ISDS

*Emerging Health Threats Journal* 2011, **4**:s53. **doi:** 10.3134/ehtj.10.053
© 2011 C Safta *et al.*; licensee Emerging Health Threats Journal.
www.eht-journal.org

ABSTRACT

# Characterization of communicable disease epidemics using bayesian inversion

C Safta[1,2], J Ray[1,2], K Cheng[1,2], and D Crary[1,2]

[1]*Sandia National Laboratories, Livermore, CA, USA; and* [2]*Applied Research Associates, Inc., Arlington, VA, USA*
E-mail: csafta@sandia.gov

## Objective
We present a statistical method to characterize an epidemic of a communicable disease from a time series of patients exhibiting symptoms. Characterization is defined as estimating an unobserved, time-dependent infection rate and associated parameters that completely define the evolution of an epidemic. The problem is posed as one of Bayesian inference, where parameters are inferred with quantified uncertainty. The method is demonstrated on synthetic and historical epidemic data.

## Introduction
The evolution of a communicable disease in a human population is not entirely predictable. However, the spreading process can be assumed to vary smoothly in time. The time-dependent infection process can be linked to observations of the epidemic's evolution by convolving it with a stochastic delay model. In retrospective analyses of epidemics, when the observations are the dates of exhibition of patients' symptoms, the delay is the incubation period. In case of biosurveillance data, the delay is caused by incubation and a (hospital) visit delay, modeled as independent random variables. A model for observational error is also required. The time-dependent infection/spread rate may be inferred from observations by a deconvolution process.[1] The smooth temporal variation of the infection rate allows its representation using a low dimensional parametric model, and the inference may be performed with relatively little data. For large outbreaks, the data may be available early in the epidemic, allowing timely modeling of the outbreak. Short-term forecasts using the model could thereafter be used for medical planning.

## Methods
We extend the model by Brookmeyer and Gail, for use with biosurveillance data, by adding a model for visit delay. The model is also augmented, for use in bioterrorism scenarios, with an additive term modeling the existence of a significant number of index cases. We use the model to construct a Bayesian inverse problem for various parameters of epidemiological interest for example, spread rate parameters, index cases etc, and solve it using a Markov Chain Monte Carlo method. This procedure develops posterior distributions for the objects of inference, allowing us to quantify the uncertainty in the estimates. The inference procedure, when using biosurveillance data, can be computationally expensive as parameter estimation involves a double deconvolution. We accelerate this process by developing a surrogate model offline, which is trivially parallelizable. The surrogate model consists of a weighted sum of computationally inexpensive polynomial chaos expansions,[2] allowing the inference to proceed in a timely manner.

## Results
We demonstrate our method on data from historical plague outbreaks to study its ability to estimate infection rate properties. The methodology is validated by comparisons with published results. We also validate the inversion procedure with synthetic data from simulated plague bioattacks to gauge the inversion accuracy and computational savings when using the surrogate model. The improvement of the parameter estimates with the availability of data is also explored.

## Conclusions
Bayesian inference, using infection rate models, allows a simple way of characterizing epidemics with few observations. In conjunction with surrogate models, it may also be performed with modest computational effort.

## Acknowledgements
This paper was presented as an oral presentation at the 2010 International Society for Disease Surveillance Conference, held in Park City, UT, USA on 1–2 December 2010. This work was funded by DoD/DTRA, under contract HDTRA1-09-C-0034. The PI of the program is Ms Karen Cheng.

## References
1 Brookmeyer R, Gail M. A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic. *JASA* 1988;**83:**301–8.
2 Ghanem R, Spanos P. Stochastic Finite Elements, Springer-Verlag: New York, 1991.

OPEN ACCESS This is an Open Access article distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/2.5) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.