

**ABSTRACT**

# The influence of address errors on detecting outbreaks of campylobacteriosis

K Zinszer<sup>1,2</sup>, K Charland<sup>1,2</sup>, C Jauvin<sup>2</sup>, L Bedard<sup>3</sup>, R Allard<sup>1,3</sup>, and DL Buckeridge<sup>1,2,3</sup>

<sup>1</sup>Department of Epidemiology & Biostatistics, McGill University, Montreal, Quebec, Canada; <sup>2</sup>Clinical and Health Informatics Research Group, McGill University, Montreal, Quebec, Canada; and <sup>3</sup>Direction de Santé Publique, Agence de la Santé et des services Sociaux de Montreal, Montreal, Quebec, Canada  
 E-mail: kate.zinszer@mail.mcgill.ca

**Objective**

To examine, via a simulation study, the potential impact of residential address errors on the identification of a point source outbreak of campylobacter.

**Introduction**

Mandatory notification to public health of priority communicable diseases (CDs) is a cornerstone of disease prevention and control programs. Increasingly, the addresses of CD cases are used for spatial monitoring and cluster detection and public health may direct interventions based on the results of routine spatial surveillance. There has been little assessment of the quality of addresses in surveillance data and the impact of address errors on public health practice.

We launched a pilot study at the Montreal Public Health Department (DSP), wherein our objective was to determine the prevalence of address errors in the CD surveillance data. We identified address errors in 25% of all reported cases of communicable diseases from 1995 to 2008. We also demonstrated that address errors could bias routine public health analyses by inappropriately flagging regions as having a high or low disease incidence, with the potential of triggering misguided outbreak investigations or interventions. The final step in our analysis was to determine the impact of address errors on the spatial associations of campylobacter cases in a simulated point source outbreak.

**Methods**

We based our simulations on campylobacter case records extracted from the Montreal department of public health for the period 1995–2008. We compiled area-level counts of campylobacter cases for each of Montreal’s 102 subregions and applied an address verification algorithm to determine the validity of the primary address and to identify an alternative, valid address. An address error was defined as the street number, street name, and/or postal code being identified by the algorithm as incorrect. We then re-computed the number of cases for each subregion.

For the outbreak scenario, we added an additional 21 cases to the caseload that occurred during July and August 2008, simulating a point source outbreak of a contaminated food product. Using Bayesian hierarchical models, we estimated subregional-level risk ratios (RRs) based on the original addresses and then again after correcting the address errors. We increased the prevalence of address errors by 5% increments, to a maximum of 40% address error. High-risk regions were identified by using a cutoff probability of 0.8 for the posterior proportion and a reference threshold of 1.2.

**Results**

We found significant subregional campylobacteriosis rate differences after correcting for address error (Figure 1).

The outbreak simulation also indicated significant differences in high-risk subregions after correcting for address errors and when increasing the prevalence of address errors.

**Conclusions**

Our study demonstrates that the positional shifting of cases due to address errors can produce diverging results in spatial



**Figure 1** Map of campylobacteriosis rate differences (original address versus corrected address) per 10,000 per subregion in Montreal.

analyses. The magnitude of address errors in our simulation was reasonable, given the rates of actual address errors we found in our pilot study. Positional errors have the potential to impact cluster detection and related public health interventions.

### **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.