

Sesión Temática/Thematic sessions

Software de distribución libre para epidemiología y Salud Pública: cómo coordinar el desarrollo internacional

Free software of epidemiology and public health-how to coordinate international development

Viernes 3 de Octubre / Friday 3, October
15:00:00 a/to 16:30:00

THE OPEN EPI INITIATIVE: OPEN SOURCE WEB BROWSER SOFTWARE FOR PUBLIC HEALTH

Andrew G. Dean
Epiinformatics.com, U.S.A.

Since 1984, Epi Info software from the Centers for Disease Control and Prevention (CDC) has provided free tools for data management and epidemiologic statistics for both DOS and Microsoft Windows computers. More than 200,000 copies of the Windows version were downloaded in 2001. There are now hundreds of statistical calculators on the Internet, offering a great variety of functions. For many types of work, the problem is in distinguishing among the various offerings and coping with different data entry screens and methods of documentation. Validation methods for statistical procedures are just beginning to become available. Current trends in software include the use of the Internet and collaborative development of software by users through the process known as Open Source Development. Open Epi is an initiative partly supported by the Gates Foundation to extend the statistical capabilities of Epi Info through Open Source development on the Internet. It will provide a framework for developing epidemiologic calculators in JavaScript. The calculators will run in Windows, Linux, or Macintosh systems under web browsers. They can be used from a web site over the Internet or downloaded and run without an Internet connection or web server. The source code will be the product of a group of contributors and will be included with the programs. Open Epi will provide a framework for epidemiologic software development, with modules for entering summary data in tables, configuring the tables for different authors' preferences, providing the data to user-written statistical modules in standard format, and formatting results as web pages that can be further edited or displayed. Methods will be provided for developing help files, for validating calculations, and for language translation. The talk will include demonstrations of Open Epi and other Internet statistical software for epidemiology and a discussion of the process of Open Source development during the first nine months of the project. Possibilities for development of other types of epidemiologic software through the Open Source method will be discussed.

FREE SOFTWARE FOR EPIDEMIOLOGY AND PUBLIC HEALTH - HOW TO COORDINATE INTERNATIONAL DEVELOPMENT

Juan Carlos Fernández Merino*, Pedro Arias Bohigas**

**Servicio de Vigilancia Epidemiológica y Evaluación, D.G. Salud Pública. Consejería de Salud. Andalucía, Sevilla, Spain, **Servicio de sanidad ambiental, Instituto de Salud Pública. Comunidad de Madrid, Madrid, Spain*

At the beginning of the PC era, there were few tools of easy use for the daily work in epidemiology. Small utilities were created that allowed the use of computers in the daily work. Epi Info, the most important exponent in the free programs adapted to the epidemiology field was maybe one of these tools that, for its simplicity, it was implanted in the entire world and helped many people to be introduced in the handling of computers or in its epidemiologic instruction, being distributed thousands of copies of the last version of Epi Info for MSDOS. At the moment an important offer of commercial programs of databases, sheet and statistical or geographical analysis that allow to almost carry out all that the epidemiologist needs. However, for diverse reasons (cost, use complexity or formation), there still being producing free computer tools adapted to the necessities of the epidemiology field. The development of the Internet has increased the capacity of distribution of these programs and has added the concept of Open Source to the already classic concepts of Freeware/Public Domain Software that opens the doors to a bigger collaboration among groups/developers. The objective of the session is to present some of these projects and to open the debate on the difficulty of the development of this type of solutions, to assure its continuity and to facilitate the cooperation among developers. The speakers will be:

Andrew G. Dean ("creator" of Epi Info and director of the project during 18 years until months ago in that he retired of the CDC) at the moment he is involved in a project on Open Source for the development of a statistical calculator. (www.openepi.com)

Jens Lauritsen. When not convincing him the line in that Epi Info for Windows will be developed, based on access database, he decided to begin the development of a program for windows, EpiData based on the data format of Epi Info for DOS. The unexpected success of this project has made that its development perspective increases incorporating analytical characteristics. (www.epidata.dk)

Luis Carlos Silva Aycaguer. The EpiDat project, sponsored by OPS, it has already incorporated diverse utilities for the calculation of tabulated data. It has been a good complement to Epi Info in the treatment of populational data. In their third version it incorporates.

EPIDAT 3.0: SOFTWARE FOR THE EPIDEMIOLOGIC ANALYSIS OF GROUPED DATA

Luis Carlos Silva Aycaguer

Vicerrectoría de Investigaciones, Instituto Superior de Ciencias Médicas de la Habana, Cuba.

Epidat is a freeware software developed by public institutions and directed to epidemiologists and other health professionals that manage tabulated data. The project began in 1991. The last version is version 3. Epidat 3.0 works under Windows (16 and 32 bits) and it continues to be oriented to tabulated data. In a general way has the same objectives as previous versions: to support and to facilitate the teaching of epidemiology and to improve the analytical tools of health professionals. Version 3.0 has an important qualitative change regarding the previous versions. The work group, more stable and structured, has been a multidisciplinary team, (epidemiologists, statistics and computer professionals) from Galicia, Cuba and Pan-American Health Organization. At first it were defined the scope and the modules of the new version. The development of each module was carried out following a methodology of work: definition of the content, bibliographical search, selection of algorithms, design of screens, programming, tests and corrections. The e-mail communication has been essential among the members of the team. The beta version of each module was tested by an external reviewer, expert in that concrete topic. The help system has been developed with the collaboration of external experts. One of the major difficulties has been the impossibility of maintaining the same programmer during the whole project. The changes of the program can be summarized in relation to the general environment and those referred to the content of the modules. As for the general environment, it has been adopted the Office appearance (incorporating bars for menu and tools). On the other hand, the hypertext help system has a more didactical and critical focus, including statistical and epidemiological principals. Another important change is the results window that can be edited and saved as RTF format. It incorporate import from other format (dbf, excel or mdb) for all modules. Regarding the technical aspects it has been improved the content of the program substantially. It incorporates common specific techniques that usually are not present in programs of this kind. Briefly, it has 5 new modules: Demography, Probability distributions, Bayesian analysis, Epidemiological surveillance and Meta-analysis, and a substantial improvement of those already existent. The developer team is considering the possibility of transforming Epidat 3.0 into free software and, this way, to imply the scientific community in their development and evolution.

324

THE EPIDATA STRATEGY FOR DEVELOPMENT AND MAINTENANCE OF SIMPLE TOOLS FOR SOUND DATA IN PUBLIC HEALTH AND EPIDEMIOLOGY

Jens Lauritsen

Accident Analysis Group & EpiData Development Team, Odense University Hospital & EpiData Association, Denmark.

Working with dataentry and analysis in public health requires relevant tools and professional skills in the public health professionals. Many of these are working with data in circumstances in e.g. low income countries where modern methods such as fast computers, client server or fast internet are remote. But also in the industrialised countries it can be postulated that the most important resource is the capabilities of the professional, not of the computer tool as such. The drive to develop the EpiData software tools has therefore mainly been to maintain the working principle of Epi Info version 6 with a very basic file structure known as QES/CHK/REC files. In this working principle many persons around the world have excellent skills which are not directly transferable to other packages. But with development of graphics interfaced programs like windows most users find it hard to cope with the "dos" mode of working in EpiInfo developed until mid 1990'ties. The "small and focused" policy of EpiData started in late 1999 and has now resulted in freeware tools for dataentry and batch handling of data almost equivalent to the Epi6 datahandling, but extended with features to enhance data documentation. EpiData was completely developed outside formal institutions (apart from the early start) and is now released by the EpiData Association formed by interested persons and the core development team. An international support group "Friends of EpiData" has been established to ensure funding and proper dissemination in local public health. The international collaboration has resulted in translation to over 10 languages and there are currently about 12-1500 downloads of the main program per month from www.epi-data.dk going to around 100 countries. The future development of EpiData focuses on developing tight, compact files which can easily be transported around the world, e.g. size is less than one floppy of 1.4 Mb. Further planned functionality is 1. A basic analysis tool EpiAnalysis 2. Tools for transaction logging to meet formal data requirements. Possibly other aspects will follow such as menuing tools. Aspects for discussion in the seminar will be items such as "avoiding reinvention of the wheel", "experiences with shared responsibility without payment", "how to balance user influence and clarity", "minimal administration", "balance of support and documentation", "documentation principles", "avoiding spam on freeware sites".