5

Analytical Approaches to Studying Climate/Disease Linkages

A wide variety of approaches are used by researchers to gain an understanding of the relationships between climate and infectious disease. This includes observational-based analyses of *past* or *present* events in nature, model-based projections of possible *future* events, and cross-cutting methods such as risk assessment and integrated assessment. These techniques differ in temporal or spatial extent, source and nature of the data used, and the tools and processes of analysis. The inferences that can be drawn from the different approaches may not be similar in strength or predictive power; some are more relevant to generating hypotheses, others to testing them. Ultimately, each of these different approaches is important in informing the other. The strengths and weaknesses of all of the different approaches are discussed below.

OBSERVATIONAL AND EXPERIMENTAL STUDIES

Observational studies of the relationship between climate variations and health outcomes provide the foundation for developing theory that is used in models and, ultimately, for making forecasts about future impacts associated with climatic changes. As described in the following sections, observational studies can include retrospective and prospective analysis of natural variations, retrospective analysis of historical trends, and interregional comparisons.

Retrospective Analysis of Natural Variations

This approach treats past temporal patterns of climate variability and disease as empirical analogs of future changes. Using time-series analysis of fluctuations or extremes in specific climate variables, historical patterns are compared with those of ecosystem changes or disease outbreaks, with an appropriate time lag sometimes factored in. In recent years this approach has been widely used to study the effects of ENSO phenomenon, especially El Niño events, on infectious disease patterns.

Typically, this analytical approach is employed to look at a single region, so spatial or spatio-temporal variations in the climate/disease relationship are usually not considered. One critical limiting feature of this retrospective approach is that many years of comparable environmental and disease incidence data are usually needed to have confidence that apparent patterns are not occurring by chance. Often, disease surveillance data are inadequate, reducing the applicability of this approach. In addition, the observational nature of this method makes it difficult to separate the influences of other ecological or social changes from those of climate.

For all of these reasons one must be cautious about interpreting the findings of these analog studies and extrapolating the results beyond the specific context of any particular situation. For instance, studies of the effects of an El Niño event do show some of the ways that short-term climate variations can affect epidemic disease, but they are not necessarily a good analog of future long-term climate change. In general, though, this analytical approach does hold potential for improving forecasts of how short-term variability may alter epidemic risk; and if consistent relationships are found over a long time period or in many different places, more confidence can be gained in using these relationships to forecast future changes.

Prospective Observations of Natural Variations

Under some circumstances, surveillance of diseases may be ongoing during periods of anomalous weather events, thus allowing for "prospective" comparison of patterns of variability in disease incidence and climate. Sometimes this involves chance or good fortune in which health surveillance and climate observations at the relevant spatial and temporal scales are already being made to address other questions (e.g., Lindblade et al., 1999). In other cases, intentional focused observations can be planned when a particular climatic event is expected. During the 1998 El Niño, for example, NOAA's Office of Global Programs requested intensified disease and meteorological observations in various sites where surveillance efforts were ongoing, thus creating a prospective sampling and analysis of variability.

A potential pitfall that must be considered in such studies, however, is that one may find higher rates of disease incidence during the period under study, simply as a result of the intensified surveillance efforts. Critical to the interpretation of such observations is a historical record of disease and climate patterns with which to compare each new anomaly. Without such comparison it is impossible to draw inferences from an association of a single event. Prospective observations are needed to test whether hypothesized associations are accurate. These are all reasons why strengthened surveillance of disease incidence is critical to our capacity to analyze future health impacts from climate variability.

Retrospective Analysis of Historical Trends

This approach is similar to that involving retrospective analysis of natural variation, except that it compares the trends or slopes of change during the period of observation. Causal inference is based on the direction and strength of average change over time in both the environmental and disease incidence measures. For instance, Tulu (1996) analyzed malaria trends in the highland region of Ethiopia over the past decade and concluded that increases in nighttime temperatures expanded the altitudinal range of malaria transmission and increased the rate and duration of transmission in areas that were previously epidemic-prone.

This analytical method usually requires long time series of observations, as changes are often slow and interannual fluctuations may confound the trend. Also, the role of climate may be confounded by other trends occurring over the same time period, such as improved sanitation and public health practices; evolution of pathogenicity; land-use changes; or shifts in immunity, age structure, or mobility of the population being studied. However, if similar climate/disease trends are observed in many different regions, this could provide strong insights into underlying relationships and potential future changes.

Interregional Comparisons of Natural Patterns

Comparison of natural spatial patterns in disease incidence and climate is an approach that seeks consistent similarities and differences among regions over a specified time period. If climate patterns are related to disease incidence in a manner that is similar among some regions but different from that observed in other regions, this may help us infer how disease incidence will be affected if the climate in a particular region changes in the direction of that observed elsewhere. An example of this spatial analog approach is a study by Reeves et al. (1994) that took advantage of a 5°C temperature differential between two valleys in California to make inferences about the importance of ambient temperature on the seasonal activity of the main mosquito vector of western equine encephalitis and St. Louis encephalitis.

This construct of inferring possible temporal changes from observed spatial patterns generally assumes that other non-climate conditions are similar in the regions under study. In practice, the role of other environmental and socioeconomic variables is very difficult to identify and disentangle. Different historical patterns of environmental determinants on the disease pattern, and the likelihood of spatial auto-correlation, will complicate reasoning during this process.

Box 5-1
Interdisciplinary Research and Training Needs

The study of climate and human health linkages is highly interdisciplinary in nature, and sustained interdisciplinary research will be absolutely necessary to generate robust understanding of these linkages. While the study of climate itself involves a wide variety of disciplines, these all generally have a common basis in the physical sciences. In contrast, the study of climate/health linkages includes disciplines from the physical, biological, and social sciences. This type of collaboration is generally more difficult because there are fewer common underlying principles and research methods.

There can be numerous impediments to conducting high-quality interdisciplinary research, including difficulties in communicating across disciplines, a dearth of supporting institutional structures such as scientific societies and professional journals, inadequate or nonexistent reward systems at universities and other research institutions, funding structures that are inimical to such research, and biases that value detailed single-disciplinary studies over more generalized interdisciplinary studies.

Interdisciplinary research can be greatly facilitated by the creation and support for university centers and the development of interdisciplinary funding programs. Over the past 30 years or so, many such programs have been developed but in some cases were crippled by a lack of adequate funding. In recent years a few funding programs have been developed specifically to encourage the formation of interdisciplinary research teams, including programs related to the issue of climate and health. These efforts should be encouraged and supported. One concern, however, is that such programs are usually based on three- to five-year funding cycles, which may be too short for developing and executing complex research projects.

Development of long-term interdisciplinary programs offers opportunities for training individuals, including graduate and postdoctoral students, in a variety of methods and activities of value to interdisciplinary research such as learning the basics of disciplines outside one's primary area of expertise, dealing with different types of data analysis, learning how to effectively combine quantitative and qualitative information, and developing flexibility in methodological approaches. This type of training will help ensure that creative interdisciplinary approaches continue to be applied by scientists studying the issue of climate and human health.

Experimental Studies

Our current understanding of the dynamics of many infectious diseases reflects a long history of experimental science ranging in scale from laboratory-based reductionist studies at the molecular and organism levels to field studies at the population level. Manipulative studies are aimed at understanding the mechanism by which particular environmental variables impact parts of an infectious disease transmission system. They can help elucidate parts of a causal pathway and provide information that may be used in combination with other observations or to quantify particular associations.

An example of a field-based experimental manipulation study is an investigation of Rift Valley fever in Kenya carried out by Linthicum et al. (1999). The mosquito vector for Rift Valley fever is known to be associated with flooded areas known as "dambos"; to study the ecology of the disease, researchers intentionally flooded a dambo area and then monitored the resulting insect/virus response. An example of a laboratory-based experimental study is an investigation by Schulman and Kilborne (1963) of the environmental factors that affect influenza transmission, where it was found that the infection rate of mice exposed to the influenza virus was affected by the ambient relative humidity levels.

Interpretation of results from experimental manipulation studies is facilitated when there are multiple observations that can be compared with many "controls." This is usually possible to do in laboratory investigations, but numerous field observations are usually difficult to obtain. As with other field studies, variation in time may be important, making longer studies more likely to produce the most useful information. By controlling for some variables, experimental manipulations limit complex interactions that might actually occur under normal conditions. For these reasons, experimental environmental studies have limited utility in establishing causation or in forecasting. They can be quite valuable, however, in setting parameter values for some processes and in developing hypotheses that are better tested through other approaches.

MATHEMATICAL MODELING

Mathematical modeling can be a powerful tool for gaining insights into the dynamics of infectious disease epidemics, integrating information from laboratory and field studies, providing direction for future experimental and observational studies, evaluating monitoring and control strategies, and making predictions about future disease risk.

As discussed in Chapter 3, the SEIR framework is at the heart of many models of infectious disease transmission dynamics. Yet this is only part of a larger modeling framework that is needed to fully represent the dynamical relationships among climate, ecosystems, and infectious diseases. A comprehensive representation of these relationships needs to account for interactions among factors affecting the disease agent, the vectors and intermediate hosts of these disease agents, the human host (including SEIR parameters), and the environment. Modeling such a complex system presents an extremely challenging task that has yet to be fully met. Progress has been made in recent years, however, in the development of models that simulate parts of this overall system. The different types of modeling approaches used for such studies can be roughly classified into two categories: mechanistic models and empirical-statistical models.

Mechanistic models (or process-based models) use theoretical knowledge of underlying biophysical mechanisms to simulate the health impacts of changes in climate. The models incorporate mathematical equations to represent processes that, in theory, can be applied universally to similar systems in different environments. Mechanistic models are dynamic in that quantitative interactions among multiple variables and feedback processes can be explicitly considered. Forecasted changes in disease risk are based on current interactions of physical and biological variables; thus, most process-based models have not considered various kinds of adaptation or evolution in the many factors that determine transmission or host response. Nevertheless, this approach has been used to forecast how changing climate conditions could lead to ecological changes and affect disease patterns. This approach is useful for exploring alternative scenarios but does not easily convey the uncertainty inherent in forecasts.

Empirical-statistical models are based on relationships between climate and disease-related variables that have been estimated from observational studies. Empirical data on past patterns of variation are used to project how the studied variables may change in the future. This approach may range from applying simple indices of risk (e.g., identifying the minimum temperature threshold for malaria transmission) to using complex multivariate models that consider numerous environmental factors affecting risk. One limitation of this modeling approach is that there may be limited data points to calibrate the projections, making results difficult to validate. Generally, this approach relies little on underlying mechanisms and is not explanatory in nature; often these models are not easily able to consider interactions or other variables not included in the available observations. Regardless, these models can be useful tools for developing hypotheses that indicate where prospective observations will be useful. An advantage of this approach is that empirical-statistical models are often simpler to use and less "data demanding" than mechanistic models.

The two basic classes of models discussed above are not mutually exclusive. Some models use mechanistic approaches to represent the essence of the process under study, combined with statistical approaches to extend the process to the population level. No single approach is clearly superior (or is likely to be sufficient) for creation of reliable predictive models. The most appropriate approaches depend largely on the type of data available and the type of output needed for decision making. In general, however, there are several basic features that are desirable for any type of model used in studies of climate and infectious disease:

- *Small number of parameters.* In general, the simpler a model is, the better we understand its implications. The assumption that more complex models will more closely approximate reality is not justified in cases where there is a paucity of data on relevant "measureables." Models that use a smaller number of parameters by aggregating variables involved in the true process can narrow the range of possible outcomes without implying more knowledge than is available.
- *Well-understood dynamical behavior.* Disease transmission models can be intrinsically nonlinear and may sometimes be chaotic in behavior. Whether

Box 5-2
Modeling the Impacts of Climate Change on Vector-Borne Infectious Diseases

Numerous models have been developed to match the presence of vector species with a discrete range of meteorological parameters and to then project the effects of climate change on vector redistribution (e.g., see Rogers and Randolph, 1993; Sutherst et al., 1995; Lindsay et al., 1998). Discussed below are two recent studies of how future climate change could affect global patterns of potential malaria risk. These studies exemplify the different modeling approaches that can be used to address such questions and also illustrate the widely differing results that can be obtained.

Martens et al. (1999) developed a mechanistic model of the biological processes by which temperature affects mosquito development, feeding frequency and longevity, and the incubation period of the malarial parasite in the mosquito. The suitability of vector habitats is determined by minimum precipitation levels. The model calculates relative changes in the "transmission potential" (the reciprocal of the vector/host ratio necessary to maintain malaria transmission), and the ratio of future to present transmission potential for any given region is taken to indicate the relative degree of future malaria risk. This model predicts that climate change could lead to increases in malaria risk in high-latitude regions (e.g., in Africa and South America), and could also lead to significant increases in malaria risk (from a very low baseline) for much of the United States, Europe, and middle Asia, areas where the malaria vector is currently present but development of the parasite is inhibited by low temperature.

Rogers and Randolph (2000) explored similar questions using a multivariate statistical model. The present-day distribution of malaria was used to empirically establish how the disease is currently constrained by the mean and covariances of meteorological factors including temperature, precipitation, and atmospheric humidity. These statistical relationships were then used to predict potential malaria distribution under future climate change scenarios. Using this approach they estimated only very minor geographical extensions of the potential malaria distribution as compared to the present day; in some areas malaria was predicted to diminish. The reason for these modest changes is that the covariation of the different meteorological factors limits potential expansion of malaria transmission in many regions.

All such models have their specific disadvantages and advantages. The equations in a global model may be inappropriate for particular local conditions. While some models do attempt to take account of local/regional conditions, these studies clearly cannot include all factors that affect species distributions. For example, local geographical barriers and interaction/competition between species are important factors that determine whether species colonize the full extent of suitable habitat. Likewise, the relationship between vector-borne disease incidence and climate variables is complicated by many factors related to a community's socioeconomic status and lifestyle. There is a clear need to validate models on a local or regional scale using historical data. Unfortunately, though, historical data on vector and disease distribution are often not available, especially in many poor countries.

this reflects the actual nature of the system being modeled or is simply an unintended artifact, it is important that the dynamical behavior of a model be well understood.

- *Proper consideration of temporal/spatial scales.* A clear delineation of the temporal and spatial scales at which a model works is essential for its success. Availability of the data at appropriate scales should be taken into account when building a model. For instance, one impediment to using mechanistic models that mimic complex biological and ecological processes is the reliance on micro-environmental parameters that are seldom available in observational data.
- *Incorporation of experimental data and expert opinion.* Scientists have learned a great deal about the biology of various disease vectors through field studies and controlled laboratory experiments. It is important that this information be incorporated into models. Likewise, in cases where there are few experimental data to work with, it is useful to develop modeling approaches that can incorporate qualitative-type information and expert opinions or judgments.
- *Explanatory in spirit but predictive in behavior.* Forecasts generated by disease models will presumably be used for designing effective monitoring strategies and for making public health policy decisions. For such purposes a model that is reasonably good at prediction but somewhat less precise in explanation may still be quite useful. The model has to be appropriate for the purpose but not necessarily a perfect reflection of the complex processes occurring in nature.

There are also numerous issues that must be considered in terms of relating models to available data. Some of the particular challenges that must be commonly addressed in epidemiological modeling are the following:

- *Measurement errors.* Measurement errors are an important concern in many ecological studies. For instance, suppose one wants to develop a population dynamics model to study how mosquito abundance changes over time as a function of environmental variables. It is not possible to actually count all of the mosquitos, but there are various methods that can be used to estimate population abundance. If the model does not explicitly account for the measurement errors associated with these estimates, there can be substantial biases in the inferences drawn from the data.
- *Reporting bias.* This is an important concern in many epidemiological studies. Often, one region will have an effective surveillance/reporting mechanism for a particular disease while other regions do not. It can be difficult to compare the data yielded by these different reporting standards. Sometimes it may be possible to adjust for different reporting standards by modeling the reporting bias explicitly. It is important that reporting biases are either properly estimated or reduced by mandating uniform standards for data collection and reporting.

- *Aggregation bias.* This is another common problem in ecological and public health studies. For instance, in a study of risk factors associated with infant death rates, data analyzed at the city-block level showed a positive correlation between minority population and infant death rate, whereas the same data aggregated at the census-tract level showed a negative correlation (Richard Hoskins, personal communication). This illustrates the importance of considering spatial scale in data analysis and modeling.
- *Data availability and scale.* Although it is often preferable to build models with fine scales of resolution, this requires high-resolution observational data, which is not always available. When information on crucial variables is not available at the appropriate scale, or when data on covariates in a model are available at differing levels of resolution, this can greatly hinder the process of model fitting and validation.
- *Meta-analysis.* One approach that can be used to compensate for a paucity of epidemiological data is to conduct a "meta-analysis" of data pooled from numerous studies. This approach is difficult to apply when there are few studies that cover similar temporal/spatial scales or that focus on similar driving forces and outcomes. However, the field of meta-analysis offers strategies for dealing with some of these difficulties, and in a more qualitative sense, meta-analysis can be useful for looking at differences among studies and providing guidance as to where more testing is needed.

Proper validation is critical for infectious disease models, due to their complexity and potential for nonlinear behavior. But given that all models are only an approximation of a real process, how do we determine whether one model is more appropriate than another? In other words, how does one choose which model is the best approximation of the truth if the truth is unknown? Model validation criteria must balance the fact that a larger number of parameters can mean less reliable parameter estimates on the one hand but more success at mimicking the observed data on the other hand.

Cross-validation procedures provide one means for testing the goodness of prediction. In this approach one of the data points is removed, and the model tries to predict it using the rest of the data. This procedure is repeated for all of the data points, and a summary measure of the performance is reported. The model that has the smallest average squared prediction error is deemed the best in the class. This procedure accounts for the estimation cost associated with the parameters as well as the benefit in prediction due to the complexity of the model.

Developing predictive epidemiological models is an important goal, because "experimental" forecasts of disease outbreaks can provide a valuable means to test our understanding of the linkages among climate, ecosystems, and infectious disease. It is important to recognize, however, that even if a predictive model is feasible from a scientific perspective, that does not necessarily make it appropri-

Box 5-3
Comparison with Other Types of Climate Impact Modeling Studies

It is instructive to compare the models used to study climate impacts on infectious diseases with the models used to study climate impacts in other realms. For instance, simulations of climate impacts on agricultural crops and on natural ecosystems typically use dynamic, mechanistic models in which detailed processes are simplified using empirical relationships and aggregation of variables. The fundamentals of these models are generally considered valid over a wide range of contexts, but location-specific input is often required for some parameters. Infectious disease models often use this same general approach, and yet for a variety of reasons, can pose a much more complicated scientific challenge. For instance, infectious disease models must account for numerous social processes that are not constrained by well-defined physical laws and are often highly location specific. Also, in many cases, key relationships between environmental parameters and the parameters characterizing disease agents, vectors, and hosts are not well quantified, either because there are insufficient data to empirically define these relationships, or because the underlying mechanisms linking these parameters are not well understood.

ate as a real-world management tool. As discussed in Chapter 7, there can be risks and substantial costs associated with issuing disease warnings and mobilizing intervention strategies. Before decision-makers could use model predictions as a basis for any such actions, the reliability and precision of these predictions would need to be thoroughly understood.

RISK ASSESSMENT FRAMEWORKS

Risk assessment can be defined as the characterization and estimation of potential adverse health effects associated with exposure of an individual or population to hazards. Risk assessment is used by the regulatory community for assessing environmental contaminants and the risks they pose to human and ecosystem health. The National Research Council (NRC, 1983) has defined a four-stage process for risk assessment: hazard identification, dose-response, exposure assessment, and risk characterization. This approach has been used by the U.S. Environmental Protection Agency to develop controls for key waterborne disease agents (U.S. EPA, 1989; Regli et al., 1991; Rose and Gerba, 1991), and has been used to examine the probability of infection and disease resulting from exposure to a variety of pathogenic microorganisms, primarily focusing on fecal-oral and waterborne pathogens (Haas et al., 1999).

Hazard identification refers to both the identification of the microbial agent and the spectrum of human illnesses associated with the specific microorganism (which can range in seriousness from asymptomatic infections to death). Clinical and surveillance data are used to describe what microorganisms are causing what diseases, and quantitative analyses are undertaken to determine the spectra of disease outcomes.

Dose-response refers to the mathematical characterization of the relationship between the dose administered and the probability of infection or disease in the exposed population. In dose-response studies, various doses of specific microorganisms are given to sets of human volunteers, and the percentage of individuals infected at each dose is fit to a best-fit curve. Pathogen/host dose-response data for numerous pathogens have been summarized by Haas et al. (1999).

Dose-response relationships may vary depending on the type of microorganism and the type of exposure. For many pathogens the dose-response relationship is linear only at low doses, and at some point the concentrations are high enough to cause a flattening out of the curve, which reflects very high levels of probability of infection and would result in a large number of symptomatic cases. Most dose-response studies focus on water-borne disease agents, with exposure through an ingestion route. Determining a dose-response relationship for vector-borne diseases may be more difficult. For instance, since high doses of viruses or parasites are likely delivered in a single mosquito bite, the response may simply be dichotomous.

Exposure assessment is aimed at determining the size and nature of the population exposed and the duration of its exposure. This is governed by numerous aspects of a population's overall vulnerability, for instance, how much time is spent in environments where the vector is present, whether homes provide adequate protection from mosquitos and other vectors (e.g., with window screens), and whether there is access to clean drinking water, vaccines, and other public health protections. Exposure assessment also includes characterizing the microorganisms' occurrence (concentration), prevalence (how often the microorganisms are found), and distribution in space and over time, which in turn requires assessing the environmental factors that influence the microorganisms' survival, bioconcentration, and transport.

Quantitative risk characterization is aimed at estimating the magnitude of the public health problem and understanding the variability and uncertainty of the risk. Techniques such as Monte Carlo analysis are used to examine the distribution of exposures and outcomes from the individual to the population level. Dose-response and exposure estimates can be used as input for epidemiological models to describe the transmission in a population and take into account such factors as incubation times, secondary spread, and immunity (Eisenberg et al., 1996).

This framework can potentially be used to study the influence of climate and other environmental factors on disease risk, primarily through their influence on exposure. The first step would be to define and describe the environmental factors that are related directly to the pathogen of concern and the disease outcome. The next step is to evaluate how environmental factors influence the concentration, distribution, prevalence, viability and/or virulence of the pathogen. It is likely that several "probability functions" would be of interest—for instance, the probability that any given dose (e.g., glass of water, mosquito bite) contains the pathogen; the exposure frequency in some unit of time (e.g., number of glasses of water per day, number of mosquito bites per hour); and the distribution of the concentrations of pathogens in the dose(s). In some cases, studies can be designed to relate environmental factors directly to disease or probability of infection, such as the example mentioned earlier where mice were exposed to influenza virus at varying levels of relative humidity to examine how this affected infection rates.

The case of *Vibrio cholerae* provides a useful example of the types of information that might go into an exposure assessment on a climate-sensitive pathogen. *V. cholerae* is spread by the fecal-oral route through contaminated water and food, and as discussed in Chapter 4, there is evidence that *V. cholerae* naturally occurs in estuarine environments in association with copepods. It is hypothesized that the bacteria are then transmitted to humans by copepods getting into water supplies in coastal communities or possibly the food chain. Colwell (1996) has shown that *V. cholerae* can be associated with these zooplankton, and this relationship exhibits a seasonal pattern that can be linked to the seasonal occurrence of cases of disease. Two of the primary environmental variables influencing this complex ecology of pathogenic *Vibrios* are sea surface salinity and temperature. Table 5-1 lists some of the environmental factors that

TABLE 5-1 Environmental and Climate Factors Associated with Exposure to *Vibrio cholerae*

Types of Data/Factors to be Studied	Possible Approaches
Chlorophyll a and/or turbidity associated with rainfall events	Measured by remote sensing
Influence of rainfall and runoff on salinity and nutrients leading to algal blooms. Influence of temperature on phytoplankton growth.	Population dynamic modeling
Influence of phytoplankton on zooplankton dynamics and succession.	Population dynamic modeling
Influence of temperature and salinity on the growth of Vibrio in the copepod	Bacterial growth curves
Concentration of the bacteria in the copepod	Plankton sampling and immunofluorescence microscopy
Numbers of copepods transmitted upstream and in a glass of water	Straining and examination of the water supply

can affect exposure to *vibrios* and the opportunities to monitor and study these factors.

A dose-response model for *V. cholerae* has been developed by Hornick et al., (1971), and this information could be used together with an exposure assessment to estimate the probability of infection and the potential numbers of cases of illness in a population. Such studies, however, must account for the fact that the surveillance and reporting of *Vibrio* cases of illness generally underestimate the true level of infection and risk to the population: and also that time lags of up to several months can exist between environmental changes (such as a rainfall event) and final exposure to the pathogen.

Risk assessment can help quantify the relationships between environmental factors, exposure, and illness. This framework can be used to help determine the combination of factors that contribute to risk and that would be most useful to monitor for the purpose of providing early warnings.

INTEGRATED ASSESSMENT

Integrated Assessment (IA) can be defined as a structured process of using knowledge from various scientific disciplines and/or stakeholders such that integrated insights are made available to decision-makers (Rotmans, 1998). IA has emerged as a new approach for meeting two central challenges of global change research: (1) adequately characterizing the complex interactions and feedback mechanisms among the various facets of global change; and (2) providing support for public decision making through a framework for testing the effectiveness of various policy strategies (Rotmans and van Asselt, 1999). IA provides a coherent framework for working through the causal chain from climate dynamics to climate impacts to policy response strategies. It is an iterative process wherein insights from the scientific and stakeholder communities are communicated to the decision-making community, and, in turn, the evolving informational needs of decision makers provide input for future research.

A wide variety of research methods fall under the rubrick of IA. Included are analytical methods traditionally rooted in the natural sciences such as risk analysis and mathematical modeling. Also included are a variety of "participatory" methods rooted in the social sciences, which aim to involve non-scientists as stakeholders in the process and to facilitate stakeholder-scientist interactions.

Chan et al. (1999) developed a conceptual IA framework to identify key linkages among the different systems affected by climate change and that in turn affect epidemiological outcomes; these include changes in transmission biology, ecological changes, and sociological changes. They found that only some of the linkages within this framework have received much attention from the research community. For instance, there have been numerous studies to estimate the direct impacts of temperature on transmission biology, but comparatively little work has been done to integrate climate-related ecological and sociological factors. As noted by the authors, this type of analysis can help identify the different pathways through which infectious diseases are affected by climate and ecological changes, but it does not necessarily help us understand which factors are most critical in any particular context.

There have been numerous IA mathematical models developed for studying global climate change (see Schneider, 1997, or Parson and Fischer-Vanden, 1997, for an overview), including models for assessing climate-related health risks (Martens, 1998). These models try to quantitatively describe as much as possible of the relevant cause-effect relationships (vertical integration) and interactions between different processes (horizontal integration), including feedbacks and societal adaptations. These models integrate global climate change scenarios with local socioeconomic and environmental factors into a coherent modeling framework based on variables describing climate, vectors, parasites, human populations, and health impact. Since infectious disease transmission dynamics are complex systems that can display spontaneous or socially based adaptive responses, some researchers have attempted to develop algorithms that incorporate a capacity for adaptive change and "learning" to simulate such processes (Janssen and Martens, 1997; Sethi and Jain, 1991). It is inevitable that some natural and social phenomena will be oversimplified with such approaches; however, integrated models can provide a useful complement to more focused models that provide highly detailed representation of these complex processes.

As recently summarized in Rotmans and Van Asselt (1999), the primary advantages of IA methods are that they can help to:

- put complex issues in a broader context by exploring their interrelations with other issues;
- assess potential response options to complex problems, including cost-benefit analyses;
- identify and clarify different sources of uncertainty in the cause-effect chain of a complex problem;
- assist in decision making by putting uncertainties into the framework of risk analysis;
- set priorities for future research by identifying and prioritizing knowledge gaps;
- enhance communication between scientists of many disciplines and between scientists and decision makers.

The weaknesses of integrated assessment methods (in particular, of IA models) include the following:

- *High level of integration.* Many processes of relevance to infectious disease dynamics occur at a micro level, far below the spatial and temporal aggregation of current IA models.

- *Inadequate treatment of uncertainties*. IA models are prone to an accumulation of uncertainties that arise from incomplete knowledge about future climatic changes and health-climate relationships, and the inherent unpredictability of future geopolitical, socioeconomic, demographic, and technological evolution.
- *Absence of stochastic behavior*. Most IA models describe processes in a continuous deterministic manner, excluding the potential effects of extreme conditions or "chaotic" events that may significantly influence the long-term behavior of a system.
- *Limited calibration and validation*. The high level of aggregation implies an inherent lack of empirical variables and parameters, and current datasets are often too small and/or unreliable to apply.

SURVEILLANCE/OBSERVATIONAL DATA NEEDS

Epidemiological Data

All of the analytical methods discussed in the previous sections depend highly on the availability of surveillance and observational data. One of the most critical obstacles to improving our understanding of climate/disease linkages is the lack of high-quality epidemiological data on disease incidence for many locations. These data are needed to establish an empirical basis for assessing climate influences and to develop and validate predictive models. Researchers need long-term datasets in order to establish a baseline against which to detect anomalous changes. Likewise, researchers need datasets with broad geographical coverage in order to do comparative studies of how climate affects transmission of disease agents in different contexts.

Currently, most major public health institutions rely primarily on "passive" disease surveillance systems. For instance, U.S. government agencies such as the Centers for Disease Control and Prevention (CDC) and Department of Defense (DoD) rely on state/local public health personnel to submit surveillance data voluntarily (mostly from the records of hospitals and physicians offices); the situation is similar on the international level, where the World Health Organization relies on individual countries to collect and report national data on disease occurrence. In many locations, data are simply not collected since there is no mandate to do so. In other locations, data that are collected are not publicly reported out of a fear that reports of an outbreak will cause a loss of tourism or other economic hardships. This leads to problems of reporting bias, in which disease outbreaks appear worse in some regions simply because more thorough surveillance and reporting are done in these areas. Also, with no universal guidelines or control over how data are collected, it is very difficult to assure quality control or to compare datasets from different sources.

Awareness of these problems and growing concerns about emerging infectious diseases have prompted efforts to develop new programs such as DoD's

Global Emerging Infections Surveillance and Response System. This program and others within the CDC are taking a more active role in monitoring disease incidence by providing standardized criteria for collecting data and developing electronic databases to facilitate rapid standard reporting and sharing of these data. Such efforts should be encouraged on both national and international levels.

Another impediment is that disease surveillance data collected for particular scientific studies are usually considered to be the property of the researchers who collected them and only become available if and when published in the scientific literature. There are a variety of steps that can be taken (and in some cases are already in place) to encourage investigators to make their datasets publicly available. For example, funding agencies can require that research proposals include a plan for data dissemination, scientific journals can make data accessibility a condition for publication of a study, and submission of data to a central database could be counted as an official publication for the researcher.

Technological developments such as the advent of the Internet and steadily decreasing costs of computing and communications are providing exciting new opportunities for sharing data among researchers all over the world. Yet at the same time there are political and economic developments that jeopardize the open exchange of data among scientists, especially across national boundaries. Data exchange is affected by governmental concerns related to national security, foreign policy, and international trade and an increasing push toward commercialization of electronic databases. These constraints may present serious difficulties in attempts to study regional- and global-scale infectious disease patterns.

A more daunting challenge is the fact that many countries lack the resources to maintain even the most basic population-based disease surveillance programs. For the foreseeable future, the establishment of comprehensive epidemiological surveillance is not likely to be an attainable goal in such regions. It may be possible, though, to target resources for intensive surveillance in specific locations where the populations are most vulnerable to infectious disease threats. Such programs may require substantial assistance from international organizations, but in order to maintain surveillance systems over the long term, it is important to develop each country's internal capacity and intellectual/technological infrastructure.

It must also be recognized that for many diseases even the best surveillance systems will capture only a fraction of the infections that occur in a given population. For instance, in developing countries most individuals who develop dengue fever never seek care and those who do are rarely tested to confirm or rule out dengue. Likewise, individuals with cholera or malaria in many developing countries often do not seek medical care. For the purpose of understanding climate/disease linkages, however, much can still be learned by studying *relative* changes and patterns of disease transmission. Surveillance programs that capture only a fraction of the actual disease incidence can still provide useful information as long as consistent data collection methodologies are used over time or across regions.

Environmental Observational Data

In order to relate changes in disease incidence to environmental changes, epidemiological surveillance must be coordinated with meteorological and ecological observations. Ground-based measurements of meteorological parameters such as daily mean temperature, daily accumulated precipitation, and relative humidity are collected at thousands of sites around the world. There are still some regions of the world, however, where weather observations are quite limited or where there have been significant declines in meteorological reporting in recent years. It is important that these networks be maintained, as they provide vital information both about climate/disease linkages and about climate variability itself.

For the past several decades, meteorological observations have been processed and archived on a global basis through a system of world/regional meteorological data centers. The World Weather Watch, part of the United Nations' World Meteorological Organization, is responsible for planning and coordination of the system, but the national meteorological system of each member nation is responsible for actually operating the data collection and processing centers. This type of international system for collecting and sharing scientific data may provide a good working model for the public health community to follow.

Routine observations of ecological parameters such as soil moisture, vegetative cover, and sea surface temperature are rare, but new remote-sensing technologies are rapidly expanding the opportunities to monitor many of these parameters on a global scale (as discussed further in the following section). However, disease dynamics are predominantly driven by environmental factors at fine spatial and temporal scales (e.g., mosquitos are affected more by temperature in their immediate microenvironment than by mean ambient conditions), and thus methods must be developed for relating remote-sensing data to conditions occurring on these smaller scales.

For both environmental monitoring and public health surveillance, a central challenge is maintaining observations of key parameters over the course of years to decades. Such extended time series are needed in order to establish a baseline against which one can detect long-term trends or patterns of variability.

Remote-Sensing Surveillance Tools

In recent years investigators have begun using remotely sensed data collected by satellite-borne instruments to study environmental factors that can influence disease transmission risk for particular regions. Over the next several years dozens of new instruments will be launched that will provide high-resolution monitoring for a wide variety of environmental parameters relevant to disease risk (Beck et al., 2000).

One of the instruments being used in this capacity is the Advanced Very High Resolution Radiometer (AVHRR), which is carried aboard NOAA-series satellites. AVHRR can be used to calculate a Normalized Difference Vegetation Index (NDVI, or "greenness index"), a tool for monitoring changes in vegetation growth (Justice et al., 1985). This is a widely useful parameter to monitor, since nearly all vector-borne diseases are linked to the vegetated environment during some aspect of their transmission cycle. AVHRR has been in continuous operation since 1981, and this multi-decade time series makes it possible to calculate statistically significant variations from the long-term mean value. AVHRR provides data with high enough resolution (1 kilometer) to be useful for studies on fine spatial scales. Scientists have used NDVI data to identify areas of potential Rift Valley fever activity in regions of eastern Africa (Linthicum et al., 1999), and similar techniques have been used to study other diseases such as trypanosomiasis, schistosomiasis, and malaria.

Soil moisture is an important parameter for evaluating the suitability of a particular habitat for disease vectors such as mosquito larvae, ticks, and snails. NDVI provides an indirect proxy for soil moisture, but instruments that can directly measure soil moisture (e.g., microwave sensors, synthetic aperture radar) also are in operation. Instruments that can "see through" heavy vegetation will be especially useful since, at present, soil moisture for regions such as tropical rain forests can be obtained only by surface sensors.

Remote-sensing data can also be used to aid the study of water-borne disease. For instance, Lobitz et al. (2000) have studied the relationship between cholera outbreaks and sea surface temperature (from the NOAA/AVHRR instrument), sea surface height (from the TOPEX /Poseidon radar altimeter), and chlorophyll concentration (from the OrbView-2/SeaWIFS instrument). Satellite instruments such as SeaWifs are also used to provide information about harmful algal blooms and toxic elements that can infect shellfish.

In recent years several major programs have been established to help foster the use of remote sensing technologies by the health care community, such as MALSAT (Environment Information Systems for Malaria and Meningitis) of the Liverpool School of Tropical Medicine, the Environment and Health Initiative of NASA's Earth Science Enterprise, the INTREPID (International Research Partnership for Infectious Diseases) program, and NASA's Center for Health Applications of Aerospace Related Technologies (CHAART). Through a series of workshops involving scientists from the environmental science and health care communities, CHAART has compiled a comprehensive list of remotely sensed factors that can help in the study of disease (see Table 5-2) and the current/planned instruments that will measure these parameters.

Geographic Information Systems (GIS), a framework for analyzing geographically referenced data, can greatly facilitate integration of remotely sensed parameters with health data. For instance, maps of potential vector habitat obtained by analysis of remote-sensing data can be "layered" with geographically referenced data on land use, human population, and so forth, to create maps of disease risk. For further discussion of the ways in which these technologies can be applied in the field of epidemiology, see Hay et al. (2000).

TABLE 5-2 Potential Links Between Remotely Sensed Factors and Disease

Factor	Disease	Mapping Opportunity
Vegetation/crop type	Chagas disease	Palm forest, dry and degraded woodland habitat for triatomines
	Hantavirus	Preferred food sources for host/reservoirs
	Leishmaniasis	Thick forests as vector/reservoir habitat in Americas
	Lyme disease	Preferred food sources and habitat for host/reservoirs
	Malaria	Breeding/resting/feeding habitats; crop pesticides vector resistance
	Plague	Prairie dog and other reservoir habitat
	Schistosomiasis	Agricultural association with snails, use of human fertilizer
	Trypanosomiasis	Glossina habitat (forests, around villages, depending on species)
	Yellow fever	Reservoir (monkey) habitat
Vegetation green-up	Hantavirus	Timing of food sources for rodent reservoirs
	Lyme disease	Habitat formation and movement of reservoirs, hosts, vectors
	Malaria	Timing of habitat creation
	Plague	Locating prairie dog towns
	Rift Valley fever	Rainfall
	Trypanosomiasis	Glossina survival
Ecotones	Leishmaniasis	Habitats in and around cities that support reservoir (e.g., foxes)
	Lyme disease	Habitat for deer, other hosts/reservoirs; human/vector contact risk
Deforestation	Chagas disease	New settlements in endemic-disease areas
	Malaria	Habitat creation (for vectors requiring sunlit pools)
		Habitat destruction (for vectors requiring shaded pools)
	Yellow fever	Migration of infected human workers into forests where vectors exist
		Migration of disease reservoirs (monkeys) in search of new habitat
Forest patches	Lyme disease	Habitat requirements of deer and other hosts, reservoirs
	Yellow fever	Reservoir (monkey) habitat, migration routes
Flooded forests	Malaria	Mosquito habitat
Flooding	Malaria	Mosquito habitat
	Rift Valley fever	Flooding of dambos, breeding habitat for mosquito vector
	Schistosomiasis	Habitat creation for snails
	St. Louis encephalitis	Habitat creation for mosquitos
Permanent water	Filariasis	Breeding habitat for Mansonia mosquitoes
	Malaria	Breeding habitat for mosquitos

Flooded forests	Malaria	Mosquito habitat
Flooding	Malaria	Mosquito habitat
	Rift Valley fever	Flooding of dambos, breeding habitat for mosquito vector
	Schistosomiasis	Habitat creation for snails
	St. Louis encephalitis	Habitat creation for mosquitos
Permanent water	Filariasis	Breeding habitat for Mansonia mosquitoes
	Malaria	Breeding habitat for mosquitos
	Onchocerciasis	Simulium larval habitat
	Schistosomiasis	Snail habitat
Wetlands	Cholera	Vibrio cholerae associated with inland water
	Encephalitis	Mosquito habitat
	Malaria	Mosquito habitat
	Schistosomiasis	Snail habitat
Soil moisture	Helminthiases	Worm habitat
	Lyme disease	Tick habitat
	Malaria	Vector breeding habitat
	Schistosomiasis	Snail habitat
Canals	Malaria	Dry season mosquito-breeding habitat; ponding; leaking water
	Onchocerciasis	Simulium larval habitat
	Schistosomiasis	Snail habitat
Human settlements	Diseases	Source of infected humans; populations at risk for transmission
Urban features	Chagas disease	Dwellings that provide habitat for triatomines
	Dengue fever	Urban mosquito habitats
	Filariasis	Urban mosquito habitats
	Leishmaniasis	Housing quality
Ocean color	Cholera	Phytoplankton blooms; nutrients, sediments
Sea surface temperature	Cholera	Plankton blooms (cold water upwelling in marine environment)
Sea surface height	Cholera	Inland movement of Vibrio-contaminated tidal water

SOURCE: adapted from Beck et al., 2000