

SAGES: A Suite of Freely-Available Software Tools for Electronic Disease Surveillance in Resource-Limited Settings

Sheri L. Lewis^{1*}, Brian H. Feighner¹, Wayne A. Loschen¹, Richard A. Wojcik¹, Joseph F. Skora¹, Jacqueline S. Coberly¹, David L. Blazes²

1 National Security Technology Department, The Johns Hopkins University Applied Physics Laboratory, Laurel, Maryland, United States of America, **2** Division of GEIS Operations, Armed Forces Health Surveillance Center, Silver Spring, Maryland, United States of America

Abstract

Public health surveillance is undergoing a revolution driven by advances in the field of information technology. Many countries have experienced vast improvements in the collection, ingestion, analysis, visualization, and dissemination of public health data. Resource-limited countries have lagged behind due to challenges in information technology infrastructure, public health resources, and the costs of proprietary software. The *Suite for Automated Global Electronic bioSurveillance* (SAGES) is a collection of modular, flexible, freely-available software tools for electronic disease surveillance in resource-limited settings. One or more SAGES tools may be used in concert with existing surveillance applications or the SAGES tools may be used *en masse* for an end-to-end biosurveillance capability. This flexibility allows for the development of an inexpensive, customized, and sustainable disease surveillance system. The ability to rapidly assess anomalous disease activity may lead to more efficient use of limited resources and better compliance with World Health Organization International Health Regulations.

Citation: Lewis SL, Feighner BH, Loschen WA, Wojcik RA, Skora JF, et al. (2011) SAGES: A Suite of Freely-Available Software Tools for Electronic Disease Surveillance in Resource-Limited Settings. PLoS ONE 6(5): e19750. doi:10.1371/journal.pone.0019750

Editor: Abdissalan Mohamed Noor, Kenya Medical Research Institute - Wellcome Trust Research Programme, Kenya

Received: November 23, 2010; **Accepted:** April 15, 2011; **Published:** May 10, 2011

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

Funding: Funding for this work has been provided by the Armed Forces Health Surveillance Center (<http://www.afhsc.mil>). The funders had no role in the study design, data collection and analysis, or decision to publish. The funders did review the manuscript prior to submission.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: Sheri.Lewis@jhuapl.edu

Introduction

History and Background

Emerging and re-emerging infectious diseases are a serious threat to global public health. [1,2] The World Health Organization (WHO) has identified more than 1100 epidemic events worldwide in the last five years alone. [3] Recently, the emergence of the novel 2009 influenza A (H1N1) virus and the SARS coronavirus have demonstrated how rapidly pathogens can spread worldwide. [4] This infectious disease threat, combined with a concern over man-made biological or chemical events, spurred WHO to update their International Health Regulations (IHR) in 2005. [5] The new 2005 IHR, a legally binding instrument for all 194 WHO member countries, significantly expanded the scope of reportable conditions and are intended to help prevent and respond to global public health threats. SAGES, an electronic biosurveillance initiative described herein, aims to improve local public health surveillance and IHR compliance with particular emphasis on resource-limited settings.

Electronic disease surveillance, particularly syndromic surveillance, holds promise to improve health security in resource-limited environments. [6,7] Such systems have become versatile tools in health departments in the United States. [8] Epidemiologists using electronic disease surveillance not only have the potential to detect anomalous disease activity earlier than traditional laboratory-based surveillance, but they also have the ability to monitor the

health of their community in the face of a known threat. [9,10,11] More than a decade ago, in collaboration with the US Department of Defense (DoD), the Johns Hopkins University Applied Physics Laboratory (JHU/APL) developed the *Electronic Surveillance System for the Early Notification of Community-based Epidemics* (ESSENCE). ESSENCE collects, processes, and analyzes non-traditional data sources (i.e. chief complaints from hospital emergency departments, school absentee data, poison control center calls, over-the-counter pharmaceutical sales, etc) to identify anomalous disease activity in a community. The data can be queried, analyzed, and visualized both temporally and spatially by the end user. [10] ESSENCE is currently being utilized by the US Department of Defense, US Veterans Health Administration, and numerous state and local health departments in the US. [11]

The current SAGES initiative leverages the experience gained in the development of ESSENCE, and the analysis and visualization components of SAGES are built with the same features in mind. This paper will describe the key features and potential uses of the SAGES product in developing or resource-limited settings.

Methods

Selected Advances - SAGES

The Global Emerging Infections Surveillance and Response System, a division of the Armed Forces Health Surveillance

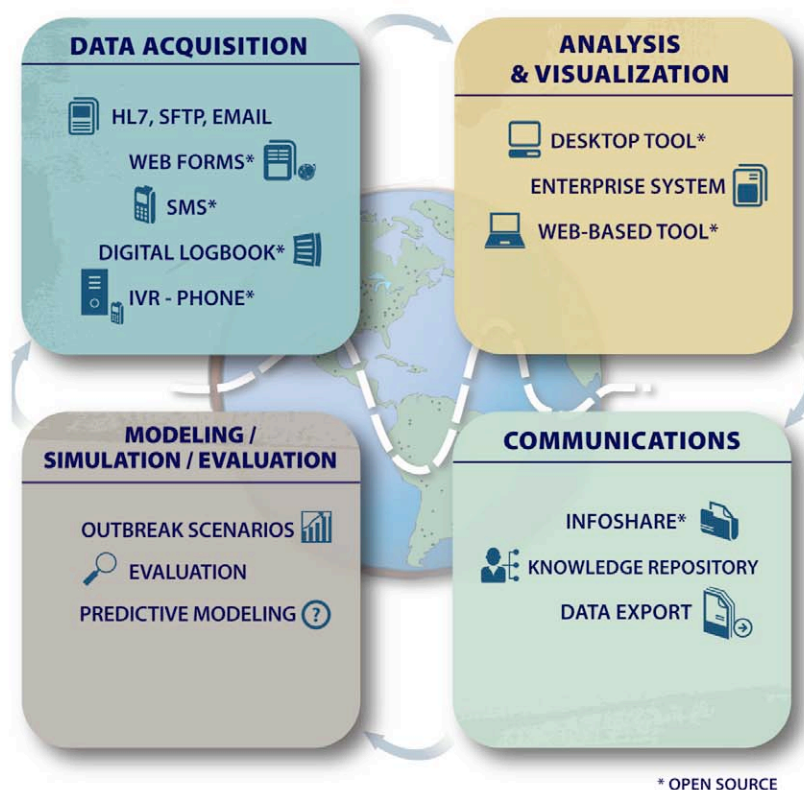
Center, is committed to enhancing electronic disease surveillance capacity in resource-limited settings around the world. To this end, they have entered into a robust collaboration with the JHU/APL to create SAGES. Cognizant of work underway on individual surveillance systems components, e.g., collection of data by cell phones, we have focused our efforts on the integration of inexpensive and/or freely available, interoperable software tools that facilitate regional public health collaborations. [12]

SAGES tools are organized into four categories: 1) data collection, 2) analysis & visualization, 3) communications, and 4) modeling/simulation/evaluation. (Figure 1) Within each category, SAGES offers a variety of tools compatible with surveillance needs and different types or levels of information technology infrastructure. In addition to the flexibility of tool selection, there is flexibility in the sense that the analysis tools do not require a fixed database format. For example, rather than requiring an existing database to adapt to the tool, the SAGES database tools adapt to the format of all Java database compliant formats. Lastly, the SAGES tools are built in a modular nature, which allows for the user to select one or more tools to enhance an existing surveillance system or use the tools *en masse* for an end-to-end electronic disease surveillance capability. Thus, each locality can select tools from SAGES based upon their needs, capabilities, and existing systems to create a customized electronic disease surveillance system.

Data Acquisition. Rapid data acquisition is arguably the most challenging aspect of establishing a successful electronic

disease surveillance system. [9,11] In resource-limited settings, it is imperative to select the technology that is both easy to incorporate into existing health services and sustainable with little or no additional financial investment. The approach should allow customizable data collection, enable multiple data streams collected in different ways, and be scalable based upon needs. [6,7] Data collection tools included within SAGES are web forms, short message service texting programs, digital logbooks, and interactive voice response systems. Data entered into SAGES can be validated by the stakeholder. Where appropriate, other collection methods such as email and secure file transfer protocol can be applied as well. The data acquisition methods in SAGES are readily adaptable to evolving standards in both minimum data sets for disease surveillance and routine diagnosis and care.

Analysis & Visualization. As previously discussed, the SAGES analysis and visualization tools are built upon the features and functionality of the more mature enterprise ESSENCE system. The enterprise ESSENCE system requires a high speed internet connection, relies on automated data streams, and uses proprietary software for the display of data. [10,11,13] The SAGES web-based application and the desktop application are both freely-available tools that provide similar functionality to countries with limited public health resources. All ESSENCE applications contain alerting algorithms developed by JHU/APL and the Early Aberration Reporting System (EARS) algorithms



APL Funding provided by
Armed Forces Health Surveillance Center,
Division of GEIS Operations



Figure 1. SAGES Suite of Tools. The Suite for Automated Global Electronic bioSurveillance (SAGES) is a collection of modular, flexible, freely-available software tools for electronic disease surveillance in resource-limited settings. One or more SAGES tools may be used in concert with existing surveillance applications or the SAGES tools may be used *en masse* for an end-to-end biosurveillance capability.
doi:10.1371/journal.pone.0019750.g001

developed by the U.S. Centers for Disease Control and Prevention (CDC) to identify anomalous events. Users also have the ability to add additional algorithms as desired.

The desktop tool, known as ESSENCE Desktop Edition (EDE), is a single-user stand-alone analysis and visualization tool that can be installed on most computers. EDE does not need access to the internet as it ingests data files stored on the same computer. The web-based tool, known as OpenESSENCE, is a multi-user network accessible data entry, analysis, and visualization tool that enables an epidemiologist to monitor the population's health from any computer connected to that network. Available analyses for both tools depend on the nature of the data ingestion, but may include demographic characterizations, temporal and spatial analyses, display of patient level information, geographic information system mapping, anomalous event detection, and dynamic query capability. (Figures 2, 3, 4)

Communications. SAGES tools can facilitate compliance with 2005 IHR reporting requirements and allow the sharing of actionable information across jurisdictional boundaries. Sharing of patient-level data across regional boundaries is generally not realistic and often not helpful, as local public health entities are

usually best suited to interpret local events. However, once these data have been transformed into meaningful information, it may be immensely valuable to share that information with other countries in the region. Dissemination of this type of information may aid in the interpretation of regional events and helps foster better, lasting public health collaborations. SAGES includes tools for two-way communication between public health officials and graphics that are exportable into common image formats. Each SAGES user controls the type and level of detail of all information shared with each recipient ('role-based access') and also whether the information sharing is manual or automated. [14] (Figure 5) More importantly, the data are collected and stored only by the user, and remain under the sole control of the user at all times.

Modeling/Simulation/Evaluation. JHU/APL has sponsored several electronic disease surveillance exercises to train users and test surveillance system features. [15] We have developed a number of methods for developing simulated outbreaks, both natural and man-made, which can then be 'injected' into a simulated database for exercise purposes. JHU/APL has experience with agent-based infectious disease modeling

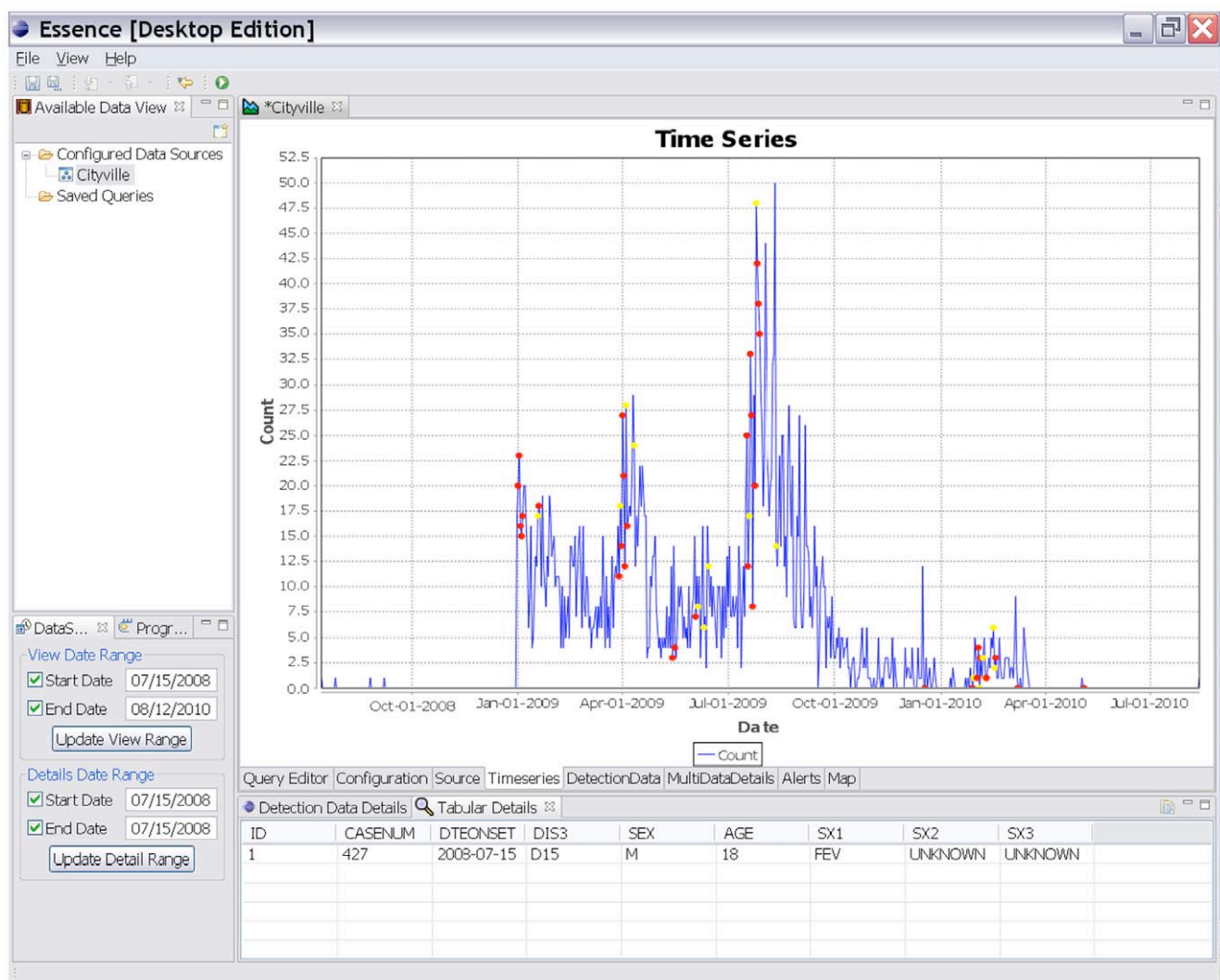


Figure 2. Times Series Display. Through *ESSENCE Desktop Edition (EDE)*, users have the ability to see time series views of their data labeled with early event detections. Detections are generated using one of the default algorithms supplied with EDE; however, users can add additional algorithms of their choosing.
doi:10.1371/journal.pone.0019750.g002

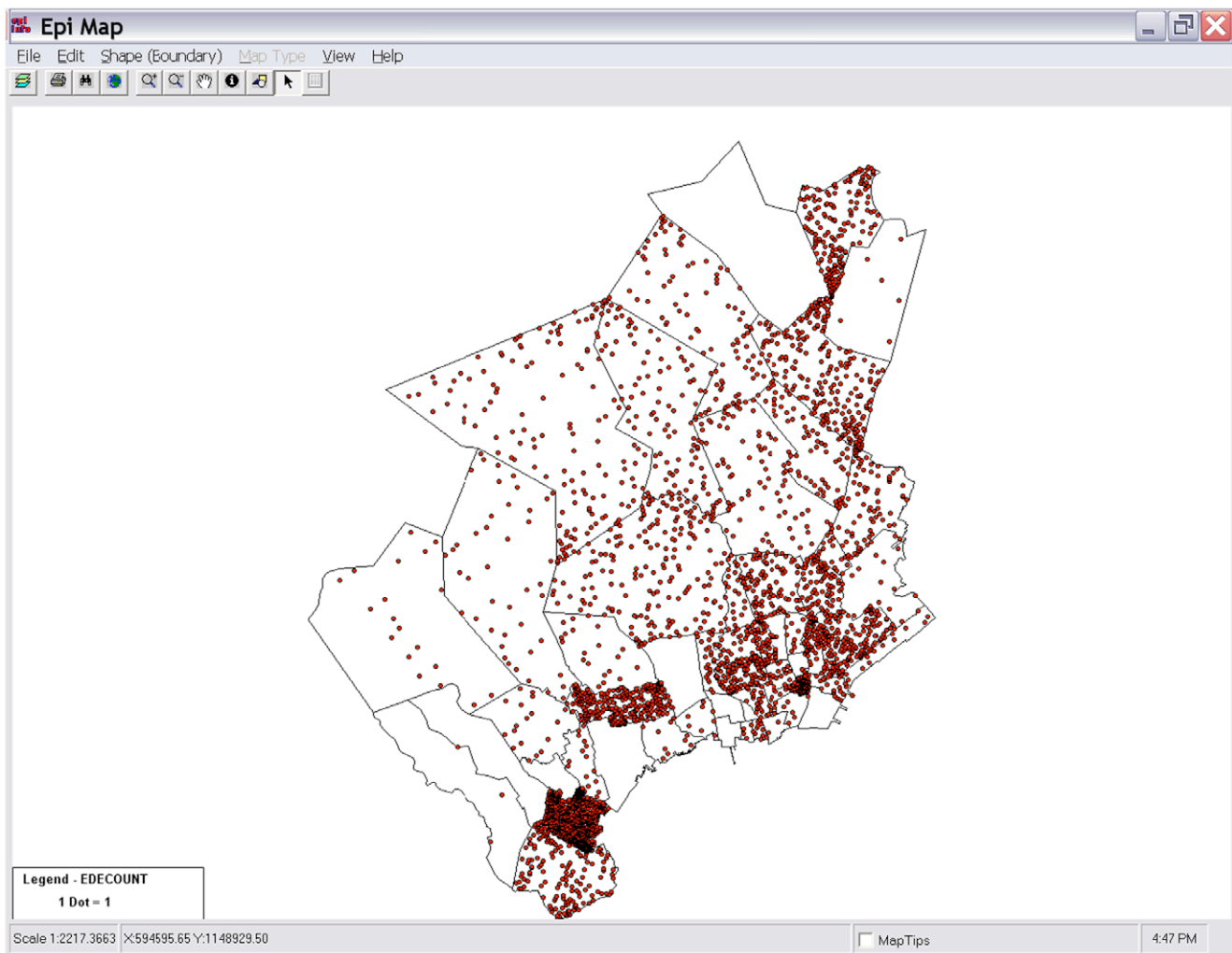


Figure 3. Mapping Utilities. *OpenESSENCE* and EDE users can map spatial data, cases, and alert levels in geographic areas defined in the data. *OpenESSENCE* embeds maps in the application using a GeoServer interface. EDE embeds maps through a uDig interface or optionally launches CDC's EpiMap application with Environmental Systems Research Institute shape files. doi:10.1371/journal.pone.0019750.g003

for pandemic influenza as well as other techniques for predictive disease modeling. [16] It is our desire to explore possible collaborations between our disease modelers and users of electronic disease surveillance systems such as SAGES.

The purpose of disease surveillance is to direct the expenditure of limited public health resources in a manner that yields the greatest return on investment. [17] Only those diseases with the highest burdens on the population should be followed and surveillance systems should be periodically evaluated to determine their usefulness. [17,18,19] At this time, SAGES does not include tools for the automated evaluation of surveillance systems. Several authors have described the evaluation of electronic disease surveillance systems. [9,17,18,19,20] We intend to perform such evaluations of SAGES surveillance systems and investigate the feasibility of developing automated evaluation tools in the future.

Results

SAGES in the Field

Pilot activities in Peru and the Philippines have provided valuable insight into how electronic disease surveillance systems can be successful in resource limited environments. The Peruvian

system uses IVR technology for data collection. The phone calls needed for IVR are very expensive in the Philippines while SMS text messages cost very little. With that in mind, an existing paper-based fever surveillance system was modified to allow basic information (age, sex, clinic, and symptoms) on patients presenting with fever to be sent to the health department daily via SMS using a standardized texting protocol and abbreviations. The SMS messages are received by a cell phone attached to a computer located at the health department and routed through a simple data cleansing program that sequesters records with incorrect formatting and sends a standard text message to the sender asking them to resend the data. The correctly formatted messages are copied to a database that is then analyzed in EDE. Some clinics quickly adapted to the SMS protocol, while others have found it more difficult to send error-free messages. The percent error decreased over time (Figure 6) suggesting that, with constructive help, most clinics were able to become proficient with the SMS protocol. As of January 2011, approximately 50% of eligible districts are sending daily fever information via text, and the data are being regularly reviewed in EDE. The remaining districts continue to send paper reports or report by phone. The primary reasons given for not using the SMS protocol were lack of personnel or

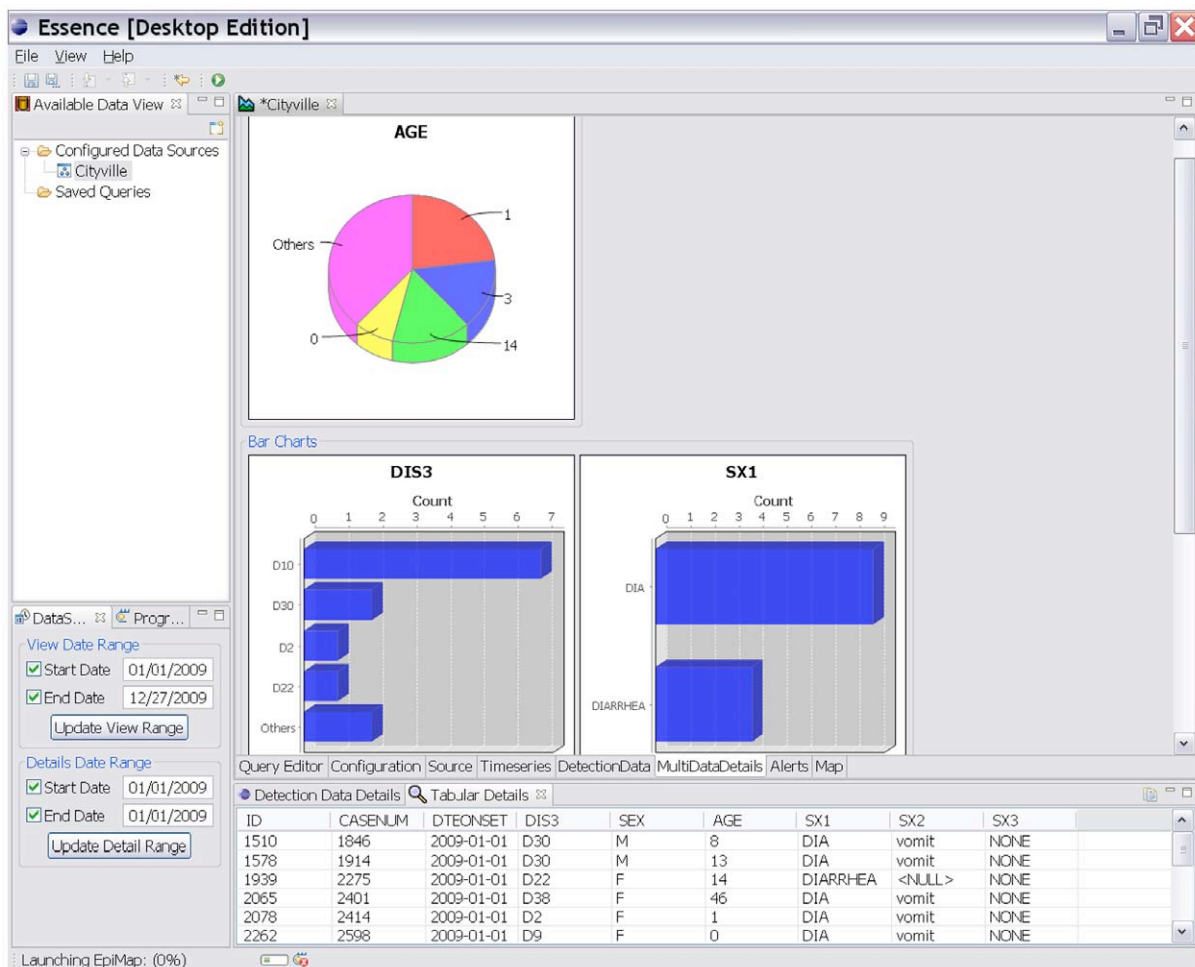


Figure 4. Chart and Graph Utilities. Open ESSENCE and EDE both have various tools for the visualization of data. All graphs and charts products may be exported into common graphic formats.
doi:10.1371/journal.pone.0019750.g004

Potential Regional Network

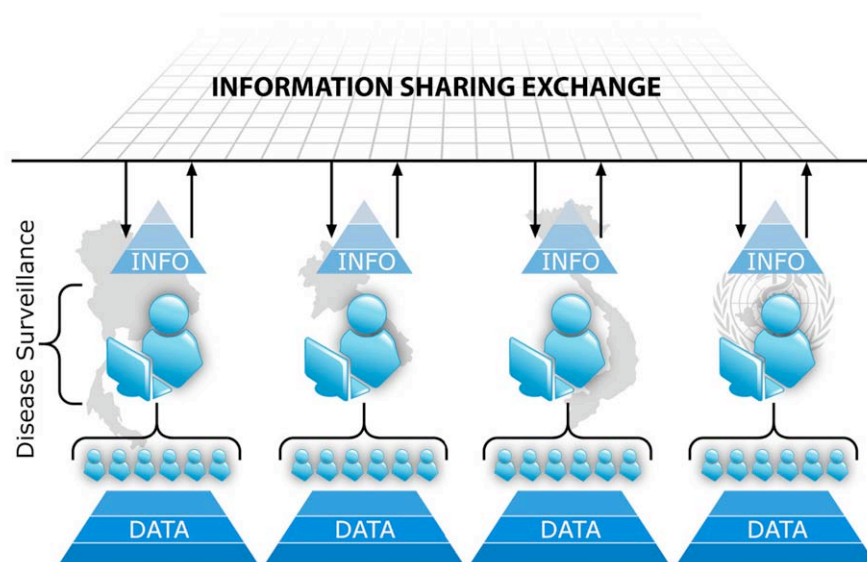


Figure 5. Role-based Information Sharing. InfoShare allows each user to determine role-based access to public health information. The data are collected and stored only by the user, and remain under the sole control of the user at all times.
doi:10.1371/journal.pone.0019750.g005

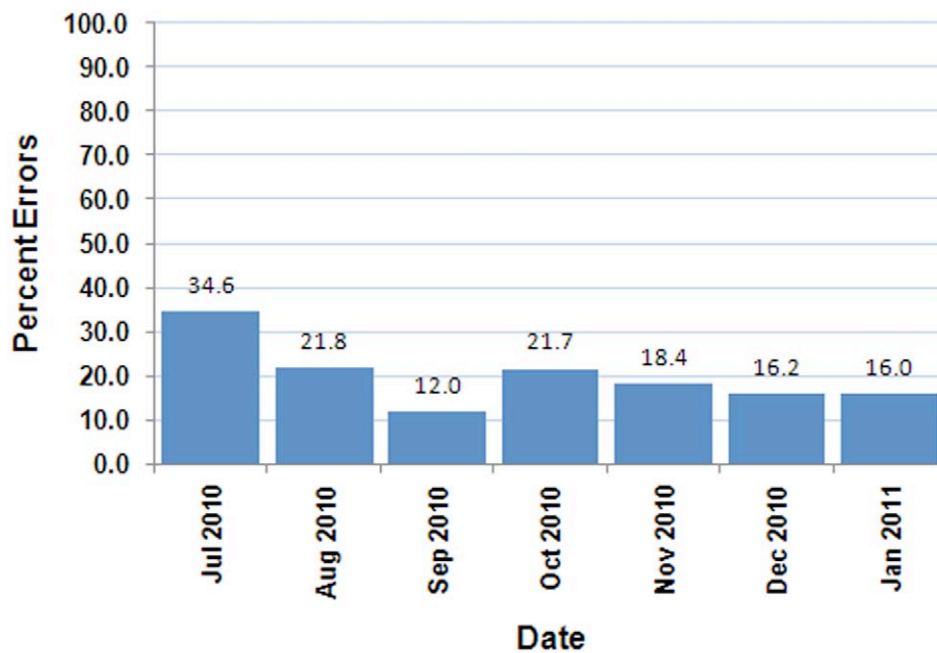


Figure 6. Percent of SMS Protocol Errors by Date. In the Philippines, the percent error of data transmitted via SMS decreased over time. doi:10.1371/journal.pone.0019750.g006

discomfort sending SMS text messages. Focused, repeated training also helped some clinics get started.

Discussion

Challenges to Development

The development and implementation of an open-source, electronic disease surveillance system in a resource limited setting faces many challenges. Assuring the smooth interoperability of all software components is sometimes difficult, as is balancing functionality with simplicity for sites with minimal infrastructure. Additionally, finding the appropriate person or group within a country to champion the implementation of a new surveillance system can be fraught with financial, political, territorial challenges. Finally, the open source nature of the software inherently has the consequence of a lack of control of modifications to the software after it is released, potentially leading to incompatible and unverified versions. Conversely, if these changes can be shared in a Wiki-like environment, the entire community using SAGES may benefit from novel improvements.

Summary

The SAGES project is intended to enhance electronic disease surveillance capacity in resource-limited settings around the world. We have combined electronic disease surveillance tools developed at JHU/APL with other freely-available, interoperable software tools to create SAGES. We believe this suite of tools will facilitate

local and regional electronic disease surveillance, regional public health collaborations, and international disease reporting. SAGES tools are currently undergoing pilot testing in locations in Southeast Asia and South America, and will be offered to other interested countries around the world. Opportunities for collaboration are currently being discussed in Central and East African sites as well.

Acknowledgments

We would like to acknowledge the invaluable assistance from our international colleagues. These include Dr. InKyu Yoon from the US Armed Forces Research Institute of Medical Sciences in Bangkok, Thailand, Dr. John Mark S. Velasco, Ms. Agnes D. Tomayao, Mr. Danny C. Obidas, Mr. Jewernest C. Casquejo, and Dr. Maria Theresa P. Alera from the US Armed Forces Research Institute of Medical Sciences Virology Research Unit, Republic of the Philippines, Dr. Fe A. Cabugao, Dr. Ilya A. Tac-an, and Ms. Durinda R. Macasocol from the Cebu City Health Office, Republic of the Philippines, and Dr. Joan M. Neyra, Dr. Delphis M. Vera, Dr. Ricardo A. Hora, and Dr. Joel Montgomery at the US Naval Medical Research Center Detachment in Lima, Peru.

The views expressed here are not to be construed as official or as representing the true views of the U.S. Department of the Navy or the U.S. Department of Defense.

Author Contributions

Conceived and designed the experiments: SHL BHF WAL RAW JFS JSC DLB. Wrote the paper: SHL BHF WAL RAW JFS JSC DLB.

References

1. Binder S, Levitt AM, Sacks JJ, Hughes JM (1999) Emerging infectious diseases: public health issues for the 21st century. *Science* 284: 1311–1313.
2. Morens DM, Folkers GK, Fauci AS (2004) The challenge of emerging and re-emerging infectious diseases. *Nature* 430: 242–249.
3. WHO (2007) The world health report 2007 - a safer future: global public health security in the 21st century. Geneva: World Health Organization.
4. Hollingsworth TD, Ferguson NM, Anderson RM (2007) Frequent travelers and rate of spread of epidemics. *Emerging Infectious Disease* 13: 1288–1294.
5. WHO (2005) International health regulations (2005). Geneva: World Health Organization.
6. Chretien JP, Burkom HS, Sedyaningsih ER, Larasati RP, Lescano AG, et al. (2008) Syndromic surveillance: adapting innovations to developing settings. *PLoS Medicine* 5.
7. Soto G, Araujo-Castillo RV, Neyra J, Fernandez M, Leturia C, et al. (2007) Challenges in the implementation of an electronic surveillance system in a resource-limited setting: Alerta, in Peru. *BMC Proceedings* 2.

8. CDC (2007) Status of state electronic disease surveillance systems - United States, 2007. *MMWR* 58: 804–807.
9. Jajosky RA, Groseclose SL (2004) Evaluation of reporting timeliness of public health surveillance systems for infectious diseases. *BMC Public Health* 4.
10. Lombardo J, Burkom H, Elbert E, Magruder S, Happel-Lewis S, et al. (2003) A system overview of the electronic surveillance system for the early notification of community-based epidemics (ESSENCE II). *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 80: i32–42.
11. Lombardo JS, Ross D (2007) Disease surveillance: a public health priority. In: Lombardo JS, Buckeridge DL, eds. *Disease Surveillance: A Public Health Informatics Approach*. Hoboken: John Wiley & Sons, Inc. pp 1–39.
12. Ashar R, Lewis SL, Blazes DL, Chretien JP (2010) Applying information and communications technologies to collect health data from remote settings: a systematic assessment of current technologies. *Journal of Biomedical Informatics* 43: 332–341.
13. Lombardo JS, Burkom H, Pavlin J (2004) ESSENCE II and the framework for evaluating syndromic surveillance systems. *MMWR* 53: 159–165.
14. Loschen W, Coberly JS, Sniegowski C, Holtry R, Sikes ML, et al. (2007) Event communication in a regional disease surveillance system. *AMIA Annu Symp Proc*. pp 483–7.
15. Happel-Lewis S, Cutchis PN, Babin SM, Burkom H (2003) Simulated release of plague in Montgomery County, Maryland. *JHU Applied Physics Laboratory Technical Digest* 24: 354–359.
16. Feighner BH, Chretien JP, Murphy SP, Skora JF, Coberly JS, et al. (2009) The pandemic influenza policy model: a planning tool for military public health officials. *Military Medicine* 174: 557–565.
17. Teutsch SM, Churchill RE, eds (1994) *Principles and Practice of Public Health Surveillance*. New York: Oxford University Press.
18. Buchler JW, Hopkins RS, Overhage JM, Sosin DM, Tong V (2004) Framework for evaluating public health surveillance systems for early detection of outbreaks. *MMWR RR-05*: 1–11.
19. Klaucke DN, Buchler JW, Thacker SB, Parrish RG, Trowbridge FL, et al. (1988) Guidelines for evaluating surveillance systems. *MMWR* 37: 1–18.
20. Sosin DM (2003) Draft framework for evaluating syndromic surveillance systems. *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 80: i8–13.