Risk Analysis, Vol. 16, No. 4, 1996

Quantifying Water Pathogen Risk in an Epidemiological Framework

Joseph N. Eisenberg,[1,3] Edmund Y. W. Seto,[1] Adam W. Olivieri,[2] and Robert C. Spear[1]

Received September 20, 1995; revised February 27, 1996

Traditionally, microbial risk assessors have used point estimates to evaluate the probability that an individual will become infected. We developed a quantitative approach that shifts the risk characterization perspective from point estimate to distributional estimate, and from individual to population. To this end, we first designed and implemented a dynamic model that tracks traditional epidemiological variables such as the number of susceptible, infected, diseased, and immune, and environmental variables such as pathogen density. Second, we used a simulation methodology that explicitly acknowledges the uncertainty and variability associated with the data. Specifically, the approach consists of assigning probability distributions to each parameter, sampling from these distributions for Monte Carlo simulations, and using a binary classification to assess the output of each simulation. A case study is presented that explores the uncertainties in assessing the risk of giardiasis when swimming in a recreational impoundment using reclaimed water. Using literature-based information to assign parameters ranges, our analysis demonstrated that the parameter describing the shedding of pathogens by infected swimmers was the factor that contributed most to the uncertainty in risk. The importance of other parameters was dependent on reducing the *a priori* range of this shedding parameter. By constraining the shedding parameter to its lower subrange, treatment efficiency was the parameter most important in predicting whether a simulation resulted in prevalences above or below non outbreak levels. Whereas parameters associated with human exposure were important when the shedding parameter was constrained to a higher subrange. This Monte Carlo simulation technique identified conditions in which outbreaks and/or nonoutbreaks are likely and identified the parameters that most contributed to the uncertainty associated with a risk prediction.

KEY WORDS: Microbial risk characterization; epidemiological model; Monte Carlo simulations; uncertainty and variability.

1. INTRODUCTION

To assess risks from biological, chemical, and physical agents in the environment, public health agencies have traditionally relied on epidemiology to provide a direct empirical assessment on risk. However, current exposures to environmental chemicals and pathogens are often quite low, and empirical studies can no longer produce sufficiently sensitive information to be the sole means of assessing these risks.[1] Therefore, methodologies for this assessment have increasingly relied on indirect measures of risk by using analytical models for the estimation of the intensity of human exposure and the probability of human response from this exposure.

Attempts to provide a quantitative framework for the assessment of human health risks associated with the ingestion of waterborne pathogens have generally fo-

[1] 140 Warren Hall, School of Public Health, University of California, Berkeley, California 94720.
[2] Eisenberg, Olivieri and Associates, 1400 Jackson St., Oakland, California 94612.
[3] To whom all correspondence should be addressed.

0272-4332/96/0800-0549$09.50/1

cused on static models that calculated the probability of individual infection or disease as a result of a single exposure event.[2–5] These models, all of the same generic form, are based on dose–response data which are used to fit a standard distribution function such as an exponential or beta function. This model structure does not provide ways to incorporate epidemiological data such as incubation period, immune status, duration of disease, and the rate of symptomatic development, or exposure data such as processes affecting the pathogen concentration. These are all factors important in the disease process and necessary to be able to track variables such as the number of susceptible, infected, diseased, and immune within a population group.

To take advantage of the available infectious disease and dose–response data, we took a population perspective in the development of a mathematical model that characterizes the human disease risk of waterborne pathogen exposure. Based on the host/microbe interaction, this approach makes explicit the mechanistic aspects of the infectious disease process and provides a structure from which data are gathered. The existing dose–response model[23] was imbedded into an epidemiological framework, relying on a large base of literature describing the use of dynamic population models in the study of epidemics.[6] These dynamic population models emphasize the importance of how the susceptible, infected, diseased, or immune status of individuals within a defined population group vary over time. In addition to these four epidemiologically-based variables, our model incorporates a state variable to account for the dynamics of pathogen concentration at the site of exposure.

To provide a quantitative description of an infectious disease process requires a model that consists of a large number of parameters and state variables. A methodological problem with obtaining information from such a complex model is that the high levels of uncertainty and variability inherent in environmental processes preclude the use of traditional parameter estimation techniques. In general, biological systems have a high degree of variability due to both genetic and other differences between individuals and environmental factors that are not explicitly modeled. In addition, data collected from biological systems contains uncertainty that primarily arises from the high cost of experimentation. The impact of high levels of uncertainty and variability is reflected in the type of data seen in the literature. For example, in an assessment of giardiasis, Veazie *et al.*[7] report an average disease duration of 14.8 days with a range of 1–120. Shaw *et al.*[8] report that in half of the cases the disease lasted 7 days and a fourth more than 30 days. Dykes *et al.*[9] report a mean of 22 days with a range of 10–40 days. Likewise, prevalence rates range from 2–7% in Western European countries, to 10–13% in Oregon, to 8–40% in South America, Middle East, and South East Asia.[10] These rates depend not only on the geographical location, but also on the methods used to collect the data.

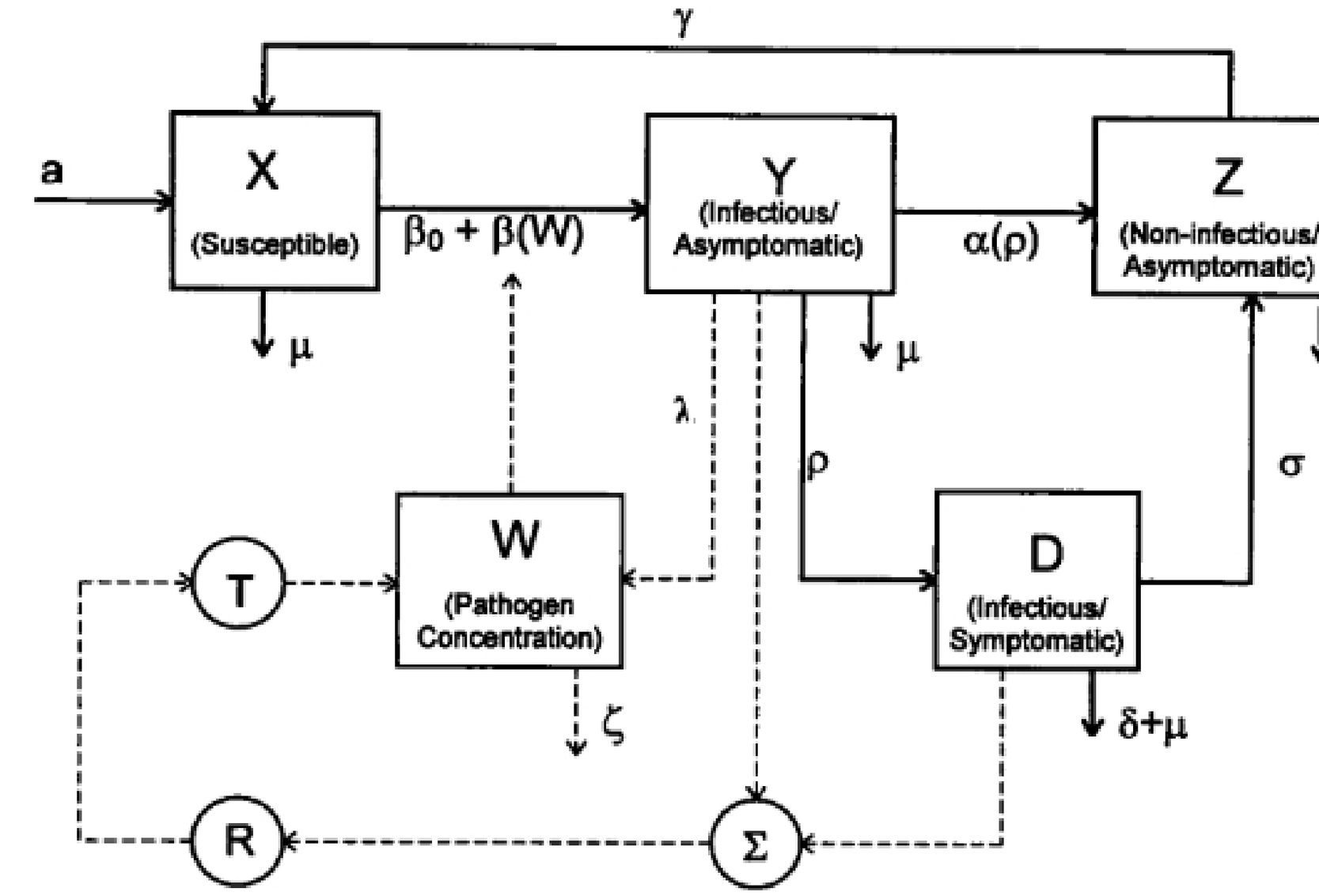
To obtain maximal information from uncertain and variable data, a distributional approach to risk characterization is needed. Therefore, rather than characterizing the risk of human exposures using a point estimate calculation of individual risk, Cooper *et al.*[11] estimated the probability distribution of the number of infected people in the exposed population. We have extended this approach by assigning probability distributions to an expanded list of model parameters, obtaining distributional outputs of average daily prevalence through the use of Monte Carlo techniques, and providing information on parameters important in producing uncertainty in the risk assessment. As will be seen, an outcome of this study was the finding that, even in the presence of uncertainty and variability, there is a significant amount of information obtainable in both the data collected from experiments and the mechanistic knowledge of the system. Moreover, the results of these simulations inform us on the level of uncertainty in risk, as well as identifying parameters which drive uncertainty and require better definition.

The analysis concerns estimation of the human risk of giardiasis via ingestion of reclaimed water in a recreational swimming impoundment filled with water reclaimed from community sewage. The focus was on a comparative study based on two scenarios: one in which water is not the exposure vehicle and one in which water is an exposure vehicle. Other pathogens and water reuse scenarios were explored in a previous report.[12]

2. METHODS

2.1. Model Description

The structure of the model is illustrated in Fig. 1. The model is comprised of five state variables and 13 parameters. Four of the state variables represent the human population in four epidemiological groups: susceptible individuals (**X**), infectious/asymptomatic individuals (**Y**), non-infectious/asymptomatic individuals (**Z**), and infectious/symptomatic individuals (**D**). The remaining



States:

X	Number of susceptible individuals.
Y	Number of infectious/asymptomatic individuals.
Z	Number of non-infectious/asymptomatic individuals.
D	Number of infectious/symptomatic individuals.
W	Concentration of pathogens in reclaimed water.

Parameters:

ρ	Fraction in state Y who move to state D per d.
α	Fraction in state Y who move to state Z per d.
σ	Fraction in state D who move to state Z per d.
γ	Fraction in state Z who move to state X per d.
δ	Fraction in state D who die due to modeled disease per d.
μ	Fraction who die from natural causes per d.
λ	Number of pathogen shed per L of water in swimming area per d per infectious/asymptomatic individual.
β_0	Baseline transmission rate.
β	Infection rate due to ingestion of pathogen in reclaimed water per d.
R	Number of pathogen per L of wastewater entering treatment plant per infectious individual.
T	Fraction of pathogen remaining per d after water treatment and dilution.
ζ	Fraction of pathogen in water which become non- viable per d.
a	Number of new susceptible individuals via migration per d.

Fig. 1. Model diagram.

state variable, **W**, keeps track of the concentration of pathogen in the water to which the population is exposed. The movement of individuals from one state to another and the concentration of pathogen are governed by the following set of five differential equations:

$$\frac{dX}{dt} = a + \gamma Z - \mu X - \beta_0 X - \beta(W)X$$

$$\frac{dY}{dt} = \beta_0 X + \beta(W)X - \mu Y - \rho Y - \alpha(\rho)Y$$