

Utilization of Automated Keyword Search to Identify E-Scooter Injuries in the Emergency Department

Review began 10/27/2021

Review ended 11/10/2021

Published 11/13/2021

© Copyright 2021

Pourmand et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Ali Pourmand¹, Keith S. Boniface¹, Katherine Douglass¹, Colton Hood¹, Sarah E. Frasure¹, Jeremy Barnett¹, Kunj Bhatt¹, Neal Sikka¹

1. Emergency Medicine, George Washington University School of Medicine and Health Sciences, Washington, DC, USA

Corresponding author: Sarah E. Frasure, sarahfrasure@yahoo.com

Abstract

Background and objective

Accurate identification and categorization of injuries from medical records can be challenging, yet it is important for injury epidemiology and prevention efforts. Coding systems such as the International Classification of Diseases (ICD) have well-known limitations. Utilizing computer-based techniques such as natural language processing (NLP) can help augment the identification and categorization of diseases in electronic health records. We used a Python program to search the text to identify cases of scooter injuries that presented to our emergency department (ED).

Materials and methods

This retrospective chart review was conducted between March 2017 and June 2019 in a single, urban academic ED with approximately 80,000 annual visits. The physician documentation was stored as combined PDF files by date. A Python program was developed to search the text from 186,987 encounters to find the string "scoot" and to extract the 100 characters before and after the phrase to facilitate a manual review of this subset of charts.

Results

A total of 890 charts were identified using the Python program, of which 235 (26.4%) were confirmed as e-scooter cases. Patients had an average age of 36 years and 53% were male. In 81.7% of cases, the patients reported a fall from the scooter and only 1.7% reported wearing a helmet during the event. The most commonly injured body areas were the upper extremity (57.9%), head (42.1%), and lower extremity (36.2%). The most frequently consulted specialists were orthopedic and trauma surgeons with 28% of cases requiring a consult. In our population, 9.4% of patients required admission to the hospital.

Conclusions

The number of results and data returned by the Python program was easy to manage and made it easier to identify charts for abstraction. The charts obtained allowed us to understand the nature and demographics of e-scooter injuries in our ED. E-scooters continue to be a popular mode of transportation, and understanding injury patterns related to them may inform and guide opportunities for policy and prevention.

Categories: Public Health, Epidemiology/Public Health, Trauma

Keywords: traumatic injuries, python text search, emergency department, scooter, natural language processing

Introduction

Dockless e-scooters saw tremendous growth in 2017 and rapidly gained popularity as a convenient and environmentally-friendly form of transportation [1]. Several reports of injuries related to e-scooters followed soon after as riders began presenting to emergency departments (EDs) and acute care settings around the country [2,3]. A clear understanding of injury epidemiology is important to inform policy decisions, including aspects such as speed regulations, safe locations for riding, and safety practices for riders. Timely identification of emerging injuries and illnesses is an essential component of injury surveillance. Prevention programs may be delayed due to poor and fragmented data collected from clinical systems, such as electronic health records [4]. The International Classification of Diseases (ICD) was developed by the World Health Organization (WHO) and has been adopted by countries to standardize the documentation of clinical diagnoses for a variety of purposes. This classification system is used to identify the trends, causes, and outcomes of medical cases, but the use of ICD codes as a primary source of data collection can lead to gaps in information, particularly related to the mechanism of injury [5]. Under extraordinary circumstances, such as in response to the ongoing coronavirus disease 2019 (COVID-19) pandemic, ICD-10 codes have been quickly updated and adopted; however, this is not typical [6]. For

How to cite this article

Pourmand A, Boniface K S, Douglass K, et al. (November 13, 2021) Utilization of Automated Keyword Search to Identify E-Scooter Injuries in the Emergency Department. Cureus 13(11): e19539. DOI 10.7759/cureus.19539

example, for injuries related to e-scooters, the ICD-10-CM V00.8441A “Fall from standing electric scooter, initial encounter” was not added until October 2020 [7]. Potentially dangerous injuries could be inconsistently reported by the emergency medicine community through the improper use of ICD-10 codes [8] and both the underutilization of as well as the erroneous use of ICD-10 codes poses a challenge to attaining a comprehensive understanding of specific domains such as work-related injuries in the ED [9].

Natural language processing (NLP) refers to the field that combines computer science, artificial intelligence, and linguistics to enable computers to understand or even create human languages [10]. Clinicians documenting patient care in the electronic health record generate significant amounts of free text, and NLP can be used to bring meaning and structure for further analysis [11]. Processing free text documentation has been used in the past to accurately recognize simple wording in medical language to identify potentially underreported diseases when analyzed only by ICD-10 classification [12]. In combination with ICD codes, it can help identify uncommon but serious diseases [13] and free text can be searched to identify conditions independent of the ICD code [14].

Like other EDs, our urban tertiary care center witnessed increasing cases of injuries related to e-scooters soon after their introduction within the community. Being aware of the limitations and challenges of extracting injury information from our electronic health record, we undertook this study to better understand the epidemiology of e-scooter-related injuries. Our primary objective was to identify e-scooter injuries by using a Python computer program to search the text of our ED physician documentation. Our secondary objective was to characterize the patient demographics and injury patterns associated with e-scooter use that result in ED visits.

This article was previously presented as an abstract at the ACEP Research Forum, October 26-29, 2020 (Douglass K, Sikka N, Boniface K, Bhatt K, McCarville P, Pourmand A. Epidemiological Analysis of E-Scooter Injuries among Patients Presenting to the Emergency Department. ACEP Research Forum; October 26-29, 2020).

Materials And Methods

This retrospective chart review was conducted between March 2017 and June 2019 in a single, urban academic ED with approximately 80,000 annual visits. The physician documentation was stored as combined PDF files by date. A Python program was used to prepare and search the text from 186,987 encounters to identify charts. First, the text was converted to lower case, and scooter injuries were identified using the standard library string method to locate occurrences of “scoot” and to extract the 100 characters before and after this location. This string along with the unique identifier of its encounter was written into a comma-separated value (CSV) file to facilitate a manual review of these identified charts. An error file was generated to list any files that could not be opened or processed. We included all patients identified as presenting to the ED with an e-scooter-related injury.

Charts identified in the keyword search were reviewed by research assistants against our inclusion criteria. Research assistants were trained in data abstraction, and progress was monitored through regular meetings. A standard data abstraction form was developed by the research team including patient-centered epidemiologic data, injuries sustained, treatments rendered, and circumstances surrounding the injury. No personally identifiable information was included. The form was pilot tested, finalized, and entered into a Research Electronic Data Capture (REDCap) database for secure collection and storage. Study data were collected and managed using REDCap electronic data capture tools hosted at (BLINDED). REDCap is a secure, web-based software platform designed to support data capture for research studies, providing (1) an intuitive interface for validated data capture, (2) audit trails for tracking data manipulation and export procedures, (3) automated export procedures for seamless data downloads to common statistical packages, and (4) procedures for data integration and interoperability with external sources [15,16].

Of note, 10% of charts were randomly selected for abstraction by a second blinded researcher to assess inter-observer reliability. Data were analyzed using descriptive and summary statistics. The study was reviewed and deemed exempt from approval by our institutional review board.

Results

A total of 186,987 electronic charts comprising 1,153,882 pages were searched. The search was completed in 9.26 hours. It took an average of 0.029 seconds to search each page and 0.178 seconds to search each chart. The search identified 890 charts that contained the string “scoot”. A total of 235 (26.4%) e-scooter cases were identified upon manual review of the search results. The error file contained no entries.

Table 1 presents the characteristics of the patients injured in e-scooter crashes. Because our hospital catchment area includes Washington DC, Maryland, and Virginia, the majority of patients resided in one of these jurisdictions. In 81.7% of injuries, the patient reported a fall from the scooter, and only 1.7% reported wearing a helmet during the event. The most commonly injured body areas were the upper extremity (57.9%), head (42.1%), and lower extremity (36.2%) with the injuries described as abrasion (82.5%), fracture (39.1%), and laceration (31.5%). X-ray and CT scans were performed on 69.4% and 34.5% of patients

respectively. Specialists, most commonly orthopedic and trauma surgeons, were consulted in 28% of cases; 9.4% of patients required hospital admission.

Characteristics	Values
Patient sex/gender, n (%)	
Male	124 (52.8%)
Female	111 (47.2%)
Patient age, years	
Mean	36.02
Median (IQR)	31 (23)
Range	7–89
State of residence, n (%)	
Washington, DC	116 (49.4%)
Maryland	14 (6.0%)
Virginia	22 (9.4%)
Other	83 (35.3%)
Position on the scooter, n (%)	
Riding alone on the scooter	206 (87.7%)
Riding double, in the front	5 (2.1%)
Riding double, in the back	1 (0.4%)
Injured by someone riding on a scooter	4 (1.7%)
Otherwise injured related to a scooter	7 (3.0%)
Not recorded	12 (5.1%)
Location of the scooter at the time of the incident, n (%)	
Street	107 (45.5%)
Sidewalk	19 (8.1%)
Bike lane	1 (0.4%)
Not recorded	108 (46.0%)
Circumstances of the incident, n (%)	
Collision	35 (14.9%)
Related to uneven surface	35 (14.9%)
Lost control	32 (13.6%)
Fell off the device	192 (81.7%)
Other	16 (6.8%)
Not recorded	6 (2.6%)

Reported alcohol or drug usage, n (%)	
Alcohol	19 (8.1%)
Drug use	1 (0.4%)

TABLE 1: Patient characteristics (n=235)

IQR: interquartile range

Table 2 describes the injury characteristics.

Characteristics	Values
Area of injury/physical exam finding, n (%)	
Head	99 (42.1%)
Upper extremity	136 (57.9%)
Lower extremity	85 (36.2%)
Trunk	13 (5.5%)
Other	9 (3.8%)
Type of injury*, n (%)	
Abrasion	123 (52.3%)
Contusion	55 (23.4%)
Concussion	13 (5.5%)
Laceration	74 (31.5%)
Fracture	92 (39.1%)
Dislocation	4 (1.7%)
Other	32 (13.6%)
Imaging procedures in ED*, n (%)	
X-ray	163 (69.4%)
CT	81 (34.5%)
MRI	1 (0.4%)
Ultrasound	6 (2.6%)
None	32 (13.6%)
Type of specialty consultation*, n (%)	
Orthopedic surgery	36 (15.3%)
Neurosurgery	8 (3.4%)
Trauma surgery	21 (8.9%)
ENT	12 (5.1%)
Ophthalmology	5 (2.1%)

Neurology	2 (0.9%)
Plastic surgery	1 (0.4%)
Patient disposition, n (%)	
Home	215 (91.5%)
Admitted to floor	17 (7.2%)
Admitted to ICU	3 (1.3%)

TABLE 2: Injury characteristics

*Of note, some patients fit into multiple categories, and hence percentages may add up to >100% and numbers may exceed 235

CT: computed tomography; ED: emergency department; ENT: ears, nose, and throat; ICU: intensive care unit; MRI: magnetic resonance imaging

Of the 235 patient charts reviewed, 185 (78.7%) included an external cause code related to the mechanism of injury (“e-code”). Of those charts with an e-code, 53% of those codes included the word “scooter” (Table 3). The other 47% included various mechanisms of injury, many related to falls or motor vehicle collisions.

ICD-10 code	Number of patients (percentage)
V00141A (fall from scooter)	27 (14.6%)
V00831A (fall from motorized mobility scooter)	50 (27.0%)
W051XXA (fall from non-moving, non-motorized scooter)	10 (5.4%)
V00832A (motorized mobility scooter colliding with stationary object)	2 (1.1%)
W052XXA (fall from non-moving motorized mobility scooter)	5 (2.7%)
V00148A [other scooter (non-motorized) accident]	2 (1.1%)
V00142A [scooter (non-motorized) colliding with stationary object]	1 (0.5%)
V00838A (other accident with motorized mobility scooter)	1 (0.5%)
Non-scooter codes	87 (47.0%)

TABLE 3: Mechanism of injury codes

ICD: International Classification of Diseases

Discussion

Real-time identification of emerging patterns of disease is an ongoing challenge encountered in injury research and prevention and other emerging diseases. Reliance on ICD-10 coding to incorporate both mechanism of injury and actual injury patterns into electronic medical records offers an incomplete solution, although textual components of the chart will often include important factors omitted in the codes. As described in this study, only 52% of the reviewed charts had ICD-10 codes including the word scooter, and none of the scooter-related codes were appropriate for the new e-scooter device. The 2021 version of ICD-10 includes an extensive array of updated codes related to e-scooters [17], although clinician coding will always have some form of expected deficiency [18]. This study described a pilot effort to identify incidents of e-scooter injury using text searching to improve and expedite the determination of the epidemiologic profile of injuries related to an evolving transportation technology.

Meanwhile, e-scooters have continued to grow in popularity [19] due to their environmental benefits while continuing to be a source of injuries worldwide. Urban centers may see different patterns of injury depending on various characteristics such as road infrastructure, availability of bike lanes, topography, and level of tourism. For example, in our study, over 35% of patients with scooter-related injuries reside outside our local area. A better understanding of injury patterns will help to improve policymaking and prevention

strategies. In our study, 45.5% of scooter crashes occurred on the street, which may influence future policy changes in trafficking dynamics with e-scooters in injury prevention. We also found that 8.1% of our e-scooter riders reported alcohol use while operating the e-scooter, which may increase the likelihood of crashes as well as the number and severity of injuries. Our data detail a relatively low acuity injury pattern, as 91.5% of patients were discharged home from the ED. However, the vast majority of patients (86.4%) required imaging studies of some kind, and 28% required specialty consultation. More detailed information regarding rider factors, such as experience in riding and familiarity with routes, and environmental factors such as dangerous intersections will be important to fully understand opportunities for injury prevention.

The review of the charts was conducted at a single center, limiting the population catchment, which may make the generalization of the results difficult. Epidemiological data may be more difficult to find based on the charting style of different medical professionals. We did not identify cases using ICD-10 codes but used a search string and surrounding characters to locate charts; further research will be needed to evaluate the effectiveness of the tool in retrospective chart reviews in the ED. Multi-center studies in the future will be effective in providing deeper insights into the injury pattern, population, and ED treatment patterns.

Conclusions

The Python program used in this study was able to identify charts related to scooter injuries based on the search for the string "scoot" in the ED physician documentation text. The search returned a manageable number of charts, in addition to 100 words surrounding the location of the target string to enable the rapid identification and abstraction of e-scooter injuries. Data collection from these charts provided an understanding of the demographics and circumstances of e-scooter-related injuries presenting to our ED in the District of Columbia. E-scooters continue to be a popular mode of travel in urban centers as well as a major source of injuries that present to the ED. Future research should combine analysis of the text with ICD-10 codes to maximize the identification of potential injuries. Lastly, our study method combined with ICD search results should help us identify and categorize injury documentation, thereby allowing us to create a training set for the future development of more advanced injury identification using machine learning techniques such as NLP.

Additional Information

Disclosures

Human subjects: Consent was obtained or waived by all participants in this study. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Kobayashi LM, Williams E, Brown CV, et al.: The e-merging e-pidemic of e-scooters. *Trauma Surg Acute Care Open*. 2019, 4:e000337. [10.1136/tsaco-2019-000337](https://doi.org/10.1136/tsaco-2019-000337)
2. Badeau A, Carman C, Newman M, Steenblik J, Carlson M, Madsen T: Emergency department visits for electric scooter-related injuries after introduction of an urban rental program. *Am J Emerg Med*. 2019, 37:1551-3. [10.1016/j.ajem.2019.05.003](https://doi.org/10.1016/j.ajem.2019.05.003)
3. Trivedi TK, Liu C, Antonio AL, et al.: Injuries associated with standing electric scooter use. *JAMA Netw Open*. 2019, 2:e187381. [10.1001/jamanetworkopen.2018.7381](https://doi.org/10.1001/jamanetworkopen.2018.7381)
4. Centers for Disease Control and Prevention. Data science strategy for injury and violence prevention. (2020). Accessed: November 12, 2021: https://www.cdc.gov/injury/pdfs/data-science/Data-Science-Strategy_FINAL_508.pdf.
5. Clery MJ, Hudson PJ, Moore JC, Mercer Kollar LM, Wu DT: Exploring injury intentionality and mechanism via ICD-10-CM injury codes and self-reported injury in a large, urban emergency department. *Inj Prev*. 2021, 27:i62-5. [10.1136/injuryprev-2019-043508](https://doi.org/10.1136/injuryprev-2019-043508)
6. Kadri SS, Gundrum J, Warner S, Cao Z, Babiker A, Klompas M, Rosenthal N: Uptake and accuracy of the diagnosis code for COVID-19 among US hospitalizations. *JAMA*. 2020, 324:2553-4. [10.1001/jama.2020.20323](https://doi.org/10.1001/jama.2020.20323)
7. 2021 ICD-10-CM Diagnosis Code V00. 841A: Fall from standing electric scooter, initial encounter. (2021). Accessed: April 22, 2021: <http://www.ICD10data.com>.
8. Lavoie-Gagne O, Siow M, Harkin W, Flores AR, Girard PJ, Schwartz AK, Kent WT: Characterization of electric scooter injuries over 27 months at an urban level 1 trauma center. *Am J Emerg Med*. 2021, 45:129-36. [10.1016/j.ajem.2021.02.019](https://doi.org/10.1016/j.ajem.2021.02.019)
9. Harduar Morano L, Richardson D, Proescholdbell S: Descriptive evaluation of methods for identifying work-related emergency department injury visits. *Am J Ind Med*. 2019, 62:568-79. [10.1002/ajim.22984](https://doi.org/10.1002/ajim.22984)
10. Hirschberg J, Manning CD: Advances in natural language processing. *Science*. 2015, 349:261-6. [10.1126/science.aaa8685](https://doi.org/10.1126/science.aaa8685)
11. Kreimeyer K, Foster M, Pandey A, et al.: Natural language processing systems for capturing and standardizing unstructured clinical information: a systematic review. *J Biomed Inform*. 2017, 73:14-29.

- [10.1016/j.jbi.2017.07.012](https://doi.org/10.1016/j.jbi.2017.07.012)
12. Kaspar M, Fette G, Güder G, et al.: Underestimated prevalence of heart failure in hospital inpatients: a comparison of ICD codes and discharge letter information. *Clin Res Cardiol*. 2018, 107:778-87. [10.1007/s00392-018-1245-z](https://doi.org/10.1007/s00392-018-1245-z)
 13. Udelsman B, Chien I, Ouchi K, Brizzi K, Tulsy JA, Lindvall C: Needle in a haystack: natural language processing to identify serious illness. *J Palliat Med*. 2019, 22:179-82. [10.1089/jpm.2018.0294](https://doi.org/10.1089/jpm.2018.0294)
 14. Nellamattathil M, Amber I: An evaluation of scooter injury and injury patterns following widespread adoption of E-scooters in a major metropolitan area. *Clin Imaging*. 2020, 60:200-3. [10.1016/j.clinimag.2019.12.012](https://doi.org/10.1016/j.clinimag.2019.12.012)
 15. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG: Research electronic data capture (REDCap)-- a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009, 42:377-81. [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)
 16. Harris PA, Taylor R, Minor BL, et al.: The REDCap consortium: building an international community of software platform partners. *J Biomed Inform*. 2019, 95:103208. [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)
 17. 2021 ICD-10-CM. (2021). Accessed: June 8, 2021: <https://www.cms.gov/medicare/icd-10/2021-icd-10-cm>.
 18. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM: Measuring diagnoses: ICD code accuracy. *Health Serv Res*. 2005, 40:1620-39. [10.1111/j.1475-6775.2005.00444.x](https://doi.org/10.1111/j.1475-6775.2005.00444.x)
 19. U.S. scooter ridership surged in 2019. Now what? . (2020). Accessed: May 27, 2021: <http://www.bloomberg.com/news/articles/2020-08-27/how-big-was-2019-s-scooter-boom-and-what-s-next>.