机CPU 80286 以上;

内存 ≥ 1 MB 至少有一个软盘驱动器;

硬盘驱动器至少10 MB 根据数据量而增加;

显示器彩色VGA 或更高分辨率;

打印机LQ1500等、HP激光打印机、绘图仪等。

§15.2.2 软件配置要求

DOS 3.3 及以上版本。

UCDOS 2.01及以上版本或西山DOS 5.0及以上版本。

§15.2.3 安装或复制

Epi Info 第5版是用Turbo Pascal语言及汇编语言编写的。源程序约占7.5兆，经编译后共105个文件，容量约2兆(1984613字节)。经压缩后存放于4张360K 5 1/4英寸软盘中。不经恢复不能使用。因而需要用Install程序进行安装。汉化版使用方法与原版基本相同。

安装或复制时需要将安装盘置于相应驱动器中，(根据安装盘的种类)键入Install并回车即出现提示。按提示操作即可以人机对话方式装入全部程序。或者复制另一安装盘以推广本软件。

安装之后，在所安装的磁盘中即已建立了子目录\EPI5、\RSURV以及\SURV。

如果在相应磁盘中已经安装过Epi Info软件，则在安装时会逐一询问是否要复盖原来的文件，这样比较化费时间，不如先把原有的所有\EPI5中的文件删去，然后删去子目录后再进行安装。如果你在\EPI5子目录中已建立了你自己的一些调查表或记录文件，而你希望更新Epi Info且要保存所有的调查表和记录文件，则可以把第一张盘插入A:驱动器，在\EPI5子目录中给出命令A:WIPE_E5即可。Epi Info安装到最后时，程序要求，如果此软件不是来自CDC或USD. INC，用户可向他登记，登记后可以在以后得到他们无偿提供的有关Epi Info的信息。

§15.3 使用

由于Epi Info具有中文的教学软件、Help文件、实例及大量菜单及提示，因而学习和使用十分方便。

首先进入中文DOS，然后用C:>CD\EPI5进入子目录，即出现提示C:\EPI5j。可用以下三种方法中任一种运行程序：

1. 直接启动所需程序，这样可以节省内存。如键入C:\EPI5> STATCALC即可启动统计计算程序。各程序格式如下：

命令	命令行参数
ENTER	<.REC文件> {<.QES文件>}
ANALYSIS	<程序文件名> <数据文件名> /BW /8
CHECK	<.REC数据文件名>
STATCALC	{打印机或j文件名>}
CONVERT	<.REC文件><新文件>< #1-12>
IMPORT	<.REC文件><外来文件>< #1-4>
MERGE	<f1><f2><f3>< #1-4> {<对应字段><Y/N>}
VALIDATE	<.REC文件1><.REC文件2><关键词>

命令行参数是可选的，若键入命令时没有指定，系统则给予提示。设有一个全国性调查，各省上报数据文件为“省名.REC”，用DOS批文件直接把它们转到SAS中处理，批文件的内容为：

```
convert ANHUI ANHUI 2 y
convert BEIJING BEIJING 2 y
convert GUANGXI GUANGXI 2 y
convert HUNAN HUNAN 2 y
```

...

可见其简便。

- 键入EPI, 启动主控程序即出现主菜单。主菜单中每一主要程序名之中有一另一种颜色的字符(如黄色字符串STATCALC 中的S 为蓝色字符)。键入该字符即可运行该主要程序。
- 在主菜单中有一光标条(cursor bar) 位于第一个主程序名上, 用↑和↓键可以移动此光标到其它主程序名。按回车键即可启动此程序。此时如按空格键, 则可在菜单下方提示框中直接输入命令。如光标在EPED 主程序上, 按空格键后, 即可在提示框中EPED 后键入文件名, 按回车键后即可对此文件进行编辑。

Epi Info 各主程序在屏幕上方列出各功能键的作用、可用功能键进行选择。如在EPED 按<F2>键, 可出现文件功能子菜单。Epi Info 各主程序都有逐级子菜单可列于屏幕上。各子菜单也可用上述第2、3种方法进行选择。在文字处理等程序也可用控制符进行操作, 如EPED 中常可用与Wordstar 类似的控制符操作。

用<ESC> 键可退到上一级菜单, 用<F10>键退到主菜单或退出Epi Info。使用中并可随时进入或退出DOS, 使用DOS 命令。

汉字输入比较麻烦, 在Epi Info 中常可以利用文件名清单或变量名清单用光标条进行选择。如在EPED 中用F2 功能键中“打开本窗口文件”功能, 在要求输入文件名时输入*.*, 或在路径后输入*.*, 回车后即显示相应子目录中文件名清单, 用光标条选择即输入了文件名。在ANALYSIS 中, 调用某一记录文件后用<F3> 功能键即可列出字段名清单。同样, 用光标条选择并按回车后即可输入此变量名, 方便了汉字文件名或变量名的输入。

【例15.1】肿瘤报告卡的处理。CANCER.QES中汉化Epi Info中处理肿瘤报告卡的例子, 可直接用于屏幕数据输入, 这里主要是介绍它所用的数据格式。其中###、###.##、<mm/dd/yy>及----等限定了数据录入为整型、实型、日期型和字符型格式。

【例15.2】汉化Epi Info 中假想了一个居民饮水与疾病调查, 调查地区为北京市、上海市、江苏省和浙江省, 调查的基本单位是户, 每个被调查的人可能有许多次随访, 所以数据建立时应设家庭编号和个人编号关键字。利用CHECK 可以建立几个数据集的关系, 在每个问卷文件中已进行相关的说明, 文件名在方括号(【】) 内。这几个文件也可以经其它文字处理软件生成。

【HOUSE.QES】

家庭记录

记录连接说明. 每个家庭记录以识别符“家庭编号”与一个或多个个人记录连接, 在数据输入时, “家庭”文件为主文件. 在分析时, “随访”文件是主文件, 而个人以“个人编号”与之相联系, 然后家庭以“家庭编号”字段与合并了的“随访-个人”记录相联系.

家庭编号<idnum>

地址_____

《肿瘤调查表》

<input type="checkbox"/>	上海市肿瘤病例报告卡	总号_____
		<ICD编码> ###.#
门诊号_____		更正诊断报告栏
住院号_____		(原报告诊断有误时填写)
患者姓名____ 性别__ 实足年龄###		原诊断.....
出生年月<mm/dd/yy> 民族.....		原诊断日期.....
职业(具体职务).....		
工作单位_____		诊断依据#
常住户口在....1:本市市区2:本市郊区		1:临床6:病理(继发)
地址_____		2:X,超声波7:病理(原发)
诊断.....		内窥镜,CT
病理学类型-----		3:手术,尸检8:尸检
诊断日期<mm/dd/yy>		(无病理) (有病理)
报告单位### 报告日期<mm/dd/yy>		4:生化,免疫9:不详
D <mm/dd/yy> 死亡原因## 报告医师....		5:细胞学,血片0:死亡补发病

沪卫生局批准

县市_____ 省市自治区_____

卧室数: ##

自来水: <Y>

【PERSON.QES】

个人记录

记录连接说明. 个人记录以"家庭编号"识别符和家庭记录联系,
并以"个人编号"识别符与随访记录相联系.

个人编号<idnum> 家庭编号#####

姓名-----

年龄### 性别__ 患病<Y>

【VISIT.QES】

随访

说明在Epi Info中三个文件的连接. 这一随访记录由"个人"文件中的个人建立. 每个人可有几次随访. 这两个文件中的记录

以"个人编号"字段互相连接.

个人编号#####

随访日期<mm/dd/yy>

随访记录_____

```
【HOUSE.CHK】
家庭编号
    KEY UNIQUE
END
    省市自治区

    Legal
        北京
        上海
        江苏
        浙江
    END
END
    自来水

    ENTER
    ENTER PERSON.REC 家庭编号
END
【PERSON.CHK】

个人编号
    KEY
END
    患病

    ENTER visit 个人编号
END
【VISIT.CHK】

个人编号
    KEY
END
```

以上各.CHK文件中汉字开始END结束的内容是对对该字段的说明。实际调查的项目可能更多，逻辑关系也会更复杂，这个例子仅仅是一个示范。准备好上述几个文件后就可以用ENTER录入了，录入时ENTER自动在几个数据文件中跳转。

【例15.3】下是1991-1994年医院满意度调查的例子，在“数据处理与综合应用”一章对之有更多的介绍。

卫生部“纠风调查”病人问卷

同志,您好!欢迎您参加卫生部组织的民意测验,我们希望通过您的真诚回答,反映出群众对卫生系统行业作风的意见和看法,本问卷不记单位和姓名,我们保证对您的回答保密。感谢您的合作!

医院名称NAME: _____

医院编号{ID1}: #####

病人编号{ID2}: #####

病人类别{SS}: #

- | | | |
|------------------------------|-------|----|
| 一您这次在哪级医院就医? | {N1} | # |
| 二您这次在哪科就医? | {N2} | ## |
| 三您的主要职业: | {N3} | ## |
| 四您的医疗费用支付方式为 | {N4} | # |
| 五您对医生服务态度是否满意? | {N5} | # |
| 六您认为医生技术 | {N6} | # |
| 七您对护士的服务态度是否满意? | {N7} | # |
| 八您认为护士技术 | {N8} | # |
| 九您对这所医院的医疗质量是否信任? | {N9} | # |
| 十您对挂号处工作人员的服务态度是否满意? | {NA} | # |
| 十一您对药房工作人员的服务态度是否满意? | {NB} | # |
| 十二您对医院膳食是否满意? | {NC} | # |
| 十三您对医院的环境卫生条件是否满意? | {ND} | # |
| 十四您这次就医期间是否托关系? | {NE0} | # |
| (一)中间人是否本院职工 | {NE1} | # |
| (二)中间收礼? | {NE2} | # |
| (三)托人价值 | {NE3} | # |
| 十五您这次就医期间是否送钱物? | {NF0} | # |
| (一)送礼价值为 | {NF1} | # |
| (二)送礼原因是 | {NF2} | # |
| (三)送礼态度是 | {NF3} | # |
| 十六您在这次就医过程中有无宴请 | {NG0} | # |
| (一)宴请价值为 | {NG1} | # |
| (二)宴请原因是 | {NG2} | # |
| (三)宴请态度是 | {NG3} | # |
| 十七您对医院最不满意的是什么?有什么{建议}(文字叙述) | | |

由于问卷项目必须符合一定的逻辑关系或条件,应在CHK型文件中加以说明,如合法范围、汉字标号以及跳转与重复等。使用汉之标号便于录入过程中进行特定项目的核实,使用跳转则避免了可能存在的数据库的不合理。设文件名为SURV.CHK,其内容如下:

```

NAME
  Repeated
END

SS
  Comment Legal
    1 门诊
    2 住院
    3 出院
  END
END

N1
  Comment Legal
    1 省级
    2 地、市级
    3 县级
  END
END

N2
  Comment Legal
    1 内科
    2 外科
    3 妇产科
    4 儿科
    5 中医科
    6 眼科
    7 耳鼻喉科
    8 口腔科
    9 皮肤科
    10 其它
  END
END

ND
  Comment Legal
    1 满意
    2 较满意
    3 一般
    4 不满意
    5 很不满意
  END
END

NE0
  Comment Legal
    1 是
    2 否
  END
  Jumps
    2 NF0
  END
END

NE1
  Comment Legal
    1 是
    2 否
  END
END

NE2
  Comment Legal
    1 是
    2 否
  END
END

```


N3	Comment Legal	NE3	Comment Legal
	1 工人		1 50元以下
	2 农民		2 50-
	3 军人		3 100-
	4 教师、医务人员或科验		4 200-
	5 企事业机关干部		5 500-
	6 个体从业者		END
	7 商业从业者		END
	8 无职业者		NF0
	9 学生		Comment Legal
	10 其它		1 是
	END		2 否
	END		END
			Jumps
N4	Comment Legal		2 NG0
	1 公费		END
	2 劳保		END
	3 半自费		END
	4 自费		NF1
	5 商业性医疗保险		Comment Legal
	END		1 50元以下
	END		2 50-
			3 100-
			4 200-
N5	Comment Legal		5 500-
	1 满意		END
	2 较满意		END
	3 一般		NF2
	4 不满意		Comment Legal
	5 很不满意		1 出于感激
	END		2 想得到方便和照顾
	END		3 担心医务人员工作不认真
			4 受它人影响
N6	Comment Legal		5 医务人员暗示的
	1 好		6 医务人员直接索要的
	2 较好		END
	3 一般		END
	4 不好		NF3
	5 很不好		Comment Legal
	END		1 拒绝收
	END		2 事后全部退还

```

N7
  Comment Legal
    1 满意
    2 较满意
    3 一般
    4 不满意
    5 很不满意
    6 没接触
  END
END

N8
  Comment Legal
    1 好
    2 较好
    3 一般
    4 不好
    5 很不好
    6 没接触
  END
END

N9
  Comment Legal
    1 信任
    2 较信任
    3 不信任
    4 很不信任
    5 说不好
  END
END

NA
  Comment Legal
    1 满意
    2 较满意
    3 一般
    4 不满意
    5 很不满意
    6 没接触
  END
END

    3 收礼后照价付了钱
    4 受礼后付了部分钱
    5 推辞过，但还是收下了
    5 没有拒绝就收下了
  END
END

NG0
  Comment Legal
    1 是
    2 否
  END
  Jumps
    2 建议1
  END
END

NG1
  Comment Legal
    1 50元以下
    2 50-
    3 100-
    4 200-
  END
END

NG2
  Comment Legal
    1 出于感激
    2 想得到方便和照顾
    3 担心医务人员工作不认真
    4 受它人影响
    5 医务人员暗示的
    6 医务人员直接索要的
  END
END

NG3
  Comment Legal
    1 拒绝收
    2 事后全部退还
    3 收礼后照价付了钱
    4 受礼后付了部分钱

```

```

NB
  Comment Legal
    1 满意
    2 较满意
    3 一般
    4 不满意
    5 很不满意
    6 没接触
  END
END

NC
  Comment Legal
    1 满意
    2 较满意
    3 一般
    4 不满意
    5 很不满意
    6 不清楚
  END
END

```

为了节省篇幅，后面一部分与前面部分同时列出。

第一项设为重复，因此对特定医院的病人，汉字名只需输入一次就够了。其它项目有相应的合法值和说明，只要在ENTER 下打入F9，就可以行到相应的标号提示，最后一些项目当出现“否”时，其下面的内容略去不输，比dBASE 自动填零进了一步。

