

# Bayesian data integration for small area estimation of pathogen prevalence dynamics from pooled and individual data

Braden Scherting  
Dept. Mathematical Sciences  
Montana State University  
MT, USA  
`bradenscherting@montana.edu`

Alison Peel  
Centre for Planetary Health and Food Security  
Griffith University  
Queensland, AU

Raina Plowright  
Dept. Microbiology and Immunology  
Montana State University  
MT, USA

Andrew Hoegh  
Dept. Mathematical Sciences  
Montana State University  
MT, USA

March 31, 2022

## Abstract

Estimating the prevalence of a disease is necessary for evaluating and mitigating risks of its transmission within or between populations. Estimates that consider how prevalence changes with time provide more information about these risks but are difficult to obtain due to the necessary survey intensity and commensurate testing costs. We propose pooling and jointly testing multiple samples to reduce testing costs and use a novel nonparametric, hierarchical Bayesian data integration model to infer population prevalence from the pooled test results. This approach is shown to reduce uncertainty compared to individual testing at the same budget and to produce similar estimates compared to individual testing at a much higher budget through three synthetic studies and two case studies of natural infection data. Furthermore, the data integration procedure enables estimates of prevalence dynamics for sub-populations of the study through small area estimation by combining information from pooled and individual data.

**Keywords:** small area estimation, survey methods, Bayesian methods, time series, Gaussian process