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

Nicole E Kogan 1,2; Isabelle Bolon, DVM, PhD 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 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:
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 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.

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
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.

Keywords
Wet market, Digital epidemiology, TripAdvisor, Mapping, Web scraping, Foodborne disease, Social network, Tourism
Introduction

Traditional food markets (e.g. 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 reach is unknown given many episodes go unreported [6].

Due to wet markets’ cultural importance, there exists extensive, relevant, 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: to show the touristic interest of wet markets on the Web and to generate the first map of the distribution of wet markets of touristic interest and their associated adverse health events using tourist-generated social network data.

Methods

Establishing the Data Source

A Google search performed on different food market types was conducted (see Multimedia Appendix 1). For each designation, the first 150 results were scraped and characterized based on relevance to tourism. 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.

Web Scraping

“Wet market” (most tourism-related Google Search results), 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 pre-process wet market data for mapping, specifically converting wet market locations to latitudes and longitudes.

Text Mining

For each wet market, the TripAdvisor “Reviews” section was parsed for mentions of keywords most often associated with foodborne disease: “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 into a document. When possible, comments were assigned a GeoSentinel diagnosis code [11], performed by mapping keyword combinations and indicators of symptom duration from comments to disease descriptions provided by GeoSentinel.

Results

Touristic Importance of Wet Markets

89 of 150 Google Search results (59.3%) for “wet market” were classifiable as tourism-related websites, the highest proportion among all food market designations.
Global Wet Market Distribution

“Wet market” yielded 1,090 attractions on TripAdvisor, 393 (36.06%) of which were confirmed. 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 where such reports were lacking. A heat map of market density (grey) is added to show areas with high market density. The background country border is sourced from Natural Earth vector data (naturalearthdata.com) projected in the World Robinson coordinate reference system on QGIS 2.18.2

Wet market reviews and syndromic information by visitors

All 393 wet markets were reviewed on TripAdvisor. 57 (14.5%) contained information on adverse health events experienced by wet market visitors; 95 of 98 reviews (97%) with mention of such events could be assigned a GeoSentinel diagnosis code (Figure 2). Class 525 (“acute gastroenteritis < 12 hrs., food poisoning”) was the most common diagnosis, accounting for 51 of 95 syndromes (54%); it was most descriptive of travelers to South and North American wet markets (accounting for 33 class 525 occurrences [65%]; note 6 occurrences [12%] for travelers to Asian markets), with normalization by the number of continental wet markets. Classes 525, 287 (“acute gastroenteritis > 12 hrs.”), and 132 (“diarrhea, acute unspecified”) comprised 88 of 98 diagnoses (90%).
Figure 2. Frequency distribution of 98 TripAdvisor reviewer diagnoses following GeoSentinel encoding (525 = acute gastroenteritis < 12 hrs. [food poisoning]; 287 = acute gastroenteritis > 12 hrs.; 132 = diarrhea, acute unspecified; N/A = symptoms present, illness unascertainable; 191 = respiratory tract infection [upper]; 167 = irritable bowel syndrome (IBS), post-infectious; 194 = scabies; 216 = rash, urticaria, or angioedema)

**Discussion**

This study shows the touristic importance of wet markets and to our knowledge, provides the first global map of wet markets' locations and related syndromic information communicated on TripAdvisor by their visitors.

Our visualization reveals that wet markets are heterogeneously distributed across continents. Asia is a hotspot, both in number of markets and adverse associated health events (i.e. acute gastroenteritis, although when normalizing the latter by the former, South American wet markets showed the highest proportion of GeoSentinel class 525 diagnoses). In contrast, Africa and South America – continents known for live animal markets – showed sparse wet market presence. This could be partly explained by the Chinese etymology of “wet market” [12], implying the designation is more applicable to Asian markets than to similar vendors on other continents; use of region-specific terminology could have yielded denser mapping and should be used in future work.

Little is known about wet market threats to health partly due to underreporting of adverse health events experienced by visitors [13]. New approaches exploiting online social networks, which facilitate instantaneous reporting, offer promising opportunities to detect these missing cases and improve epidemiological monitoring across wet markets. 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. iwaspoisoned.com is equally notable in its effort to crowdsource information, though it is limited in scope outside the United States. Wet markets, though, are more challenging establishments because of their location in more developing nations and their operation under minimal 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.

To confirm our results and explore others, larger datasets are needed to cover wet markets in all geographical regions. This could feed predictive models for improved epidemic forecasting and contribute to development of diagnostic tools based on syndromic surveillance and artificial intelligence. Platforms such as TripAdvisor could 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 Summerschool Grant from the Zeno Karl Schindler Foundation.

Authors’ Contributions
RRdC, IB, and NEK designed the study. NR, GA, JLFM, MMM and SPM guided the implementation. NR and NK developed the spatial analyses. GA paired GeoSentinel diagnosis codes with TripAdvisor reviews. All coauthors reviewed the manuscript.

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
   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,
   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.
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


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