Using Geographic Information Systems and Decision Support Systems for the Prediction, Prevention, and Control of Vector-Borne Diseases

Lars Eisen¹ and Rebecca J. Eisen²

¹Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, Colorado 80523; email: lars.eisen@colostate.edu
²Division of Vector-Borne Infectious Diseases, Coordinating Center for Infectious Diseases, Centers for Disease Control and Prevention, Fort Collins, Colorado 80522; email: dyn2@cdc.gov

Key Words
GIS, modeling, risk map, Lyme borreliosis, plague, malaria, dengue, West Nile virus disease

Abstract
Emerging and resurging vector-borne diseases cause significant morbidity and mortality, especially in the developing world. We focus on how advances in mapping, Geographic Information System, and Decision Support System technologies, and progress in spatial and space-time modeling, can be harnessed to prevent and control these diseases. Major themes, which are addressed using examples from tick-borne Lyme borreliosis; flea-borne plague; and mosquito-borne dengue, malaria, and West Nile virus disease, include (a) selection of spatial and space-time modeling techniques, (b) importance of using high-quality and biologically or epidemiologically relevant data, (c) incorporation of new technologies into operational vector and disease control programs, (d) transfer of map-based information to stakeholders, and (e) adaptation of technology solutions for use in resource-poor environments. We see great potential for the use of new technologies and approaches to more effectively target limited surveillance, prevention, and control resources and to reduce vector-borne and other infectious diseases.
INTRODUCTION

Emerging and resurging vector-borne diseases (VBDs) cause significant morbidity and mortality, especially in the developing world (42). VBDs account for 7 of 10 neglected infectious diseases that disproportionately affect poor and marginalized populations and therefore have been targeted for control programs by the World Health Organization (WHO) Special Program for Research and Training in Tropical Diseases (http://apps.who.int/tdr/). These diseases include, among others, malaria, with an estimated 247 million cases and nearly a million deaths in 2006, and dengue, with up to 50 million dengue infections and 500,000 cases of severe dengue hemorrhagic fever estimated to occur each year (119–120). Furthermore, new VBDs have emerged and have become established in developed regions of the world. Examples of VBDs that now are a fact of life in such areas, and highly unlikely to be eliminated, include West Nile virus (WNV) disease in North America and Lyme borreliosis in Asia, Europe, and North America (62, 100).

Technological advances over the last decades with relevance to VBDs include the emergence of molecular techniques for vector species identification and pathogen detection and identification, and a rapid evolution in hardware and software options to support data collection, management, and analysis. These advances are now dramatically changing our capacity to predict, prevent, and control VBDs. In this review, we focus on how advances in mapping, Geographic Information System (GIS), Remote Sensing (RS), and Decision Support System (DSS) technologies, and progress in the fields of spatial and space-time modeling, can be harnessed to reduce the burden that VBDs inflict on humans. Major themes to be explored include (a) selection of appropriate spatial and space-time modeling techniques, (b) the importance of using high-quality and biologically and/or epidemiologically relevant data, (c) incorporation of new technologies and approaches into operational vector and disease control programs, (d) transfer of map-based information to the stakeholder community, and (e) adaptation of technology solutions for use in resource-poor environments. These will be addressed using examples from a broad range of VBDs including tick-borne Lyme borreliosis, flea-borne plague, and mosquito-borne dengue, malaria, and WNV disease (Table 1). It should be noted that the literature is too extensive for exhaustive reviews of all related published papers; therefore, we present only selected, representative publications as examples.

SPATIAL AND SPACE-TIME RISK MODELS AND THEIR APPLICATIONS AS PUBLIC HEALTH TOOLS

Statistical modeling techniques are commonly incorporated into a GIS framework to (a) identify spatial and space-time patterns of vectors and VBD cases, (b) improve our understanding of how environmental factors affect arthropod vectors and influence transmission of vector-borne pathogens, and (c) predict future changes in spatial risk of exposure to vectors and vector-borne pathogens in response to shifting land use or climatic patterns. The ultimate goal of these activities is to reduce disease burdens by generating information that empowers the public to take protective action and helps public health agencies to allocate limited prevention, surveillance, and control resources to best effect.

Introduction to Spatial Risk Models

Spatial risk models are defined, in the context of this review, as GIS-based statistical models used to estimate vector presence or abundance, or VBD case presence or incidence, within a particular geographical area. Model outputs typically are displayed in map format (Figure 1). Basic spatial modeling approaches include (a) interpolation based on spatial dependence in vector or VBD data and (b) extrapolation based on associations between vector or VBD data and environmental or socioeconomic predictor variables. Importantly, this allows for
Table 1 Characteristics of vector-borne diseases used as examples in the review

<table>
<thead>
<tr>
<th>Disease</th>
<th>Causative agent(s)</th>
<th>Primary vectors</th>
<th>Primary vertebrate reservoirs</th>
<th>Primary current disease foci</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mosquito-borne viral disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dengue</td>
<td>Dengue virus</td>
<td><em>Aedes aegypti, Ae. albopictus</em></td>
<td>Humans</td>
<td>Subtropics and tropics, especially in Asia and the Americas&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>West Nile virus disease</td>
<td>West Nile virus</td>
<td><em>Culex</em> spp.</td>
<td>Birds</td>
<td>North America&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Tick-borne bacterial disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lyme borreliosis</td>
<td><em>Borrelia burgdorferi sensu lato</em></td>
<td><em>Ixodes</em> spp.</td>
<td>Rodents</td>
<td>Temperate areas in Asia, Europe, and North America</td>
</tr>
<tr>
<td>Flea-borne bacterial disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plague</td>
<td><em>Yersinia pestis</em></td>
<td>Various flea species</td>
<td>Rodents</td>
<td>Africa</td>
</tr>
<tr>
<td>Mosquito-borne parasitic disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malaria</td>
<td><em>Plasmodium</em> spp.</td>
<td><em>Anopheles</em> spp.</td>
<td>Humans</td>
<td>Subtropics and tropics in Africa, the Americas, and Asia</td>
</tr>
</tbody>
</table>

<sup>a</sup>Disease burden in Africa is poorly understood.

Spatial Risk Interpolation Models

Spatial dependence for vector abundance or VBD case count or incidence is frequently observed at fine spatial scales (58). For example, areas with high vector abundance or high disease incidence often border on other areas with high vector abundance or high disease incidence, and the similarity in the response variable decreases with increasing distance. In such instances, kriging or other types of interpolation models are used to produce smooth interpolated maps of the response variable (12, 20, 22, 32, 41, 58, 64, 80, 85, 99). High-end GIS software packages, such as ArcGIS®, have extensive capacity for interpolation modeling.

Although interpolation methods are useful for transforming point-based data into smooth risk surfaces that can be used to infer risk in areas that were not sampled, they are most useful at fine spatial scales and often are unreliable beyond the geographical area within which point data were gathered (20, 58). Because of this limitation, other techniques that depend on identifying environmental predictors of vector abundance or VBD incidence are used to extrapolate risk surfaces beyond the local areas where vector or VBD data were collected. In
Spatial Risk Extrapolation: Model Development Based on Generalized Linear Models

Spatial patterns of vector abundance or VBD incidence often can be predicted on the basis of their associations with environmental or socioeconomic variables, such as land use, soil type, temperature-related factors, or rainfall. In this scenario, the GIS software is first used to extract spatially explicit data for environmental factors of interest for the point locations (e.g., mosquito trap locations) or geographical areas (e.g., boundary units for which VBD incidence was calculated) where data were collected. Thereafter, a predictive model is developed in a statistical software package and the model equation is then applied in the GIS, for example, using the Raster Calculator in ArcGIS, to extrapolate a surface for the risk measure of interest. This basic approach has multiple benefits including (a) the potential for identifying environmental or socioeconomic predictors for risk of exposure to vectors or vector-borne pathogens, (b) development of continuous spatial surfaces that present estimates of risk for exposure to vectors or vector-borne pathogens and can be delivered to stakeholders in readily understandable map
format, and (c) use of map outputs to guide decisions regarding allocation of surveillance, prevention, and control resources.

Commonly used techniques for the model development step include linear or Poisson regression models for continuous response variables, binomial logistic regression models for binary response variables, and multinomial or ordinal logistic regression models for multicategorical response variables. There also is increasing use of other techniques such as generalized additive models, which extend the generalized linear models to include nonparametric fits, and Bayesian approaches, which include a more rigorous accounting of uncertainty compared with models based on frequency probability (19, 51, 97, 98). Generalized linear models also may include terms that account for spatial dependence in the response variable. Below, we provide examples of the use of generalized linear models to identify factors predictive of elevated risk for exposure to vectors or vector-borne pathogens, and explore some issues relating to extrapolation of spatial risk surfaces from these models.

Examples of Identification of Risk Factors for Exposure to Vectors or Vector-Borne Pathogens Based on Use of Generalized Linear Models

Generalized linear models are effective tools to identify factors that are associated with elevated vector abundance or VBD incidence. Logistic regression modeling based on epidemiologic data was used to identify environmental predictors of elevated risk for human plague in the southwestern United States and Uganda; this consistently identified elevation as a primary risk predictor together with vegetation type, moisture, and temperature (29–31, 33, 124). Furthermore, use of linear, Poisson, and logistic regression modeling has revealed that high abundances of *Ixodes pacificus* and *I. scapularis*, which are key tick vectors of Lyme borreliosis spirochetes in North America, can be predicted by GIS- or RS-based environmental factors related to elevation, slope of the landscape, vegetation type, soil type, temperature, and moisture (10, 12, 20, 28, 38, 43).

The introduction of WNV into North America in 1999 spurred a series of similar modeling exercises for abundance of *Culex* WNV vectors and incidence of human WNV disease, which revealed that environmental predictive factors for elevated mosquito vector abundance or WNV disease incidence among the human population include availability of water sources, elevation, vegetation type, and temperature-related factors (11, 21, 121, 123). Furthermore, a detailed study from the greater Chicago area revealed that in urban settings factors predicting high WNV disease incidence among the human population can include environmental factors (presence of vegetation) as well as socioeconomic factors (income, age, housing age) and presence or absence of mosquito control activities (93).

Generalized linear models have been used extensively to identify environmental factors predictive of elevated abundances of anopheline malaria vectors and malaria prevalence in humans in sub-Saharan Africa (3, 7, 19, 46, 53, 60, 98, 99, 109, 126), and a few examples of model results are given below. Logistic regression modeling was used to predict the spatial abundance pattern of two malaria vectors (*Anopheles gambiae* sensu stricto and *Anopheles arabiensis*) in Mali based on vegetation indices, soil features, distance to water, temperature, and rainfall (98). In the same country, a logistic regression modeling approach was used to identify the normalized difference vegetation index (NDVI), distance to water, temperature, and rainfall as predictors of malaria prevalence in children (60). The model predictions in the latter study were further refined by incorporating a spatial interpolation (kriging) of the model residuals. Studies from other African countries have produced similar results; prevalence of malaria infection in humans was associated with rainfall, temperature, and elevation in Botswana (19), and elevation alone predicted 73% of households where an occupant had splenomegaly associated with a malaria infection (an indicator of repeated attacks and
prolonged exposure to malaria parasites) in the Usambara Mountains of Tanzania (3).

The models described above for anopheline vectors and malaria provide good examples of instances in which the environmental predictors make biological sense based on our understanding of the vectors. Water sources, ranging from rice fields to cattle hoof prints, provide development sites for the immature life stages of the anopheline mosquito vectors, and temperature, which decreases with increasing elevation, affects the development time for immatures, the length of the gonotrophic cycle in the female mosquito, and the extrinsic incubation period for the malaria parasite within the vector (46, 66). This type of clear linkage between readily understandable predictor variables with obvious biological relevance and entomological or epidemiologic model outcomes is an important factor for decisions to use model results to guide operational vector and disease control program activities.

There also is a growing literature for use of generalized linear models to identify factors that can predict elevated spatial risk of exposure to dengue virus or to the primary vector, *Aedes aegypti*. Because this endophilic and endophagic mosquito is closely associated with human habitation and utilizes containers in the peridomestic environment as larval development sites, fine-scale modeling efforts in urban areas to identify predictors of elevated vector abundance or dengue infection rate need to incorporate factors relating to socioeconomic conditions such as presence of piped water, housing characteristics, and income (8, 75, 91, 107). In more rural settings in Thailand, land cover around the home was a useful predictive factor for dengue virus exposure, perhaps related to the presence of another dengue virus vector, *Ae. albopictus* (110, 112). There also have been some models developed for larger scales. Dengue risk in Taiwan was found to increase with average annual temperatures above 18°C and the degree of urbanization (125), and a study from Puerto Rico showed that the relative strength of temperature versus precipitation as predictors for dengue virus transmission varied between different parts of the island (55). One complicating factor for development of spatial risk models based on epidemiologic dengue data is that the disease can be caused by four different serotypes of dengue virus and that infection with one serotype does not provide long-term cross-protection against the other serotypes. This results in a dynamic situation with potential for rapid fluctuations in serotype-specific levels of susceptibility among the human population, which can confound spatial modeling efforts.

**Spatial Risk Extrapolation: Production of Risk Surfaces Based on Use of Generalized Linear Models**

Extrapolation of continuous spatial surfaces estimating the level of risk for exposure to vectors or vector-borne pathogens, based on application of a model equation in a GIS, was incorporated into many of the studies described above (see example for plague in the southwestern United States in Figure 1). One important but poorly studied aspect of this activity is the selection of the geographical area over which the risk model can reliably be extrapolated. A commonsense approach is to restrict the extrapolation of a model to geographical areas that fall within the data range of the ecological or climatic predictor variables used to develop the model. This was illustrated in a study on the WNV vector *Culex tarsalis* in Colorado, where a mosquito abundance model was developed on the basis of an association with cooling degree days (CDD) for the eastern slope of the Rocky Mountains (121). The model then was scaled up to different parts of the state of Colorado, and correlations between the extent of area with high predicted *Cx. tarsalis* abundance and WNV disease incidence in humans were examined at the census tract scale. This revealed a positive correlation between model-predicted areas with high vector abundance and WNV disease incidence in the western, mountainous part of the state, which has a CDD range similar to the model development area and
to which extrapolation thus is appropriate. In contrast, a negative correlation was observed in the eastern Colorado plains, where the CDD range is distinctly different from the model development area; extrapolation therefore should not be assumed to produce reliable results.

When model extrapolation is restricted to areas with ecological and climatic characteristics similar to those of the model development area, this approach can be a powerful tool to gain insights into levels of risk within areas where surveillance data are lacking or unreliable. As an example, in the most recent revision of WHO’s International Health Regulations (117), only the most severe form of plague (pneumonic plague) is an internationally notifiable disease whereas other manifestations are notifiable only from nonendemic localities. Compliance with these regulations thus requires a clear understanding of the locations of plague foci, and WHO recommends use of GIS technology and modeling to refine our knowledge of these risk areas and to cost-effectively target surveillance resources (118). In the West Nile region of Uganda, a linear regression modeling approach revealed that incidence of plague cases was higher above 1,300 m and in parishes with higher surface temperature and greater land cover variety, and with certain remotely sensed indicators of soil and vegetation type (124). This spatial risk model, and an additional model for the same area that focuses on the finer village and subvillage scale and therefore is better suited for allocation of scarce plague control resources (31), has great potential for extrapolation from Uganda’s West Nile region into neighboring areas of the Democratic Republic of Congo with similar ecological, socioeconomic, and climatic characteristics but where plague surveillance activities are lacking (Figure 2).

Spatial Risk Assessments Based on Machine Learning Algorithms or Dynamic Simulation Models

Low confidence in data related to absence of vectors or VBD cases is often an obstacle to the use of generalized linear models. In this case, presence-only machine learning (rule-based) algorithms or dynamic simulation models serve as alternatives (84, 101, 103). Examples include various ecological niche modeling algorithms such as genetic algorithm for rule-set prediction (GARP) (101) and MAXENT, a machine learning algorithm based on maximum entropy (84), and simulation models such as CLIMEX (http://www.climatemodel.com/index.html).

Ecological niche models have, for example, been used to model distributions of anopheline malaria vectors at continental scales using archived species occurrence data (74, 81). Likewise, risk of human plague in the United States and in Africa has been modeled using GARP (77–79) and MAXENT (50). In instances where model outputs of generalized linear models and GARP could be directly compared, the former typically provided a more restricted area that was expected to pose elevated risk of plague (29, 33, 77, 79, 124). The more restricted risk surfaces that resulted from the generalized linear models are well suited for prevention and control resource allocation, whereas the broader ecological niche that is captured by GARP may be useful for identifying new areas where vectors or VBDs are likely to emerge.

One stated benefit of the dynamic simulation model CLIMEX, which has been used to model areas in Australia at risk for malaria under different climate change scenarios (102), is that absence data in areas estimated as suitable from presence data are attributed as unknowns, which prevents the model from restricting its parameter values to simulating only the presence records and stimulates a search for other explanations (103). This type of model is particularly useful for anticipating future expansions of vector or pathogen distributions, as opposed to current distributions that are often accurately portrayed using generalized linear models (103). Another recently emerged simulation model is the stochastic and spatially explicit Skeeter Buster, which focuses on container-breeding dengue virus mosquito vectors and operates at the spatial scale of individual

www.annualreviews.org • GIS and Vector-Borne Disease Prevention and Control 47
Space-Time Risk Models

Space-time risk models are used to explore spatial clustering of vectors or VBD cases and are useful for identifying changing risk patterns. This is especially relevant for VBDs such as dengue, which is characterized by outbreaks with rapidly building dengue case numbers and explosive spread within affected areas. Space-time modeling may aid in identifying...
underlying factors that regulate the spread of vectors or the occurrence of VBD cases and may be more sensitive than purely spatial or temporal models in detecting local outbreaks (40, 63). Therefore, outputs of these models may be used as early-warning systems and could guide vector control or surveillance activities. Commonly used methods include space-time permutation scan statistics (e.g., SaTScan) (63, 113, 116), Knox tests (61), generalized additive mixed models (69), and Bayesian hierarchical regression models (39, 65).

For dengue, space-time modeling consistently demonstrates that dengue cases are clustered in space and time (4, 13, 57, 71, 76, 92). Part of the challenge to ultimately forecast space-time patterns of a dengue outbreak, and rapidly implement vector control response to curb the outbreak, is that the virus spreads in two different clustering diffusion patterns: (a) a predictable contiguous pattern of dengue diffusion over short distances from an index case that probably is driven by a combination of movement by infected mosquito vectors and short-range movement of infectious humans, and (b) an unpredictable relocation pattern of dengue diffusion with cases jumping to new, unconnected areas and initiating new local dengue foci, most likely resulting from longer-range movements of infectious humans (57, 96).

For malaria, space-time permutation scan statistics were incorporated into a malaria early-warning system to detect local malaria clusters and to target vector control and health education activities in South Africa’s Mpumalanga Province (16). They were also used to detect temporally stable malaria clusters in Ethiopia (83). In Kenya, regression analyses were used to identify meteorological factors and remotely sensed vegetation indices that were predictive of mean monthly percentages of total annual malaria admissions; model thresholds were then extrapolated spatially in a clinical disease seasonality map to define the number of months per location with expected malaria admissions (48). Similarly, a malaria seasonality risk map for Zimbabwe was created using a space-time regression model within a Bayesian framework (68).

Space-time models also have been applied to Lyme borreliosis and WNV disease in the United States to gain a better understanding of how the geographic coverages of these diseases may change over time. Waller et al. (114) used hierarchical linear models to describe an increasing county-level Lyme borreliosis incidence pattern for the central northeastern Atlantic coast with branching to the north and west and stable or slightly decreasing incidence in counties located in western New York State. Within the state of New York, Chen et al. (15) implemented a Bayesian hierarchical regression approach to reveal a northerly progression from 1990 to 2000 in the location of counties reporting the highest Lyme borreliosis incidence. The emergence of WNV in the United States prompted the development of the Dynamic Continuous-Area Space-Time (DYCAST) system, which uses a localized Knox test to capture the space-time interaction of WNV surveillance data such as dead birds (108).

THE IMPORTANCE OF HIGH-QUALITY AND BIOLOGICALLY OR EPIDEMIOLOGICALLY RELEVANT DATA

Perhaps the most important thing to remember when embarking on a modeling project is that models are only as good as the data upon which they are based. Furthermore, the data requirements for a meaningful model with public health utility are dependent on first determining the intent of the modeling exercise, e.g., hypothesis testing, identifying knowledge gaps, providing direction for surveillance and control efforts, or evaluating effectiveness of an intervention (59). For example, modeling the probability of plague case occurrence at regional or even continental scales may be useful to broadly determine the need for surveillance activities (77–79), but finer-resolution models are required for targeting prevention and control through more focused risk assessments and control interventions.
resources implemented at local spatial scales (29–30, 33, 124).

Risk of human exposure to vector-borne pathogens may be assessed on the basis of modeling the spatial distribution of human disease cases using epidemiologic data, or on the basis of vector data. The advantages and disadvantages of using vector data versus epidemiologic data in spatial risk modeling were reviewed recently (27) and are summarized briefly here. Benefits of training spatial risk models on epidemiologic data include (a) that a human VBD case is unequivocally linked with exposure to the disease agent and (b) there is potential for identifying socioeconomic factors associated with exposure risk. However, there are numerous drawbacks to using epidemiologic data, including (a) not all VBDs are notifiable, (b) case definitions and reporting practices change over space and time, (c) socioeconomic factors influence care-seeking behavior, and (d) for some VBDs, such as dengue and WNV disease, asymptomatic infections are common. Furthermore, location of pathogen exposure is often not investigated or reported and instead the location of residence is used as a surrogate, often without a detailed understanding of the spatial dimension of pathogen transmission to humans. To account for patient privacy, outputs from models based on epidemiologic data often are displayed at coarse spatial scales (e.g., county of residence in the United States) and fail to account for fine-scale variability in risk patterns (25). Finally, spatial models based on epidemiologic data cannot be used to assess risk of exposure on publicly owned lands where humans do not reside but where visitors may be at risk for recreational exposure to vector-borne pathogens.

Modeling spatial risk on the basis of vector data can be advantageous because (a) it allows for fine-scale spatial precision in the vector collection location, (b) many vectors transmit multiple pathogens and therefore assessing abundance of a single vector may be informative for risk assessments of several VBDs, (c) many VBDs are not notifiable and therefore epidemiologic data are lacking, and (d) models can be developed for public or private land because the vector abundance estimate is independent of a human population base. Limitations of basing spatial risk models on vector data include (a) the cost and effort associated with field collections and pathogen detection and (b) the lack of a direct correlation between abundance of vectors or infected vectors and human disease (prevention activities such as use of repellents may limit contact between humans and vectors even in areas with high vector abundance).

The quality of data for predictor variables is equally important. As noted previously, risk models for vectors or VBDs have been based on associations with socioeconomic conditions or environmental factors such as elevation, soil type, vegetation type, land cover, and climatic or meteorological variables. In the United States, GIS-based data layers containing a wide range of socioeconomic variables can be accessed from the U.S. Census Bureau and fine-resolution GIS-based data layers for elevation, soil type, hydrological features, vegetation type, and land use are available from sources such as the U.S. Geological Survey and state Gap Analysis Program projects. Furthermore, GIS-based climatic and meteorological data layers are readily available from Oregon State University’s PRISM climate group at spatial resolutions in the 2- to 4-km range and are commonly reported at monthly intervals. Although such classified data layers are often lacking from developing countries, RS data for vegetation indices (NDVI, greenness), soil and water characteristics (brightness, wetness), and some meteorological variables can be derived using images from satellites such as Landsat (15- to 60-m resolution, covers entire earth every 16 days) and MODIS (moderate-resolution imaging spectroradiometer; 250- to 1,000-m resolution, covers entire earth every 1 to 2 days). Use of RS data for modeling risk of exposure to vectors and VBDs was reviewed previously (90). Finally, when compiling the predictor variable data layers, they should represent epidemiologically relevant spatial and temporal scales in order to produce model outputs of direct use in prevention and control efforts (54, 94).
Potential stumbling blocks to consider for use of GIS and RS data in risk modeling for vectors and VBDs include lack of fine-scale GIS-based data for zoonotic pathogen reservoirs (e.g., for modeling risk of exposure to the etiologic agents of Lyme borreliosis, plague, and WNV disease in the United States), the challenge of acquiring cloud-free satellite imagery in the tropics, and the need for an enhanced meteorological observation network in developing countries to develop better GIS-based climate data.

OPERATIONAL USE OF GEOGRAPHICAL INFORMATION SYSTEM TECHNOLOGY IN VECTOR AND DISEASE CONTROL PROGRAMS

Numerous reviews have broadly addressed the use of GIS/RS and spatial and space-time modeling approaches in the field of VBDs (5, 27, 58, 90). However, the critically important issue of the potential for such technologies and methodologies to be used for operational surveillance and control of VBDs has not received the attention it deserves, especially for neglected tropical diseases. Eisen & Lozano-Fuentes (26) recently reviewed this topic for dengue and concluded that there is tremendous potential for moving mapping and modeling approaches from the research arena to practical applications that can enhance operational vector and dengue control. Similar conclusions were drawn for malaria in a review by Saxena et al. (95). Below, we provide some examples of GIS technology used in operational vector and disease control.

In South Africa, GIS technology was incorporated into operational malaria control programs in the 1990s and is now used for a variety of purposes including malaria case mapping and monitoring of vector control coverage (9, 72). Australia and Singapore are making extensive use of GIS in their operational vector and dengue control programs. For example, Queensland Health in Australia employs GIS in ongoing mapping of dengue case locations in relation to spatial coverage of implemented vector control to help determine if response activities have adequate spatial coverage (86–87). Singapore’s National Environment Agency uses GIS in a wide range of operational vector and dengue control activities including tracking of dengue case locations, vector mosquito surveillance, and monitoring of vector control coverage (1, 105–106). Nicaragua’s Ministry of Health also is starting to use GIS and mapping software in vector and dengue control activities such as dengue case mapping and mosquito vector surveillance (2, 14).

In the United States, the emergence of WNV in 1999 resulted in a new national GIS-based surveillance system for arboviral diseases: ArboNET. This system compiles surveillance data for a wide range of arboviral diseases (WNV disease, St. Louis encephalitis, eastern equine encephalitis, western equine encephalitis, La Crosse encephalitis, and Powassan encephalitis) and includes data for disease in humans and domestic animals as well as surveillance data for infection in vertebrates (e.g., sentinel chickens or wild birds) and vectors (e.g., testing of Culex mosquitoes for WNV). Map-based outputs from ArboNET are made available online by the Centers for Disease Control and Prevention (http://www.cdc.gov/ncidod/dvbid/westnile/surv&control.htm) and the U.S. Geological Survey (http://diseasemaps.usgs.gov/).

To achieve increased use of mapping and GIS technologies in operational vector and disease control, it is critically important for control programs to share their experiences with GIS and other emerging technologies through publications and other information delivery mechanisms. Although it may be difficult to assign time to such undertakings in the midst of the day-to-day control activities, it must be stressed that operational vector and disease control programs around the world have gained invaluable experiences regarding the benefits and drawbacks of using GIS and other emerging technologies that need to be shared with their counterparts in other disease-endemic areas.
TRANSFER OF MAP-BASED INFORMATION FOR VECTOR-BORNE DISEASES TO THE STAKEHOLDER COMMUNITY

Basic options to present information for spatial risk of VBDs in map formats to the stakeholder community include point locations for disease cases, aggregation of disease case counts or disease incidence to administrative boundary units, or smoothing. A map showing individual case point locations is obviously the most precise way to present spatial disease data. However, this has distinct disadvantages including (a) the possibility that the address of residence is not the site of pathogen exposure, (b) a lack of accounting for population size, and (c) in some countries, including the United States, strict regulations to guide the use of patient health information. To avoid privacy issues, it is common practice to aggregate disease case counts or disease incidence to administrative boundaries in the display of spatial risk patterns.

GIS software and new mapping software, such as Google Earth™ (http://earth.google.com/), now provide capacity to generate risk maps in a variety of formats including overlays on satellite imagery and dynamic illustrations of space-time patterns that can be played as movie clips (6, 26, 67). This is accompanied by explosive development in the field of Web-based information delivery, which now provides an effective medium to distribute risk maps to a wide range of stakeholders including the medical community, vector control practitioners, policy makers, and the public at large (36, 37). The Malaria Atlas Project (http://www.map.ox.ac.uk/) is one example of effective online map delivery, including maps of the spatial limits of Plasmodium falciparum transmission (45, 47). Another is the MARA/ARMA project with online delivery of a wide range of malaria-related maps (http://www.mara.org.za/).

As illustrated in Figure 3 for WNV disease in Colorado, maps showing case counts and disease incidence for different spatial boundary units can be used as tools to target limited prevention, surveillance, and control resources to high-risk areas for WNV exposure, and to inform the public about local risk levels (122). For example, a mosquito control program aiming to implement control activities to suppress vector mosquitoes and reduce the burden of WNV disease likely will be most interested in finding out where high numbers of WNV disease cases occur in order to focus expensive prevention efforts. On the other hand, a member of the public seeking information to help determine his/her personal risk of exposure to WNV, and the need for use of personal protective measures such as repellents, will be more interested in spatial risk estimates based on WNV disease incidence (which account for population size).

However, with this new technological capacity to present spatial risk patterns comes a series of questions regarding how it should be used responsibly in public health (54, 94). One key question relates to the spatial scale at which risk maps based on epidemiologic data should be presented to best balance the needs to provide spatially explicit and accurate risk information while protecting patient privacy. Using finer spatial scales (and smaller population bases) can result in more informative maps but also may result in analysis artifacts and misleading risk maps. Further studies are urgently needed for important VBDs to determine the benefits and drawbacks of presenting risk maps at different spatial scales. In addition, we need to gain a better understanding of what type of information different types of stakeholders feel

Figure 3
West Nile virus disease case counts and incidences per 100,000 person-years in Colorado assorted by county, census tract, and zip code based on combined data for 2003 and 2007 and classified as no cases reported, or by quartile for the spatial boundary units reporting cases. This figure was previously published in the American Journal of Tropical Medicine and Hygiene (122).
### Total cases

<table>
<thead>
<tr>
<th>Quartile: (Range of cases for county, census tract, and zip code)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st ([1–3], (1), (1–2))</td>
</tr>
<tr>
<td>2nd ([4–11], (2–3), (3–6))</td>
</tr>
<tr>
<td>3rd ([12–67], (4–7), (7–15))</td>
</tr>
<tr>
<td>4th ([&gt;67], (&gt;7), (&gt;15))</td>
</tr>
<tr>
<td>No cases reported</td>
</tr>
</tbody>
</table>

### Incidence per 100,000 person-years

<table>
<thead>
<tr>
<th>Quartile: (Range of incidence for county, census tract, and zip code)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st ([1–11.7], (0–15.6), (2.8–14.1))</td>
</tr>
<tr>
<td>2nd ([11.8–26.9], (15.7–31.4), (14.2–39.2))</td>
</tr>
<tr>
<td>3rd ([27–88.4], (31.5–70.9), (39.3–106.4))</td>
</tr>
<tr>
<td>4th ([&gt;88.4], (&gt;70.9), (&gt;106.4))</td>
</tr>
<tr>
<td>Incidence of 0</td>
</tr>
</tbody>
</table>
that they require, and to determine optimal map and text formats to ensure that the message we aim to transmit in fact is clear to the user (6). Finally, there is a critical need to determine if presenting risk maps to the public has any impact, positive or negative, on preventive behaviors relating to VBDs.

TECHNOLOGIES FOR COLLECTION, MANAGEMENT, ANALYSIS, AND DISPLAY OF VECTOR AND DISEASE DATA

Management and analysis of data can be a complex undertaking and is greatly facilitated by use of database and statistical analysis software. A basic practical example involves entering and manipulating data in a Microsoft Excel spreadsheet and then exporting the data to a statistical software package for analysis. This, however, becomes challenging when large amounts and diverse types of data are handled. We then instead can make use of a relational database software package, e.g., Microsoft Access, Microsoft SQL Server, or an open source alternative such as PostgreSQL (.http://www.postgresql.org/). The primary benefits of relational databases are that they can be customized with regards to data entry and that the relationships defined within the database allows for effective querying and extraction of data. GIS software typically includes a relational database management system to facilitate entry, storage, extraction, and visualization of data.

With the current expansion in mobile data capturing technology, we now also have the opportunity to move the stage of electronic data capture all the way down to the initial data capturing session in the field (2, 56, 111). This can be accomplished using a laptop or netbook computer, a personal digital assistant (PDA), or even a smart phone. The basic workflow involves capture of data on the mobile electronic device followed by download into a central database by means of direct connection or transmission of data over the Internet and/or cell phone network. 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 location at which the data were entered. Mobile data-capturing technology is a field where we expect to see tremendous advances in the coming decade.

Data analysis capacity also has improved dramatically in recent decades. Numerous software packages are now available to support statistical analyses, to process GIS and RS data, and to visualize maps. An exhaustive review of such software is beyond the scope of this paper, but there are some developments worth noting. First, at-cost software packages are rapidly being complemented by freely available alternatives, many of which also are open source with access to the source code. Examples of software that can be downloaded and used at no cost include the statistical package R (.http://www.r-project.org/), the previously mentioned relational database software PostgreSQL and the mapping software Google Earth™, and the space-time pattern analysis software SaTScan™ (.http://www.satscan.org/). Second, the availability of environmental GIS and RS data and demographic and socioeconomic GIS data is rapidly increasing and there is a positive trend toward such data being made freely available for download. Third, there is a drive toward developing DSS software packages for operational control of VBDs, including neglected tropical diseases such as dengue, malaria, and human African trypanosomiasis, that combine user-friendly data entry with data analysis and visualization capacity (17–18, 24, 49, 67, 88, 104).

ADAPTATION OF TECHNOLOGY SOLUTIONS FOR USE IN RESOURCE-POOR ENVIRONMENTS: DECISION SUPPORT SYSTEMS

Adaptation of technology solutions for use in resource-constrained disease-endemic
environments that experience the most severe VBD burdens must be made part of the new frontier in VBD research. For example, use of cell phones for rapid and inexpensive information transfer has great potential for implementation in malaria- and dengue-endemic areas with poor Internet access but well-developed cell phone networks (56). Development of DSS software packages for malaria and dengue that are composed entirely of software components that can be distributed to users without licensing costs, which ensures that they can be implemented in resource-poor environments, is another positive example (17, 24, 49). These systems incorporate a wide range of data from entomological surveillance, disease case surveillance, vector and disease control intervention monitoring, and stock control. They also include a GIS backbone and reporting tools that allow the user to produce a wide range of outputs including tables, graphs, and maps. Key benefits of implementing a DSS include (a) improved capacity for electronic data storage; (b) compilation of a wide range of data in a single system, which allows the user to produce outputs combining different types of data such as coverage of vector control in relation to disease case incidence; and (c) improved capacity for monitoring and evaluation of control program performance.

**SUMMARY POINTS**

1. Advances in mapping, GIS, and DSS technologies, and progress in the fields of spatial and space-time modeling, provide new opportunities to prevent and control emerging and resurging VBDs.
2. Benefits of spatial and space-time risk modeling include identification of risk patterns for exposure to vectors and vector-borne pathogens, and an improved understanding of how socioeconomic and environmental factors affect the vectors and influence transmission of their associated pathogens.
3. Perhaps the most important thing to remember when embarking on a mapping or modeling project is that map or model outputs are only as good as the data on which they are based.
4. GIS-based spatial and space-time risk modeling have proven effective tools to develop risk surfaces (maps) to inform policy makers, control programs, and the public.
5. There needs to be a stronger emphasis on moving GIS technology and modeling approaches from the research arena into operational vector and disease control programs.
6. Clear linkage between readily understandable predictor variables with obvious biological relevance and entomological or epidemiologic model outcomes is an important factor for decisions to use model results to guide operational vector and disease control program activities.
7. We need to determine what type of map-based information different stakeholders require in order to make practical use of the maps, and to determine if presenting risk maps to the public has any impact, positive or negative, on preventive behaviors.
8. Adaptation of technology solutions for use in resource-constrained environments that experience the most severe disease burdens must be made part of the new frontier in VBD research.
FUTURE ISSUES

1. In addition to the strong current focus on improving statistical modeling techniques, there is a need to allocate resources for development of the high-quality data sets for vector and epidemiologic data without which development of high-quality models and risk maps is impossible. This includes determination of probable pathogen exposure sites for VBD patients to complement information for residence location.

2. Concerted efforts are needed to ensure ready and inexpensive access for the academic and public health communities in developing countries to both high-end GIS software and high-quality GIS/RS-based data for socioeconomic and environmental factors as well as administrative boundaries and natural features.

3. The research community is very adept at using GIS/RS-based data to develop predictive models for spatial or space-time patterns of VBDs and to display these as risk maps. There is, however, a disconnect between this model and risk map production process and the practical use of the models and risk maps for prevention and control purposes. To bridge this gap, studies are urgently needed to determine how stakeholders make use of model findings and map-based risk information.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This review was developed, in part, based on funding from the Innovative Vector Control Consortium and the National Institutes of Allergy and Infectious Diseases (contract N01-AI-25489).

LITERATURE CITED


Contents

Bemisia tabaci: A Statement of Species Status
Paul J. De Barro, Shu-Sheng Liu, Laura M. Boykin, and Adam B. Dinsdale .......... 1

Insect Seminal Fluid Proteins: Identification and Function
Frank W. Avila, Laura K. Sirot, Brooke A. LaFlamme, C. Dustin Rubinstein, and
Mariana F. Wolfner ................................................................. 21

Using Geographic Information Systems and Decision Support Systems
for the Prediction, Prevention, and Control of Vector-Borne Diseases
Lars Eisen and Rebecca J. Eisen ................................................. 41

Salivary Gland Hypertrophy Viruses: A Novel Group of Insect
Pathogenic Viruses
Verena-Ulrike Lietze, Adly M.M. Abd-Alla, Marc J.B. Vreysen, Christopher J. Geden, and Drion G. Boucias .................. 63

Insect-Resistant Genetically Modified Rice in China: From Research
to Commercialization
Mao Chen, Anthony Shelton, and Gong-yin Ye ................................................. 81

Energetics of Insect Diapause
Daniel A. Hahn and David L. Denlinger .............................................. 103

Arthropods of Medicoveterinary Importance in Zoos
Peter H. Adler, Holly C. Tuten, and Mark P. Nelder .......................... 123

Climate Change and Evolutionary Adaptations at Species’
Range Margins
Jane K. Hill, Hannah M. Griffiths, and Chris D. Thomas .............. 143

Ecological Role of Volatiles Produced by Plants in Response
to Damage by Herbivorous Insects
J. Daniel Hare ................................................................. 161

Native and Exotic Pests of Eucalyptus: A Worldwide Perspective
Timothy D. Paine, Martin J. Steinbauer, and Simon A. Lawson ................. 181
Urticating Hairs in Arthropods: Their Nature and Medical Significance
Andrea Battisti, Göran Holm, Bengt Fagrell, and Stig Larsson ........................................ 203

The Alfalfa Leafcutting Bee, *Megachile rotundata*: The World’s Most
Intensively Managed Solitary Bee
Theresa L. Pitts-Singer and James H. Cane ................................................................. 221

Vision and Visual Navigation in Nocturnal Insects
Eric Warrant and Marie Dacke .................................................................................... 239

The Role of Phytopathogenicity in Bark Beetle–Fungal Symbioses:
A Challenge to the Classic Paradigm
Diana L. Six and Michael J. Wingfield ........................................................................ 255

Micky D. Eubanks, Michael J. Raupp, and Deborah L. Finke .................................... 273

The Role of Resources and Risks in Regulating Wild Bee Populations
Tai H. Roulston and Karen Goodell ........................................................................... 293

Venom Proteins from Endoparasitoid Wasps and Their Role
in Host-Parasite Interactions
Sassan Asgari and David B. Rivers .............................................................................. 313

Recent Insights from Radar Studies of Insect Flight
Jason W. Chapman, V. Alistair Drake, and Don R. Reynolds ..................................... 337

Arthropod-Borne Diseases Associated with Political and Social Disorder
Philippe Brouqui ............................................................................................................. 357

Ecology and Management of the Soybean Aphid in North America
David W. Ragsdale, Douglas A. Landis, Jacques Brodeur, George E. Heimpel,
and Nicolas Desneux ................................................................................................... 375

A Roadmap for Bridging Basic and Applied Research
in Forensic Entomology

Visual Cognition in Social Insects
Aurore Avargues-Weber, Nina Deisig, and Martin Giurfa .......................................... 423

Evolution of Sexual Dimorphism in the Lepidoptera
Cerisse E. Allen, Bas J. Zwaan, and Paul M. Brakefield ............................................. 445

Forest Habitat Conservation in Africa Using Commercially Important
Insects
Suresh Kumar Raina, Esther Kioko, Ole Zethner, and Susie Wren ......................... 465

Systematics and Evolution of Heteroptera: 25 Years of Progress
Christiane Weirauch and Randall T. Schub ............................................................... 487