

The EpiSurveyor Project

“Faster, easier, cleaner, cheaper, standardized, shareable data.”

DATA DYNE

“There can be no science without measurement.”
– Lord Kelvin, 1891

Contents

Introduction	1
The Big Problem: Not Enough Good Data	2
The Benefits of Computers (and Why Nobody is Using Them)	2
DataDyne EpiSurveyor: Lowering the Barriers to High-Quality Data	3
Project Components	4
Current Project Status	4
Contact Information	4

Introduction



Volunteer collecting health data in Aceh using EpiSurveyor software, January 2005

DataDyne’s two founders, Joel Selanikio and Rose Donna, have combined technological, epidemiological, and clinical medical expertise to become pioneers and leaders in the use of handheld computers to gather information in developing country projects. Using commercial hardware and software we have conducted surveys for CDC, the American Red Cross, the US Army, the International Rescue Committee, and many other clients, and in locations from Bolivia to Ghana to tsunami-ravaged Sumatra. Our experience is that handheld data collection has the power to revolutionize developing country data collection, but that currently available software presents a barrier of cost and complexity that is insurmountable except through the use of expensive software and expensive consultants. This places the technology, and its promise of low-cost but high-quality data, out of reach of all but the most well-funded programs.

By creating simple yet powerful software, making it affordable to all, actively disseminating it, and providing technical support, we can overcome the most important current barriers to a data-driven model of developing country public health.

DataDyne’s EpiSurveyor project centers on development of a desktop/handheld software suite allowing for:

- The easy **creation of data collection instruments** (“surveys”) on a Windows desktop or laptop computer, or the import of instruments created with CDC’s Epi Info software.
- The **use of those instruments to collect data** on a Palm handheld computer
- The **transfer of collected data back to the desktop** for analysis
- The **development of a library of standard data collection instruments** collected from CDC, WHO, NGOs, and translated into EpiSurveyor format
- The **training of NGOs and other organizations** in how to use the technology, and good field epidemiological techniques, for program assessment and other data-dependent activities.

The Big Problem: Not Enough Good Data

If we agree with Lord Kelvin that “there can be no science without measurement”, how should we rate the current state of public health science in developing countries? Data (“measurement”) should drive our understanding of developing country health, helping us to understand what programs are needed, what problems paramount, and what interventions effective, but that data – from baseline indicators to program evaluations, disease surveillance, and rapid assessments – is famously difficult and expensive to obtain.

It is widely understood, but rarely discussed, that because of these factors high-quality data is the exception rather than the rule in developing country public health. Even existing safeguards, such as “double data entry,” have substantial and usually unreported problems.

It is less widely understood that because continued funding increasingly depends on being able to produce high-quality program evaluation data, these same factors of difficulty and expense bias the funding process in favor of programs with large evaluation budgets and against less-well-funded programs – which may be effective but are unable to get the data to prove it.

“The Millennium Development Goals have accelerated demand for data and highlighted limitations in their availability and quality. Significant numbers of countries do not have enough data to track changes in poverty, child malnutrition and HIV/AIDS prevalence, and in most countries serious data quality issues exist in measuring maternal mortality and access to water and sanitation.”

– from developmentgoals.org

An Example: Bias Against Quality Evaluation of Small Projects

In many projects, the evaluation budget is usually set at about 10% of project costs. Current technology (or lack of technology) means that a high-quality assessment costs at least \$50,000, thus limiting such assessments to projects costing \$500,000 or more. Low-cost (usually community-based) projects do not have adequate funding for high quality assessments and the lack of assessment means that there is little published data comparing high-cost to low-cost projects.

Thus, evidence-based donors are potentially biased towards high-cost projects. The EpiSurveyor project can lower the cost of assessment several-fold, thereby bringing high quality assessments to low-cost projects and allowing donors to make more informed decisions about project funding.

The Benefits of Computers (and Why Nobody is Using Them)

Most people would be surprised to find that in the 21st century the vast majority of health data gathered in developing countries is still collected on paper, without the use of computers. And this is despite the fact that data collection on highly portable handheld computers has been technically possible for several years, with demonstrated advantages including increased speed and quality and decreased cost (see brief entitled “Advantages of Using Handheld Computers for Data Collection”).

So why haven’t handheld computers been widely used for data collection? The answer is not in the hardware – like all electronics, handhelds get better, faster, and cheaper every year – but rather in the software: the currently available software to put questionnaires onto a handheld and transfer collected

data back to a desktop computer is complicated enough to require computer programming skills. Prices for such software can run in the thousands, and with the cost of a computer the ultimate cost is typically tens of thousands of dollars. Per project.

Clearly this cost and complexity are a substantial barrier to adoption of the technology – and as with data collection in general, the smaller the program the higher the relative barrier.

EpiSurveyor: Lowering the Barriers to High-Quality Data

The DataDyne EpiSurveyor project is lowering the barriers to collection of high-quality data by creating inexpensive, easy-to-use software for data collection on handhelds: if the cost and difficulty of collecting data are drastically reduced, data is more likely to be collected.

Beyond the obvious benefit of having more data to guide our programs and our understanding, we believe the creation of a public domain common data collection platform will have other, far-reaching effects:

- **More and faster analysis** – Because the data collected with EpiSurveyor will be digital from the moment of collection and because digital data is much easier to analyze, we believe that analysis is much more likely to be done, and done promptly (without having to wait months for data entry). Faster data means the ability to more quickly respond to changing circumstances.
- **Lowered cost of high-quality program assessment** –The current system puts good program assessment – for malaria, TB, HIV, measles and other problems – out of the reach of all but the most well-funded programs. As mentioned earlier, this inability to afford high-quality assessment handicaps less well-funded programs in the competition for funding – regardless of the effectiveness of their interventions.
- **Promotion of standardization and meta-analysis** – right now, there is no universal system for sharing data collection instruments. EpiSurveyor survey files, which as electronic files are easy to catalog, download, share, and use, can provide a common platform for sharing those data collection instruments that represent best practices. We believe that if a standard “CDC Maternal Health Evaluation Form” exists in a searchable online database, for example, public health practitioners looking to evaluate maternal health are likely to “take the path of least resistance” and utilize that form. Even if minor modifications are made to allow for local circumstances, it will still be much more likely that data collected in a variety of programs will be structured similarly, a prerequisite for meta-analysis.



Interviewer uses handheld to capture measles vaccination information, Zambia 2003.

Project Components

1. Software development: DataDyne EpiSurveyor
 - a. Designer (desktop component)
 - b. Survey Engine (handheld component)
 - c. Porting of EpiSurveyor to other platforms, including UNIX, Mac, and PocketPC
2. Training activities
 - a. Development of training materials for both live and online learning
 - b. Live seminars to teach both use of the software and essential epidemiology for data collection
 - c. Online learning
3. Website for:
 - a. Distribution of software
 - b. Distribution of training materials
 - c. Forum for user discussion and collaboration
 - d. Online training
 - e. Tracking and responding to bug reports, suggestions, etc
 - f. Promoting use of software
 - g. Cataloging standard data-collection instruments in EpiSurveyor format
 - i. Actively collecting data collection instruments from key organizations
 - ii. Translating the collected instruments into EpiSurveyor format
 - iii. Creating searchable online library of instruments

Current Project Status

Using grants from the infoDev consortium (www.infodiv.org), the UN Foundation (www.unfoundation.org), and private funding from the project director, we are now poised to release a beta version of the software to the public health community by October 2005, and have developed a website to allow software downloads and registration of users (www.datadynegroup.com).

We are currently seeking operational funding for a two-year plan to:

- Continue software development, respond to beta-user feedback, and promote the software
- Actively train potential users in the NGO community and elsewhere
- Begin digitizing, categorizing, and cataloging “best practice” data collection instruments from CDC, WHO, UNICEF and others to form the core of a searchable online database

Contact Information

Joel Selanikio, MD – Project Director

jselanikio@datadynegroup.com

(202) 468-7227