

ORIGINAL RESEARCH

Open Access



# Using social media in Kenya to quantify road safety: an analysis of novel data

J. Austin Lee<sup>1\*</sup>, Lyndsey Armes<sup>2</sup> and Benjamin W. Wachira<sup>3</sup>

## Abstract

**Background:** Road traffic injuries are a large and growing cause of morbidity and mortality in low- and middle-income countries, especially in Africa. Systematic data collection for traffic incidents in Kenya is lacking and in many low- and middle-income countries available data sources are disparate or missing altogether. Many Kenyans use social media platforms, including Twitter; many road traffic incidents are publicly reported on the microblog platform. This study is a prospective cohort analysis of all tweets related to road traffic incidents in Kenya over a 24-month period (February 2019 to January 2021).

**Results:** A substantial number of unique road incidents (3882) from across Kenya were recorded during the 24-month study period. The details available for each incident are widely variable, as reported and posted on Twitter. Particular times of day and days of the week had a higher incidence of reported road traffic incidents. A total of 2043 injuries and 1503 fatalities were recorded.

**Conclusions:** Twitter and other digital social media platforms can provide a novel source for road traffic incident and injury data in a low- and middle-income country. The data collected allows for the potential identification of local and national trends and provides opportunities to advocate for improved roadways and health systems for the emergent care from road traffic incidents and associated traumatic injuries.

**Keywords:** Road safety, Road traffic, Trauma, Kenya, Social media, Twitter

## Background

Road traffic injuries are a growing cause of morbidity and mortality in low- and middle-income countries (LMICs), with a particularly high burden in Africa, including in Kenya [1–3]. In Kenya road traffic incidents are among the leading causes of both morbidity and mortality [4–9]. Road traffic-associated injuries are particularly common among users of motorcycles and public transportation (matatus/mini-bus), as well as among pedestrians [4, 8, 10]. Broadly, across age groups, Kenyans have poor seatbelt and helmet utilization [11]. Among a cohort of head injured patients in a Kenyan emergency department

(ED), none reported seatbelt or helmet use at the time of injury [12]. Furthermore, the Kenyan public has recognized the key role vehicles and over crowded public transit play in increasing risk of injury on Kenyan roadways [13, 14].

Yet, systematic data collection for traffic incidents is broadly lacking across Africa, as well as in Kenya. The ability to target road and traffic safety improvements with tailored solutions requires an understanding of the burden of disease and the current strengths and weaknesses. Robust and accurate statistics around road traffic deaths, hospital data, population surveys, or police reports are often not available [1, 2]. Prior work in Kenya has identified the need to “improve the collection and availability of accurate [road traffic injury] data” in the country [15]. There is recognized underreporting, variance from recognized international standards, and overall inadequate

\*Correspondence: [jaustinlee@brown.edu](mailto:jaustinlee@brown.edu)

<sup>1</sup> Department of Emergency Medicine, Brown University Warren Alpert Medical School, Providence, RI, USA  
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

data collection around road traffic incidents in Kenya [16]. Previously, the World Health Organization has estimated that nearly 80% of road traffic fatalities were unreported in prior Kenyan government data [17].

The majority of Kenyans have access to mobile telephones (91% penetration per capita, compared to 80% across Africa) and many Kenyans use social media (an average of nearly 3 hours per day), including an estimated 50% of Kenyans who use the popular microblog and social media platform Twitter [18]. Prior research has evaluated crowdsourced data from mobile phone data and social media platforms such as Twitter to provide timely incident detection [19]. However, the preponderance of prior work in the space has been performed in high-income countries with robust and publicly available crash data, utilized real time GPS or accelerometer data, or had evaluated unique newsworthy events, as opposed to routine social media postings [20–22].

This study aims to evaluate publicly available social media posts regarding road traffic incidents in Kenya. We hypothesize that the data from Twitter can be used as one of the few publicly available sources of road traffic incident data in Kenya given the paucity of available government or other reliable data sources. Furthermore, we detail the epidemiology of Kenyan road traffic incidents using information contained in Twitter posts.

## Methods

### Study design and setting

This was an observation, prospective cohort study, wherein all identified road traffic incidents from February 1, 2019, to January 31, 2021, were included for analysis. Research staff from the

Emergency Medicine Kenya Foundation's (EMKF) The Injury Prevention and Safety Initiative (TIPSI) actively monitor Twitter daily for reports of road traffic incidents in Kenya (Fig. 1; these examples are from outside of the study period). TIPSI staff then retweet each unique incident onto their Twitter feed (@TIPSIKenya) and take care to ensure no duplicate events are included. Subsequently, research staff abstract key data from each traffic incident from the TIPSI Twitter timeline. Events involving non-motor vehicle incidents, such as boats and planes, were excluded from the analysis.

### Outcomes

Data regarding time of day and extent of injuries were included for analysis when available, based on the details included in a tweet. Information regarding casualties (number involved in an incident), injuries (number reported to have been injured), and fatalities (those who are reported to have died) were captured when available and generally were from what was known or available on

scene. Data that is determined some time after an incident, like eventual outcomes such as death for individuals who received hospitalization, was not available and as such not included. Casualties were further divided into whether or not the incident involved vulnerable roadway users (such as motorcyclists, cyclists, and pedestrians).

### Analysis

EMKF TIPSI staff monitor Twitter and abstract data into a Microsoft Excel® database. All statistical analyses were performed using STATA Statistical Software Release 15 (StataCorp LLC, College Station, TX).

### Results

The data captured a substantial number of unique road incidents ( $n=3882$ ) from across Kenya during the 24-month study period (Table 1). There was month to month variation in the number of incidents (median monthly incidents: 152), with May 2019 having the maximal incidence ( $n=251$ , 6.47%). Fridays (15.6%), Saturdays (16.8%), and Sundays (15.4%), account for the highest proportion of incidents (Fig. 2). The majority ( $n=3529$ , 90.9%) of data have an associated time of occurrence (Fig. 3), and the highest proportion of incidents occurred at 7am (7.2%) and 6pm (6.8%). In particular, rush hour traffic (6:00–8:59 AM, 4:00–8:59 PM) accounts for only 8 hours of the day, but 49.4% (1745/3529) of Twitter-reported road incidents (with an associated time of occurrence) happen during those hours.

Twitter reports included information about the number of vehicles involved in each incident, revealing that single vehicle incidents were most common (64.47%). There were 1278 (32.9%) tweets that contained information about the casualties from (number of individuals involved in) an incident. Of the tweets reporting types of casualties, 441 (34.51%) involved at least one vulnerable road user. Additionally, a total of 510 tweets (13.1%) reported known human injuries, with a cumulative total of 2043 injuries identified during the data collection period. A further 696 (17.9%) tweets reported at least one fatality, and a cumulative total of 1503 fatalities were recorded.

### Discussion

The methods outlined here allowed for the creation of a dataset with information about unique road incidents in Kenya, a topic of public health interest that has historically lacked sufficient reporting mechanisms. The work by the EMKF TIPSI team to monitor Twitter for unique road traffic incidents in Kenya indicates that this methodology of tracking is a viable source of epidemiologic data in a lower resources setting. It is of particular value given the lack of other publicly available data sources.



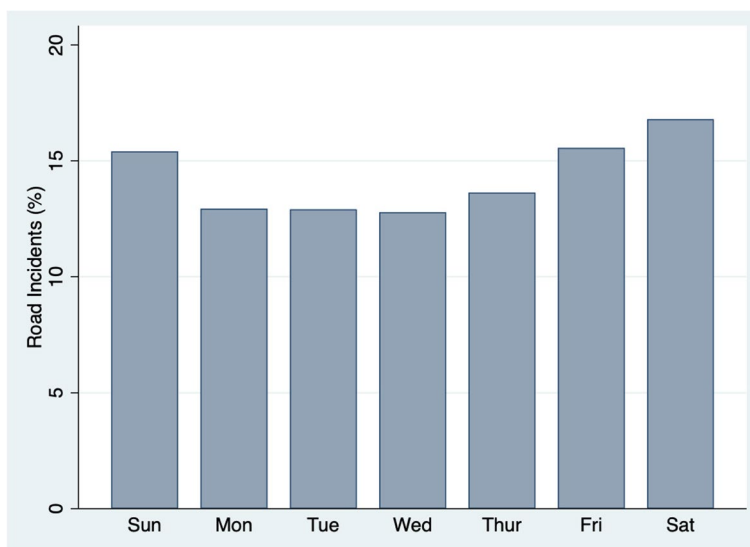
Fig. 1 Examples of tweets from @TIPSIKenya

**Table 1** Characteristics of Twitter-reported road traffic incidents in Kenya (February 2019 to January 2021)

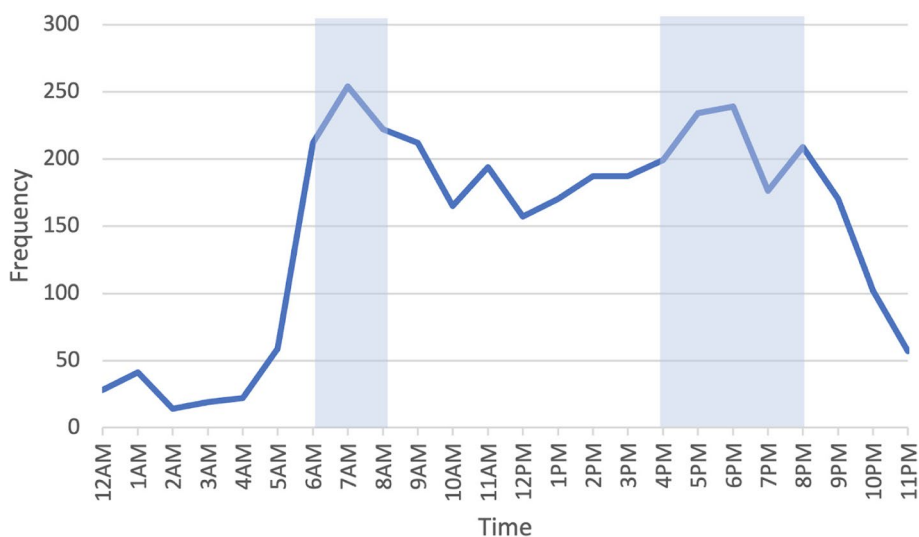
| Characteristic                 | Incidents (n = 3882)<br>No. (%) |
|--------------------------------|---------------------------------|
| <b>Time of day</b>             |                                 |
| Rush hour (6a–8:59a, 4p–8:59p) | 1745 (45.0%)                    |
| Non-rush hour                  | 1784 (46.0%)                    |
| Missing                        | 353 (9.0%)                      |
| <b>Number of vehicles</b>      |                                 |
| 1                              | 2218 (57.1%)                    |
| 2                              | 1069 (27.6%)                    |
| 3+                             | 137 (3.5%)                      |
| Missing                        | 458 (11.8%)                     |

Available data from the Kenyan National Transport and Safety Authority indicate higher numbers of fatalities from road traffic incidents than was identified through this dataset [23]. This analysis, with a lower fatality calculation, is somewhat expected given that information gleaned from Twitter presumably only contains information reported from the scene and does not follow up hospital or longer-term morbidity/mortality. Furthermore, though many are, certainly not all road traffic incidents in Kenya are documented by Twitter users.

Previous analysis of a single Kenyan Twitter user (@Ma3route) was focused on geospatial analysis of road traffic incidents and showed significant clustering of incidents at several high risk locations and intersections [24]. This Twitter user is frequently found in



**Fig. 2** Incidents, by day of the week



**Fig. 3** Incidents, by time of day, with highlighted rush hours

the @TIPSIKenya feed. Interestingly, one analysis of a small subset of @Ma3route data showed that the majority of Tweets with geolocation data were found to be geographically accurate when verified in person, though the authors noted that most tweets about road traffic incidents do not have any associated geolocation data [25].

**Limitations**

Several limitations are noted. There are challenges that come due to the manual process of monitoring Twitter

for information and incidents, as well as a manual data abstraction process, which leave room for the possibility of missed incidents. Additionally, many Kenyans are active users of Twitter, but many more Kenyans do not use the platform at all. Furthermore, the TIPSI team does not have strict incident inclusion/exclusion criteria. Rather, the goal is to catalog any incident and all available data without duplicating previously noted incidents. The initial data source (tweets) are microblogs not meant to be complete or comprehensive public health data. As such, there is no standard format to road traffic



data reported on the platform, and may include (or have missing) any number of pieces of information including a location (or reference to a nearby landmark or intersection), the number and types of vehicles, occupants involved, the extent of injuries or fatalities, and may or may not include photographs of the scene. The disparate information is unique to every tweet as the source social-media users provide data input on an ad hoc basis.

### Implications

Road traffic accidents are a major cause of morbidity and mortality in Kenya, and while Twitter can provide some data for context, there is an urgent need to improve the quantification of the road traffic incidents in the country. Improved data would comprehensively include: geolocation, types and number of vehicles, numbers of occupants, types of injuries, the time of occurrence. Such improved data collection will make it possible to better assess the health and economic losses from road traffic incidents, and allow for improved advocacy and a more robust financial argument for improved roadways and road traffic safety. Future work should evaluate the ability for the existing Kenyan emergency medical care systems to adequately treat injuries and longer-term morbidity that results from road traffic incidents.

### Conclusion

Twitter and other digital media platforms can provide a novel source for road traffic incident and injury data. In Kenya, the EMK Foundation TIPSII team has been able to collate data through routine social media monitoring. In Kenya, and in other LMICs, cell phones are pervasive and access to social media platforms such as Twitter provide a data source that can be used for public health epidemiology and advocacy. In countries or settings with inadequate systematic data collection of road traffic incidents, Twitter and other social media platforms may provide the best available data source and can serve as an important tool for advocacy and public health improvement. Furthermore, cell phones and social media platforms are well suited to provide support and enhance timely trauma care. In Kenya, the ability to use a cell phone or social media platform to quickly identify health facilities capable of providing trauma care and orient individuals on how to quickly get to such facilities is of urgent importance.

Twitter can provide a novel method for identification and collation of road traffic incidents in Kenya and other LMICs. The data collected allows for the identification of local and national trends and provides the potential opportunities to advocate for improved roadways and health systems for the emergent care from road traffic incidents and associated traumatic injuries.

### Abbreviations

LMICs: Low- and middle-income countries; ED: Emergency department; EMKF: Emergency Medicine Kenya Foundation's; TIPSII: The Injury Prevention and Safety Initiative.

### Acknowledgements

The authors are grateful to the Emergency Medicine Kenya Foundation, particularly the research staff, for their ongoing work to improve emergency and injury care in Kenya.

### Authors' contributions

JAL and BWW designed the study. LA led the data analysis. JAL is the primary author for the manuscript, and both LA and BWW were contributing authors on the manuscript. All authors have read and approved of the final manuscript.

### Funding

This work was supported by the team at the Emergency Medicine Kenya Foundation, a Kenyan not-for-profit, non-governmental organization. This research did not receive any specific grant or funding source.

### Availability of data and materials

The datasets generated and analyzed during the current study were obtained from and are available at <https://twitter.com/TIPSIIKenya>

### Declarations

#### Ethics approval and consent to participate

Not applicable. All information for this work was obtained from publicly available sources (Twitter) and no data was abstracted for analysis that contained any personally identifiable information.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Department of Emergency Medicine, Brown University Warren Alpert Medical School, Providence, RI, USA. <sup>2</sup>School of Public Health, Brown University, Providence, RI, USA. <sup>3</sup>The Aga Khan University, Nairobi, Kenya.

Received: 23 November 2021 Accepted: 8 May 2022

Published online: 28 June 2022

### References

- Lagarde E. Road traffic injury is an escalating burden in Africa and deserves proportionate research efforts. *PLoS Med.* 2007;4(6):e170.
- Adeloye D, Thompson JY, Akanbi MA, Azuh D, Samuel V, Omoregbe N, et al. The burden of road traffic crashes, injuries and deaths in Africa: a systematic review and meta analysis. *Bull World Health Organ.* 2016;94(7):510-521A.
- Manyara CG. Combating Road Traffic Accidents in Kenya: A Challenge for an Emerging Economy. In: Koster MM, Kithinji MM, Rotich JP, editors. *Kenya After 50: Reconfiguring Education, Gender, and Policy.* New York: Palgrave Macmillan US; 2016. p. 101–22. Available from: [https://doi.org/10.1057/9781137574633\\_7](https://doi.org/10.1057/9781137574633_7). Cited 10 Dec 2020.
- Janeway H, O'Reilly G, Schmachtenberg F, Kharva N, Wachira B. Characterizing injury at a tertiary referral hospital in Kenya. *PLoS ONE.* 2019;14(7):e0220179.
- Gathecha GK, Githinji WM, Maina AK. Demographic profile and pattern of fatal injuries in Nairobi, Kenya, January–June 2014. *BMC Public Health.* 2017;17(1):34.
- Patel H, Suarez S, Shaull L, Edwards J, Altawil Z, Owuor J, et al. Patient Characteristics from an Emergency Care Center in Rural Western Kenya. *J Emerg Med.* 2019;56(1):80–6.
- Gathecha GK, Ngaruiya C, Mwai W, Kendagor A, Owondo S, Nyanjau L, et al. Prevalence and predictors of injuries in Kenya: findings from the national STEPs survey. *BMC Public Health.* 2018;18(3):1222.

8. Botchey IM, Hung YW, Bachani AM, Saidi H, Paruk F, Hyder AA. Understanding patterns of injury in Kenya: Analysis of a trauma registry data from a National Referral Hospital. *Surgery*. 2017;162(6):S54–62.
9. House DR, Nyabera SL, Yusi K, Rusyniak DE. Descriptive study of an emergency centre in Western Kenya: Challenges and opportunities. *Afr J Emerg Med*. 2014;4(1):19–24.
10. Otieno T, Woodfield JC, Bird P, Hill AG. Trauma in rural Kenya. *Injury*. 2004;35(12):1228–33.
11. Ministry of Health, Kenya. KENYA STEPwise SURVEY FOR NON COMMUNICABLE DISEASES RISK FACTORS 2015 REPORT.pdf [Internet]. Nairobi, Kenya: Kenya National Bureau of Statistics; 2015 [cited 2021 Mar 2] p. 210. Available from: <https://www.health.go.ke/wp-content/uploads/2016/04/Steps-Report-NCD-2015.pdf>
12. Myers JG, Hunold KM, Ekernas K, Wangara A, Maingi A, Mutiso V, et al. Patient characteristics of the Accident and Emergency Department of Kenyatta National Hospital, Nairobi, Kenya: a cross-sectional, prospective analysis. *BMJ Open*. 2017;7(10):e014974.
13. Odero WO, Kibosia JC. Incidence and characteristics of injuries in Eldoret. *Kenya East Afr Med J*. 1995;72(11):706–10.
14. Saidi HS, Kahoro P. Experience with road traffic accident victims at The Nairobi Hospital. *East Afr Med J*. 2001;78(8):441–4.
15. Odero W, Khayesi M, Heda PM. Road traffic injuries in Kenya: magnitude, causes and status of intervention. *Inj Control Saf Promot*. 2003;10(1–2):53–61.
16. Muguro JK, Sasaki M, Matsushita K, Njeri W. Trend analysis and fatality causes in Kenyan roads: A review of road traffic accident data between 2015 and 2020. Darwish AH, editor. *Cogent Eng*. 2020 Jan 1;7(1):1797981.
17. World Health Organization. Global status report on road safety. 2018. Available from: <https://www.who.int/publications/i/item/9789241565684>. Cited 27 Feb 2021.
18. Jumia. Jumia Mobile Report 2019 in Kenya [Internet]. [cited 2021 Nov 22]. Available from: <https://www.jumia.co.ke/sp-mobile-report/>
19. Schulz A, Ristoski P, Paulheim H. I See a Car Crash: Real-Time Detection of Small Scale Incidents in Microblogs. In: Cimiano P, Fernández M, Lopez V, Schlobach S, Völker J, editors. *The Semantic Web: ESWC 2013 Satellite Events*. Berlin, Heidelberg: Springer; 2013. p. 22–33. (Lecture Notes in Computer Science).
20. Cheng T, Wicks T. Event detection using Twitter: a spatio-temporal approach. *PLoS ONE*. 2014;9(6):e97807.
21. Bao J, Liu P, Yu H, Xu C. Incorporating twitter-based human activity information in spatial analysis of crashes in urban areas. *Accid Anal Prev*. 2017;106:358–69.
22. Callcut RA, Moore S, Wakam G, Hubbard AE, Cohen MJ. Finding the signal in the noise: Could social media be utilized for early hospital notification of multiple casualty events? *PLoS ONE*. 2017;12(10):e0186118.
23. National Transport and Safety Authority. STATUS OF ROAD SAFETY AS AT 31ST OCTOBER, 2020 [Internet]. National Transport and Safety Authority; [cited 2020 Dec 21] p. 2. Available from: <https://www.nts.go.ke/site/wp-content/uploads/2020/11/NTSA-Press-Release-on-Crash-Statistics.pdf>
24. Maudhui Newsroom. The World Bank Says Twitter Could Help End Nairobi Road Accidents [Internet]. Maudhui House. [cited 2021 Jan 11]. Available from: <https://maudhui.co.ke/the-world-bank-says-twitter-could-help-end-nairobi-road-accidents>
25. Milusheva S, Marty R, Bedoya G, Williams S, Resor E, Legovini A. Applying machine learning and geolocation techniques to social media data (Twitter) to develop a resource for urban planning. *PLoS ONE*. 2021;16(2):e0244317.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

