Wet Markets and Food Safety: TripAdvisor for Improved Global Digital Surveillance?

Nicole Erin Kogan, Isabelle Bolon, Nicolas Ray, Gabriel Alcoba, Jose Luis Fernandez-Marquez, Martin Mathias MÃ¼ller, Sharada Prasanna Mohanty, Rafael Ruiz de CastaÃ±eda

Submitted to: JMIR Public Health and Surveillance
on: July 03, 2018

Table of Contents

Wet Markets and Food Safety: TripAdvisor for Improved Global Digital Surveillance?

Nicole Erin Kogan[1,2], B.Sc.; Isabelle Bolon[2], DVM, PhD, MPH; Nicolas Ray[2,3], PhD; Gabriel Alcoba[4], MD, MPH; Jose Luis Fernandez-Marquez[5], Ph.D.; Martin Mathias MÃ¼ller[6], B.Sc.; Sharada Prasanna Mohanty[6], B.Sc.; Rafael Ruiz de CastaÃ±eda[2], Ph.D.

[1] Massachusetts Institute of Technology, , , Cambridge, United States.
[2] Institute of Global Health, Faculty of Medicine, University of Geneva, Geneva, Switzerland.
[3] Institute for Environmental Science, , University of Geneva, Geneva, Switzerland.
[4] Division of Tropical and Humanitarian Medicine, , University Hospitals of Geneva, Geneva, Switzerland.
[5] Citizen Cyberlab, Centre Universitaire d'Informatique, University of Geneva, Carouge, Switzerland.
[6] Digital Epidemiology Lab, , École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Corresponding Author:
Nicole Erin Kogan, B.Sc.
Massachusetts Institute of Technology
77 Mass Ave
Cambridge
United States
Phone: 1 4128057254
Fax: 1 000-000-0000
Email: kogan.nicole.e@gmail.com

Abstract

Background: Wet markets are critical for food security and sustainable development in their respective regions but are also associated with health risks. Due to their cultural significance, they attract numerous visitors and generate tourist-geared information on the Web (i.e. on social networks as TripAdvisor). These unexploited data can be used to create an internationally-comprehensive wet market inventory to support epidemiological surveillance and control in these settings, which to our knowledge, does not yet exist.

Objective: Using social network data, we aim to: assess the level of wet markets' touristic importance online; produce the first distribution map of wet markets of touristic interest; and identify common diseases facing visitors in these settings.

Methods: TripAdvisor was selected as the data source of this study following an analysis of food markets' touristic relevance on the web. A web scraping tool (ParseHub) was used to extract wet market names, locations, and reviews from TripAdvisor. The latter were analyzed and when possible, assigned GeoSentinel diagnosis codes. This syndromic information was overlaid onto a mapping of wet market locations.

Results: 89 of the first 150 Google Search results (59.3%) for "wet market" (July 2017) were tourism-related. Of the 1,090 hits on TripAdvisor for this keyword, 393 (36%) were confirmed wet markets; syndromic information was available for 57 of these (14.5%). The confirmed wet markets were heterogeneously distributed: Asia concentrated 246 (62.6%) of them, Europe 76 (19.3%), North America 31 (7.9%), Oceania 20 (5.1%), Africa 12 (3.1%), and South America 8 (2.0%). Analysis of reviews corresponding to these wet markets revealed the most frequently occurring disease among visitors was food poisoning, accounting for 51 of 95 diagnoses (54%). This proved most prevalent among those visiting South American markets (18 of 51 food poisoning incidents [35%]) but less for Asian markets (6 of 51 food poisoning incidents [12%]) when normalizing for wet market number.

Conclusions: To our knowledge, this study is first to map the global distribution of wet markets of touristic importance and adverse health events experienced by their visitors, highlighting the potential of social network data in global epidemiological surveillance.

(JMIR Preprints 03/07/2018:11477) DOI: https://doi.org/10.2196/preprints.11477

Preprint Settings

1) Allow access to Preprints when not published (under review or accepted) for
   (a) Open peer-review purposes.
   (b) Logged-in Users only.
✓ **(c) Anybody, anytime.**
   (d) Nobody.

3) When a final version is published in a JMIR journal
   (a) Allow download.
✓ **(b) Show abstract only.**

4) Allow access to Preprints when no longer under consideration
   (a) Logged-in Users only.
✓ **(b) Anybody, anytime.**
   (c) Nobody.

2) Accepted Setting
   (a) Logged-in Users only.
✓ **(b) Anybody, anytime.**
   (c) Same as b), plus submit as ahead-of-print version to PubMed.
   (d) Nobody.

Original Manuscript

JMIR Short Paper

Wet Markets and Food Safety: TripAdvisor for Improved Global Digital Surveillance?

Nicole E Kogan[1,2]; Isabelle Bolon, DVM, PhD, MPH[2]; Nicolas Ray, PhD[2,3]; Gabriel Alcoba, MD, MPH, DTMH[4]; Jose L Fernandez-Marquez, PhD[5]; Martin M Müller[6]; Sharada P Mohanty[6]; Rafael Ruiz de Castañeda, PhD[2]

[1] Massachusetts Institute of Technology, Cambridge, MA, USA
[2] Institute of Global Health, Faculty of Medicine, University of Geneva, Geneva, Switzerland
[3] Institute for Environmental Sciences, University of Geneva, Geneva, Switzerland
[4] Division of Tropical and Humanitarian Medicine, University Hospitals of Geneva, Geneva, Switzerland
[5] Citizen Cyberlab, Centre Universitaire d'Informatique, University of Geneva, Carouge, Switzerland
[6] Digital Epidemiology Lab, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Corresponding Author:

Rafael Ruiz de Castañeda
Institute of Global Health
Faculty of Medicine
University of Geneva
Campus Biotech G6.02
Chemin des Mines 9
Geneva, 1202
Switzerland
Phone: +41 223790951
Email: rafael.ruizdecastaneda@unige.ch

Abstract

Background

Wet markets are critical for food security and sustainable development in their respective regions. Due to their cultural significance, they attract numerous visitors and consequently generate tourist-geared information on the Web (i.e. on social networks as *TripAdvisor*). These data can be used to create a novel, international wet market inventory to support epidemiological surveillance and control in such settings, which are often associated with negative health outcomes.

Objectives

Using social network data, we aim to: assess the level of wet markets' touristic importance online; produce the first distribution map of wet markets of touristic interest; and identify common diseases facing visitors in these settings.

Methods

A *Google* search was performed on 31 food market-related keywords, with the first 150 results for each keyword evaluated based on their relevance to tourism. Of all queries, "wet market" had the highest number of tourism-related *Google Search* results; among these, *TripAdvisor* was the most frequently occurring travel information aggregator, prompting its selection as the data source for this study. A web scraping tool (*ParseHub*) was used to extract wet market names, locations, and reviews from *TripAdvisor*. The latter were searched for disease-related content, which enabled assignment of *GeoSentinel* diagnosis codes to each. This syndromic categorization was overlaid onto a mapping of wet market locations. Regional prevalence of the most commonly-occurring symptom group – food poisoning – was then determined (i.e. by dividing the number of wet markets per continent with ≥1 review containing this syndrome by the total number of wet markets on that continent with syndromic information).

Results

Of the 1,090 hits on *TripAdvisor* for "wet market", 393 (36%) conformed to the query's definition; wet markets were heterogeneously distributed: Asia concentrated 246 (62.6%) of them, Europe 76 (19.3%), North America 31 (7.9%), Oceania 20 (5.1%), Africa 12 (3.1%), and South America 8 (2.0%). Syndromic information was available for 57 of 393 wet markets (14.5%). The most frequently occurring syndrome among visitors to these wet markets was food poisoning, accounting for 51 of 95 diagnoses (53.7%). Cases of this syndrome were identified in 22 of 39 wet markets with syndromic information in Asia (56.4%), 5 of 7 in Europe (71.4%), and 5 of 7 in North America (71.4%). All wet markets in South America and Oceania reported food poisoning cases, but the number of reviews with syndromic information was very limited in these regions (n = 2).

Conclusions

The map produced illustrates the potential role for touristically-relevant social network data to support global epidemiological surveillance. This includes the possibility to approximate the global distribution of wet markets and to identify diseases (i.e. food poisoning) most prevalent in such settings.

Keywords

Wet market, Digital epidemiology, TripAdvisor, Mapping, Web scraping, Foodborne disease, Social

network, Tourism

Introduction

Traditional food markets (i.e. wet markets) play important roles in food security and local development [1]. However, they also have negative health implications. In 2003, SARS spread globally from a Chinese wet market, causing hundreds of deaths [2] and major economic losses [3]. Avian influenza has also been repeatedly associated with wet markets [1, 4]. Foodborne *Campylobacter*, *Salmonella*, *Giardia*, and *Escherichia* are most common in these settings, leading to 18 million DALYs annually [5]. These are particularly important in LMICs, but their true impact is unknown given many episodes go unreported [6].

Due to wet markets' cultural importance, there exists extensive tourist-geared information online, often on websites as *Yelp* and *TripAdvisor* that serve as forums to share experiences. While the use of these data remains unexploited for wet markets, it has yielded compelling results for restaurants. For example, iwaspoisoned.com [7] serves as an online platform where individuals can report symptoms of food poisoning alongside the offending eatery. Through citizen participation, the website has identified several foodborne disease outbreaks before traditional epidemiological methods [8].

This approach has shaped the objectives of this study, which are to show the link between wet markets and tourism as well as to exploit tourist-generated social network data to create the first map of the distribution of wet markets of touristic interest and their associated adverse health events. Rather than function as an epidemiological analysis in which foodborne disease incidence related to wet market visits is calculated, our approach aims to showcase the potential of online social networks to pick up potentially overlooked instances of disease.

Methods

A *Google* search was performed on different food market types (see Multimedia Appendix 1). For each keyword, the first 150 results were scraped and characterized based on their relevance to tourism (i.e. presence of tourist-geared content and promotion of an area to potential visitors). Of all food market types, "wet market" was linked to the highest proportion of tourism-related websites (89 of 150 [59.3%] *Google Search* results). This connection motivated its use in this study. Any social networking websites were flagged and their touristic importance assessed. Of these, *TripAdvisor* appeared most frequently and had the most comprehensive wet market-related information, prompting its selection as this study's data source.

The term "wet market" was inputted into *TripAdvisor* (July 2017). Wet market names and locations were harvested using *ParseHub* [9]. Irrelevant results (e.g. waterparks with "wet" in their names) were removed manually. A *Python 2.7* script integrating geocoding library *GeoPy* [10] was developed to convert wet market locations to geographic coordinates for mapping.

For each wet market, the *TripAdvisor* "Reviews" section was parsed for mentions of keywords most often associated with foodborne disease, a list corroborated by the NIH NIDDK's webpage on foodborne illness [11]: "diarrhea", "vomit/vomiting/vomited", "food poisoning", "stomach ache", "headache", "nausea/nauseous", "upset stomach", "sick", "ill", and "dizzy." Comments containing at least one of these were manually extracted. When possible, comments were assigned a *GeoSentinel* diagnosis code [12] through medical expert analysis using syndrome keywords and indicators of symptom duration. Wet market and syndromic distribution were analyzed using descriptive statistics.

The study represents a passive analysis of information on *TripAdvisor*. The investigators did not participate in nor were involved in *TripAdvisor* communications (i.e. there was no posting), so the analysis should not be considered active internet-based research requiring human subject consent.

Results

"Wet market" yielded 1,090 attractions on *TripAdvisor*, 393 (36.06%) of which aligned with the term's definition. Mapping revealed Asia as the region with the greatest wet market density, accounting for 246 (62.6%) of 393 wet markets reviewed (Figure 1). The second-most wet market-dense region was Europe (76 [19.3%]), followed by North America (31 [7.9%]), Oceania (20 [5.1%]), Africa (12 [3.1%]), and South America (8 [2.0%]).

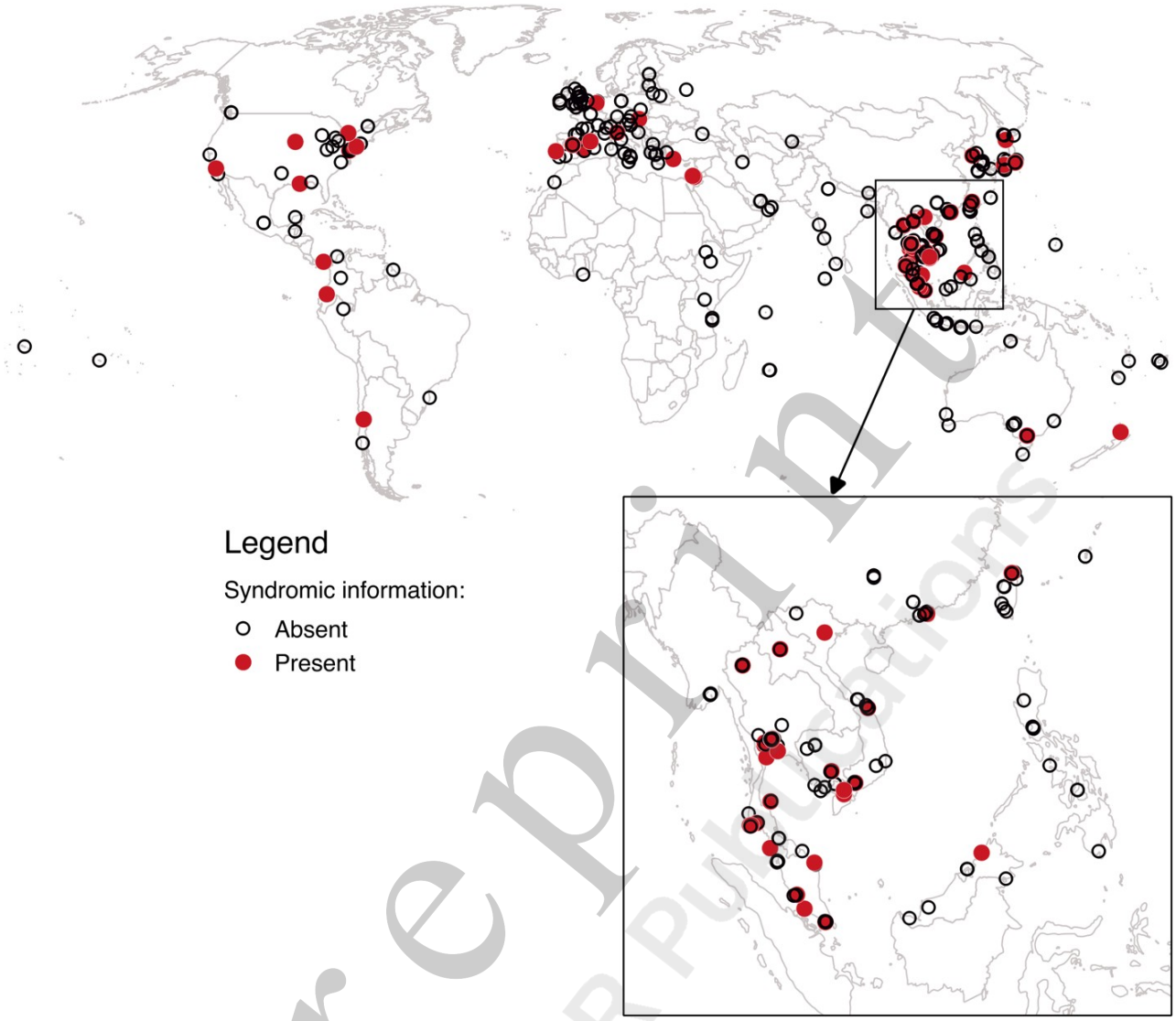

Figure 1. Locations of *TripAdvisor*-sourced wet markets, with zoomed inset in Southeast Asia. Red circles represent wet markets where visitors reported adverse health events, while empty circles denote locations where such reports were lacking. The background country border is sourced from Natural Earth vector data (naturalearthdata.com) projected in the World Robinson coordinate reference system using QGIS 2.18.2.

All 393 wet markets were reviewed on *TripAdvisor*. 57 (14.5%) contained at least one review with syndromic information, broken down regionally as 39 Asian (68.4%), 7 European (17.9%), 7 North American (17.9%), 2 South American (5.1%), and 2 Oceanic (5.1%). In total, these yielded 98 reviews with syndromic information, of which Asian markets accounted for 73 (74.5%). “Acute gastroenteritis < 12 hrs., food poisoning” was the most common diagnosis in wet markets globally (51 of 95 syndrome references [54%], discarding three reviews corresponding to unascertainable illness) (Figure 2). Asia, where the review volume was particularly noticeable, had 22 of 39 [56.4%] wet markets with this specific syndrome. The syndromes “Acute gastroenteritis > 12 hrs.”, “Acute gastroenteritis < 12 hrs.”, and “Diarrhea, acute unspecified” comprised 90 of 95 non-“N/A” diagnoses (95%) from the 57 wet markets.

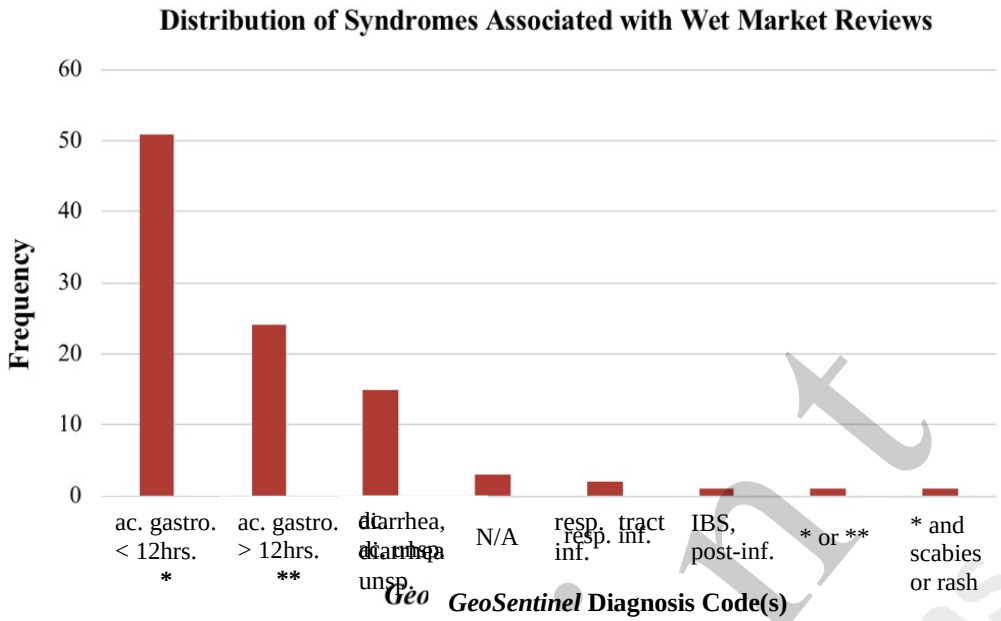

Figure 2. Frequency distribution of 98 TripAdvisor reviewer diagnoses following *GeoSentinel* encoding (ac. = acute, gastro. = gastroenteritis, unsp. = unspecified, resp. tract inf. = respiratory tract infection, IBS = Irritable Bowel Syndrome, N/A = symptoms present but illness unascertainable)

Discussion

This study shows the touristic importance of wet markets and provides, to our knowledge, the first global map of wet markets' locations and related syndromic information communicated on *TripAdvisor* by their visitors. Our map reveals that wet markets are heterogeneously distributed across continents. Asia is a hotspot, both in number of markets and absolute number of adverse associated health events. When controlling by the number of continental wet markets with syndromic information, South America and Oceania exhibited the highest proportion of *GeoSentinel* "acute gastroenteritis < 12 hrs., food poisoning" diagnoses. However, both of these regions were wet market-sparse (n = 2 wet markets with syndromic information in each) compared to a continent like Asia.

Use of region-specific terminology for designing markets could have yielded denser mapping and should be used in future work. Furthermore, the choice to look only at *TripAdvisor* reviews in English may have impacted the wet market distribution shown in the map. One example is the density of wet markets, which seems to be in preferred destinations of English-speaking tourists; another example is the concentration of wet markets in Africa is largely in English-speaking countries. This could prompt search on the platform using other languages to enhance the results.

Little is known about wet market health threats to locals and travelers partly due to underreporting of adverse health events experienced by visitors [13]. While social networking platforms do not paint a complete picture of wet markets (e.g. total number of regional wet markets, how many yearly visitors), they can provide opportunities to instantaneously report and detect these elusive cases – primarily traveler-related ones – and improve epidemiological monitoring within these settings.

Reviews from *Yelp*, a *TripAdvisor*-like platform connecting citizens with local businesses, have been leveraged by New York public health authorities to detect restaurant-related foodborne disease events [14]; interestingly, fewer than 2% of individuals with an alleged illness explicitly mention reporting their case to a medical professional [14]. *Twitter* has been implemented in St. Louis, Missouri to detect food poisoning cases [15] but its potential is restricted by short post length. The website *iwaspoisoned.com* is equally notable in its effort to crowdsource information, though it is limited in scope outside the United States. Wet markets are more challenging establishments because of their typical location in LMICs and their operation under often limited regulations. This necessitates tapping into other online communities – social platforms like *TripAdvisor* – to glean information on visitor health.

Traveler's diarrhea is a common disorder affecting tourists visiting developing countries [16] and is generally associated with consumption of foods prepared under unhygienic condition (common in wet markets). Our analysis shows that acute gastroenteritis (food poisoning) and diarrhea were the most frequent illnesses among wet market visitors. However, this must be carefully considered from medical and epidemiological perspectives since the result is based on rarely-corroborated, online descriptions. We also cannot exclude that reported symptoms could have been caused by an event prior to or after (but unrelated to) a wet market visit.

As use of *TripAdvisor* in isolate has accounted for only a fraction of wet markets of touristic importance and their associated health risks, larger datasets are needed to confirm the results presented here and to explore others. Even so, the pipeline we present is significant for travel medicine and epidemiology. It could ultimately contribute to predictive models for improved epidemic forecasting and to the development of diagnostic tools based on syndromic surveillance and artificial intelligence. *TripAdvisor* could crosstalk with other social networks (e.g. *Yelp*) for maximal information coverage and partner with other initiatives for more structured collection of wet market-related health information in near-real time. In this way, we can gain an improved understanding of global wet markets and their associated health risks while also ensuring their safer promotion.

Acknowledgements

NEK was supported by the MIT International Science and Technology Initiatives (MISTI) and by a Summer school Grant from the Zeno Karl Schindler Foundation. RRdC was supported by Fondation Louis-Jeantet, which we thank very much. We also thank Prof. A. Flahault for supporting this work at the Institute of Global Health. We thank two anonymous reviewers for their constructive comments.

Authors' Contributions

RRdC, IB, and NEK designed the study. NR, GA, JLFM, MMM and SPM guided the implementation. NR and NEK developed the spatial analyses. GA paired *GeoSentinel* diagnosis codes with *TripAdvisor* reviews. NEK wrote the initial draft of the manuscript, and all coauthors reviewed and improved it.

Conflicts of Interest

None declared.

Abbreviations

SARS: Severe Acute Respiratory Syndrome
LMIC: low/middle income country

Multimedia Appendix 1

List of market types whose relevance to tourism was assessed in an initial *Google Search* screen.

Data and code availability

The database of wet markets and associated Python scripts are available upon request.

References

1. World Health Organization. A guide to healthy food markets. 2006.

http://www.who.int/foodsafety/publications/capacity/healthymarket_guide.pdf?ua=1. Archived at: http://www.webcitation.org/70UclvVLz

2. Wu T, Perrings C, Kinzig A, Collins JP, Minteer BA, Daszak P. Economic growth, urbanization, globalization, and the risks of emerging infectious diseases in China: A review. Ambio 2017 Feb;46(1):18-29. PMID:27392678
3. Lee JW, McKibbin WJ. Estimating the global economic costs of SARS. In: Knobler S, Mahmoud A, Lemon S, editors. Learning from SARS: Preparing for the next disease outbreak: Workshop summary. Washington DC: National Academies Press; 2004. ISBN:0309091543
4. Wu J, Lu J, Faria NR, ..., Ke C. Effect of live poultry market interventions on Influenza A (H7N9) virus, Guangdong, China. Emerg Infect Dis 2016 Dec;22(12):2104-2112. PMID:27869613
5. Grace D. Food safety in low and middle income countries. Int J Environ Res Public Health Aug 2015; 12:10490-10507. Doi:10.3390/ijerph120910490
6. Bernardo TM, Rajic A, Young I, Robiadek K, Pham MT, Funk JA. Scoping review on search queries and social media for disease surveillance: a chronology of innovation. J Med Internet Res 2013 Jul;15(7):e147. PMID:23896182
7. Quade P. 2009. Iwaspoisoned.com. https://iwaspoisoned.com. Archived at: http://www.webcitation.org/70Uh5XbE0
8. Neimark J. 2017. “I was poisoned”: can crowdsourcing food illnesses help stop outbreaks? https://www.npr.org/sections/thesalt/2017/11/30/565769194/i-was-poisoned-can-crowdsourcing-food-illnesses-help-stop-outbreaks. Archived at: http://www.webcitation.org/70Ujfiyin
9. Fomina A, Prelich P, Toarca S. 2015. ParseHub. https://www.parsehub.com/. Archived at: http://www.webcitation.org/70bzDzEXQ
10. Esmukov K. 2006. Geopy. https://github.com/geopy/geopy. Archived at: http://www.webcitation.org/70UIpq9sy
11. NIH NIDDK. 2014. Foodborne Illnesses. https://www.niddk.nih.gov/health-information/digestive-diseases/foodborne-illnesses. Archived at: http://www.webcitation.org/74jmK4zPz
12. International Society of Travel Medicine. N/A. GeoSentinel Document Center. http://www.istm.org/geosentineldocumentcente2. Archived at: http://www.webcitation.org/70XWu0RXn
13. Nsoesie EO, Kluberg SA, Brownstein JS. Online reports of foodborne illness capture foods implicated in official foodborne outbreak reports. Prev Med 2014 Oct;67:264-269. PMID:25124281
14. Ostrov BF. 2015. Using Yelp to track food poisoning outbreaks. https://www.usatoday.com/story/news/2015/10/24/kaiser-using-yelp-to-track-food-poisoning-outbreaks/74469136/. Archived at: http://www.webcitation.org/70bzpGieK
15. Harris JK, Hinyard L, Beatty K, ..., Brownstein JS. Evaluating the implementation of a Twitter-based foodborne illness reporting tool in the City of St. Louis Department of Health. Int J Environ Res Public Health 2018 Apr;15(5):N/A. PMID:29695038
16. Porter CK, Olson S, Hall A, Riddle MS. Traveler’s diarrhea: An update on the incidence, etiology, and risk in military deployments and similar travel populations. Mil Med 2017 Sep;182(S2):4-10. PMID:28885918

Supplementary Files

Figures

Figure 1. Locations of TripAdvisor-sourced wet markets, with zoomed inset in Southeast Asia. Red circles represent wet markets where visitors reported adverse health events, while empty circles denote locations where such reports were lacking. The background country border is sourced from Natural Earth vector data (naturalearthdata.com) projected in the World Robinson coordinate reference system using QGIS 2.18.2.

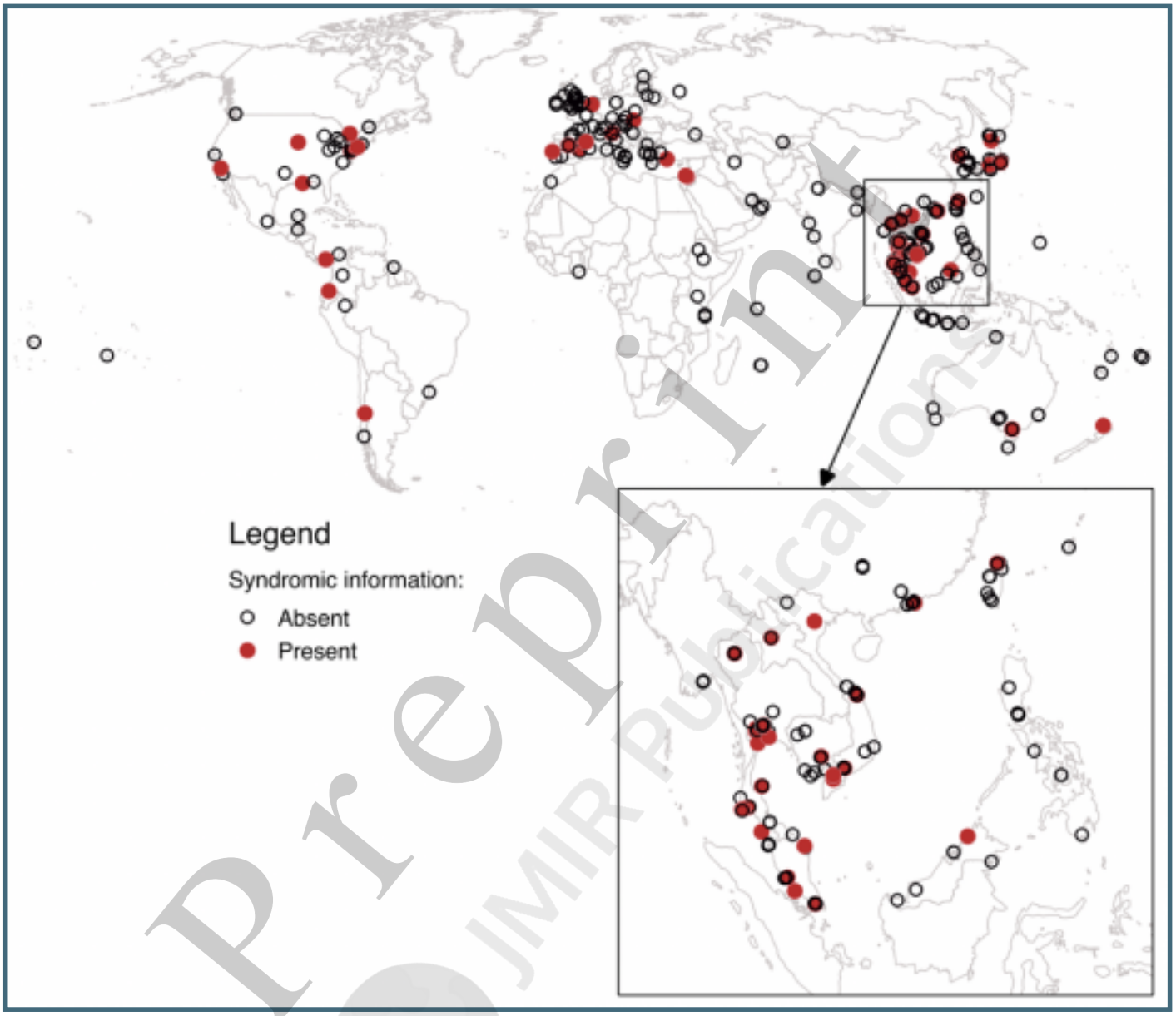

Figure 2. Frequency distribution of 98 TripAdvisor reviewer diagnoses following GeoSentinel encoding (ac. = acute, gastro. = gastroenteritis, unsp. = unspecified, resp. tract inf. = respiratory tract infection, IBS = Irritable Bowel Syndrome, N/A = symptoms present but illness unascertainable).