EPIDATA: A POTENTIAL TOOL FOR PANDEMICS AND LARGE SCALE OUTBREAKS

Camille Achonu

CD Surveillance Network Meeting

October 29th, 2010

Overview

- APHEO EpiData Project
- Data needs & sources in a pandemic/large scale outbreak
- EpiData
  - History
  - Description
  - Utility

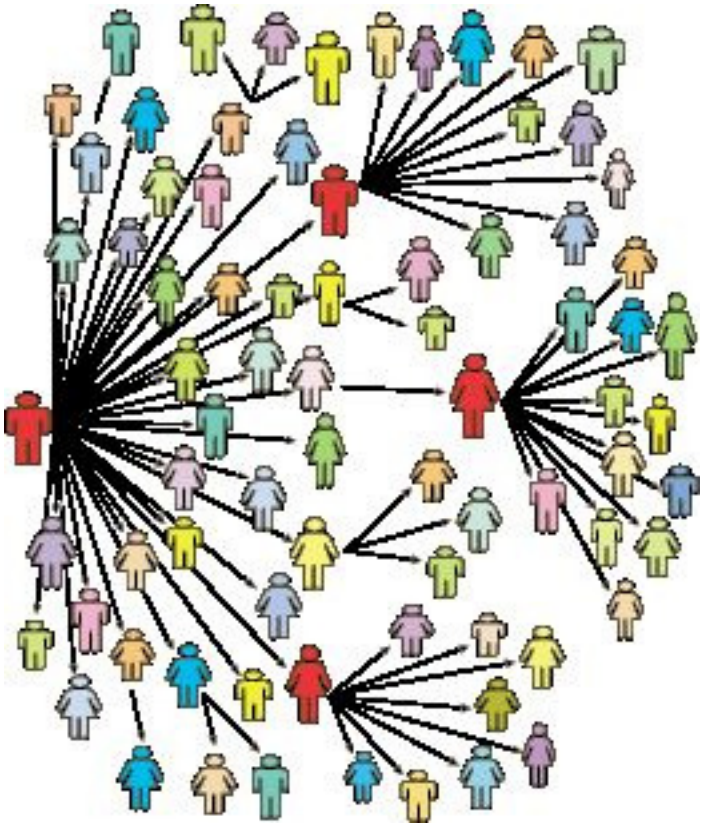
APHEO EpiData Project

- Funded by Public Health Agency of Canada since 2007
- Purpose: APHEO will support and influence the further development of EpiData to improve its utility in the field and as a training tool.
- Project Goals:
  - Improved EpiData functions that are well suited to the practice of public health in Canada
  - Production of test versions and documentation of EpiData software and Canadian case studies to illustrate the management and analysis of data in the field
  - The software, documentation and case studies will be used by public health practitioners and trainers in Canada to complement existing surveillance software, analysis software and training tools.
- EpiData Project website: http://www.apheo.ca/index.php?pid=47

EpiData

Data Challenges in a Pandemic/Large Scale Outbreak

- Source/disease may be unknown or unfamiliar
- Existing systems fail to capture adequate or relevant information
- Changing information needs/data requirements
- Need for rapid access to data to inform decision-making



EpiData

Sources of data

- Population surveys of behavioural risk factors e.g RRFSS, CCHS
- Surveillance - case reporting e.g. iPHIS
- Targeted studies
- Databases to capture interventions e.g. vaccinations

EpiData

Options for Adhoc Database Systems

- Spreadsheets (e.g. Excel)
  - Prone to error, data corruption, & mismanagement
  - Lacks data controls, limited programmability
  - Suitable only for small projects
  - Good for last step data cleaning
- Commercial database programs (e.g. Oracle, Access)
  - Powerful and widely available
  - Slower to develop
  - More expertise required
- Web-based data entry (e.g. SurveyMonkey)
  - Easy to design
  - Limited QA checks and analytical capability
  - Privacy concerns associated with collecting nominal data
- Public domain programs (e.g. EpiData, EpiInfo)
  - Controlled data entry, good programmability
  - Quick & easy to create and adapt

Data Management

- Planning data needs
  - Study design (cohort, case-control, cross-sectional survey)
  - Person, place, time variables
  - Exposure & outcome variables
- Data processing
  - Data entry control
  - Validation and checks
  - Data cleaning and variable transformation
- Data backup and storage
- System documentation

What is EpiData?

- Public domain software package for epidemiologists, public health professionals and researchers
- Two components:
  - EpiData Entry (3.1)
  - EpiData Analysis (v.2.1)
- Suitable for outbreak investigation and survey datasets
- User-friendly interface
  - Pull down menu

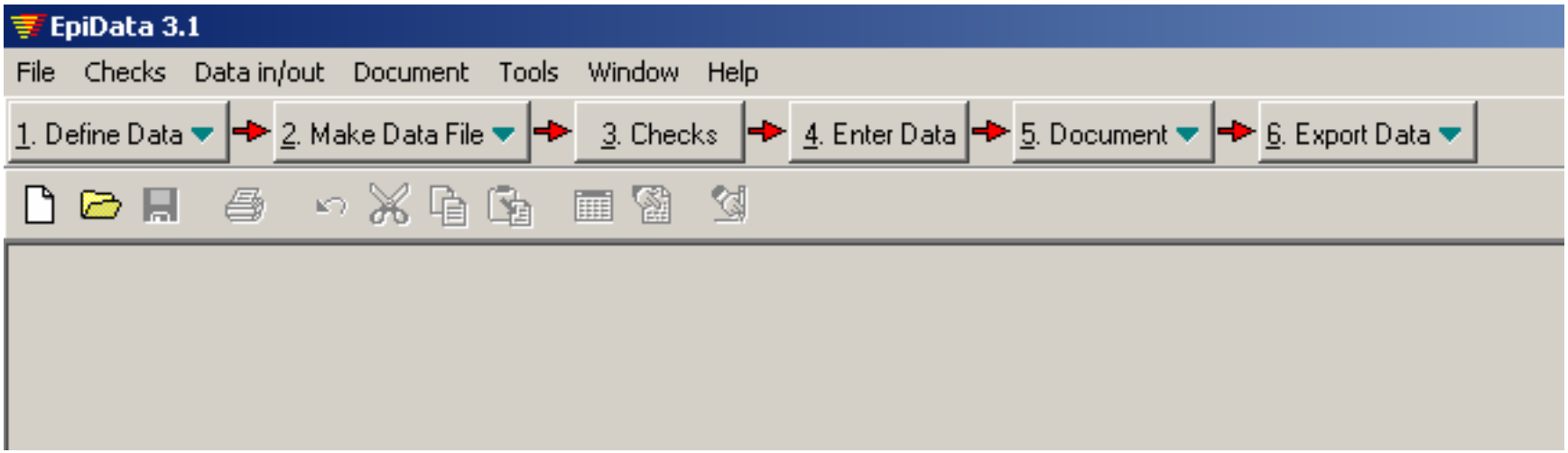
EpiData Entry

- Data entry based on a questionnaire
- Data documentation
- Protect the database from input errors by writing a check program
- Import and export data
- Error detection features. E.g. double entry verification
- Support for relational databases

EpiData

EpiData Entry in Steps

- Define data (QES file)
- Make data file (REC file)
- Checks (CHK file)
- Enter data
- Document
- Export data



EpiData Analysis

- Descriptive statistics
  - Summary statistics, Confidence intervals
- Defining/modifying data
- Produce graphs & tables
- Comprehensive data management: recode variables, define missing values, label values and label variables.
- Multivariable analysis
- Linear regression, correlations
- Stratified analysis

EpiData

EpiData – Not just for outbreaks

- Durham Region Health Unit – 'DineSafe Durham' Evaluation
- WHO - STEPwise approach to surveillance (STEPS)

EpiData

Limitations of EpiData

- Does not meet provincial reporting requirements
  - Additional data entry into iPHIS required
- Does not support multiple, simultaneous users
- No mapping capability
- Not ideal for extremely large datasets

Advantages of EpiData

- ☑ Publicly available, free
- ☑ Existing templates are easily modified
  - ☑ Include in outbreak toolkit
- ☑ New data entry form created from scratch rapidly
- ☑ Minimal training required, user friendly
- ☑ Easy to import and export data in other file formats
- ☑ Consistency checks are easy to apply

EpiData

The Future of EpiData

- New: EpiData XML file format "EPX" will combine and replace the current qes, chk and rec files into one physical file
- More access controls
- The re-written software will consist of three parts
  - a. EpiData Data Manager
    b. EpiData Data Entry Client
    c. EpiData Analysis
  - Public tests of re-written software ongoing

Resources

1. EpiData website: www.epidata.dk
2. EpiData Introduction Guide: A Canadian Example (http://www.apheo.ca/index.php?pid=47#Project Documents)
3. Bennett, et al. (2001). *Data Management for Surveys and Trials. A Practical Primer Using EpiData*. The EpiData Documentation Project: http://www.epidata.dk/downloads/dmepidata.pdf

APHEO EpiData Expert Panel

- Brenda Guarda - Project Manager brenmin@rogers.com
- Camille Achonu - Ontario Agency for Health Protection and Promotion camille.achonu@oahpp.ca
- Anne Arthur - Toronto Public Health aarthur@toronto.ca
- John Barbaro - Simcoe Muskoka District Health Unit john.barbaro@smdhu.org
- Jason Garay - York Region Community and Health Services Department jason.garay@york.ca
- Kathy Moran - Durham Region Health Department kathy.moran@durham.ca

Questions

- What functionality is essential to have in managing and analyzing data during a pandemic/large scale outbreak?
- How can EpiData be enhanced to fill these needs?