

International Journal of Population Data Science

Journal Website: www.ijpds.org



The use of Online Consumer reviews in the Surveillance and Detection of Infectious Gastrointestinal Disease Outbreaks

Nicola Love¹, Elad Yom Tov², Ingemar Cox³, Hilary Kirkbride¹, and Alex Elliot¹

¹UK Health Security Agency, UK

²Bar-Ilan University, Israel

³University College London, UK

Introduction & Background

Gastrointestinal (GI) infections are an important cause of morbidity in England, with over 17 million cases estimated annually, leading to 1 million healthcare consultations and ~90,000 confirmed laboratory diagnoses. Transmission is through contaminated food or water, or contact with infected individuals, animals, or the environment. International travel to countries with compromised water, sanitation and hygiene is also a key risk factor.

The UK Health Security Agency (UKHSA) monitors GI infections through multiple surveillance systems, capturing laboratory confirmed diagnoses and syndromic presentations. This supports outbreak detection and informs public health action. However, shifts in healthcare seeking behaviour may influence the effectiveness of these systems. This study assesses whether online reviews could support UKHSA surveillance, using outbreaks associated with all-inclusive international travel as a case study.

Objectives & Approach

Using a Google Maps review dataset we identified hotel reviews posted between 1 January 2024 and 28 April 2025 relating to 20 destinations popular with UK tourists. A filtering model extracted reviews containing one of 16 keywords describing common GI symptoms. Over the same period, we used a UKHSA case management system to identify hotels frequented by individuals who had a laboratory confirmed GI infection or food poisoning notification and had travelled during their incubation period. We used a capture–recapture approach to compare hotels identified in each dataset.

Relevance to Digital Footprints

This study highlights how digital footprint data can supplement infectious disease surveillance and support early outbreak detection leading to improved outcomes for patients.

Results

We identified 376,688 reviews of which 169 were manually labelled, with strong model performance (ROC-AUC:0.89). The keyword search identified 645 reviews containing GI-related terms, referencing 360 unique hotels (~0.1% hotels included). The number of hotels identified from UKHSA systems was low, largely because information was unavailable. Overlap between datasets was observed, particularly where multiple confirmed cases reported the same hotel. However, each dataset also identified unique hotels, indicating that neither source was individually comprehensive.

Conclusions & Implications

Online consumer reviews show potential for identifying premises linked to GI outbreaks. Such approaches could supplement existing surveillance systems and support more timely public health responses. There could also be other public health applications for such a digital footprint data source and methodology.

Keywords

infectious diseases, consumer review data, travel health, gastrointestinal infections

