Journal of Transportation Safety & Security

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/utss20

Young drivers and cellphone distraction: Pattern recognition from fatal crashes

Md Mahmud Hossain, Huaguo Zhou, Subasish Das, Xiaoduan Sun & Ahmed Hossain

To cite this article: Md Mahmud Hossain, Huaguo Zhou, Subasish Das, Xiaoduan Sun & Ahmed Hossain (2022): Young drivers and cellphone distraction: Pattern recognition from fatal crashes, Journal of Transportation Safety & Security, DOI: 10.1080/19439962.2022.2048763

To link to this article: https://doi.org/10.1080/19439962.2022.2048763

Published online: 28 Mar 2022.

JOURNAL OF TRANSPORTATION SAFETY & SECURITY
https://doi.org/10.1080/19439962.2022.2048763

Young drivers and cellphone distraction: Pattern recognition from fatal crashes

Md Mahmud Hossain[a], Huaguo Zhou[a], Subasish Das[b], Xiaoduan Sun[c], and Ahmed Hossain[c]

[a]Department of Civil and Environmental Engineering, Auburn University, Auburn, AL, USA; [b]Texas A&M Transportation Institute, Bryan, TX, USA; [c]Department of Civil Engineering, University of Louisiana at Lafayette, Lafayette, LA, USA

ABSTRACT

More than 30% of cellphone-distracted fatal crashes occurred to drivers younger than 25-years-old in 2018, even though they constitute less than 12% of total licensed drivers in the U.S. Using joint correspondence analysis (JCA), this study analyzed six years (2014–2019) of cellphone-related fatal crashes involving young drivers based on the data from the Fatality Analysis Reporting System (FARS). This unsupervised learning algorithm can graphically display the co-occurrence of variable categories in a lower-dimensional space by effectively summarizing the knowledge of a complex crash dataset. The Boruta algorithm was applied to select the relevant features from the preliminary crash dataset. The empirical results of JCA manifest a few interesting fatal crash patterns. For example, young male drivers in light trucks were involved in deadly collisions while performing specific cellphone activities (other than talking and listening), cellphone-related fatal crashes occurred to young females with prior crash records, and so on. Apart from alcohol and drug involvement, this study identified young drivers' additional risk-taking maneuvers while engaged in cellphone usage, including: disregarding traffic signs and signals, speeding, and unrestrained driving. The associations could guide the safety officials and policymakers in developing appropriate engineering, education, and enforcement strategies when dealing with cellphone-distracted young drivers.

KEYWORDS

Young driver; cellphone distraction; correspondence analysis; fatal crash; crash data analysis

1. Introduction

The rapid growth of smartphone users has escalated the frequency of cellphone usage while driving. From roadside surveys in Virginia, the Insurance Institute for Highway Safety (IIHS) reported that the proportion of drivers manipulating a cellphone while driving increased from 2.3% in 2014 to 3.4% in 2018 (Insurance Institute for Highway Safety (IIHS), 2019). The greater involvement of drivers in such risky behavior raises

CONTACT Md Mahmud Hossain mahmud@auburn.edu Department of Civil and Environmental Engineering, Auburn University, Auburn, AL 36849-5412, USA.

cellphone prevalence as a critical determinant of distracted driving. Moreover, the steady evidence of deadly crashes associated with in-vehicle cellphone use galvanizes the public attention in this issue. More than 350 drivers using a cellphone while driving were involved in fatal collisions in 2018, representing 13% of distracted drivers in the U.S. (NHTSA, 2020). This estimated percentage remains relatively identical in the five years (2014–2018) of national statistics (NHTSA, 2020). However, the actual proportion is expected to be higher than reported, as distracted driving is often underreported due to a limited source of relevant information at the crash spot (NSC, 2013). To minimize drivers' engagement in non-driving tasks, several states in the U.S. have enacted distracted driving legislation that prohibits the use of wireless telecommunication devices during driving. Most of the policymakers concentrate on either the mechanism of cellphone distraction (texting or other handheld options), or the high-risk driver population classes (young and novice drivers) (Li et al., 2020). However, the effectiveness of such state-based cellphone laws is still inconclusive for adolescent drivers (Zhu, Rudisill, Heeringa, Swedler, & Redelmeier, 2016). The National Highway Traffic Safety Administration (NHTSA) noted that the percentage of young drivers aged 16–24 years manipulating phones soared from 1.1% in 2009 to 4.2% in 2018 (NHTSA, 2019a), found from the National Occupant Protection Use Survey (NOPUS). In addition, this driver cohort accounted for more than 30% of cellphone-distracted fatal crashes in 2018, although they constituted less than 12% of all licensed drivers in the U.S.

The high level of phone involvement and lower risk perception induces young drivers as a vulnerable group in distracted driving. This novice driver group usually shows a greater propensity toward risk-taking behaviors compared to older drivers because of less experience (Mayhew, Simpson, & Pak, 2003), sensation-seeking disposition (Jonah, Thiessen, & Au-Yeung, 2001), and ongoing cognitive development (Keating & Halpern-Felsher, 2008). Adolescents often overestimate their vehicle controlling skills, which increases their participation in numerous cellphone-related tasks (e.g., talking, texting, listening). However, these secondary activities have already proven vulnerable to safe driving due to their association with any combination of visual (taking eye glances away from roadways), manual (removing hands from the steering wheel), and cognitive (disregarding critical information of surrounding driving environment) distractions (Jannusch, Shannon, Völler, Murphy, & Mullins, 2021; Nasr Esfahani, Arvin, Song, & Sze, 2021). Also, repeated usage of cellphones affects overall driving performances by increasing response times, shortening headways, and increasing deviations of lateral control (Haque & Washington, 2015).

Using a cellphone before a crash does not specify that the incident was caused solely by cellphone-distracted driving. Each crash reflects a sequence of events covering driver, road, environment, and crash-related factors (Das et al., 2022; Hossain, Rahman, Sun, & Mitran, 2021). The effect of one contributory factor can be more ruminative when combined with others. For example, in relation to cellphone-related crashes, unrestrained driving on high-speed roads may be more vulnerable than that on low-speed roadways. Although the effect of distraction varies among driver population classes (Guo et al., 2017), no previous studies have been carried out to reveal the shrouded interrelations among contributory factors that caused collisions involving drivers younger than 25 years, distracted by cellular phones. This study utilizes six years (2014–2019) of police-reported fatal crashes from the Fatality Analysis Reporting System (FARS) to identify the related associations by applying a unique dimension reduction method known as joint correspondence analysis (JCA).

2. Literature review

The enormity of cellphone distraction has been reflected in research interests over the last decade. Numerous driving simulators, on-road field surveys, and naturalistic driving studies (NDS) have been conducted to describe the links between modes of cellphone usage and their related risk of crash involvement in terms of driving performance. cellphone conversation increases drivers' attention lapses (Beede & Kass, 2006), which subsequently impairs the risk-perception ability and appropriate reactions necessary when facing any hazardous events (Haque & Washington, 2015). Conversely, visual-manual tasks on cellphones (e.g., typing text messages) negatively affect nearly all aspects of safe driving, which significantly elevates the collision risk (Caird, Simmons, Wiley, Johnston, & Horrey, 2018; Owens et al., 2018), sometimes even more than alcohol-intoxicated driving (Klauer et al., 2006).

Young drivers have been emphasized for their considerably higher risk of crashes or near-crash episodes in a variety of cellphone-related activities (e.g., talking, texting, dialing, reaching for the device) behind the wheel (Klauer et al., 2014; NHTSA, 2012; Rahman, Hossain, Mitran, & Sun, 2021). Young male drivers are reported for greater exposure in cellphone usage (Jannusch et al., 2021; Shaaban, Gaweesh, & Ahmed, 2018), which can be readily explained as a substantial body of literature has verified male teenagers' higher risky driving tendencies compared to female drivers (Byrnes, Miller, & Schafer, 1999; Shope & Bingham, 2008). Adolescents who repeatedly speak on phones are more likely to be involved in risky driving behaviors such as disregarding traffic signals (Haque, Ohlhauser, Washington, & Boyle, 2013), changing lanes incessantly (Zhao, Reimer, Mehler, D'Ambrosio, & Coughlin, 2013), and speeding (Jannusch et al., 2021). Besides, multiple studies have documented a combined aspect of impairment (under the influence of alcohol or drug) and cellphone usage during driving among young drivers (Jannusch et al., 2021; Li, Bower, Zhu, & Board, 2019).

Table 1 represents the key findings of a few studies that specifically focus on the cellphone usage of young drivers. The purpose of this summarization is to discern the associated contributing factors linked with the attributes of police-investigated crash reports.

More than a decade earlier, Huang, Stutts, and Hunter (2003) analyzed six years of North Carolina crash data to examine the contributory factors associated with cellphone-related crashes by comparing them with non-cellphone collisions. The authors argued that drivers who use a cellphone were more frequently involved in rear-end crashes, and more likely to drive in urban areas, often during mid-day and afternoon hours. Regarding cellphone conversation prior to a crash, most of the drivers were males, going straight, and driving sport utility vehicles (SUVs). Later, a telephone survey by Beck, Yan, and Wang (2007) investigated dispositional and behavioral factors of cellphone drivers to underline their self-reported risk-taking maneuvers. The study identified lower seatbelt usage among cellphone users who had previous crash or violation records. A national phone survey was conducted by NHTSA to further update the assessment on driving attitudes and behaviors allied with distracted drivers (Tison, Chaudhary, & Cosgrove, 2011). The respondents showed more affinity toward talking on a cellphone when stopped or traveling at lower speeds. Rahman, Sun, Sun, and Shan (2020) performed a case study in Louisiana to understand how the association of multiple factors influence driver's cellphone use and found rurality, road geometry, and crash time as crucial factors in differentiating driving patterns with respect to in-vehicle cellphone usage.

The association between cellphone-distracted driving and environmental characteristics such as day of the week, lighting conditions, and the weather has not been well elucidated. Young, Rudin-Brown, and Lenné (2010) compared the rates of handheld and hands-free cellphone usage by time of day and day of the week, whereas Vivoda, Eby, St. Louis, and Kostyniuk (2008) only concentrated on distinguishing the changes of rates within the nighttime hours. Both observational studies conveyed no significant dissimilarities. In relation to weather, studies focused on the distribution of cellphone users by the variable categories (McDonald & Sommers, 2015; Wang, Xu, Asmelash, Xing, & Lee, 2020) rather than on understanding the crash risk of cellphone-distracted driving in adverse weather conditions.

Table 1. Summary of selected young driver cellphone studies.

Study	Study classification, location, approach	Key findings
Hosking et al. (2009)	Experimental-based, Melbourne (Australia), Descriptive analysis	• While texting, young drivers spent more time looking away from the roads, therefore, they often missed the regular traffic signs and signals.
Atchley et al. (2011)	Survey-based, Kansas (USA), Structural equation model	• The choice of initiating texting while driving was significantly influenced by weather, road lighting, and highway class.
Tucker et al. (2015)	Survey-based, Ontario (Canada), Descriptive analysis	• No significant gender difference had been observed in talking on a cellphone while driving. • Young male drivers were often reported for violating the posted speed limits when engaged in texting during driving.
Shaaban et al. (2018)	Survey-based, Qatar, Structural equation model	• Young drivers who had previous cellphone-related crash records were more likely to be involved in cellphone use while driving. • With longer distraction duration and more driving experience, young drivers showed a greater tendency of cellphone involvement.
Jannusch et al. (2021)	Survey-based, Germany, Correspondence analysis	• Speaking on cellphones was more frequent among young male drivers. • Young drivers who reported talking on a phone were more likely to be engaged in speeding and intoxicated driving.
Li et al. (2019)	Survey-based, USA, Poisson regression model	• Adolescents who were older and reported alcohol-involved driving were more likely to be involved in cellphone usage behind the wheel.
Walshe, Winston, Betancourt, Arena, and Romer (2018)	Survey-based, Philadelphia (USA), Descriptive analysis	• Young drivers' cellphone usage was associated with prior citation records and intentional violations while driving.

In brief, a wide array of experimental research has demonstrated how certain aspects of driving performance change with cellphone distraction, which indirectly evaluates the relative crash risk (Beede & Kass, 2006; Caird et al., 2018; Haque & Washington, 2015; Hosking, Young, & Regan, 2009). Also, survey-based and observational studies have sought to connect drivers' cellphone behavior with a multitude of factors (Atchley, Atwood, & Boulton, 2011; Beck et al., 2007; Jannusch et al., 2021; Rahman et al., 2020; Tison et al., 2011). However, these studies provide sparse knowledge on the road and environment-related factors that could influence cellphone involvement during driving. On the contrary, NDS and crash-based studies have predominantly included common road crashes (minor injury or near-crash events). Therefore, contributing factors to serious crash incidents and their associations remain unclear. In addition, the majority of them applied either simple descriptive analyses or traditional parametric models (e.g., risk ratio) to recognize the effects of individual risk factors on a response variable (Huang et al., 2003; Klauer et al., 2006; Owens et al., 2018; Shaaban et al., 2018; Tucker, Pek, Morrish, & Ruf, 2015). However, these analytical approaches have been criticized for their predefined assumptions, which can lead to biased and incompatible results (Montella, Mauriello, Pernetti, & Rella Riccardi, 2021; Nafis, Alluri, Wu, & Kibria, 2021). In relation to cellphone-distracted driving, the national crash statistics have consistently highlighted young drivers as a highly at-risk age group. However, no previous studies have investigated police-reported cellphone collisions involving them to discern the associated risk factors. The comprehensive literature review has solicited more emphasis and extensive research in characterizing the complex nature of the related fatal crash incidents.

According to the National Safety Council (NCS), crash deaths in samples where drivers were on the phone are seriously underreported. However, there is no robust evidence on the portion of underreporting in comparison to the actual cellphone-involved crash incidents. Also, this underreporting may vary by crash type and geography. Therefore, designing a comparative study (cellphone versus non-cellphone collisions) may not be a suitable research strategy. In this regard, extracting the cellphone-related crashes from an established representative of police-reported road crashes could minimize the effect of underreporting while aiming to detect real-world crash patterns. This study utilizes the FARS crash database that includes a broad spectrum of crash, driver, vehicle, road, and environmental variables involved in each deadly collision. JCA is applied to discover the fatal crash patterns of young drivers distracted by cellphones. Researchers have already adopted JCA to subdue the shortcomings of conventional statistical modeling (e.g., the covariates are mutually exclusive) while investigating critical traffic safety issues (Das, Jha, Fitzpatrick, Brewer, & Shimu, 2019; Hossain, Sun, Mitran, & Rahman, 2021). This unsupervised learning algorithm can graphically represent the interdependencies among crash contributing factors from a multidimensional dataset without imposing any predefined hypotheses. Since underreporting of cellphone-related crashes is still a key concern, prioritizing critical crash attributes from the associations could be beneficial in reducing the related collisions and fatalities. Additionally, the combination of factors identified from JCA depicts real-world deadly crash scenarios that can assist safety officials in designing effective crash mitigation strategies.

3. Methodology

Correspondence analysis (CA) is a popular multivariate technique that utilizes nominal variables to reflect their interrelations in a lower-dimensional plane. Multiple correspondence analysis (MCA) and JCA are two prevalent applications of CA for exploratory analysis. In contrast to MCA, JCA is a well-defined pattern recognition method as it performs natural generalization of variables by a least square approach (Burt matrix) instead of a multiple indicator matrix (Camiz, Matematica, & Universit, 2013; Hossain, Sun et al., 2021). The model considers the joint display of cross-tabulations among all attainable pairs of categorical factors in a dataset (Greenacre, 1988). However, the application of a Burt Matrix has certain shortcomings in the diagonal block-entries of the total variation. The problem can be adjusted by estimating the maximum likelihood for an r-way distribution in JCA (Vermunt & Anderson, 2005). The idea of such an alternative approach, suggested by Choulakian (1988), is to evaluate the goodness-of-fit of the models by employing standard chi-squared tests through three-way contingency tables.

Assume that categorical variables $Y_1, Y_2, \ldots, Y_r$ form a traditional cell proportion $\vartheta^{Y_1 Y_2 \ldots Y_r}_{y_1 y_2 \ldots y_r}$ in the r-way contingency table. For each variable pair Y_r and Y_s, the bivariate marginal distributions $\vartheta^{Y_r Y_s}_{y_r y_s}$ of a U-dimensional JCA model can be defined as:

$$\vartheta^{Y_r Y_s}_{y_r y_s} = \vartheta^{Y_r}_{y_r} \vartheta^{Y_s}_{y_s} \{1 + \sum_{m=1}^{U} \varepsilon_u \omega^{Y_r}_{u y_r} \omega^{Y_s}_{u y_s}\}$$

where $r \neq s$

Here,

$\vartheta^{Y_r}_{y_s}$ = univariate cell distribution in the U-way cross-tables formed by variable Y_r,

$\omega^{Y_r}_{u y_r}$= quantification of attribute y_r for variable Y_r,

ε_u = singular value of correlations between the variables in dimension u

A multivariate correlation model (MCM) has been suggested to extend the JCA model for the r-way tables (Vermunt & Anderson, 2005).

$$\vartheta^{Y_1 Y_2 \ldots Y_r}_{y_1 y_2 \ldots y_r} = \prod\nolimits_{r=1}^{R} \vartheta^{Y_r}_{y_r} \{1 + \sum_{u=1}^{M} \sum_{r=1}^{R} \sum_{s=r+1}^{R} \gamma_u \pi^{Y_r}_{u y_r} \pi^{Y_s}_{u y_s}\}$$

Here,

γ_u = canonical correlation in dimension u,

$\pi^{Y_r}_{u y_r}$= quantification of attribute y_r of variable Y_r for dimension u.

It is noteworthy that this estimation will provide the best fitting when all $\vartheta^{Y_1 Y_2 \ldots Y_r}_{y_1 y_2 \ldots y_r}$ are within the acceptable range.

In JCA, cloud means a combination of points (categories) in a lower-dimensional (first plane) space. Figure 1 exhibits a simple representation of cloud formation. Three variables e, f, and g contain three (e_1, e_2, e_3), four (f_1, f_2, f_3, f_4), and four (g_1, g_2, g_3, g_4) attributes, respectively. Categories with similar distribution are closely spaced, which specifies a significant correlation (Das et al., 2019; Hossain, Sun et al., 2021). In the figure, the relative proximity of the coordinates of four attributes (e_2, f_1, g_2, g_3) that generate a combination cloud is apparent.

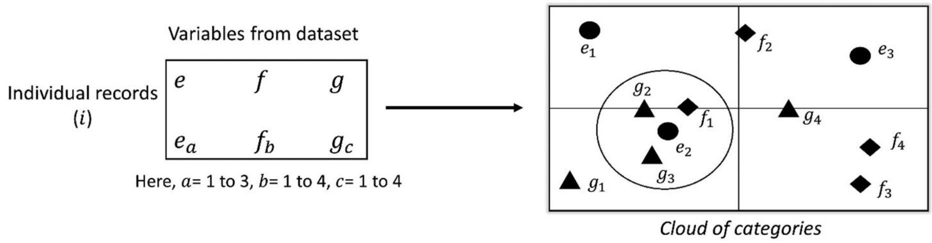

Figure 1. Joint correspondence analysis plots.

This study performs JCA using the "ca" package on the statistical software 'R' to analyze fatal crashes of young drivers involved in cellphone usage before the incidents.

4. Study data

4.1. Data preparation

The FARS is a nationwide database that stores all fatal motor vehicle crashes in the U.S. from 1975, including the District of Columbia and Puerto Rico. The archive merely includes the deadly collisions on public roads that resulted in the fatality of a driver, passenger, or non-motorists within the first thirty days of the crash incident. The FARS database contains multiple tables to reposit all fatal crash information at the crash, vehicle, and person levels. Since 2010, the FARS has included three distinct cellphone distractors to secure distracted driving information more comprehensively (NHTSA, 2019b). The categories are as follows.

- Talking or listening to cellular phone (hand-held or hand free).
- Manipulating cellular phone (dialing/texting/browsing on a cellphone).
- Other cellphone related (locating/reaching for/answering cellular phone or any other cellphone usage other than previous two categories).

Figure 2 exhibits the analytical framework of this study with data pre-processing techniques. This research extracted cellphone-related fatal crash information from the FARS for 2014 to 2019. Nationwide, some form of Graduated Driver Licensing (GDL) system has been implemented to reduce novice young drivers' crash risk by gradually increasing their exposure to complex driving scenarios (Baker, Chen, & Li, 2011). A typical GDL program includes a learner stage, an intermediate stage, and a full privilege stage. However, the specific components of GDL provisions vary by state, and teenagers in the probationary stages of the GDL program operate vehicles in numerous controlled environments (e.g., supervised driving, nighttime restriction, teen passenger restriction). Therefore, for analytics

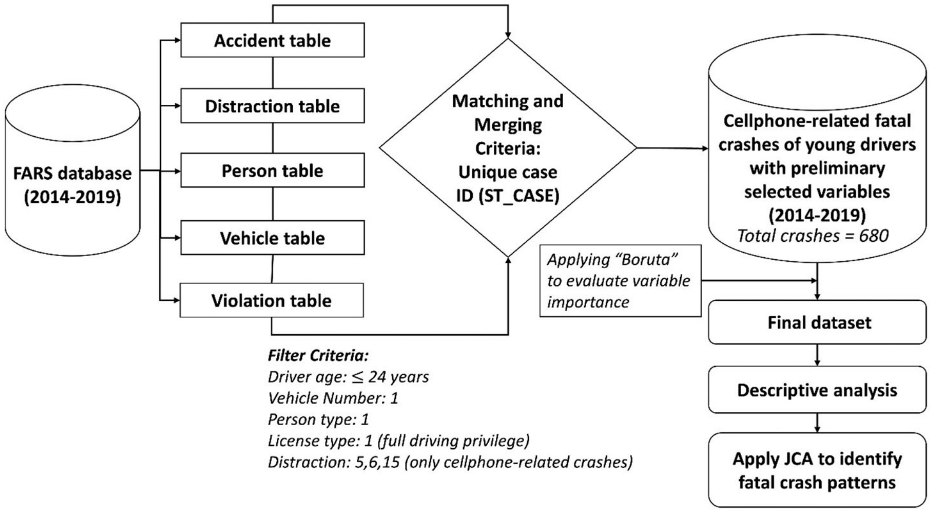

Figure 2. Analytical framework of the study.

and interpretive purposes, the researchers filtered out the fully licensed young driver crashes from the original crash database. Also, only driver information of vehicle 1 (major liability holder in a crash) was considered. Extricating the information regarding the level of liability of other drivers (vehicle 2, 3, etc.), with the aim of identifying all responsible young drivers involved in multiple-vehicle crashes was found to be unfeasible.

The primary dataset contains 680 unique, cellphone-distracted, 16 to 24 years old driver fatal crashes with sixteen driver, environment, road, and crash-related factors. Occupant characteristics are not considered in this study. The variables were selected based on previous cellphone-related studies and the availability of contributory factors in the FARS database. Table 2 represents the initially selected variables with categories. An extensive literature review and engineering judgment facilitated the categorization of variable classes. For example, there was no significant and consistent difference in crash frequencies between 1:00 am to 12:59 pm (Figure 3). On the contrary, cellphone-related crashes were increased by the end of the afternoon hour (12:00–12:59 pm), at a peak from 3:00 pm to 3:59 pm, and decreased at the beginning of evening hours (5:00–5:59 pm). A relatively similar pattern had been observed from 6:00 pm to 12:59 am. Therefore, Crash time was categorized into four different periods — 7:00 am to 12:59 pm, 1:00 to 5:59 pm, 6:00 pm to 12:59 am, and 1:00 to 6:59 am — according to the hourly fatal crash frequencies. Day of the week was classified following the NHTSA classification (weekday: Monday 6:00 am to Friday 5:59 pm and weekend: Friday 6:00 pm to Monday 5:59 am). Rurality and trafficway were merged to form Road type. Drivers who are aged 19 years or less and allowed to drive are generally recognized as teen

Table 2. Preliminary selected variables with attributes.

Variable	Attribute
Driver age	16 to 19 years, 20 to 24 years
Driver gender	Male, female
Previous crash record	No, yes
Alcohol/drug involvement	No, yes
Restraint usage	Properly used, improperly used, not used, others
Cellphone usage	Talking or listening, manipulating, others
Violation	No violation, careless driving, disregarding traffic signs/signals, improper action/ turning, others
Vehicle type	Passenger car, light truck (van/SUV/pickup-truck), others
Crash time	7:00am to 12:59pm, 1:00 to 5:59pm, 6:00pm to 12:59am, 1:00 to 6:59am
Day of the week	Weekday, weekend
Lighting condition	Daylight, dark-lighted, dark-not-lighted, others
Weather	Clear, cloudy, rain, others
Road type	Rural two-way divided (RUTWD), rural two-way undivided (RUTWUD), urban two-way divided (URTWD), urban two-way undivided (URTWUD), others
Posted speed limit	30-35 mph (48-56 km/h), 40-45 mph (64-72 km/h), 50-55 mph (80-88 km/h), $\geq$ 60 mph ($\geq$ 96 km/h), others
Roadway geometry	Straight segment, curve segment, intersection, intersection on curve
Crash type	Single vehicle, angle, head-on, rear-end, others

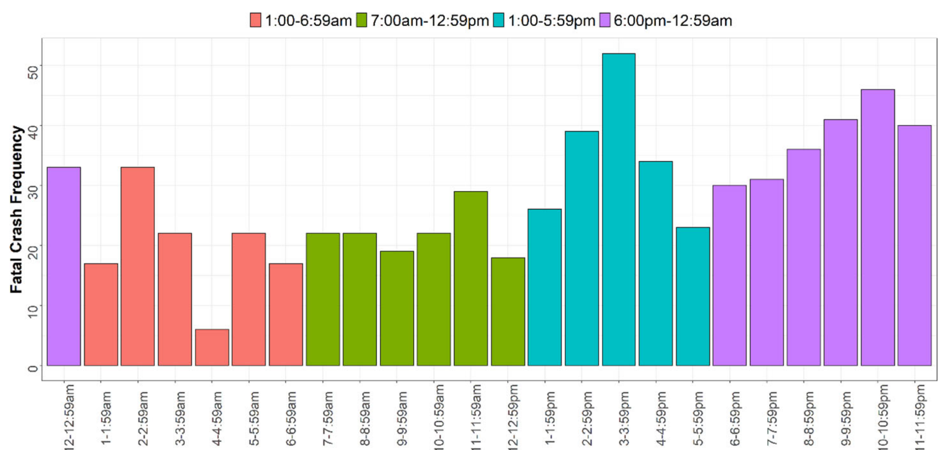

Figure 3. Hourly distribution of fatal crashes of cellphone-distracted young drivers.

drivers. Therefore, Driver age was grouped into two attributes (16 to 19 years and 20 to 24 years).

4.2. Variable importance

This study applied the Boruta algorithm using the "Boruta" package in "R" to evaluate the importance of selected features or variables. The conventional feature selection methods (e.g., random forest algorithm) comply with a minimal optimal approach where they depend on a small subset of variables, which induces a minimal error on a selected classifier. This occurs by setting an over-pruned version of the input dataset, which in turn, eliminates a few pertinent features (Rudnicki & Kursa, 2010).

However, the wrapper algorithm Boruta employs the "all-relevant" feature section technique, which captures both strongly and weakly-relevant features connected with the outcome variable (e.g., numerous cellphone activities) (Das, Khan, & Ahmed, 2020; Rudnicki & Kursa, 2010). Identifying all relevant features from a multidimensional dataset is preferable in correspondence analysis aimed to understand the associations among key contributory factors rather than developing a model of high predictive accuracy (Nilsson, Peña, Björkegren, & Tegnér, 2007). Moreover, Boruta is a serial algorithm that can parallel utilize the Random Forest (RF) algorithm to iteratively compare the importance of variables with that of shadow features, formed by shuffling the actual dataset (Rudnicki & Kursa, 2010). That means the features compete with a randomized version of them instead of competing among themselves. The shadows are re-created in each iteration. Features with higher importance than the shadow ones are called confirmed or important variables. On the contrary, the less important ones are discarded as dropped or unimportant variables. Figure 4 displays Boruta results on the preliminary crash dataset. Blue boxplots represent the minimal, average, and maximum Z score of a shadow feature (starting from the left). The green and red boxes correspond to confirmed and dropped variables, respectively. Driver age is identified as an unimportant feature in the algorithm. That means clustering younger drivers into additional two groups will not provide additional insights. Therefore, the factor is not considered for the final analysis.

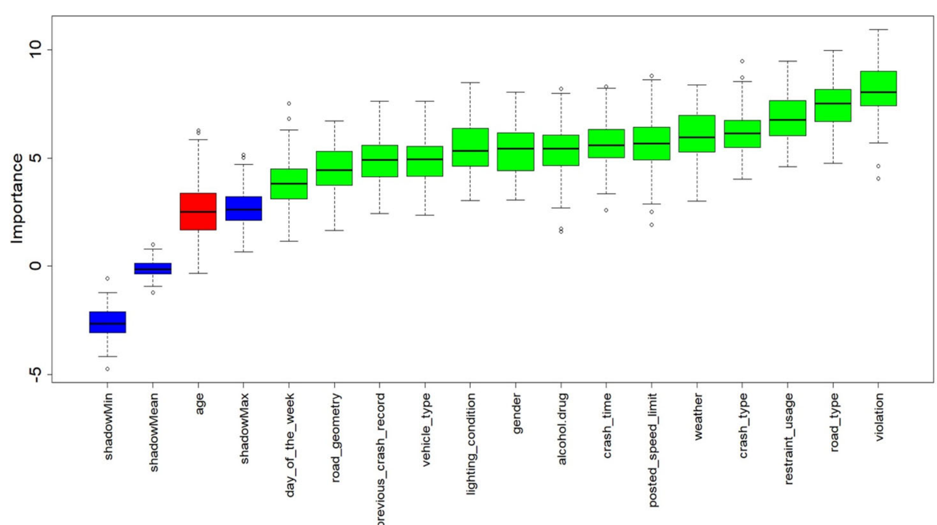

Figure 4. Variable importance using the Boruta algorithm.

4.3. Descriptive analysis

Table 3 shows the percentage of variable categories by cellphone distraction types in regard to the total fatal crashes of the final dataset. Around 38.97% of young drivers were engaged in manipulating a cellphone, as texting is one prevalent alternative to making phone calls (Jannusch et al., 2021). The proportion of male cellphone drivers was higher than females across all distribution classes, which is parallel with previous studies (Huang et al., 2003; Jannusch et al., 2021). Out of 680 fatal crashes with young drivers, 30.74% had reported alcohol and drug involvement along with cellphone-distracted driving. These intoxicated drivers were more likely to partake in manipulating (31.70%) rather than speaking or listening (26.75%). Despite enforcement of seatbelt laws, 33.09% of young drivers were unrestrained or improperly belted. Such offense was prevalent in manipulating (36.60%) when stratified by the modes of cellphone use. Most of the fatal crashes were associated with violations such as improper action/turning (15.29%), careless driving (11.62%), and speeding (11.03%). Also, a higher proportion of drivers disregarded traffic signs and signals (18.47%) when talking or listening to cellphones. In terms of vehicle type, the distributions showed a persistent pattern- passenger car (53.38%), followed by light truck (40.88%), and others (5.74%). However, a slightly higher magnitude was observed in talking or listening (42.04%) among young drivers operating light trucks.

About 37.94% of cellphone-related crashes occurred from 6:00 pm to 12:59 am, followed by 25.59% from 1:00 to 5:59 pm, 19.41% from 7:00 am to 12:59 pm, and 17.21% from 1:00 to 6:59 am. This pattern pinpoints that young drivers become more frequent or careless in cellphone usage as the day progresses. Fatal crash data of youth show usual trends in days of the week and the weather as the magnitudes were disproportionally high during weekdays (61.47%) and clear weather (68.53%). This skewness can be annotated as drivers are more likely to use cellphones in clear weather and on business days (Walsh, White, Watson, & Hyde, 2007). In terms of dark with streetlighting conditions, a higher proportion was observed when engaged in talking or listening on cellphones prior to the crashes (21.66%) compared to total cellphone-related fatal crashes (18.82%).

Two-way undivided roadways showed greater fatal crash occurrences (69.41%) when RUTWUD and URTWUD were combined. A majority of them took place in rural areas. About 35.15% of crash episodes were at a posted speed limit of 50–55 mph, accompanied by 26.32% on a posted speed limit of 40–45 mph. Cellphone drivers are more likely to divert their attention away from driving when traveling in light traffic or on roadways with higher speed limits (Jannusch et al., 2021). At intersections, young drivers are more frequent in fatal crash involvement when interacting with

Table 3. Statistics of contributory factors in the final dataset.

Variable	Category	Total fatal crashes (680) %	Fatal crashes by the modes of cellphone usage: Talking/ Listening (157) %	Manipulating (265) %	Others (258) %
Driver characteristics					
Gender	Male	57.65	57.32	58.11	57.36
	Female	42.35	42.68	41.89	42.64
Previous crash record	No	71.03	72.61	72.83	68.22
	Yes	28.97	27.39	27.17	31.78
Alcohol/drug involvement	No	69.26	73.25	68.30	67.83
	Yes	30.74	26.75	31.70	32.17
Restraint usage	Properly used	58.68	53.50	57.74	62.79
	Improperly used	3.38	7.64	3.02	1.16
	Not used	29.71	28.66	33.58	26.36
	Others	8.24	10.19	5.66	9.69
Violation	No violation	28.38	29.30	26.79	29.46
	Careless driving	11.62	8.92	9.43	15.50
	Disregarding traffic Signs/signals	9.85	18.47	7.55	6.98
	Improper Action/turning	15.29	16.56	18.11	11.63
	Speeding	11.03	7.01	13.96	10.47
	Others	23.82	19.75	24.15	25.97
Vehicle type	Passenger car	53.38	49.68	55.09	53.88
	Light truck	40.88	42.04	40.38	40.70
	Others	5.74	8.28	4.53	5.43
Environment characteristics					
Crash time	7:00am-12:59pm	19.41	17.83	19.25	16.67
	1:00-5:59pm	25.59	26.75	24.91	29.07
	6:00pm-12:59am	37.94	38.85	39.25	36.05
	1:00-6:59am	17.21	16.56	16.60	18.22
Day of the week	Weekday	61.47	61.78	60.75	62.02
	Weekend	38.53	38.22	39.25	37.98
Lighting condition	Daylight	48.82	46.50	48.30	50.78
	Dark-lighted	18.82	21.66	18.87	17.05
	Dark-not-lighted	28.09	27.39	28.68	27.91
	Others	4.26	4.46	4.15	60.08
Weather	Clear	68.53	61.78	71.70	69.38
	Cloudy	17.06	22.93	14.34	16.28
	Rain	8.38	7.01	8.30	9.30
	Others	6.03	8.28	5.66	5.04
Road characteristics					
Road type	RUTWD	10.59	10.19	9.06	12.40
	RUTWUD	38.53	41.40	36.60	38.76
	URTWD	16.76	18.47	20.75	11.63
	URTWUD	30.88	26.11	30.94	33.72
	Others	3.24	3.82	2.64	3.49
Posted speed limit	30-35 mph (48-56 km/h)	12.79	14.01	13.21	11.63
	40-45 mph (64-72 km/h)	26.32	21.66	30.19	25.19
	50-55 mph (80-88 km/h)	35.15	38.85	34.72	33.33
	$\geq$ 60 mph ($\geq$ 96 km/h)	21.03	19.11	18.11	25.19
	Others	4.71	6.37	3.77	4.65
Road geometry	Straight segment	60.74	59.24	62.26	60.08
	Curve segment	18.38	16.56	19.62	18.22
	Intersection	18.68	24.20	14.72	19.38
	Intersection on curve	2.21	0.00	3.40	2.33
Crash characteristics					
Crash type	Single vehicle	50.88	51.59	50.57	50.78
	Angle	15.29	19.75	14.72	13.18
	Head-on	17.94	14.01	19.25	18.99
	Rear-end	13.53	12.10	13.21	14.73
	Others	2.35	2.55	2.26	2.33

a cellphone involving talking or listening (24.20%). A disproportionate prevalence of single vehicle crashes has been reflected in the distribution table (50.88%).

5. Results and discussion

The purpose of applying JCA is to graphically represent the complementary attributes in a lower-dimensional plane by effectively summarizing the knowledge of a complex crash dataset. This multivariate analysis technique concentrates on off-diagonal contingency tables that provide a better estimation of data visualization elements (Das et al., 2019; Hossain, Sun et al., 2021). A smaller Euclidean distance of categories implies a significant interdependency. The attributes with closer proximity are regarded as a cloud (discussed in the methodology section). All attributes of the analyzing dataset were initially represented in one map, which made the graph difficult to distinguish the clouds. Therefore, two popular data visualization packages "ggplot2" and "ggrepel" were utilized to develop readily understandable JCA plots by defining the axes limit (Hossain, Sun et al., 2021).

Table 4 exhibits the mass and inertia of each variable attribute (recognized as a column) in JCA. Each of the quantities was multiplied by 1,000. In general, column mass is the summation of all the frequencies in that column divided by the sum of all the frequencies in the contingency table (Nenadic & Greenacre, 2007). Larger mass values of a column imply that the column has a higher relative frequency. For example, "no previous crash record" had the largest column mass (0.474) as the highest percentage of crashes belonged to this category (71.03%). Cell inertia is the chi-squared value in the cell divided by the total frequency for the contingency table. The inertia of a column means the sum of the cell inertias for that column (Nenadic & Greenacre, 2007). A higher value generally indicates a stronger association. For example, "dark-lighted condition" had the highest inertia (0.006), which means a stronger correlation with crash observations compared to other attributes. The top five variable categories with respect to column inertia were dark-lighted, intersection, disregarding traffic signs and signals, angle crash, and improper restraint usage.

In the analysis, the crash observations were initially presented in 16 dimensions. The eigenvalue (range 0 to 1) was estimated to specify how much attribute information each dimension accounts for. The dimensions were sequenced based on the eigenvalues. Figure 5 shows the top 10 dimensions by eigenvalues. The first two dimensions explained greater category information compared to any other dimension- 0.0164 and 0.0132, respectively. As JCA applies an iterative algorithm to best fit the off-diagonal submatrices of the Burt matrix, the explication doesn't have rigidly nested dimensions (Greenacre,

Table 4. Mass and inertia of each variable category in JCA.

Attribute	Mass*	Inertia*	Attribute	Mass*	Inertia*
Gender			*Day of the week*		
Male	38.4	2.1	Weekday	41.0	2.1
Female	28.2	2.9	Weekend	25.7	3.4
Previous crash record			*Lighting condition*		
No	47.4	1.4	Daylight	32.5	4.3
Yes	19.3	3.5	Dark-lighted	12.5	6.0
Alcohol/drug			Dark-not-lighted	18.7	5.1
No	46.2	1.7	Others	2.8	4.6
Yes	20.5	3.8	*Weather*		
Cellphone distraction			Clear	45.7	1.5
Talking/listening	15.4	3.8	Cloudy	11.4	4.1
Manipulating	26.0	2.9	Rain	5.6	4.5
Others	25.3	3.0	Others	4.0	4.5
Restraint usage			*Road type*		
Properly used	39.1	2.2	RUTWD	7.1	5.1
Improperly used	2.3	5.3	RUTWUD	25.7	3.9
Not used	19.8	4.2	URTWD	11.2	4.6
Others	5.5	5.0	URTWUD	20.6	4.2
Violation			Others	2.2	4.7
No violation	18.9	3.5	*Posted speed limit*		
Careless driving	7.7	4.5	30-35mph	8.5	4.7
Disregarding traffic signs/signals	6.6	5.8	40-45mph	17.5	4.1
Improper action/turning	10.2	5.2	50-55mph	23.4	3.6
Speeding	7.4	4.3	>=60mph	14.0	5.0
Others	15.9	3.7	Others	3.1	4.7
Vehicle type			*Road geometry*		
Passenger car	35.6	2.4	Straight segment	40.5	2.2
Light truck	27.3	2.9	Curve segment	12.3	4.4
Others	3.8	4.7	Intersection	12.5	5.8
Crash time			Curve on intersection	1.5	4.6
7:00 am–12:59 pm	13.2	5.0	*Crash type*		
1:00–5:59 pm	14.5	4.7	Single vehicle	33.9	3.0
6:00 pm–12:59 am	25.3	4.2	Angle	10.2	5.6
1:00–6:59 am	13.6	4.3	Head-on	12.0	5.0
			Rear-end	9.0	4.6
			Others	1.6	4.7

*Quantities were multiplied by 1,000.

Nenadic, Friendly, & Nenadic, 2020). Therefore, the optimal percentage of variance is computed for the chosen dimensionality (the two principal axes represent larger eigenvalues) rather than for each dimension. In this study, diagonal inertia calculated from eigenvalues was 0.008. The first two dimensions explained 50.4% of the total variance. The number of iterations was 54.

Figure 6 shows the JCA factor map exhibiting all 55 attributes of the final dataset. Only one cloud can be recognized from that plotting; therefore, Figure 7 was developed to identify the rest of the associations. The analysis displayed seven meaningful clouds that illustrate the association of categories featuring fatal cellphone-related crashes of young drivers. To establish the identified clouds, each of the associations has been discussed with reference to previous literature findings.

Cloud 1 (violation = disregarding traffic signs and signals, road geometry = intersection, crash type = angle)

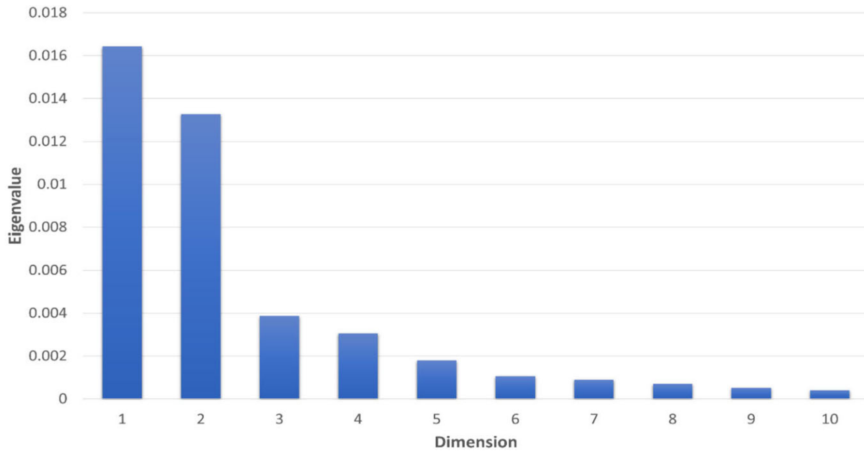

Figure 5. Top 10 dimensions by eigenvalues.

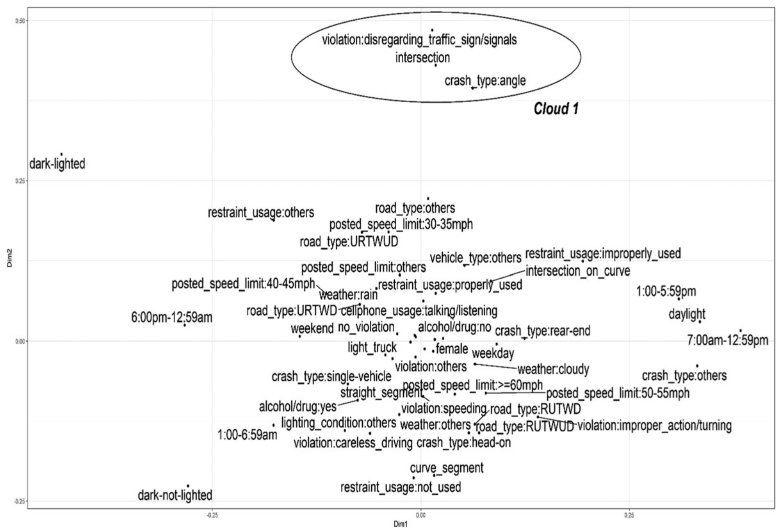

Figure 6. Joint correspondence analysis factor map with Cloud 1.

Cloud 1 indicates that fatal angle crashes of young drivers distracted by cellphones are associated with disobeying traffic signs and signals at intersections (Figure 6). One previous study in Missouri reported a positive correlation between cellphone distraction and angular crash risk (Ghazizadeh & Boyle, 2009). Cellphone use in complex driving circumstances such as intersections is already proven to be hazardous for novice drivers (Klauer et al., 2006). Young drivers are more likely to run through the stop sign or red light while speaking on a cellphone (Beck et al., 2007; Haque et al., 2013), a cumulative effect of attention diversion and driving inexperience.

Cloud 2 (alcohol or drug involvement = yes, crash type = single vehicle)

Cloud 2 describes young drivers' cellphone use under the influence of alcohol and drugs that resulted in deadly, single vehicle collisions (Figure

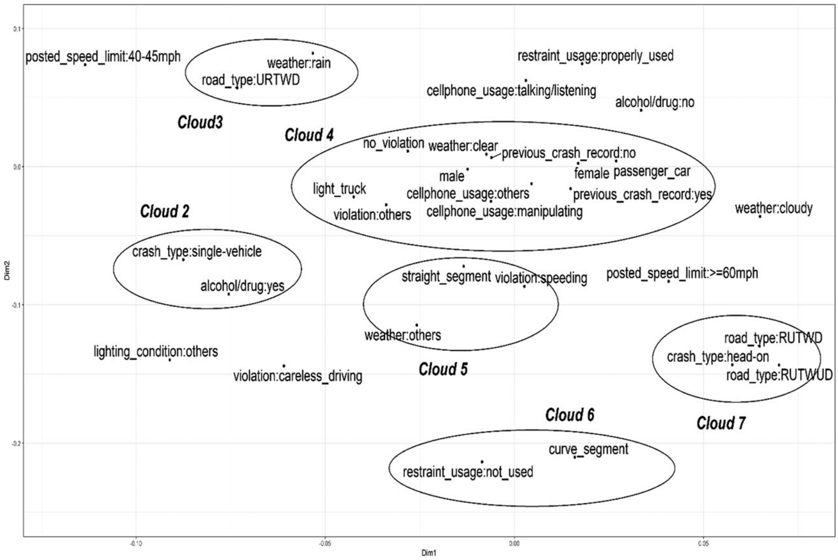

Figure 7. Joint correspondence analysis factor map with Cloud 2 to Cloud 7.

7). Intoxicated young novice drivers are often engaged in talking to a cellphone, which can lead to critical driving errors with potentially fatal outcomes (Jannusch et al., 2021). Due to less experience in drinking and driving, adolescents have a higher risk of injury crashes even at a lower intoxication level (Mayhew, Donelson, Beirness, & Simpson, 1986). On the contrary, alcohol has already been underlined as a critical factor contributing to single vehicle crashes (Öström & Eriksson, 1993).

Cloud 3 (weather = rain, road type = urban two-way divided)

Cloud 3 specifies a fatal crash scenario on two-way divided roadways (with or without physical separation) in urban areas with regard to young drivers using a cellphone during rainy weather conditions. Safe driving on urban roads is critical for novice drivers due to the surrounding complexities generated from heavier traffic. Precipitation reduces driver's visibility and tire-road friction; therefore, drivers have to respond quicker than usual in critical driving situations (Lobo, Ferreira, Iglesias, & Couto, 2019). In these circumstances, young drivers become more exposed to the risk of serious crashes when distracted by a cellphone.

Cloud 4 (gender = male, gender = female, previous crash record = no, previous crash record = yes, cellphone usage = manipulating, cellphone usage = others, violation = no violation, violation = others, vehicle type = passenger car, vehicle type = light truck, weather = clear)

Cloud 4 represents two separate deadly crash scenarios. The first sequence of events describes fatal crashes of young male drivers in light trucks while engaged in cellphone activities other than talking or listening. Overall, texting, dialing, reaching for, or answering the cellphone are more vulnerable than talking or listening as these tasks are affiliated with visual-manual distraction (Owens et al., 2018). Besides, concerning adolescents, male drivers and light truck motorists have been specified for their greater inclination toward aggressive driving (Paleti, Eluru, & Bhat, 2010). The second scenario indicates fatal cellphone-related crashes of young female drivers who have a crash history. One earlier young driver study in Michigan argued that past crash and traffic offenses are a significant predictor of subsequent crashes, notably for female drivers (Elliott, Waller, Raghunathan, & Shope, 2001).

Cloud 5 (violation = speeding, road geometry = straight segment, weather = others)

Cloud 5 specifies the speeding tendency of young cellphone drivers on straight roadway segments in adverse weather conditions (other than cloudy and rain). Speeding in inclement weather (e.g., snow, fog) intensifies the chances of severe injury crashes due to inadequate skid resistance and poor visibility (Edwards, 1998).

Cloud 6 (restraint usage = not used, road geometry = curve segment)

Cloud 6 displays the combined effect of cellphone distraction and unrestrained driving on curve road segments. Speed-related collisions are more frequent on curve segments owing to improper approaching speed on the entry points. At curves, Charlton (2004) found a significant speed deviation with respect to safe driving speed when the driver performs attention-demanding secondary tasks on cellphones. In addition, multiple studies manifested the severe consequences of driving without any protection system, particularly for young drivers (Rahman et al., 2021; Shaaban & Abdelwarith, 2020; Vachal et al., 2009).

Cloud 7 (road type = rural two-way divided, road type = rural two-way undivided, crash type = head-on)

Cloud 7 describes fatal head-on collisions on two-way roads in rural areas resulting from cellphone-distracted driving. Head-on collisions are more likely to be fatal on rural roadways (Deng, Ivan, & Gårder, 2006), as the setting occupies distinct hazardous features such as hidden driveways, narrow roads, sharp curves, and being unlighted. Young drivers often disregard these additional safety threats and pay more attention to the phone rather than oncoming traffic, which is one of the prime reasons for head-on crashes (Gårder, 2006).

6. Conclusions

6.1. Research contribution and key findings

This study employed JCA to identify the co-occurrence of attributes that influence the fatal crashes of young drivers in terms of cellphone-distracted driving. The Boruta algorithm was applied to select the relevant features from the preliminary crash dataset. As the final dataset covered a wide range of variables, JCA is a well-suited machine learning tool to offer valuable insights on the interdependencies between a significant number of variable categories. Also, JCA provides a graphical representation of all crash attributes in a lower-dimensional plane, and thereby, allows a broader audience to detect the complex associations. In contrast with traditional statistical models, this exploratory multivariate analysis technique operates without any predefined hypothesis of dependent variables and covariates. This research deals with a total of 680 fatal crashes in six years, and JCA can handle missing entries without reducing the actual size of the dataset. Two unique contributions of this study are: 1) application of the Boruta algorithm to identify the relevant contributing factors; and 2) application of a less explored CA method (JCA) to determine the key insights from these crashes.

The findings of this study exhibited that variables such as gender, previous crash record, intoxication, restraint usage, violation, vehicle type, weather, road type, and road geometry had significant contributions to the young driver, cellphone-related, fatal crash occurrences. From JCA, this research revealed a few combinations of attributes that resulted in the deadly cellphone collisions. Young male drivers in light trucks were involved in crashes while performing cellphone activities other than talking and listening, whereas young females with crash history had collisions during secondary tasks on cellphones (cloud 4). In terms of roadway type, two-way divided roads in urban areas were vulnerable during rainy weather conditions (cloud 3). On the contrary, rural two-way roads seemed to be hazardous due to head-on cellphone-distracted collisions (cloud 7). The grouping of road and environment-related factors significantly influences cellphone-related fatal crashes involving young drivers. Moreover, the associations also implied other risky driving behaviors of youth while distracted by a cellphone, for example, disregarding traffic signs and signals at intersections (cloud 1), speeding on straight segments during bad weather conditions (cloud 5), and unrestrained driving on curve roadway segments (cloud 6). The study outcomes also displayed fatal consequences of single vehicle collisions where intoxicated young drivers performed cellphone activities before the incidents (cloud 2). The meaningful combination groups shed light on the patterns of fatal cellphone-related crashes, exposed a new aspect of research in distracted driving. The findings could guide the safety officials and policymakers in developing appropriate engineering, education, and enforcement strategies when dealing with cellphone-distracted young drivers. In addition, prioritizing the key attributes from the confluence of factors can be helpful in reducing the related collisions and fatalities.

6.2. Recommendations

In the U.S., states have executed distracted driving laws to restrict cellphone usage among young and novice drivers. However, active enforcement of these regulations is becoming harder due to the blooming of distinct cellphone-related secondary tasks (George, Brown, Scholz, Scott-Parker, & Rickwood, 2018). Also, exclusively working on public awareness campaigns is unlikely to bring any behavioral change, as most youths are aware of the risk of cellphone use during driving (McDonald & Sommers, 2015). Driver education programs are a popular and structured approach targeting novice drivers to improve risk perception, alter aggressive driving maneuvers, and develop critical driving skills. However, conventional training has little effect in lessening cellphone use while driving (Delgado, Wanner, & McDonald, 2016). Several studies have recommended adding the contextual understanding of risk factors into the training contents that could depict drivers' actual crash scenarios (Arnold et al., 2019; Delgado et al., 2016). In this regard, the associations revealed in this study can be helpful to strengthen the existing educational interventions.

This research identified that road-related factors such as two-way roads in rural areas, two-way divided roads in urban areas, curve roadway segments, and intersections were associated with cellphone-related fatal crashes. Locations with high crash frequencies can be prioritized for safety assessments with respect to geometric features and functional elements. The study outcomes also confirmed cellphone-distracted young drivers' tendency toward violating driving rules such as obeying traffic signs and signals, restrictions on alcohol and drug intake, maintaining the posted speed limit, and all-time seatbelt use. Efficient roadside inspections, episodic enforcement campaigns, and strict enforcement of existing legislation are required to curb these driving offenses. In this regard, the active participation of local agencies to increase visibility toward sustained enforcement actions can be more effective (LHSC, 2019). This study highlighted the high proportion of single-vehicle collisions in the cellphone-related crash dataset. Electronic stability control (ESC) is substantially effective in reducing single-vehicle crashes as the sensor-based braking system can quickly respond to sudden instabilities (Sivinski, 2011). Few studies have stated the necessity to implement integrations of multiple countermeasures (e.g., improved education programs, strict enforcement of regulations, widespread safety awareness campaigns, etc.) against cellphone-distracted driving (Arnold et al., 2019; Delgado et al., 2016). However, the biggest challenge is to find an approach that will not only decrease the risk of cellphone use while driving, but also permit adolescents to enjoy the benefits of advanced smartphone features.

The current study is not without limitations. First, the analysis was based on the first plane with axis 1 and axis 2. The first plane only explains around 50% of the inertia. Exploration of additional planes may provide additional insights. Second, this study is limited to driver, environment, road, and crash-related factors. Future investigations can include vehicle, occupant, demographic, and situational factors that directly or indirectly influence drivers' cellphone usage. Understanding the cellphone-related crash patterns by different driver age groups can be a better approach to develop more effective and specific countermeasures. In addition, exploring the changes of these crash patterns with respect to cellphone types and technologies could be an extensive research topic. Regarding teen drivers, further cellphone research can be conducted to distinguish the effect of driving experience or the GDL stages on the corresponding interrelations between crash contributing factors. The accuracy of police-investigated crash reports in the U.S. varies with geography, which is one limitation of FARS data. Further research can be done using a more comprehensive dataset.

Acknowledgments

The authors would like to thank four anonymous reviewers for their constructive suggestions and feedback.

Declaration of conflicting interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: Md Mahmud Hossain, data preparation: Md Mahmud Hossain; analysis and interpretation of results: Md Mahmud Hossain, Subasish Das; draft manuscript preparation: Md Mahmud Hossain, Huaguo Zhou, Subasish Das, Xiaoduan Sun, Ahmed Hossain. All authors reviewed the results and approved the final version of the manuscript.

References

Arnold, L. S., Benson, A. J., Tefft, B. C., Barragan, D., Jin, L., Kolek, S., & Horrey, W. J. (2019). *Effectiveness of distracted driving countermeasures: A review of the literature* (p. 13). Washington DC, USA: AAA Foundation for Traffic Safety.

Atchley, P., Atwood, S., & Boulton, A. (2011). The choice to text and drive in younger drivers: Behavior may shape attitude. *Accident Analysis and Prevention*, *43*(1), 134–142. doi:10.1016/j.aap.2010.08.003

Baker, S. P., Chen, L.-H., & Li, G. (2011). National evaluation of graduated driver licensing programs. Teen Driv. Saf. Grad. Driv. Licens, 87–103. doi:10.1017/CBO9781107415324.004

Beck, K. H., Yan, F., & Wang, M. Q. (2007). Cell phone users, reported crash risk, unsafe driving behaviors and dispositions: A survey of motorists in Maryland. *Journal of Safety Research*, *38*(6), 683–688. doi:10.1016/j.jsr.2007.09.006

Beede, K. E., & Kass, S. J. (2006). Engrossed in conversation: The impact of cellphones on simulated driving performance. *Accident Analysis and Prevention*, *38*(2), 415–421. doi:10.1016/j.aap.2005.10.015

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, *125*(3), 367–383. doi:10.1037/0033-2909.125.3.367

Caird, J. K., Simmons, S. M., Wiley, K., Johnston, K. A., & Horrey, W. J. (2018). Does talking on a cellphone, with a passenger, or dialing affect driving performance? An updated systematic review and meta-analysis of experimental studies. *Human Factors*, *60*(1), 101–133. doi:10.1177/0018720817748145

Camiz, S., Matematica, D., & Universit, S. (2013). Multiple and joint correspondence analysis: Testing the true dimension of a study. *Moduland*, *44*, 2–33.

Charlton, S. G. (2004). Perceptual and attentional effects on drivers' speed selection at curves. *Accident Analysis and Prevention*, *36*(5), 877–884. doi:10.1016/j.aap.2003.09.003

Choulakian, V. (1988). Exploratory analysis of contingency tables by loglinear formulation and generalizations of correspondence analysis. *Psychometrika*, *53*(2), 235–250. doi:10.1007/BF02294135

Das, A., Khan, M. N., & Ahmed, M. M. (2020). Detecting lane change maneuvers using SHRP2 naturalistic driving data: A comparative study machine learning techniques. *Accident Analysis and Prevention*, *142*, 105578. doi:10.1016/j.aap.2020.105578

Das, S., Hossain, M. M., Rahman, M. A., Kong, X., Sun, X., & Mamun, G. M A. (2022). Understanding patterns of moped and seated motor scooter (50 cc or less) involved fatal crashes using cluster correspondence analysis. *Transportmetrica A: Transport Science*. doi:10.1080/23249935.2022.2029613

Das, S., Jha, K., Fitzpatrick, K., Brewer, M., & Shimu, T. H. (2019). Pattern identification from older bicyclist fatal crashes. *Transportation Research Record: Journal of the Transportation Research Board*, *2673*(6), 638–649. doi:10.1177/0361198119841863

Delgado, M. K., Wanner, K. J., & McDonald, C. (2016). Adolescent cellphone use while driving: An overview of the literature and promising future directions for prevention. *Media and Communication*, *4*(3), 79–89. doi:10.17645/mac.v4i3.536

Deng, Z., Ivan, J. N., & Gårder, P. (2006). Analysis of factors affecting the severity of head-on crashes. *Transportation Research Record: Journal of the Transportation Research Board*, *1953*(1), 137–146. doi:10.1177/0361198106195300116

Edwards, J. B. (1998). The relationship between road accident severity and recorded weather. *Journal of Safety Research*, *29*(4), 249–262. doi:10.1016/S0022-4375(98)00051-6

Elliott, M. R., Waller, P. F., Raghunathan, T. E., & Shope, J. T. (2001). Predicting offenses and crashes from young drivers' offense and crash histories. *Journal of Crash Prevention and Injury Control*, *2*(3), 167–178. doi:10.1080/10286580108902562

Gårder, P. (2006). Segment characteristics and severity of head-on crashes on two-lane rural highways in Maine. *Accident Analysis and Prevention*, *38*(4), 652–661. doi:10.1016/j.aap.2005.12.009

George, A. M., Brown, P. M., Scholz, B., Scott-Parker, B., & Rickwood, D. (2018). "I need to skip a song because it sucks": Exploring mobile phone use while driving among young adults. *Transportation Research Part F: Traffic Psychology and Behaviour*, *58*, 382–391. doi:10.1016/j.trf.2018.06.014

Ghazizadeh, M., & Boyle, L. N. (2009). Influence of driver distractions on the likelihood of rear-end, angular, and single-vehicle crashes in Missouri. *Transportation Research Record: Journal of the Transportation Research Board*, *2138*(1), 1–5. doi:10.3141/2138-01

Greenacre, M., Nenadic, O., Friendly, M., Nenadic, M. (2020). *Package 'ca'*. https://cran.r-project.org/web/packages/ca/ca.pdf

Greenacre, M. J. (1988). Correspondence analysis of multivariate categorical data by weighted least-squares. *Biometrika*, *75*(3), 457–467. doi:10.1093/biomet/75.3.457

Guo, F., Klauer, S. G., Fang, Y., Hankey, J. M., Antin, J. F., Perez, M. A., Lee, S. E., & Dingus, T. A. (2017). The effects of age on crash risk associated with driver distraction. *International Journal of Epidemiology*, *46*(1), 258–265. doi:10.1093/ije/dyw234

Haque, M. M., Ohlhauser, A. D., Washington, S., & Boyle, L. N. (2013 *Examination of distracted driving and yellow light running: Analysis of simulator data* [Paper presentation]. 92nd Annual Meeting of Transportation Research Board (TRB), Washington, DC.

Haque, M. M., & Washington, S. (2015). The impact of mobile phone distraction on the braking behaviour of young drivers: A hazard-based duration model. *Transportation Research Part C: Emerging Technologies*, *50*, 13–27. doi:10.1016/j.trc.2014.07.011

Hosking, S. G., Young, K. L., & Regan, M. A. (2009). The effects of text messaging on young drivers. *Human Factors*, *51*(4), 582–592. doi:10.1177/0018720809341575

Hossain, M. M., Rahman, M. A., Sun, X., & Mitran, E. (2021). Investigating underage alcohol-intoxicated driver crash patterns in Louisiana. *Transportation Research Record: Journal of the Transportation Research Board*, *2675*(11), 769–782. doi:10.1177/03611981211019742

Hossain, M. M., Sun, X., Mitran, E., & Rahman, M. A. (2021). Investigating fatal and injury crash patterns of teen drivers with unsupervised learning algorithms. *IATSS Research*, *45*(4), 561–573. doi:10.1016/j.iatssr.2021.07.002

Huang, H. F., Stutts, J. C., & Hunter, W. W. (2003). Characteristics of cellphone-related motor vehicle crashes in North Carolina. *Transportation Research Record: Journal of the Transportation Research Board*, *1843*(1), 10–19. doi:10.3141/1843-02

Insurance Institute for Highway Safety (IIHS). (2019). *Status report: Distracted driving, Vol. 54, No. 1*. https://www.iihs.org/api/datastoredocument/status-report/pdf/54/1

Jannusch, T., Shannon, D., Völler, M., Murphy, F., & Mullins, M. (2021). Smartphone use while driving: An investigation of young novice driver (YND) behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour*, *77*, 209–220. doi:10.1016/j.trf.2020.12.013

Jonah, B. A., Thiessen, R., & Au-Yeung, E. (2001). Sensation seeking, risky driving and behavioral adaptation. *Accident Analysis and Prevention*, *33*(5), 679–684. doi:10.1016/S0001-4575(00)00085-3

Keating, D. P., & Halpern-Felsher, B. L. (2008). Adolescent drivers. A developmental perspective on risk, proficiency, and safety. *American Journal of Preventive Medicine*, *35*(3 Suppl), S272–S277. doi:10.1016/j.amepre.2008.06.026

Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., & Dingus, T. A. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. *The New England Journal of Medicine*, *370*(1), 54–59. doi:10.1056/nejmsa1204142

Klauer, S. G., Klauer, S. G., Dingus, T. a., Dingus, T. A., Neale, V. L., Neale, V. L., Sudweeks, J. D., Sudweeks, J. D., Ramsey, D. J., & Ramsey, D. J. (2006). The impact of driver inattention on near crash/crash risk: An analysis using the 100-car naturalistic driving study data. *Analysis*, April, 226.

Li, L., Bower, J. K., Zhu, M., & Board, T. R. (2019). Calling while driving by adolescents in 7 U.S. states: Prevalence, and cellphone laws. *Proceedings of 98th Transportation Research Board Annual Meeting*, Washington DC, USA.

Li, L., Pope, C. N., Andridge, R. R., Bower, J. K., Hu, G., & Zhu, M. (2020). Cell phone laws and teens' calling while driving: Analysis of repeated cross-sectional surveys in 2013, 2015, 2017, and 2019. *Injury Epidemiology*, *7*(1), 1–9. doi:10.1186/s40621-020-00290-x

Lobo, A., Ferreira, S., Iglesias, I., & Couto, A. (2019). Urban road crashes and weather conditions: Untangling the effects. *Sustainability*, *11*(11), 3176. doi:10.3390/su11113176

Louisiana Highway Safety Commission (LHSC). (2019). *Louisiana Highway Safety Program: Federal Fiscal Year 2019*. Retrieved from https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/la_fy19_hspar.pdf

Mayhew, D. R., Donelson, A. C., Beirness, D. J., & Simpson, H. M. (1986). Youth, alcohol and relative risk of crash involvement. *Accident Analysis and Prevention*, *18*(4), 273–287. doi:10.1016/0001-4575(86)90042-4

Mayhew, D. R., Simpson, H. M., & Pak, A. (2003). Changes in collision rates among novice drivers during the first months of driving. *Accident Analysis and Prevention*, *35*(5), 683–691. doi:10.1016/S0001-4575(02)00047-7

McDonald, C. C., & Sommers, M. S. (2015). Teen drivers' perceptions of inattention and cellphone use while driving. *Traffic Injury Prevention*, *16*(sup2), S52–S58. doi:10.1080/15389588.2015.1062886

Montella, A., Mauriello, F., Pernetti, M., & Rella Riccardi, M. (2021). Rule discovery to identify patterns contributing to overrepresentation and severity of run-off-the-road crashes. *Accident Analysis and Prevention*, *155*, 106119. doi:10.1016/j.aap.2021.106119

Nafis, S. R., Alluri, P., Wu, W., & Kibria, B. M. G. (2021). Wrong-way driving crash injury analysis on arterial road networks using non-parametric data mining techniques. *Journal of Transportation Safety & Security*, doi:10.1080/19439962.2021.1960660

Nasr Esfahani, H., Arvin, R., Song, Z., & Sze, N. N. (2021). Prevalence of cellphone use while driving and its impact on driving performance, focusing on near-crash risk: A survey study in Tehran. *Journal of Transportation Safety & Security*, *13*(9), 957–977. doi:10.1080/19439962.2019.1701166

National Highway Traffic Safety Administration (NHTSA). (2012). *Young drivers report the highest level of phone involvement in crash or near-crash incidences: Traffic safety fact sheet* [Report No. DOT HS 811 611]. U.S. Department of Transportation.

National Highway Traffic Safety Administration (NHTSA). (2019a). *Driver electronic device use in 2018: Traffic safety fact sheet* [Report No. DOT HS 812 818]. National Highway Traffic Safety Administration, U.S. Department of Transportation.

National Highway Traffic Safety Administration (NHTSA). (2019b). *Distracted driving in fatal crashes, 2017: Traffic safety fact sheet* [Report No. DOT HS 812 700]. U.S. Department of Transportation.

National Highway Traffic Safety Administration (NHTSA). (2020). *Distracted driving 2018: Traffic safety fact sheet* [Report No. DOT HS 812 926]. National Highway Traffic Safety Administration, U.S. Department of Transportation.

National Safety Council (NSC). (2013). *Crashes involving cellphones: Challenges of collecting and reporting reliable crash data, National Safety Council, USA. Retrieved from https://www.nsc.org/getmedia/2e33083e-5bfb-4f21-98b7-0706bdc3af4d/nsc-under-reporting-white-paper.pdf.aspx*.

Nenadic, O., & Greenacre, M. (2007). Correspondence Analysis in R, with Two- and Three-dimensional Graphics: The ca Package. *J. Stat. Softw*, *20*, 3.

Nilsson, R., Peña, J. M., Björkegren, J., & Tegnér, J. (2007). Consistent feature selection for pattern recognition in polynomial time. *Journal of Machine Learning Research*, *8*, 589–612.

Öström, M., & Eriksson, A. (1993). Single-vehicle crashes and alcohol: A retrospective study of passenger car fatalities in Northern Sweden. *Accident Analysis and Prevention*, *25*(2), 171–176. doi:10.1016/0001-4575(93)90057-4

Owens, J. M., Dingus, T. A., Guo, F., Youjia, F., Perez, M., & McClafferty, J. (2018). *Crash risk of cellphone use while driving: A case-crossover analysis of naturalistic driving data*. Washington DC, USA: AAA Foundation for Traffic Safety.

Paleti, R., Eluru, N., & Bhat, C. R. (2010). Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. *Accident Analysis and Prevention*, *42*(6), 1839–1854. doi:10.1016/j.aap.2010.05.005

Rahman, M. A., Hossain, M. M., Mitran, E., & Sun, X. (2021). Understanding the contributing factors to young driver crashes: A comparison of crash profiles of three age groups. *Transportation Engineering*, *5*, 100076. doi: 10.1016/j.treng.2021.100076

Rahman, M. A., Sun, X., Sun, M., & Shan, D. (2020). Investigating characteristics of cellphone distraction with significance tests and association rule mining. *IATSS Research*, *45*(2), 198-209. doi:10.1016/j.iatssr.2020.09.001

Rudnicki, W. R., & Kursa, M. B. (2010). Feature selection with the Boruta package. *Journal of Statistical Software*, *36*, i11.

Shaaban, K., & Abdelwarith, K. (2020). Understanding the association between cellphone use while driving and seat belt noncompliance in Qatar using logit models. *Journal of Transportation Safety & Security*, *12*(2), 292–308. doi:10.1080/19439962.2018.1477895

Shaaban, K., Gaweesh, S., & Ahmed, M. M. (2018). Characteristics and mitigation strategies for cellphone use while driving among young drivers in Qatar. *Journal of Transport & Health*, *8*, 6–14. doi:10.1016/j.jth.2018.02.001

Shope, J. T., & Bingham, C. R. (2008). Teen driving. motor-vehicle crashes and factors that contribute. *American Journal of Preventive Medicine*, *35*(3 Suppl), S261–S271. doi:10.1016/j.amepre.2008.06.022

Sivinski, R. (2011). *Crash prevention effectiveness of light-vehicle electronic stability control: An update of the 2007 NHTSA evaluation* [Report no. DOT HS-811-486]. National Highway Traffic Safety Administration, US Department of Transportation.

Tison, J., Chaudhary, N., & Cosgrove, L. (2011). *National phone survey on distracted driving attitudes and behaviors [Report No.DOT HS 811 555*]. National Highway Traffic Safety Administration, US Department of Transportation.

Tucker, S., Pek, S., Morrish, J., & Ruf, M. (2015). Prevalence of texting while driving and other risky driving behaviors among young people in Ontario, Canada: Evidence from 2012 and 2014. *Accident Analysis and Prevention*, *84*, 144–152. doi:10.1016/j.aap.2015.07.011

Vachal, K., Research Faculty, & Malchose, D. (2009). What can we learn about North Dakota's youngest drivers from their crashes? *Accident Analysis and Prevention*, *41*(3), 617–623. doi:10.1016/j.aap.2009.02.014

Vermunt, J. K., & Anderson, C. J. (2005). Joint correspondence analysis (JCA) by maximum likelihood. *Methodology*, *1*(1), 18–26. doi:10.1027/1614-1881.1.1.18

Vivoda, J. M., Eby, D. W., St. Louis, R. M., & Kostyniuk, L. P. (2008). Cellular phone use while driving at night. *Traffic Injury Prevention*, *9*(1), 37–41. doi:10.1080/15389580701737603

Walsh, S. P., White, K. M., Watson, B., & Hyde, M. K. (2007). Psychosocial factors influencing mobile phone use while driving. Australian Transport Safety Bureau, Australia, Australian Capital Territory, Canberra.

Walshe, E. A., Winston, F. K., Betancourt, L. M., Arena, K., & Romer, D. (2018). Cell phone use while driving: A pattern of risk in young novice drivers. *Academic Journal of Pediatrics & Neonatology*, *6*(4), 555752. doi:10.19080/ajpn.2018.06.555752

Wang, X., Xu, R., Asmelash, A., Xing, Y., & Lee, C. (2020). Characteristics of driver cell-phone use and their influence on driving performance: A naturalistic driving study. *Accident Analysis and Prevention*, *148*, 105845. doi:10.1016/j.aap.2020.105845

Young, K. L., Rudin-Brown, C. M., & Lenné, M. G. (2010). Look who's talking! A roadside survey of drivers' cellphone use. *Traffic Injury Prevention*, *11*(6), 555–560. doi:10.1080/15389588.2010.499442

Zhao, N., Reimer, B., Mehler, B., D'Ambrosio, L. A., & Coughlin, J. F. (2013). Self-reported and observed risky driving behaviors among frequent and infrequent cellphone users. *Accident Analysis and Prevention*, *61*, 71–77. doi:10.1016/j.aap.2012.07.019

Zhu, M., Rudisill, T. M., Heeringa, S., Swedler, D., & Redelmeier, D. A. (2016). The association between handheld phone bans and the prevalence of handheld phone conversations among young drivers in the United States. *Annals of Epidemiology*, *26*(12), 833–837.e1. doi:10.1016/j.annepidem.2016.10.002