

# Emerging Information Technologies to Provide Improved Decision Support for Surveillance, Prevention, and Control of Vector-Borne Diseases

Saul Lozano-Fuentes<sup>1</sup>, Christopher M. Barker<sup>2,3</sup>, Marlize Coleman<sup>1,4</sup>, Michael Coleman<sup>4</sup>, Bborie Park<sup>2</sup>, William K. Reisen<sup>2,3</sup> and Lars Eisen<sup>1</sup>

<sup>1</sup>*Colorado State University, Fort Collins, Colorado*

<sup>2</sup>*University of California, Davis, California*

<sup>3</sup>*Fogarty International Center, National Institutes of Health, Bethesda, Maryland*

<sup>4</sup>*Liverpool School of Tropical Medicine, Liverpool*

<sup>1,2,3</sup>*United States of America*

<sup>4</sup>*United Kingdom*

## 1. Introduction

Vector-borne diseases, caused by pathogens transmitted by arthropods, result in significant morbidity and mortality of humans, especially in the developing world (Gratz, 1999). Malaria caused an estimated 247 million cases and nearly a million deaths in 2008, and up to 50 million dengue infections and 500,000 cases of severe dengue hemorrhagic fever are estimated to occur each year (World Health Organization, 2008, 2009). Furthermore, new vector-borne diseases have emerged and become established in developed parts of the world in recent decades. Vector-borne diseases that now are a fact of life in such areas, and unlikely to be eliminated, include West Nile virus disease in North America and Lyme borreliosis in Asia, Europe, and North America (Sonenshine, 1993; Kramer et al., 2008).

Surveillance and control of a vector-borne disease is a complex undertaking because it involves arthropod vectors, pathogens, and vertebrate amplification/reservoir hosts, each of which may be represented by one or more species. Humans may be amplification/reservoir hosts for the pathogens (e.g., dengue and malaria) or dead-end hosts (e.g., Lyme borreliosis and West Nile virus disease). Vector/disease control programs are thus faced with the challenge of handling a wide range of information including entomological surveillance data (vector collection details, vector abundance, and insecticide resistance) and pathogen-related surveillance data (infection of vectors, enzootic amplification/reservoir hosts or sentinel animals, and passively or actively acquired data on infection in humans). They also need to manage data relating to the control of the vector and/or the pathogen, potentially including the coverage in space and time of a wide range of prevention or control activities (e.g., vaccination, education campaigns, and different interventions targeting arthropod vectors or vertebrate amplification/reservoir hosts) as well as the time and amount of stock

materials (e.g., vaccines or insecticides) expended in the effort. In an ideal scenario, this is complemented by the determination of entomological and epidemiological outcome measures to assess control program performance.

Without adequate capacity for data management and analysis, a control program stands little chance of being able to assess and improve its performance. This can lead to poor decision making, including the continued use of ineffective surveillance or control methods and a failure to target available resources to areas and time periods where they would have the greatest impact. Emerging information technologies present new opportunities to reduce the burden of vector-borne diseases through improved decision support for surveillance, prevention, and control of vectors and their associated pathogens (reviewed by L. Eisen & R.J. Eisen, 2011). This ranges from improved basic capacity for data management and analysis to development of integrated systems with decision support functionalities such as custom calculations for important surveillance or control parameters, automated alerts when thresholds for key entomological or epidemiological risk measures are reached, and capacity for map-based data visualization. The flow scheme in Figure 1 illustrates how such a system can be incorporated into control program operations.

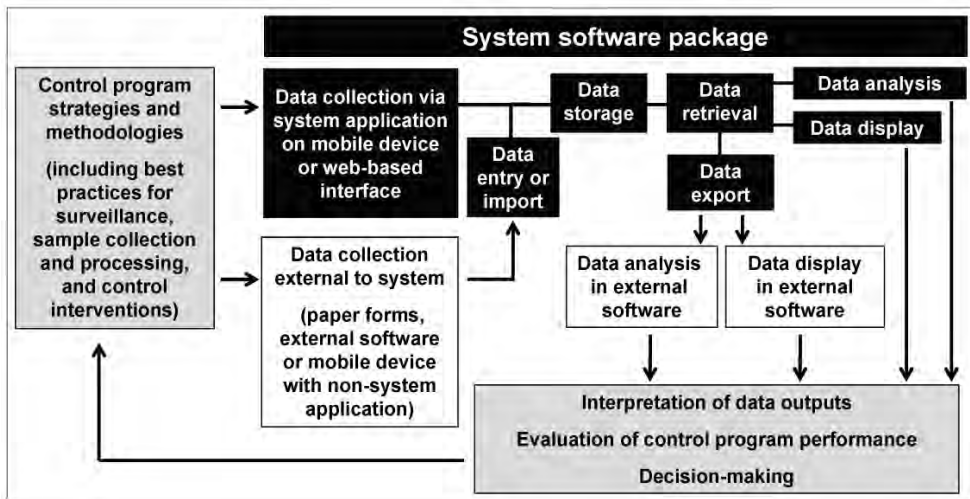


Fig. 1. Flow scheme illustrating how a data management system or decision support system can be incorporated into vector and disease control program operations.

This chapter provides a brief overview of information technologies with the potential for application to vector-borne disease surveillance, prevention, and control (section 2), and provides examples of novel decision support tools for vector-borne diseases (sections 3-4). Special emphasis is placed on information technologies that can be implemented in the resource-constrained environments that suffer the greatest burden of vector-borne diseases. Here, we use the terminology "data management system/decision support system" to avoid issues related to the definition of a decision support system and to recognize that system packages may have different levels of built-in decision support functionalities regardless of whether they are labeled data management systems or decision support systems.

## **2. Information technologies with the potential for application in data management systems/decision support systems for vector-borne diseases**

Data management and analysis are facilitated by use of database, reporting, and mapping/ Geographic Information System (GIS) software, especially when multiple applications are combined in an integrated system. For example, the recent introduction of mosquito-borne West Nile virus into North America resulted in the development of novel systems that include GIS to provide enhanced capacity for data display and to support decision-making: the Integrated System for Public Health Monitoring of West Nile Virus in Canada (Gosselin et al., 2005) and the California Vectorborne Disease Surveillance Gateway (section 4).

Many systems include software components with high acquisition and/or licensing costs, thus preventing implementation in resource-constrained environments and limiting the potential for using the systems to address vector-borne diseases in developing countries. One way of overcoming this problem is to harness the explosion of software products that can be distributed and used without licensing costs, e.g., open source products, and to develop an integrated system based on such components (Yi et al., 2008). Both systems described below (sections 3-4) make extensive use of open source products. The infectious disease community is now starting to use freely available software, especially in resource-constrained environments where software with high licensing costs are not sustainable. Below, the mapping tool Google Earth (<http://www.google.com/earth/index.html>) is used as an example of a stand-alone application with a freely available version which now is being used as a decision support tool in public health, including vector-borne diseases (Lozano-Fuentes et al., 2008; Stensgaard et al., 2009).

Google Earth provides free access to satellite imagery and includes tools allowing for production and customization of polygons, lines, and points overlaid on the image. These features can be saved, with their spatial references, as Keyhole Markup Language (KML) files and distributed to other parties, for example as e-mail attachments. The other parties can then view the created features in Google Earth, overlaid on the same satellite image from which they were created based on the spatial reference of the features. The application also can generate dynamic time-series maps that show, in a far more intuitive way than a series of static map images, how the spatial distribution of disease cases, or other data of interest, changes over time. The features included in these dynamic time-series maps can be saved as KML files and viewed by other parties, similar to the features in a static map.

One basic use of this type of mapping tool in public health is to produce "electronic pin maps" that show the locations of disease cases overlaid on an image showing the physical environment. This reveals case clustering and may provide insights into risk factors for pathogen exposure such as aquatic mosquito habitat adjacent to urban areas. To enhance the value of a disease case pin map produced in a mapping software, it can be augmented with data for locations of health facilities or transport routes, and thus provide additional information useful to guide public health actions (Kamadjeu, 2009). Other uses for Google Earth include display of malaria parasite rates in the Malaria Atlas Project (Hay & Snow, 2006) and collection locations for insecticide-resistant vector mosquitoes (Dialynas et al., 2009).

Mapping software that provides access to high-quality imagery of the physical environment can complement GIS software in settings where spatial data layers are unavailable but imagery that can be accessed through the mapping software is of high quality. Lozano-Fuentes et al. (2008) demonstrated how an image available through Google Earth, combined with the basic editing tools included in the application, can be used to first develop a basic

city infrastructure representation and then use this representation to display disease case locations, in this instance for dengue. Because KML files can be converted into shapefiles for use in GIS software, and vice versa, there is great potential for combined use of GIS and Google Earth. Kamadjeu (2009) imported district boundaries originally developed in a GIS into Google Earth and then developed a choropleth map for polio vaccination coverage by district which was overlaid by the location of polio cases. In Nicaragua, a base map for town infrastructure developed in Google Earth was imported into a GIS and used, together with data for dengue case locations and development sites for the mosquito vector, to support dengue control operations (Chang et al., 2009). These examples highlight the potential for creative use of emerging information technologies as stand-alone decision support tools.

Mobile data capture is an emerging information technology with potential for incorporation into a data management system/decision support system. It provides the opportunity, through mobile computers (laptops or netbooks), personal digital assistants (PDAs; also known as palmtop computers), remote sensors, or even cell phones, to move the stage of electronic data capture all the way down to the initial data capturing session in the field or laboratory. Mobile computing has been potentiated with the advent of faster, cheaper low-power processors and more robust wireless data transmission technologies. The technical limitations of mobile devices are constantly changing and mobile internet access is now achieved in most large urban areas of the world, including those in developing countries.

The basic workflow for mobile data capture involves initial capture of data on the mobile electronic device followed by upload into a central data repository. Uploading data can be done by direct connection or by transmission of data over wireless networks (Wi-Fi or cell phone networks). In an ideal scenario, the data-capturing device also has capacity to act as a Global Positioning System receiver and thus generate data for the spatial location where data were entered (Vanden Eng et al., 2007). When a software application on the mobile device is directly compatible with a data management system/decision support system, the mobile device essentially becomes part of the system and the data collection executed on the device becomes part of the work flow where the user interfaces directly with the system (Figure 1). Additional relevant data, e.g., from subsequent laboratory diagnostic tests, visualizations, or analyses, later can be entered into the system through other means.

There is strong interest in using mobile data capture in public health including vector-borne disease surveillance; recent studies have evaluated the use of PDAs for data collection during household surveys for malaria or bed net use (Vanden Eng et al., 2007; Ahmed & Zerihun, 2010) and collection of data for suspected dengue patients in clinical studies in Nicaragua (Aviles et al., 2008). The potential for use of cell phones as mobile data capturing devices was evaluated for infectious disease surveillance in Peru (Johnson & Blazes, 2007) and for malaria surveillance and monitoring in Thailand (Meankaew et al., 2010).

### **3. A multi-disease data management system/decision support system for tropical vector-borne diseases**

The Innovative Vector Control Consortium (IVCC) recognized the potential for using information technologies to improve vector and disease control program performance and ultimately reduce the burden of tropical vector-borne diseases such as dengue and malaria (Hemingway et al., 2006). This resulted in an initiative that led to the development of the software package described herein: a multi-disease data management system/decision support system (hereafter, the system) with current capacity for dengue and malaria, and

with potential for addition of other important tropical vector-borne diseases such as Chagas disease, human African trypanosomiasis, leishmaniasis, lymphatic filariasis, and onchocerciasis (L. Eisen et al., 2011). To some extent, the system builds upon previous experience with development and implementation of data management systems for malaria in southern Africa (Booman et al., 2000, 2003; Martin et al., 2002; Marlice Coleman et al., 2008). Key goals for the system include: (1) ensuring that it can be distributed without the user incurring licensing costs, (2) producing a flexible system that can be adapted to local circumstances by the user with no or minimal involvement of software developers, (3) achieving a user-friendly system with capacity to support data entry, storage, and querying, and production of maps and reports, as well as including decision support functionalities such as custom calculations and automated alerts, and (4) delivering a system capable of enhancing the user's ability to carry out surveillance, engage in evidence-based decision making, monitor interventions, and evaluate control program performance.

### **3.1 System development, architecture, requirements, installation, and licensing**

The system was developed in an iterative process with close contact between developers and subject matter experts including operational field input from Malawi, Mexico, Mozambique, South Africa, and Zambia public health partners. System functionalities were assessed by positive and negative testing by an internal testing team. Additional testing needs to include pilot implementations in different operational settings with naïve users.

The system was developed with a 3-tiered architecture (data tier–application/business logic tier–presentation tier) and is comprised exclusively of software components which can be distributed without licensing costs. The data tier includes a PostgreSQL database (<http://www.postgresql.org/>) enhanced with PostGIS (<http://postgis.refractor.net/>) to support spatial data. The business logic tier includes Java (<http://www.java.com/en/>) and Apache Tomcat (<http://tomcat.apache.org/>). The presentation tier uses Firefox (<http://www.mozilla.com/en-US/firefox/ie.html>), and is augmented by applications to support production of reports, BIRT (<http://www.eclipse.org/birt/phoenix/>), and maps, GeoServer (<http://geoserver.org/display/GEOS/Welcome>) and OpenLayers (<http://openlayers.org/>). To link the tiers, and to allow the user to make changes that automatically are reflected across the 3-tiered architecture, the system includes TerraFrame's Runway SDK application (<http://www.terraframe.com/products/runwaysdk>).

The system requires a minimum of 2 GB RAM, 100 GB storage, and 2.0 GHz CPU to operate on a stand-alone machine (projected cost of \$500-600 per stand-alone desktop in 2011). The target operating system is Microsoft Windows XP (Microsoft Corporation, Redmond, Washington, U.S.A.) but the system has been informally tested on and shown to function also for Windows Vista, Windows 7, Apple Mac OSX (Apple Inc., Cupertino, California, U.S.A.), and Ubuntu Desktop (<http://www.ubuntu.com/>). The system can be implemented on a single stand-alone computer, in a client/server environment where the client (user) logs in remotely to the system, or as a pool of installations with one master and one or multiple dependent installations (the system includes synchronization functionality).

The system installation package includes the system itself, the system manual, and stand-alone versions of OpenOffice (<http://www.openoffice.org/>), the reporting tool BIRT, to allow the user to create customized report templates which then can be imported into the

system, FWtools (<http://fwtools.maptools.org/>) to assist the user in transforming spatial data into well known text (WKT) format, and QCal which calculates dose/time response curves for insecticide resistance bioassays (<http://sourceforge.net/projects/irmaproj/>). Distribution and licensing is currently executed through the IVCC (<http://www.ivcc.com/>). There is no cost associated with the license (the system is a royalty free, licensed software).

### 3.2 Adaptability of the system

One key goal was to produce a flexible system that can be adapted to local circumstances by the user with no or minimal involvement of software developers. Key points of system adaptability are described below.

The system currently handles dengue and malaria. Selection of the disease in which to work is done through a menu item called Disease. Selecting a disease of interest in the Disease menu results in the user being presented with a default menu for the selected disease. This default menu can be re-configured by the user, including: (1) incorporation of functionalities that are present in the system but not as a default for the disease of interest and (2) changes to menu label names. This provides an economy of scope as new diseases downstream can be added to the system at decreased cost through re-use of already existing functionalities.

The system also includes the capacity to customize user roles and their permissions. It is delivered with a set of default roles but these are completely configurable in that the user can change the names of existing roles, create new roles, and, importantly, define separate permissions for each role to access or work with different system functionalities (Write, Read, or None/No access). The permissions can be further refined by generating individual log-in names and passwords for each person using the system and then assigning one or multiple roles to a given individual. This helps to restrict access to sensitive information such as data for individual patients. Furthermore, the system allows the user to define, by disease of interest, the status of many data entry fields, i.e., whether they are mandatory or non-mandatory, and also to select, by disease of interest and role, whether to show or hide a given data entry field (Figure 2). Exceptions are fields which are system mandatory and thus cannot be made non-mandatory or hidden. A data entry field also may be given different display label names in the different disease menus. This is accomplished with a localization functionality which also can be used to develop display labels, by disease, for a language other than the default English (i.e., languages or dialects based on the Latin character set).

The system includes three user-configurable information trees: (1) a controlled vocabulary term tree, (2) a universal tree for key spatial concepts, and (3) a geographical entity tree, where each entity is an instance of a universal (e.g., the geo entity United States is an instance of the universal Country).

The term tree, based on ontological principles following the Open Biomedical Ontologies (Smith et al., 2007), is used in the system to define options in pop-up select lists for data entry fields and for pre-configured entries for rows and/or columns in data entry tables (Figure 3B). An ontology is a set of standardized and logically defined terms (controlled vocabulary) and their inter-relationships. The controlled vocabulary term tree is built on a single ontological relationship, *is\_a*; for example virus isolation *is\_a* laboratory test for dengue virus (Figure 3A).

The system is delivered with a default term tree and each data entry field or row/column configuration in a data entry table that is populated from the term tree has a pre-configured root term (Figure 2) that defines what is included in the select list for the data entry field or which terms that are used to define table rows/columns. Both the term tree itself and the selection of root terms are completely configurable by the user, including the ability to make terms active or inactive by disease (Figure 3A). The term tree has multiple benefits including

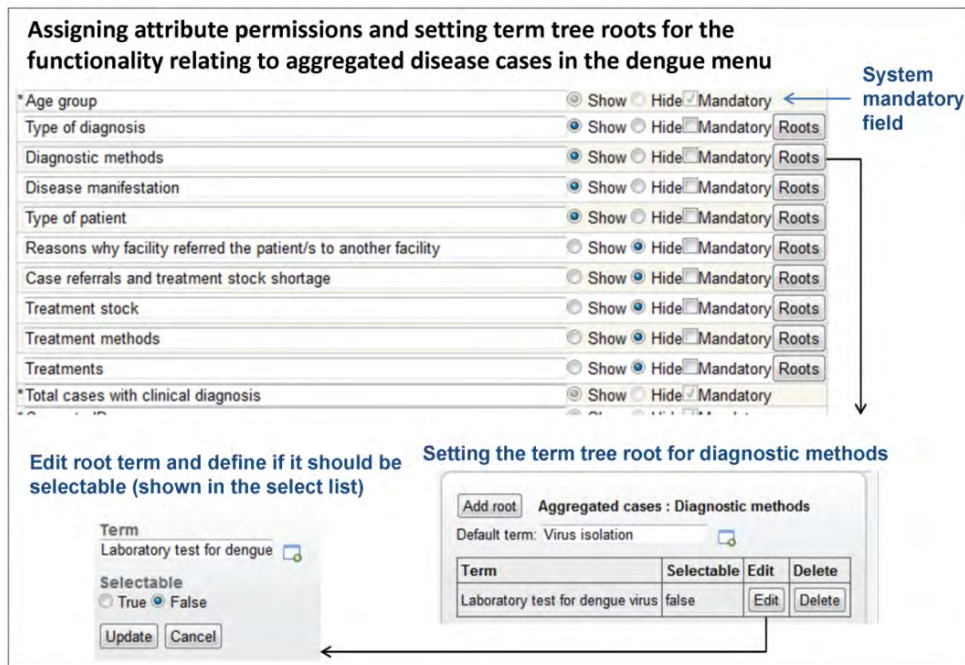


Fig. 2. Administration functionality to assign attribute permissions and set term tree roots. Adapted from a figure presented previously by L. Eisen et al. (2011).

the use of standardized terms. Each term is given a name, display label, and ID, and the user also can provide a definition for the term (Figure 3A). Use of standardized terms provides potential for sharing of data with related database or ontology initiatives. The system's term tree includes terms related to insecticide resistance that were derived from the Mosquito Insecticide Resistance Ontology and are used in the IRbase global database for insecticide resistance in mosquito vectors (Dialynas et al., 2009). It also includes terms drawn from the malaria ontology IDOMAL (Topalis et al., 2010). The term tree provides capacity to dynamically change both the number of rows/columns that appear in a data entry table and their respective header labels (Figure 3B) simply by making changes to the term tree content under the root term that is used to populate that specific data entry table. This type of dynamic data entry table provides exceptional potential for system adaptation to local conditions without the involvement of software developers. Finally, based on the is\_a ontological relationship of the term tree, data can be aggregated to higher levels of the term tree in the system's internal data querying tools.

Universals, as used in this system, are key spatial concepts that can be defined with regards to how they are used to support system functionalities. A small set of universals are required for specific system functionalities (health facility, collection site, sentinel site, spray zone, stock depot, and surface) and these typically will be complemented in the system by a set of user-created universals, such as country, state, county, settlement, etc.

### A) Term tree structure

### B) Using the term tree in data entry screens

**Data entry field with pop-up select list: Laboratory test for dengue virus**

**Data entry table with pre-populated rows: Laboratory test for dengue virus**

Diagnostic methods used		
Diagnostic methods	Total tests	Total positive tests
IgG antibody capture ELISA		
IgM antibody capture ELISA		
NS1 antigen capture ELISA		
RT-PCR assay		
Virus isolation		

Fig. 3. Partially opened term tree structure for dengue menu and use of configurable term tree lists in data entry screens. A) Example of term tree section leading to laboratory tests for dengue virus. B) Using the term tree section for laboratory tests for dengue virus in data entry fields and tables. Adapted from a figure presented previously by L. Eisen et al. (2011).

The geographical entity tree provides a representation of the area in which the system is implemented and, for geo entities in the system which have spatial data associated with them, can be used for mapping. Geo entities have the properties of universal type, entity status (active/inactive), name and ID, type of geometry (point, polygon, etc.), and spatial data (in WKT format). Each entity is related to other entities by means of a located\_in relationship defined by the total inclusion of one entity inside another: for example Colorado is located in the United States. The system default is a single root entity called Earth under which the user can build a locally relevant geographical entity tree. Once this is configured, the user can add or delete geo entities and edit the information for existing ones. One benefit of the geographical entity tree, derived from its located\_in relationship



structure, is the potential for aggregating data to coarser spatial resolutions in the system's internal data querying tools.

### 3.3 Data entry, data query, and reporting/mapping

To minimize data entry error, data entry fields make extensive use of hard-coded select lists or radio buttons, geo entities selected from the geographical entity tree, dates selected from pop-up calendars, and terms selected from pop-up select lists from the term tree. Data querying is done through a set of unique system tools referred to as query builders, linked to specific data input screens, where the user can define a specific data query (Figure 4).

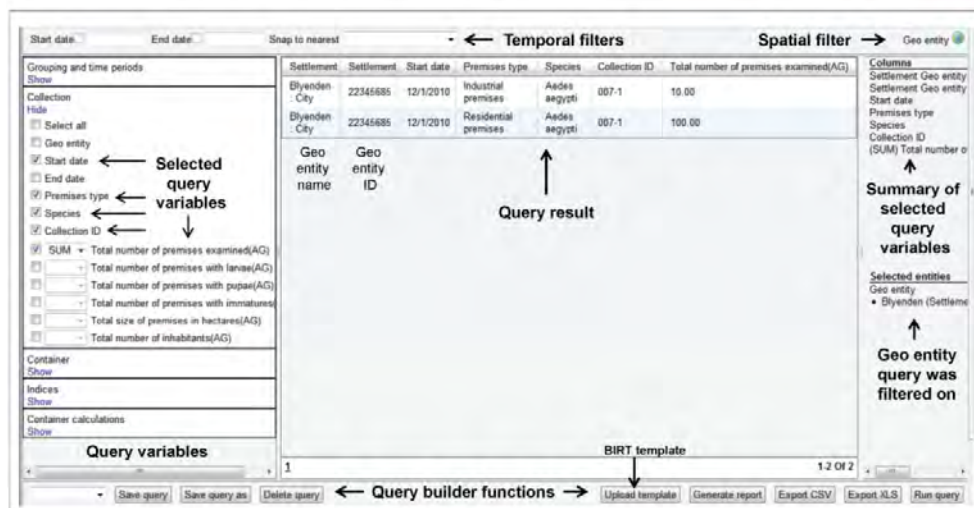


Fig. 4. Example of query builder screen. Adapted from a figure presented previously by L. Eisen et al. (2011).

All query builders include the capacity to filter a query on start and end dates and geo entities from the geographical entity tree (top pane in query builder; Figure 4). Additional filtering of a query can be done on specific variable fields (left pane in query builder, Figure 4) corresponding to the data entry fields in the relevant data input screens; this can include terms from a term tree root, values from hard-coded select lists or numerical values or ranges. Many of the query builders also include pre-defined custom calculations (section 3.4.1). The query builders also include options to export query results as .csv or .xls files, to save and re-use specific querying field combinations that are executed on a regular basis, and to upload pre-configured report templates (from BIRT) and use these to produce standardized reports (bottom pane of query builder; Figure 4). Mapping is directly linked to the query builders as the system's map generation process makes use of information that is saved in the query builders as specific named query results. Maps can combine data that are generated through different query builders, e.g., for intervention coverages and disease case locations or disease incidence, and overlaid on a map base layer showing locations of households, administrative boundaries, etc. (Figure 5).

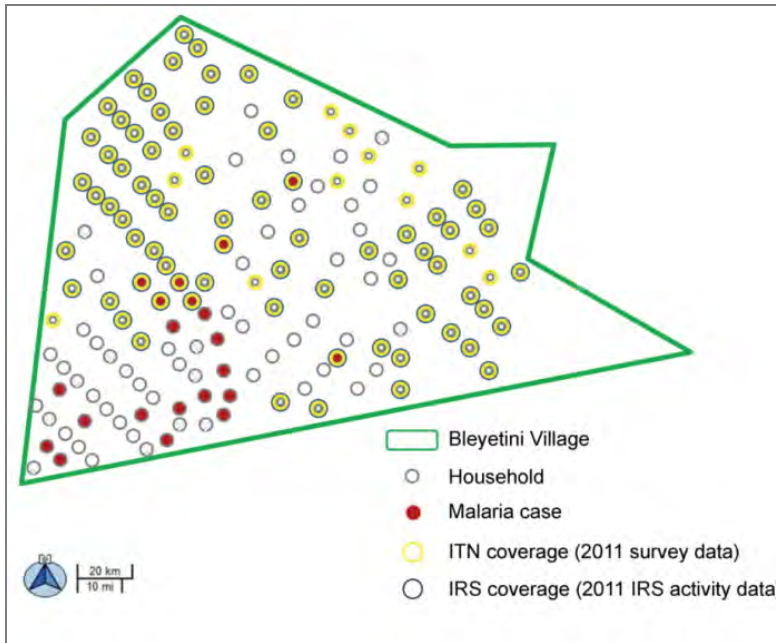


Fig. 5. Example of map output. The base layer shows locations of households within a village and this is overlaid with data for disease cases and coverage by control interventions including insecticide-treated nets (ITN) and indoor residual spraying (IRS).

### 3.4 Examples of system decision support functionalities

One key goal was to produce a system capable of enhancing the user's ability to carry out continuous surveillance, engage in evidence-based decision making, monitor interventions, and evaluate control program performance. The following sections provide examples relevant to dengue control programs of system decision support functionalities, including pre-defined custom calculations relating to specific surveillance or control parameters and automated alerts when system thresholds for key entomological or epidemiological risk measures are reached.

#### 3.4.1 Custom calculations in query builders

The system provides decision support through pre-defined custom calculations that are included in query builders and address issues of operational relevance for a vector/disease control program. Examples of custom calculations relevant to dengue are provided below.

Entomological surveillance relating to dengue is peculiar in that the immature stages (larvae and pupae) of key dengue virus vectors, especially *Aedes aegypti*, exploit a wide range of containers (e.g., water storage containers, tires, bottles, cans, etc.) that accumulate in the peridomestic environment as development sites (Focks, 2003). This has led to a strong focus in vector and dengue control on reducing availability of containers, especially the container types that locally are most productive for immatures, through environmental sanitation to

remove trash containers and treatment of necessary containers with biological or chemical control agents to kill immatures (World Health Organization, 2009). The system therefore includes specific functionalities for collection of data on immatures from containers. Local variability of important container types is addressed by user definition of a locally relevant set of container types in the term tree for use in data entry screens (as term tree-driven pop-up select lists or prepopulated rows in data entry tables) and corresponding query builders. To enhance the capacity of the system to support decisions regarding the need for vector control actions, the specific query builder which handles immatures by container type includes pre-defined custom calculations for commonly used immature abundance indices that are used in operational settings to determine whether or not control actions should be executed. Additional custom calculations are included to help the user determine which container types contribute the most to production of mosquito vector immatures and thus are especially important to target for control.

Other examples of pre-defined custom calculations in query builders in the system's dengue menu that provide information directly useful for making operational decisions regarding surveillance and control activities include: (1) incidence of disease cases and case fatality rate in the case surveillance query builders and (2) percentages of available and visited premises within a given geographical area and time period for which prevention or control activities were carried out (this can also be broken down by prevention/control method, as defined by the user in the term tree), and percentage of visited premises that were not treated (this can also be broken down by reason for non-treatment, as defined by the user in the term tree), in the intervention monitoring query builder. The intervention monitoring query builder also illustrates the value of the term tree in the system. The user can specify, through the section of the term tree that is used to dynamically create the columns (and their labels) in the data entry table for prevention or control activities that data are to be collected against, any combination of different prevention or control methods and then use the intervention monitoring query builder to produce a breakdown by method for their spatial coverages in a given area during a specific time period.

### **3.4.2 Automated alerts**

Automated alerts that are triggered when threshold values are reached are perhaps the clearest examples of decision support in the system. Alerts are currently included for: (1) abundance indices for container-inhabiting mosquito immatures and (2) disease cases. In both cases thresholds can be configured by the user to suit local conditions so that alerts are not excessive to the point of being meaningless due to lack of resources to respond to them. The system can provide alerts as on-screen pop-ups (Figure 6B) and/or e-mail notifications.

Thresholds for abundance indices for container-inhabiting mosquito immatures are set as fixed numbers and up to 13 different indices can be activated by entering threshold values against them in a configuration screen (Figure 6A). An alert that the threshold value was reached for a given index is triggered on entry of a data record for immatures by container type (defined by geo entity, time period, premises type, and mosquito species). When the user saves a data record, the system automatically calculates the 13 abundance indices and compares the results to the configured threshold values. Alerts are then provided for the indices where the threshold values were reached (Figure 6B).

**A) Configure threshold value(s)**

Display label	Value	
Breteau Index for immatures		<a href="#">Edit</a>
Breteau Index for larvae	5.00	<a href="#">Edit</a>
Breteau Index for pupae		<a href="#">Edit</a>
Container Index for immatures		<a href="#">Edit</a>
Container Index for larvae		<a href="#">Edit</a>
Container Index for pupae		<a href="#">Edit</a>
House Index for immatures		<a href="#">Edit</a>
House Index for larvae		<a href="#">Edit</a>
House Index for pupae		<a href="#">Edit</a>
Pupae per hectare		<a href="#">Edit</a>
Pupae per person		<a href="#">Edit</a>
Pupae per premise		<a href="#">Edit</a>
Pupal Index		<a href="#">Edit</a>

**B) Alert that the threshold value for an index was reached or exceeded for a given data entry record**

Threshold for Breteau Index for larvae was exceeded in Blyenden (Settlement : City) - 22345685 for Residential premises (MDSS:0000135) and *Aedes aegypti* (MIRO:40000518) during the period from Wed Dec 01 00:00:00 PST 2010 to Wed Dec 01 00:00:00 PST 2010.

[Export](#)

Fig. 6. Thresholds and automated alerts for abundance indices for container-inhabiting mosquito immatures. A) Threshold value configuration screen. B) Example of automated on-screen pop-up alert that the threshold value for a given index was broken for a given data entry record (defined by geo entity, time period, premises type, and mosquito species).

Setting alert threshold values for disease cases, which is done by disease and presumed source of infection (a geo entity representing an administrative boundary unit) or health facility (where the case was reported), is more complicated because there are several options for the user to configure the threshold value calculation. Thresholds can be calculated using different algorithms including mean + 1.5 SD, mean + 2 SD, modified binomial 95%, modified binomial 99%, and the upper third quartile (Marlize Coleman et al., 2008), and two of these can be included as separate threshold alert levels 1 and 2. Based on what type of historical data have been entered into the system, the threshold values can be calculated from data for aggregated cases, individual cases, or individual and aggregated cases combined. The system also allows the user to: (1) define the number of weeks preceding and following an epidemiological week that are included in the threshold value calculation for epidemiological week, (2) define the number of previous pathogen transmission seasons, which can be contained within a single year or span two years, for which to include data for the threshold value calculation, and (3) provide a weight for each pathogen transmission season to address the effect of outliers with unusually low or high disease case loads. To address scenarios where an initial clinical diagnosis often may prove incorrect and many clinically diagnosed patients do not return to provide convalescent samples for confirmatory tests of pathogen exposure, the system automatically calculates which percentage of clinically diagnosed cases should be included in the threshold value calculation based on historical data in the system for confirmed positive cases (for dengue; cases where laboratory tests for dengue virus or dengue virus exposure were positive) versus confirmed negative cases (for dengue; cases where laboratory tests for dengue virus and dengue virus

exposure were negative). Finally, the system provides the option to manually enter or edit threshold values, by geo entity, transmission season, and epidemiological week, for thresholds alert levels 1 and 2.

The system tracks a current case count, by disease, which is updated every time a new individual case is entered into the system. An alert is triggered on the case entry for which the threshold value is reached for a given disease in a given geo entity or health facility for a given time period (the system allows for selection of standard epidemiological week or a sliding week that includes cases occurring from six days prior to the date of onset of symptoms for the case that is entered into the system). If disease cases are entered into the system on a timely basis, these alerts can be used to facilitate outbreak response. To address the scenario outlined above where an initial clinical diagnosis often may prove incorrect (and laboratory confirmation is slow or often lacking), the alert functionality can be modified by the user defining which percentage of clinically diagnosed cases that should be used for calculation of the current case count. Because the percentage is set manually, it can be changed temporarily in response to extraordinary events such as outbreaks of other diseases resulting in increased levels of clinical misdiagnosis.

Notably, the development of tools for rapid detection of outbreaks/epidemics supports the long-term goal of global malaria eradication and the current drive for elimination in countries and regions of low malaria endemicity. Scaling up control efforts and improving surveillance practices are critical objectives for these undertakings to succeed. Sensitive tools to timely identify an unusual increase in disease cases followed by prompt outbreak response will support such efforts.

### **3.5 Range of the information handled in the system's dengue menu**

In addition to the generic system modules for administration and GIS, the latter of which includes the geographical entity tree and the functionality for map generation, the system's dengue menu includes modules dealing with case surveillance, entomological surveillance, intervention planning, intervention monitoring, and stock control. Case surveillance includes separate functional components for: (1) aggregated disease case data and (2) data for individual disease cases. Aggregated disease case data are captured by time period, geo entity or health facility, and age group. The core of the data entry functionality is comprised of entries for numbers of cases with clinical diagnosis, confirmed positive diagnosis, and confirmed negative diagnosis together with a series of data entry tables relating to type of diagnosis, disease manifestation, and type of patient where both the rows (e.g., specific disease manifestations such as dengue fever and dengue hemorrhagic fever) and the columns (e.g., providing breakdowns by gender, dengue virus serotype, and locally acquired versus imported cases) are generated dynamically from the term tree and thus can be readily configured by the user to be relevant in the local setting. Data entry for individual disease cases include basic patient data as well as symptomology, laboratory diagnostic data, and administrative information such as whether a case was detected through passive or active surveillance and, in the case of passively detected cases, data for the health facility and attending physician.

The module for entomological surveillance includes functional components which can handle data relating to non-container based collections of mosquito vectors, for example collection of adults by traps or active collection with aspirators, as well as container-based surveillance data for immatures which was mentioned previously. In the latter case, the

system includes separate data entry screens for data collected by individual container versus data collapsed to user-defined container types. This is complemented by functionalities relating to capture of data for assays conducted on mosquito collections, including assays for pathogen detection, assays to determine killing efficacy of insecticides, and insecticide resistance bioassays, biochemical assays, and molecular genetic assays.

Intervention planning is restricted to a planning calculator tool which helps the user determine the man-power or amount of insecticide product needed to complete an intervention. Intervention monitoring deals with coverage, in space and time, of different types of control interventions. This includes functionalities to handle data relating to person-days and amount of insecticide product used for the intervention as well as data for intervention coverage, by user-configured control methods, collected on a premises-to-premises basis or aggregated to larger spatial units such as blocks or neighborhoods. Finally, the system's stock control module helps the user to track stock levels in different storage locations and also to track cost of stock. Locally relevant stock items are configured by the user in the term tree.

### **3.6 System limitations and plans for future improvements**

The system's potential for adaptation by the user to local circumstances without the involvement of software developers results in increased complexity of the system, most notably for the system administrator when it is first installed and configured. The system includes extensive capacity for data import, including import spreadsheets that are tailored to specific functionalities. However, the import process is, for data quality purposes, unforgiving when it comes to poor quality data. Furthermore, because the system lacks capacity for mass-deletion of data, import of large amounts of data must be considered very carefully to avoid time-consuming data deletion exercises. The system was developed to support operational control programs and therefore has very limited statistical and spatial analysis capacity. Statistical operations that are directly supported are restricted to: (1) query builder calculations of sums, averages, and minimum and maximum values and (2) pre-defined query builder custom calculations that relate to specific system functionalities, such as disease case incidence or vector abundance indices. Other statistical operations require the user to export data for subsequent import into a statistical software package. Finally, the system supports basic mapping functions but essentially lacks spatial analysis capacity. To achieve this, the user needs to export a shapefile from the system for subsequent import into GIS software with spatial analysis capacity.

Plans for future system improvements include: (1) making the system directly compatible with hand-held mobile data capturing devices, especially PDAs and smart phones, (2) developing an over-arching query builder to make it easier to combine data from different parts of the system, (3) developing additional user-configurable functionalities such as configurable surveys, and (4) expanding the system to include other important vector-borne diseases in addition to dengue and malaria.

## **4. The California Vectorborne Disease Surveillance Gateway: A data management system/decision support system for vector control programs and public health agencies**

California's arbovirus surveillance program is a collaborative project between the network of vector control agencies that make up the Mosquito and Vector Control Association of

California, the California Department of Public Health, and the University of California at Davis (UC Davis). These agencies are autonomous but are linked by a cooperative agreement that defines their respective roles in vector-borne disease surveillance, and together, they are charged with surveillance and control of mosquitoes and arboviruses to protect the health of California's human population of more than 38 million (~12% of U.S. population).

In recent decades, record-keeping has graduated from paper-only systems to electronic data storage, initially weekly spreadsheets -- essentially electronic versions of paper forms -- and presently relational databases that link data over long time periods and permit efficient data storage, data analysis, and reporting. The transition was facilitated by National Oceanic and Atmospheric Administration-funded projects that supported the conversion of California's paper-based surveillance records from the last 50 years to electronic form, now stored in databases maintained at UC Davis. Once this centralized data repository was established, there was a need to provide participating agencies with access to the data and to provide mechanisms for ongoing data input and retrieval. In 2006, the first step was taken toward these goals with the launch of the first version of the California Vectorborne Disease Surveillance Gateway (hereafter, the Gateway). The Gateway was designed for three overarching purposes: (1) to provide a user-friendly system for the storage of surveillance data, (2) to facilitate more efficient data exchange among the collaborating agencies, including the diagnostic laboratories, and (3) to provide tools for analysis and visualization of surveillance results and the calculation of risk estimates to support control decisions.

Vector control in California is conducted within local districts, and each agency is independently operated and funded by local property taxes. Budgets vary widely, from a single salary and control supplies in rural areas to multi-million dollar budgets in urban areas with more complex control needs. These differences in budgets, work forces, and control strategies result in a diverse array of data management solutions among agencies, and because of the autonomy of the agencies, implementation of a centralized software solution for statewide data management hinges on its ability to accommodate their needs and provide tools that enhance their surveillance and control activities. The system described below has been in place as the software decision support tool used for mosquito control in California since 2006 and has combined large-scale monitoring of encephalitis virus enzootic amplification with immediate reporting, analysis, and visualization.

#### **4.1 System architecture, requirements, and licensing**

The Gateway is designed for use by local vector control agencies to store, manage, and analyze data collected through their surveillance activities. Agencies are the primary administrative units, and the Gateway was designed to scale from an individual agency to groups of collaborating local agencies and to accommodate the needs of a hierarchy of agencies at many levels. Here, we focus on California's implementation of the system, where it is used to coordinate and streamline surveillance activities and provide summaries and calculations to be utilized by the local agencies, as well as other state and national public health agencies, such as the California Department of Public Health and the United States Centers for Disease Control and Prevention.

The system was written using a modified Model-View-Controller (MVC) architecture to facilitate rapid development and customization by the local user. All components of the

system are readily available at no cost, can be adapted for specific use, and nearly all are open-source. Data are stored in a PostGIS-enabled PostgreSQL database while server-side code is written in PHP (<http://www.php.net/>). HTML templates using jQuery (<http://jquery.com/>) and Google Maps (<http://maps.google.com/>) comprise the client code and are accessible through standard web browsers. Use of Google Maps requires compliance with the Google Maps Terms of Service.

The minimum hardware requirements for running the system on a single computer are 1 GB RAM, 40 GB storage, and 1.4 GHz CPU. Though the system can be installed and operated on a single computer, it is strongly recommended that the system be installed in a client-server environment with the database and server running in a UNIX-like environment, although installation on Windows is possible. On the client side, access to the Gateway occurs through a web-based interface. Software installation is not required, and any operating system can be used, as long as the internet browser is current and supports established web development standards. Operating systems tested include Microsoft Windows XP and 7, Apple Mac OSX, and Ubuntu Desktop. Web browsers tested include Microsoft Internet Explorer 8 (<http://www.microsoft.com/windows/internet-explorer/default.aspx>), Mozilla Firefox 3.6, Apple Safari 5 (<http://www.apple.com/safari/>), and Google Chrome 8 (<http://www.google.com/chrome/>).

The Gateway system is available as a compressed file archive and can be obtained by contacting the administrator of the CalSurv website (<http://www.calsurv.org>) or the UC Davis Center for Vectorborne Diseases' website (<http://cvcc.ucdavis.edu>). The system is licensed under the GNU General Public License version 3 and has no costs associated with distribution and use of the software.

#### **4.2 System security and extensibility**

Data integrity is ensured in the Gateway by having all records include identifying information of the user who created or modified the information. A background audit log tracks every record added or changed in the system. To eliminate the possibility of accidentally deleting important data, the Gateway has no ability to completely remove information but can remove a record from view, which also excludes the record from reporting and analysis. With the identification information stored with each Gateway record, end users are restricted to only the records of their agency, thereby ensuring an agency's privacy. For California's instance of the Gateway, the system is housed at the UC Davis Center for Vectorborne Diseases in a climate-controlled, locked server room protected by several layered firewalls that regulate network traffic and, in addition to the university network structure, prevent unwanted access. Temperature and humidity are monitored remotely, with thresholds set to trigger notification e-mails to a system administrator if conditions exceed allowable tolerances. A back-up server with a RAID 5 drive array and a LTO5 tape drive provides short- and long-term storage for all Gateway data. Back-ups occur hourly, daily, weekly, monthly, and yearly to the RAID 5 array and tapes. Monthly and yearly tapes are stored offsite.

Agencies are the principal administrative unit for managing users' access to the Gateway, and each agency assigns permissions to its users that regulate their ability to view, add, or modify data. Each agency designates at least one agency manager, which is a senior person who manages privileges for all other users within the agency. Users receive permission to



access the agency's data from their manager, and privileges are based on the user's category, which is one of the following: view-only, diagnostic, user, or agency manager.

The Gateway's architecture (i.e., code and database schema) is not confined to a particular structure, and the system is designed to be broadly applicable for vector surveillance and control programs in any location. Users of the system can be assigned to one or more agencies, with no limits on the number of agencies or assigned users. Agencies also may share data with other agencies (e.g., their neighbors) by providing them with a "user" or "view-only" account. Other extensible features include arthropod names that are structured according to taxonomic rank from phylum through species levels, and the system is extensible to include new taxa as required. Spatial data are stored using a consistent, global spatial reference system (WGS 1984), and these data can be transformed using PostGIS into any other spatial reference system as needed for individual applications.

### 4.3 Data input and output

Two mechanisms are provided for entering data into the Gateway -- direct record-by-record entry through a web-based front-end, or bulk import of record sets after entry into a local data management system. The first and most common method for entering data is direct input using the Gateway's data entry forms. Forms are provided for field data (surveillance site locations, arthropod collections, sentinel chicken samples, and dead bird reports) and results of laboratory testing for arboviruses (mosquitoes, sentinel chickens, and dead birds). Accuracy of entry is maximized by the use of drop-down menus, pop-up calendars, and messages or mouse-over tips clarifying the meanings of individual fields.

Keyboard-only data entry is also supported, and the Gateway suggests possible values as the user types part of a field's value, with the range of possibilities narrowing with each keystroke. Once the desired value has been identified, the user can tab to the next field. Samples of mosquitoes (i.e., mosquito pools) to be submitted for virus testing can be entered directly below the collection data, which links the pools with the collection information and avoids redundant entry (Figure 7). Later, diagnostic test results for mosquito pools and sentinel chickens can be added directly without re-entry of collection information. Many of California's vector control agencies use the Gateway as their primary means for data management, but a number of other agencies also maintain in-house data solutions that range from generic spreadsheet or database software to fully customized programs. To avoid redundant data entry, the Gateway provides mechanisms for bulk data import from these systems. Data that can be imported include surveillance site locations, field data on sentinel animals and arthropod collections or pools, and laboratory test results. Structures and plain-text formats are specified for each data type, along with a sample data set. Once the data are in the appropriate format, the user simply selects the type of data and the file to be imported, and a preview of the data is shown before they are uploaded into the Gateway.

For users who prefer storing a copy of their data locally or want to import data back into their local data management systems, the Gateway offers the ability to export any data set for download. Before export, the data may be filtered by criteria relevant for the data type, such as agency, date range, site list, mosquito trap type, or mosquito species. Several formats can be selected for the export file, such as OpenDocument spreadsheet (<http://opendocument.xml.org/>), Microsoft Excel, plain text, or HTML.

**New Collection**

\*Trap Type

\*Collection Date  Trap started the evening of 06/24/2010.

\*# of Traps

\*# of Nights/Trap

\*Trap(s) at Site

**Latitude**

If the Latitude is in the southern hemisphere, the value must be prefixed by the minus '-' sign.

**Longitude**

If the Longitude is in the western hemisphere, the value must be prefixed by the minus '-' sign.

**Identified By**

Trap(s) ran with no problems?

**Comments**

**Count of Species**

	Species	Males	Females - Mixed	Females - Unfed	Females - Bloodfed	Females - Gravid	Total Females	Unknown Sex
1	Culex tarsalis			73	1		74	
2	Culex pipiens	2	32				32	
3							0	
		2	32	73	1	0	106	0

▼ Pools

**New Pools**

	Pool #	Other Sites	Species	Sex/Condition	# in Pool	Comments
1	145		Culex tarsalis	Females - Mixed	50	...
2	146		Culex tarsalis	Females - Mixed	24	...
3	147		Culex pipiens	Females - Mixed	32	...
4	148				50	...

Fig. 7. Example showing the Gateway’s data input screen for mosquito collections, including entry of mosquito pools associated with the collection that are to be tested for arboviruses in the lower pane.

#### 4.4 Examples of Gateway functionality

In the following sections, we describe several of the Gateway’s features that add value to the underlying data sets and provide decision support for vector control agencies.

##### 4.4.1 Calculators for mosquito abundance and arboviral prevalence

The Gateway provides two calculators that are used to assess the current year’s (or any other selected year’s) mosquito abundance and virus activity (Figure 8). Each calculator has filtering options for agency, time interval, and site groupings, including the spatial features discussed below in section 4.4.3. Other choices include mosquito species and sex, mosquito

trap type, and virus, depending on the calculator. If multiple selections are made for the filters, users are provided an option to treat each of the choices individually in the calculations, which makes calculations for several species or spatial features quite easy. In addition to the requested time period, averages are calculated for comparison of mosquito abundance to the prior 5-year period. Mosquito infection prevalence and 95% confidence intervals are estimated using maximum likelihood estimate methods (Biggerstaff, 2006). After running any of the calculations, the results can be downloaded to the user’s computer or graphed as shown in Figure 8.

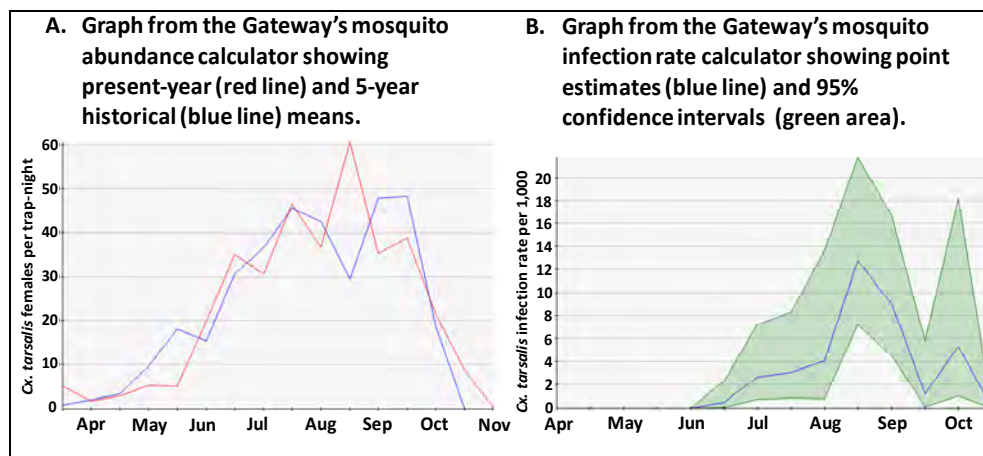


Fig. 8. Output graphs, aggregated by half-month, from the Gateway’s calculators. A) Female mosquitoes per trap-night. B) Infection rate per 1,000 tested mosquitoes.

#### 4.4.2 Risk of West Nile virus (WNV) transmission to humans

The California Department of Public Health publishes the California Mosquito-borne Virus Surveillance and Response Plan (California Department of Public Health et al., 2010) to provide guidance for vector control and public health agencies and to specify appropriate interventions if the risk for transmission of WNV to humans escalates. The plan includes an assessment of arboviral transmission risk based on climate and the components of enzootic surveillance systems, including mosquito abundance, WNV infection prevalence, avian serology, and diagnostic testing of dead birds (Barker et al., 2010). Each factor is assigned a risk value from 1-5 and all available factors are averaged to obtain an overall risk level.

The Gateway automates these risk calculations, taking advantage of the stored historical surveillance data for estimation of baseline mosquito abundance and temperature data from the NASA Terrestrial Observation and Prediction System (TOPS) (Jolly et al., 2005; Nemani et al., 2007) available through a collaboration with Ames Research Center. Calculations are scripted to run automatically at the conclusion of each half-month during the surveillance season and graphs are automatically e-mailed as PDF files to agency managers throughout California showing risk estimates overall and for individual surveillance factors (Figure 9).

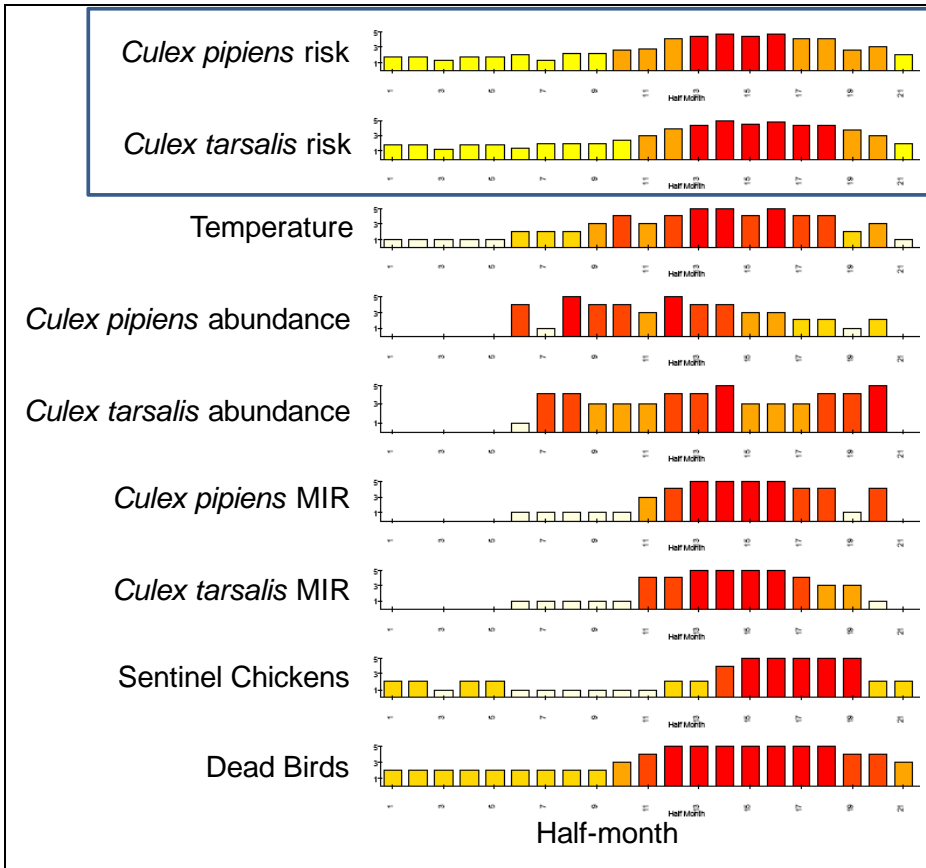


Fig. 9. Example from a PDF file showing an agency's calculated risk for each half-month. The top two graphs indicate overall risk based on the two key mosquito vectors in California, and remaining graphs show calculated risk for individual surveillance factors.

#### 4.4.3 Spatial features

Perhaps the most exciting new capability of Gateway 2.0 (launched in early 2010) has been the addition of spatial features. Most aspects of surveillance are inherently spatial, and calculations (section 4.4.1) are frequently aggregated over spatial units. In previous versions of the Gateway, calculations at the sub-agency level had been limited to the use of site groups, which required tedious assignment of surveillance site codes to groups that could be used for calculations. Site groups remain a flexible option, but new spatial features make this process easier by capturing sites that fall within a user-defined area without identifying the site codes a priori.

Spatial features are collections of one or more polygons of any shape that are defined by a user and shared (or not) with other users in the agency. The features are created through a graphical interface that allows users to point and click with the mouse to define one or more polygons for each feature (Figure 10) or by import of an ESRI (Redlands, California, U.S.A.)

shapefile that defines the features. After features have been created and saved, they appear along with individual sites, site groups, and agencies as choices for spatial filtering in all of the Gateway’s calculators. Each agency’s boundary is provided as a spatial feature by default.

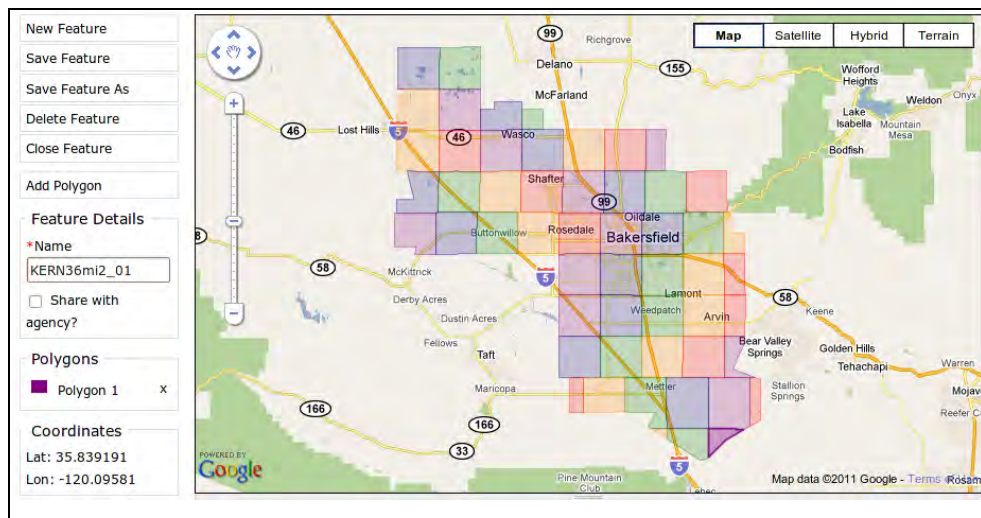


Fig. 10. Gateway editor showing spatial features that correspond to townships defined by the Public Land Survey System within the Kern Mosquito and Vector Control District.

#### 4.5 System limitations

The Gateway is designed primarily for use by one or more agencies in an environment where the software and underlying database would reside on a central server with clients utilizing web-based interfaces through an internet browser. This results in a requirement for a reliable network connection, which can be a limitation in some environments. Benefits of this centralization are that corrections can be made or new features added quickly, without the need for periodic distribution and installation of software by the system’s users. Also, operating system compatibility is maximized by the web-based interface. However, this centralization also limits the tools available for data input, output, and analysis to those created by the developer(s) with access to the central server, and there is no facility available for development of tools by end users. As a result, the success of the system for a particular implementation and its value to users will rely on developers being responsive to users’ needs for a given system. The Gateway’s data export mechanisms also allow users to download their agency’s data for additional analysis using other software.

The Gateway’s calculators of mosquito abundance and arbovirus prevalence are highly flexible in terms of their options for aggregating and filtering data according to space, time, or other attributes, such as mosquito species or sex. Each filtering option also typically allows for selecting multiple choices and treating each choice individually for calculations. If several of these grouping variables are applied simultaneously, this can result in an extreme number of dimensions for calculations, which would require a large amount of compute time. To address this potential problem, Gateway calculators have hard-coded limits on the

number of dimensions for calculations, and if calculations exceed a pre-specified threshold for total compute time, an error message is returned on the user's screen, and a message is sent to the system administrator. Hard-coded limits and notification triggers may be set by the system administrator as needed for a particular implementation of the Gateway.

#### **4.6 Future of the Gateway**

Several important improvements are planned for the Gateway, including enhanced interactive mapping, management of pesticide use and resistance data, and development of an interactive risk calculator and forecasting module to complement the existing tools. Currently, the Gateway provides interactive maps that are updated in real-time as new data are added to the Gateway. These are run on ArcIMS (ESRI) and Geocortex IMF (Geocortex, Victoria, British Columbia, Canada), but this system is being transitioned to a Google Maps-based system later in 2011. The key advantages to the new system will be: (1) free software that eliminates licensing costs, (2) automatically updated base layers maintained by Google, and (3) increased freedom to customize the user interface for queries and other interactions with the underlying PostgreSQL/PostGIS databases.

In addition to surveillance, vector control agencies must carefully track and report their use of pesticides. This aspect of control programs is receiving increased emphasis due to new regulations aimed at preventing pollutant discharge into waterways, and user surveys are currently underway to assess the data management needs of California's vector control agencies in this important area. Depending on the results of the surveys, new modules will be added to the Gateway to collect information on pesticide use and resistance that is consistent with the requirements of the United States Environmental Protection Agency and state or county-level agencies, such as the California Department of Pesticide Regulation.

The existing tools on the Gateway support vector control decisions by adding value to California's surveillance data. To improve assessments of concurrent arboviral transmission risk (section 4.4.2) as the surveillance season progresses, a new calculator will be added to utilize the aggregation and filtering options from existing calculators and provide finer-grained calculation of arboviral transmission risk at user-specified spatial and temporal scales. Better lead-time is also needed for planning interventions, and we are currently working to extend historical datasets using hierarchical Bayesian models for forecasting mosquito abundance and arbovirus transmission. These models explicitly account for spatial and temporal structure in the data, and climate and land cover predictors have been evaluated for the varied ecological regions of California. Once validation is complete, these will be used to provide seasons-in-advance regional forecasts via the Gateway.

### **5. Conclusions**

Emerging information technologies present new opportunities to reduce the burden of vector-borne diseases through improved decision support for surveillance and control of vectors and their associated pathogens. Such technologies range from stand-alone mapping applications, for example Google Earth, to integrated data management system/decision support system software packages, such as those described in sections 3-4, and mobile data capturing devices. Incorporation of novel technology solutions into operational control programs will lead to improved data management and analysis capacity and thus provide knowledge to support evidence-based decision-making that enhances the control program's ability to carry out surveillance, monitor interventions, and improve program performance

including the targeting of limited intervention resources where they would have the biggest impact. The rapid growth in software applications which can be used without licensing costs is now setting the stage for development of integrated data management system/decision support system software packages that can be readily distributed to resource-poor environments and used to reduce the terrible burden of tropical vector-borne diseases.

## 6. Acknowledgments

Funding for the multi-disease system for tropical vector-borne diseases was provided (to LE, MC, MC, and SLF) by the Innovative Vector Control Consortium and the United States National Institutes of Allergy and Infectious Diseases (1R03AI083254-01). Personnel from many organizations contributed, including the system development and testing teams at Colorado State University, the Innovative Vector Control Consortium/Liverpool School of Tropical Medicine, the Medical Research Council of South Africa, Akros Research, and TerraFrame, Inc. We also thank field site partners in Mexico and Africa for their invaluable support: Universidad Autónoma de Yucatán, México, Servicios de Salud de Yucatán, México, the Malaria Alert Centre, Malawi, the National Malaria Control Programme, Mozambique, and the National Malaria Control Centre, Zambia.

Funding for the Gateway was provided (to CMB, BP, and WKR) by NASA Decision Support through Earth-Sun Science Research Results (RM08-6044 for NNA06CN02A) and the United States Centers for Disease Control and Prevention's program on Climate Change: Environmental Impact on Human Health (5U01EH000418). CMB and WKR also acknowledge financial support from the Research and Policy for Infectious Disease Dynamics (RAPIDD) program of the Science and Technology Directorate, United States Department of Homeland Security, and the Fogarty International Center, United States National Institutes of Health. Important contributions were made by the member agencies of the Mosquito and Vector Control Association of California and the Vector-borne Disease Section of the California Department of Public Health, and by Forrest Melton and Andrew Michaelis at California State University, Monterey Bay and NASA Ames Research Center.

## 7. References

- Ahmed, S.M. & Zerihun, A. (2010). Possession and Usage of Insecticidal Bed Nets Among the People of Uganda: Is BRAC Uganda Health Programme Pursuing a Pro-Poor Path? *PLoS One*, Vol. 5, e12660.
- Aviles, W.; Ortega, O.; Kuan, G.; Coloma, J. & Harris, E. (2008). Quantitative Assessment of the Benefits of Specific Information Technologies Applied to Clinical Studies in Developing Countries. *American Journal of Tropical Medicine and Hygiene*, Vol. 78 , pp. 311-315.
- Barker, C.M.; Kramer, V.L. & Reisen, W.K. (2010). Decision Support System for Mosquito and Arbovirus Control in California. Available from <http://www.earthzine.org/2010/09/24/decision-support-system-for-mosquito-and-arbovirus-control-in-california/>
- Biggerstaff, B.J. (2006). *PooledInfRate, Version 3.0: A Microsoft Excel Add-In to Compute Prevalence Estimates from Pooled Samples*. United States Centers for Disease Control and Prevention, Fort Collins, Colorado, United States of America.

- Booman, M.; Durrheim, D.N.; La Grange, K.; Martin, C.; Mabuza, A.M.; Zitha, A.; Mbokazi, F.M.; Fraser, C. & Sharp, B.L. (2000). Using a Geographical Information System to Plan a Malaria Control Programme in South Africa. *Bulletin of the World Health Organization* Vol 78, pp. 1438-1444.
- Booman, M.; Sharp, B.L.; Martin, C.L.; Manjate, B.; La Grange, J.J. & Durrheim, D.N. (2003). Enhancing Malaria Control Using a Computerised Management System in Southern Africa. *Malaria Journal* Vol. 2, 13.
- California Department of Public Health, Mosquito and Vector Control Association of California, and University of California. (2010). *California Mosquito-Borne Virus Surveillance and Response Plan*. California Department of Public Health, Sacramento, California, United States of America.
- Chang, A.Y.; Parrales, M.E.; Jimenez, J.; Sobieszczyk, M.E.; Hammer, S.M.; Copenhaver, D.J. & Kulkarni, R.P. (2009). Combining Google Earth and GIS Mapping Technologies in a Dengue Surveillance System for Developing Countries. *International Journal of Health Geographics* Vol. 8, 49.
- Coleman, M.; Coleman, M.; Mabuza, A.M.; Kok, G.; Coetzee, M. & Durrheim, D.N.D. (2008). Evaluation of an Operational Malaria Outbreak Identification and Response System in Mpumalanga Province, South Africa. *Malaria Journal*, Vol 7, 69.
- Dialynas, E.; Topalis, P.; Vontas, J. & Louis, C. (2009). MIRO and IRbase: IT Tools for the Epidemiological Monitoring of Insecticide Resistance in Mosquito Disease Vectors. *PLoS Neglected Tropical Diseases*, Vol. 3, e465.
- Eisen, L. & Eisen, R.J. (2011). Using Geographic Information Systems and Decision Support Systems for the Prediction, Prevention, and Control of Vector-Borne Diseases. *Annual Review of Entomology*, Vol. 56, pp. 41-61.
- Eisen, L.; Coleman, M.; Lozano-Fuentes, S.; McEachen, N.; Orlans, M. & Coleman, M. (2011). Multi-Disease Data Management System Platform for Vector-Borne Diseases. *PLoS Neglected Tropical Diseases* Vol 5 (3), e1016. doi:10.1371/journal.pntd.0001016.
- Focks, D.A. (2003). *A Review of Entomological Sampling Methods and Indicators for Dengue Vectors*, World Health Organization Special Programme for Research and Training in Tropical Diseases, Geneva, Switzerland.
- Gosselin, P.; Lebel, G.; Rivest, S. & Douville-Fradet, M. (2005). The Integrated System for Public Health Monitoring of West Nile Virus (ISPHM-WNV): A Real-Time GIS for Surveillance and Decision-Making. *International Journal of Health Geographics* Vol. 4, 21.
- Gratz, N.G. (1999). Emerging and Resurging Vector-Borne Diseases. *Annual Review of Entomology*, Vol. 44, pp. 51-75.
- Hay, S.I. & Snow, R.W. (2006). The Malaria Atlas Project: Developing Global Maps of Malaria Risk. *PLoS Medicine*, Vol. 3, e473.
- Hemingway, J.; Beaty, B.J.; Rowland, M.; Scott, T.W. & Sharp, B.L. (2006). The Innovative Vector Control Consortium: Improved Control of Mosquito-Borne Diseases. *Parasitology Today*, Vol. 22, pp. 308-312.
- Johnson, P.R. & Blazes, D.L. (2007). Using Cell Phone Technology for Infectious Disease Surveillance in Low-Resource Environments: A Case Study from Peru. In *Global Infectious Disease Surveillance and Detection: Assessing the Challenges – Finding Solutions*. Institute of Medicine Workshop Summary, S.M. Lemon, M.A. Hamburg, P.F.



- Sparling, E.R. Choffnes & A. Mack, (Eds.), pp. 136-157, The National Academies Press, ISBN 9780309111140, Washington, D.C., United States of America.
- Jolly, W.M.; Graham, J.M.; Michaelis, A.; Nemani, R. & Running, S.W. (2005). A Flexible, Integrated System for Generating Meteorological Surfaces Derived from Point Sources Across Multiple Geographic Scales. *Environmental Modelling and Software* Vol. 20, pp. 873-882.
- Kamadjeu, R. (2009). Tracking the Polio Virus Down the Congo River: A Case Study on the Use of Google Earth in Public Health Planning and Mapping. *International Journal of Health Geographics* Vol. 8, 4.
- Kramer, L.D.; Styer, L.M. & Ebel, G.D. (2008). A Global Perspective on the Epidemiology of West Nile Virus. *Annual Review of Entomology*, Vol. 53, pp. 61-81.
- Lozano-Fuentes, S.; Elizondo-Quiroga, D.; Farfan-Ale, J.A.; Loroño-Pino, M.A.; Garcia-Rejon, J.; Gomez-Carro, S.; Lira-Zumbardo, V.; Najera-Vazquez, R.; Fernandez-Salas, I.; Calderon-Martinez, J.; Dominguez-Galera, M.; Mis-Avila, P.; Coleman, M.; Morris, N.; Moore, C.G.; Beaty, B.J. & Eisen, L. (2008). Use of Google Earth to Strengthen Public Health Capacity and Facilitate Management of Vector-Borne Diseases in Resource-Poor Environments. *Bulletin of the World Health Organization*, Vol. 86, pp. 718-725.
- Martin, C.; Curtis, B.; Fraser, C. & Sharp, B. (2002). The Use of a GIS-Based Malaria Information System for Malaria Research and Control in South Africa. *Health & Place* Vol. 8, pp. 227-236.
- Meankaew, P.; Kaewkungwal, J.; Khamsiriwatchara, A.; Khunthong, P.; Singhasivanon, P. & Satimai, W. (2010). Application of Mobile-Technology for Disease and Treatment Monitoring of Malaria in the "Better Border Healthcare Programme". *Malaria Journal*, Vol. 9, 237.
- Nemani, R.; Votava, P.; Michaelis, A.; White, M.; Melton, R.; Milesi, C.; Pierce, L.; Golden, K.; Hashimoto, H.; Ichii, K.; Johnson, L.; Jolly, M.; Myneni, R.; Tague, C.; Coughlan, J. & Running, S. (2007). Remote Sensing Methodologies for Ecosystem Management. In: *Food and Water Security*, U. Aswathanarayana, (Ed.), pp. 3-19, Taylor & Francis, ISBN 9780415440189, Oxford, United Kingdom.
- Smith, B.; Ashburner, M.; Rosse, C.; Bard, J.; Bug, W.; Ceusters, W.; Goldberg, L.J.; Eilbeck, K.; Ireland, A.; Mungall, C.J.; Leontis, N.; Rocca-Serra, P.; Ruttenberg, A.; Sansone S.-A.; Scheuermann, R.H.; Shah, N.; Whetzel, P.L. & Lewis, S. (2007). The OBO Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration. *Nature Biotechnology* Vol. 25, pp. 1251-1255.
- Sonenshine, D.E. (1993). *Biology of Ticks. Volume 2*. Oxford University Press, ISBN 9780195084313, Oxford, United Kingdom.
- Stensgaard, A.S.; Saarnak, C.F.L.; Utzinger, J.; Vounatsou, P.; Simoonga, C.; Mushinge, G.; Rahbek, C.; Mohlenberg, F. & Kristensen, T.K. (2009). Virtual Globes and Geospatial Health: The Potential of New Tools in the Management and Control of Vector-Borne Diseases. *Geospatial Health*, Vol. 3, pp. 127-141.
- Topalis, P.; Mitraka, E.; Bujila, I.; Deligianni, E.; Dialynas E.; Siden-Kiamos, I.; Troye-Blomberg, M. & Louis, C. (2010). IDOMAL: An Ontology for Malaria. *Malaria Journal*, Vol. 9, 230.
- Vanden Eng, J.L.; Wolkon, A.; Frolov, A.S.; Terlouw, D.J.; Eliades, M.J.; Morgah, K.; Takpa, V.; Dare, A.; Sodahlon, Y.K.; Doumanou, Y.; Hawley, W.A. & Hightower, A.W.

- (2007). Use of Handheld Computers with Global Positioning Systems for Probability Sampling and Data Entry in Household Surveys. *American Journal of Tropical Medicine and Hygiene*, Vol 77, pp. 393-399.
- World Health Organization. (2008). *World Malaria Report 2008*, World Health Organization, ISBN 9789241563697, Geneva, Switzerland.
- World Health Organization. (2009). *Dengue: Guidelines for Diagnosis, Treatment, Prevention and Control*, World Health Organization, ISBN 9789241545006, Geneva, Switzerland.
- Yi, Q.; Hoskins, R.E.; Hillringhouse, E.A.; Sorensen, S.S.; Oberle, M.W.; Fuller, S.S.; & Wallace, J.C. (2008). Integrating Open-Source Technologies to Build Low-Cost Information Systems for Improved Access to Public Health Data. *International Journal of Health Geographics* Vol. 7, 29.