

# Forecasting high-priority surveillance regions: a socioeconomic model

E Chan<sup>1\*</sup>, T Choden<sup>1</sup>, T Brewer<sup>2</sup>, L Madoff<sup>2</sup>, M Pollack<sup>2</sup> and J Brownstein<sup>1</sup>

<sup>1</sup>Children's Hospital Informatics Program, Children's Hospital Boston, Boston, MA, USA; <sup>2</sup>ProMED-Mail, International Society for Infectious Diseases, Brookline, MA, USA

## Objective

To evaluate the association between socioeconomic factors and infectious disease outbreaks, to develop a prediction model for where future outbreaks would most likely to occur worldwide and identify priority countries for surveillance capacity building.

## Introduction

It has been suggested that changes in various socioeconomic, environmental and biological factors have been drivers of emerging and reemerging infectious diseases (1, 2), although few have assessed these relationships on a global scale (3).

## Methods

We constructed negative binomial regression models to analyze the relationship between a set of outbreaks reported by the WHO during 1996–2009 and 60 national socioeconomic variables from the World Bank's World Development Indicators database; a Human Development Index (HDI) reported by the UN Development Programme; and a Democracy socioeconomic data for each country.

## Results

Birth rate, measles immunization, urban population with access to improved sanitation facilities, life expectancy, infant mortality rate, public health expenditure, net official development by country of origin or territory and HDI were significant for latitude, as parasitic and infectious disease species richness has been for latitude, as parasitic and infectious disease species richness has been predictors in the 'univariate' analyses; the final multivariate model, with McFadden's pseudo  $R^2$  of 0.181, is shown in Table 1 and was used to predict regions at risk for future infectious disease outbreaks (Fig. 1).

## Conclusions

Public health expenditure, net ODA received, total population and HDI were significantly associated with the number of

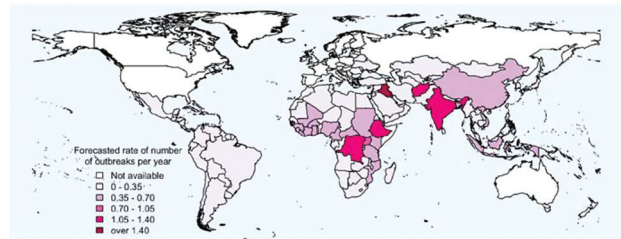


Fig. 1. Risk map for future infectious disease outbreaks forecasted by a model based on various socioeconomic factors.

outbreaks in a country, even after controlling for latitude. The final model identified regions in Africa, Asia and the Middle East as high-risk regions for future infectious disease outbreaks. The significance of ODA as a variable in the model may indicate that many of these countries may have limited capacity to support surveillance efforts, and highlight the need for assistance through international collaboration and financial support.

Sex ratio at birth, commonly called the secondary sex ratio, is easy to measure, often recorded, and only rarely subject to recall bias. More than 100 studies have investigated whether environmental or occupational exposures can change the sex ratio. The first studies on environment and the sex ratio examined acute exposures such as the nuclear bombing of Hiroshima and Nagasaki or the Great London Smog including dioxins, pesticides, and PCBs occupational exposures can change the sex ratio.

## Keywords

Probing; signal detection; resilience; high-reliability; socioeconomic data

## References

1. Morse SS. Factors and determinants of disease emergence. *Rev Sci Tech.* 2004;23:443–51.
2. Stephens DS, Moxon ER, Adams J, Altizer S, Antonovics J, Aral S, et al. Emerging and reemerging infectious diseases: A multi-disciplinary perspective. *Am J Med Sci.* 1998;315:64–75.
3. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, et al. Global trends in emerging infectious diseases. *Nature.* 2008;451:990–3.
4. Guernier V, Hochberg ME, Guegan JF. Ecology drives the worldwide distribution of human diseases. *PLoS Biol.* 2004;2:e141.

## \*E Chan

E-mail: emily.chan2@childrens.harvard.edu

Table 1. Final negative binomial regression model predicting number of outbreaks in each country from socioeconomic variables

Variable	Model coefficient (95% CI)
(Intercept)	2.893 (2.404;3.381)
Public health expenditure (Percentage of total health expenditure)	-0.010 (-0.017;-0.003)
Net ODA received (US\$)	3.941e-10 (2.171e-10;5.620e-10)
Total population	9.715e-10 (4.421e-10;1.481e-09)
Human development index	-1.501 (-2.363;-0.650)
Average latitude	-0.034 (-0.047;-0.021)

Abbreviations: CI, confidence interval; ODA, official development assistance.