Accident Analysis and Prevention 156 (2021) 106141

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Using SHRP2 NDS data to examine infrastructure and other factors contributing to older driver crashes during left turns at signalized intersections

Tracy Zafian [a], Alyssa Ryan [b,*], Ravi Agrawal [a], Siby Samuel [c], Michael Knodler [d]

[a] UMass Transportation Center, University of Massachusetts Amherst, 214 Marston Hall, 130 Natural Resources Road, Amherst, MA, 01003, USA

[b] Department of Civil & Environmental Engineering, University of Massachusetts Amherst, 34 Marston Hall, 130 Natural Resources Road, Amherst, MA, 01003, USA

[c] Department of Systems Design Engineering, University of Waterloo, Department of Systems Design Engineering, 200 University Avenue West, Waterloo, ON, N2L 3G1, Canada

[d] Department of Civil & Environmental Engineering, University of Massachusetts Amherst, 214 Marston Hall, 130 Natural Resources Road, Amherst, MA, 01003, USA

ARTICLE INFO

Keywords:
SHRP2
Naturalistic driving
Older drivers
Signalized intersection
Left turns

ABSTRACT

Drivers age 65 and over have higher rates of crashes and crash-related fatalities than other adult drivers and are especially over-represented in crashes during left turns at intersections. This research investigated the use of SHRP2 Naturalistic Driving Study (NDS) data to assess infrastructure and other factors contributing to left turn crashes at signalized intersections, and how to improve older driver safety during such turns. NDS data for trips involving signalized intersections and crash or near-crash events were obtained for two driver age groups: drivers age 65 and over (older drivers) and a sample of drivers age 30−49, along with NDS pre-screening and questionnaire data. Video scoring of all trips was performed to collect additional information on intersection and trip conditions. To identify the most influential factors of crash risk during left turns at signalized intersections, machine learning and regression models were used. The results found that in the obtained NDS dataset, there was a relatively small volume of crashes during left turns at signalized intersections. Further, model results found the statistically significant variables of crash risk for older drivers were associated more with health and cognitive factors rather than the infrastructure or design of the intersections. The results suggest that a study using only SHRP2 NDS data will not lead to definitive findings or recommendations for infrastructure changes to increase safety for older drivers at signalized intersections and during left turns. Moreover, the findings of this study indicates the need to consider other data sources and data collection methods to address this critical literature gap in older driver safety.

1. Introduction

The aging United States population poses large challenges for roadway safety. The fatal vehicle crash rate is substantially higher for older adults (Insurance Institute for Highway Safety, 2019; National Highway Traffic Safety Administration, 2020). Older drivers are also more likely to be severely injured or die in a crash due to their greater frailty compared to younger drivers (Mayhew et al., 2006). Intersections are a particular concern, as older drivers experience intersection crashes at higher rates than middle-aged drivers (Mayhew et al., 2006; Preusser et al., 1998). For example, the relative risk of an intersection fatal crash for drivers age 85 and older has been found to be 10.6 times that of drivers age 40–49 (Preusser et al., 1998). The likelihood of traffic violations at intersections is also much higher for older drivers than it is for middle-aged drivers (Garber and Srinivasan, 1991). Further, findings from a study by Staplin et al. (2012) on crash rates and individual driver self-reporting suggest that many older drivers are unaware of the consequences of aging on driving performance.

One particular intersection movement of concern with older drivers are left turns (Mayhew et al., 2006). Older drivers have a higher crash rate at intersections, particularly during left turn movements. This is due, in part, to their failure to yield the right of way to opposing traffic (Braitman et al., 2007). Older drivers have also been shown to perceive left turns onto divided highways as more difficult than other turning movements at intersections (Eck and Winn, 2002). Previous studies have suggested a number of potential explanations as to why older drivers

* Corresponding author.
E-mail address: alyssaryan@umass.edu (A. Ryan).

https://doi.org/10.1016/j.aap.2021.106141
Received 29 October 2020; Received in revised form 24 March 2021; Accepted 12 April 2021
Available online 16 April 2021

have more challenges and are more prone to crashes while turning left at intersections (e.g., Janke, 1994; McKnight and McKnight, 1999; Staplin et al., 1999). These reasons include age, as well as cognitive and physical health factors reflecting important skills for safe driving, including ability to multi-task and make decisions (Braitman et al., 2007; Chin and Zhou, 2018; Johnson, 1990; Kramer et al., 1999a; Walker et al., 1997), working memory (Zacks et al., 2000), focus and lack of distractibility (Colcombe et al., 2003; Kramer et al., 1999b; Madden and Whiting, 2004), attentional field of view (Ball, 1990; Owsley et al., 1998), ability to detect changes in visual scenes (Caird et al., 2005; Isabelle and Simon, 2020; Jensen et al., 2011; Owsley et al., 1991; Rizzo et al., 2009), ability to monitor changes in driving ability over time (Chen et al., 2020), and physical flexibility (Eby et al., 1998; Isler et al., 1997; Janke, 1994; Janke and Eberhard, 1998; McPherson et al., 1989, 1988). One specific example to consider for older driving making left turns at intersections is the left turn lane offset value that will provide unlimited sight distance, a critical component given the visual abilities of many older drivers (Staplin et al., 2001). Further, during left turns, older drivers' potential diminishing ability to share attention and turn the steering wheel sharply can compromise their ability to navigate intersections safely (Brewer et al., 2014).

Table 1
Rubric for scoring NDS videos for vehicles at signalized intersections.

Field	Value			
Initials_of_Scorer	Scorer Initials			
File_ID	Trip_ID: Unique identifier for each vehicle trip			
SCORED AS	**0**	**1**	**2**	**3**
Day_Night	Trip during non-daylight hours	Trip during daylight hours		
Weather	Clear	Not clear (rain, sleet fog)		
Intersection Type	T-intersection	4-way intersection	Other	
Dedicated Left Turn Lane at Intersection	No	Yes		
Wait at Intersection due to Queue or Traffic Signal	No Wait	Wait		
Opposing Lanes Present	No	Yes		
Oncoming Traffic Obscured	No oncoming traffic or oncoming traffic not obscured	Oncoming traffic obscured		
Vehicle Movement	No turn	Left turn	Right turn	
Traffic Signal Obscured	Not obscured	Obscured		
Traffic Signal State at time of turn/going through intersection	No signal	Green	Yellow	Red
Traffic Signal Turn Indicator (for turns)	No signal	Circular signal	Flashing signal	Arrow, not flashing
Turn Phasing for direction of turn	Not applicable/ no turn	Permissive (should yield)	Protected (have right of way)	
Gap Decision Required	No gap decision required	Gap decision Required		
Lead Vehicle to Follow through Intersection	No lead vehicle present	Lead vehicle present		
Time stamp of entering the intersection (in milliseconds)	Time at which the vehicle passed the stop line (or equivalent) to enter the intersection			
Time stop of exiting the intersection (in milliseconds, truncated to the nearest 0.1 s)	Time at which the vehicle completed the intersection; for turning vehicles, this was when the vehicle has straightened out again post-turn; for vehicles going straight through the intersection, this was when the vehicle passed the stop line (or equivalent) at the far end of the intersection			

Given these safety concerns, there is a need to investigate potential infrastructure changes that could potentially improve safety for older drivers at intersections. However, to date, older drivers' interactions with the road, vehicle, and environment have been difficult to study systematically. Driver behavior observations of participants that are aware that they are a part of a driving study are unable to represent fully naturalistic driving as participants are aware of the presence of an observer (Hutton et al., 2015). A review of the published research to date showed that the majority of current findings regarding older drivers have been identified and validated in controlled laboratory systems (e.g. in a driving simulator, cave automatic virtual environment) or via interactive field studies (open road, closed loop, operational tests, etc.) (Samuel et al., 2016). However, a rich naturalistic dataset that includes older drivers is the SHRP2 Naturalistic Driving Study (NDS) data. The NDS gathered information on trips and drivers' interactions with their vehicle, traffic environment, and roadway characteristics for over 3,400 driver participants from six United States metropolitan areas (Victor et al., 2015). The NDS intentionally included a relatively greater proportion of older drivers to better study this cohort. The NDS also involved extensive surveys and screenings of participants regarding demographics and behavioural, health and cognitive factors that can impact driving (Campbell, 2012).

Previously, NDS data have been analyzed to learn more about the safety impacts of left-turn lane offsets at intersections (Antin et al., 2015). NDS data have also been used to study the responses of through drivers in opposite direction left turn crash and near-crash events (Dinakar and Muttart, 2019). However, to the authors' knowledge, no prior studies have researched the validity of using NDS data to consider infrastructure changes to increase the safety of older drivers during left turns at signalized intersections. To fill this gap in research, this study investigated the practicality of using NDS data to examine this topic and other factors that impact intersection safety for older drivers.

2. Methods

The following section presents the data gathering process, video scoring techniques, and analysis methods used in this study.

2.1. NDS data

The initial step of this project involved obtaining the SHRP2 Naturalistic Driving Study (NDS) data required for the analysis from the Virginia Tech Transportation Institute (VTTI). The research team applied for a NDS data use license (DUL) and specified to VTTI the population being studied, a requested sample size, and the specific NDS tables and variables of principal interest. This request was fine-tuned using the sample data and data dictionaries available on the NDS InSight database web site and InSight database queries. The resulting data request included information on all trips made by drivers age 65 or over, commonly the age that is considered an ''older driver'' and of increased safety risk, which involved a signalized intersection and during which a crash or near-crash occurred. This included trips in which the driver turned left, turned right, or went straight at an intersection. Also requested was a sample of signalized intersection trips for this age group (>65) with no crash and no near-crash occurrences and comparison sample trips of signalized intersections by drivers age 30–49, including crash, near-crash, or no crash events. NDS baselines, or "no crash events", are designed to reflect "normal driving" and typical driving behaviors across the sample. These baselines were chosen via random sample stratified by participant and time driven, from the trips for each driver.

Overall, the data request to VTTI focused on variables in two main categories: driver characteristics and trip details. The driver characteristics included age, gender, driving history, driving behavior, driver knowledge, medical conditions, sleep habits, visual abilities, and cognitive abilities. The trip details included time of day, trip duration, speed, acceleration, braking, steering, event data, and event video. The requested driver characteristic data were collected through the questionnaires given to SHRP2 participants at the start of their participation in the SHRP2 study and were self-reported. SHRP2 participants were also given visual tests, including assessments of their peripheral vision and field of view, and cognitive tests. The vehicle trip and time series data (speed, acceleration, yaw, and pedal brake status) were collected through sensors installed in participants' vehicles.

Event data was requested for all trips, including crash severity, vehicle position, and maneuver data immediately before and immediately after a crash or near-crash. Event data on driver behaviors during events were also requested. Video data of each trip were requested for a 2.5 min period. For trips in which a crash or near-crash occurred, the 2.5-minute period included 2 min before the event and 30 s following the event. For baseline trips, each video began two minutes prior to the vehicle entering the signalized intersection and ended as the vehicle exited the intersection. The 2-minute period prior to an event typically allowed sufficient time to observe the drivers approaching the signalized intersection and traveling through the intersection. Additional video was provided for each vehicle which had a long wait at the intersection and did not finish traveling the intersection within the 2.5-minute timeframe.

In summary, the NDS dataset used for this study included trip videos from dashboard cameras facing out the front windshield and anonymized data tables of driver and trip characteristics. Information on 884 trips taken by 586 different drivers were obtained. Each trip had at least one signalized intersection. Of the 586 drivers, 59 % were age 65 or above (older drivers) and 41 % were in the comparison group age 30−49 (younger drivers). Overall, 46 % of the drivers were female and 54 % male. Of the 884 trips, 514 (59 %) were by older drivers. Crashes occurred during 11 % of the trips by older drivers and 7 % of trips by the younger drivers. The determination of which events were defined as crashes was made by the Virginia Tech Transportation Institute (VTTI), an instrumental leader of NDS data collection and primary processer of the raw data (Hankey et al., 2016; Virginia Tech Transportation Institute, n.d.). Overall, several contractors, university centers, and individuals were essential to the completion of this dataset (Blatt et al., 2014).

2.2. Video scoring

Forward dashboard camera video from NDS participant vehicles was obtained for 868 of the 884 trips in this study dataset. These videos were scored by trained scorers using the rubric in Table 1. This rubric focused on the characteristics of the signalized intersection and the road and traffic environments as the participant drivers approached and travelled through the intersection. Each video was scored at least once. An estimated 20 % of the videos, including 70 % of the videos with a left turn at a signalized intersection, were scored twice with different scorers for scoring verification.

2.3. Data analysis

Descriptive statistics were first used to gain an overview of the study dataset. Specifically, potential correlations between different driver and intersection characteristics were examined as were the characteristics of crashes and near-crashes. This informed the development of the modelling portion of analysis. Modelling methods were used to investigate the utility of the dataset for prediction of crash risk of older drivers performing left turns at intersections. For a detailed and robust analysis, modelling methods incorporated with machine learning techniques were used to analyze the data. Four different types of models, and supervised machine learning algorithms used in crash risk literature (Huang et al., 2020) were used in this study: Logistics Regression (Herrell, 2015), Support Vector Machine (SVM) (Hearst, 1998), Decision Tree (Freund and Mason, 1999), and Random Forest (Liaw and Wiener, 2002). Each of these modelling approaches analyses data differently to retrieve the resulting outcome. In this study, the dependent variables for each model was either the occurrence of a crash or no crash (base condition). Given the complexity and inability to account for the intensity of a near-crash case potentially resulting in a crash, these cases were not included in the final models.

Logistic regression algorithms find the best fit model for the relationship between a dichotomous characteristic of interest (dependent variable) and a set of independent variables. An example of a logistic regression model would be estimating whether a driver will crash or not crash (the dichotomous variable) given certain driver characteristics. Support Vector Machine (SVM) algorithms seek to find a hyperplane that best divides a dataset into classes; an example, for drivers approaching an intersection, one wants be able to identify whether they will drive safety through the intersection or not. From collected data, cases where drivers drove safety through the intersection are separated from the cases where drivers did not to learn the most critical factors for determining one outcome or another. Decision Trees models, on the other hand, are built in the form of a tree structure, where a dataset is broken into smaller and smaller subsets as an associated decision tree, with decision nodes and leaf nodes, is incrementally developed. Finally, Random Forest algorithms take an ensemble approach and aggregate multiple outputs made by a diverse set of predictors to obtain better results. A random forest model operates from decision trees and outputs classification of the individual trees, correcting for the habit of decision trees to overfit to their training dataset.

The machine learning models were created and run in Python, a general-purpose programming language with widespread use, with 80 % of the data assigned for training and 20 % for validation, a common split found in machine learning literature. During the analysis, the models evolved from simple models with fewer variables to more complex models with additional variables. Initially, the focus of the models was the vehicle time series data (such as speed and acceleration; hereafter referred to as vehicle data) and the video scoring data. Vehicle data were examined at 0.1 s intervals from the time participant vehicles entered the signalized intersection until they completed the intersection. All missing data were replaced by mean values, followed by the normalization of the data to the zero mean and united standard deviation, a common and accepted method in machine learning (Somasundaram and Nedunchezhian, 2011). In the next phases of the modelling, driver behavior data from the NDS participant questionnaires were incorporated into the models, and then medical questionnaire and screening data were added. Each of these variables were added sequentially to determine the most successful models for predicting crash risk. All variables were formatted and normalized before being incorporated in the models. The learning hyperparameters, or fixed training properties, for each machine learning algorithm were found using the Grid Search (Bergstra et al., 2013). The best variables were selected using the univariate feature selection method. ANOVA (Analysis of variance) F-values were used for the univariate statistical test and the k variable with highest F-values was selected to be passed through the machine learning models.

3. Results

The following section presents the descriptive statistics of the collected data and model results.

3.1. Descriptive statistics

The NDS dataset for this research contained data for 884 trips with at least one signalized intersection. An estimated 42 % of these trips (377 trips) included a crash (81 trips) or near crash (296 trips) at a signalized

Table 2
Summary of NDS data received from VTTI.

Event	All Signalized Intersection Trips		Trips with Left Turn at Signalized Intersection	
	Drivers 30–49	Drivers 65 & over	Drivers 30–49	Drivers 65 & over
Crash	26 (7 %)	55 (11 %)	13 (10 %)	26 (13 %)
Near-crash	140 (38 %)	156 (30 %)	46 (34 %)	72 (36 %)
Baseline	200 (54 %)	299 (58 %)	75 (56 %)	102 (50 %)
Non-Subject Conflict	4 (1 %)	4 (1 %)	0 (0 %)	2 (1 %)
Total	370 (100 %)	514 (100 %)	134 (100 %)	202 (100 %)

intersection as presented in Table 2. In the SHRP2 NDS data documentation, crashes are defined as any contact that the subject vehicle has with a moving or fixed object, at any speed. Crashes include "non-premediated departures of the roadway where at least one tire leaves the paved or intended travel surface of the road." Near-crashes in the documentation are defined as "any circumstance that requires a rapid evasive maneuver by the subject vehicle or any other vehicle, pedestrian, cyclist, or animal to avoid a crash." Non-subject conflicts are defined as "any incident that gets captured on video (near-crash or crash) that does not involve the subject driver."

Through the video scoring, all trips during which drivers turned left at a signalized intersection were identified. After the scoring was completed, the data for these 336 trips were separated from the main study dataset. During these trips, there were a total of 39 crashes, including 26 crashes involving older drivers (age 65 and over) and 13 crashes involving younger drivers (ages 30–49).

Within NDS crash data, four levels of crash severity are recorded (from most to the least severe): "Most severe," "Police Reportable," "Minor," and "Tire Strike, Low Risk." In this study's dataset from the NDS, there was a relatively small volume of crashes during left turn movements at signalized intersections. Further, most of the crashes were tire strikes/low risk, which occur if a driver clips a curb during a turn or slightly leaves the roadway, or minor crashes, involving physical contact with another object with minimal damage (less than $1,500). Few were police reportable, with damage of at least $1,500 or the hitting of a large animal or traffic sign, or more severe, defined as involving an injury needing doctor follow-up, airbag deployment, or vehicle towing post-crash.

In this study, of the 13 crashes of the younger driver cohort during left turns, eight (62 %) were tire strikes, four were minor crashes and one was police reportable. For the older driver cohort age (65 and over), nine of the 26 crashes (35 %) were categorized as tire strikes, ten as minor crashes (38 %), three as police reportable, and four crashes as most severe. The video scoring found that, of the older drivers' left turn crashes at signalized intersections, 35 % (9 crashes) occurred during the permissive turn phase, and that 62 % (16 crashes) took place at intersections with protected left-turn lanes during the protected turn phase.

The NDS included initial questionnaires and screenings which collect behavioral, health, and cognitive data of participants. Details on the health and cognitive challenges identified during these screenings for those drivers in this study's dataset who experienced a left-turn crash at a signalized intersection are presented in Fig. 1. An estimated 87 % of the older drivers with such crashes showed at least a mild impairment in their visual search abilities during the screenings. Impairment was assessed by measuring the time it took for participants to complete two different visual search tasks and by participants' degree of accuracy. An estimated 96 % of the older drivers with left-turn crashes made at least minor errors on the clock drawing test, a test which identifies spatial-cognitive challenges and is used to screen for dementia.

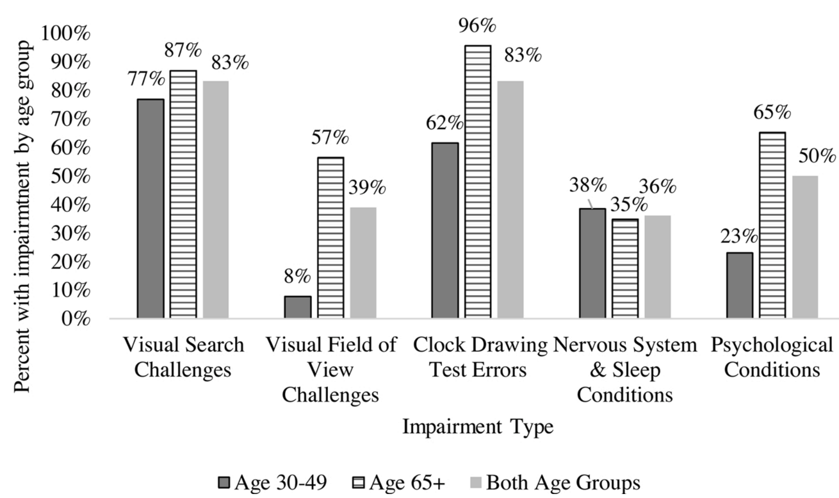

Fig. 1. Visual, cognitive, and medical challenges impacting participant drivers with left-turn crashes at signalized intersections.

Table 3
Visual, cognitive, and medical challenges impacting all participant drivers making left turns at signalized intersections.

Event	Age	Visual Search Challenges	Visual Field of View Challenges	Clock Drawing Test Errors	Nervous System Ailments	Psychological Conditions
Crashes	30–49	77 %	8 %	62 %	38 %	23 %
	65+	87 %	57 %	96 %	35 %	65 %
	Total	83 %	39 %	83 %	36 %	50 %
Near-crashes	30–49	53 %	3 %	71 %	32 %	3 %
	65+	95 %	55 %	84 %	23 %	15 %
	Total	81 %	37 %	79 %	26 %	11 %
Baseline	30–49	56 %	52 %	75 %	20 %	19 %
	65+	23 %	38 %	77 %	26 %	6 %
	Total	37 %	43 %	76 %	24 %	11 %

Table 4
Machine learning and OLS models for predicting left-turn crash risk at signalized intersections of all driver ages in study.

Significant Variables (P-values)*	R-Squared */Adjusted R-Squared*	Input	Model	Performance (Accuracy)
Day/Night (0.003) Clear/Inclement Weather (0.011) Opposing Lane Present (0.019) SD of Acceleration (0.001)	0.494/ 0.411	Vehicle & Video Data (All Variables)	Logistic Regression	Training: 89.8 % Validation: 82.7 %
Nervous System/Sleep Condition (0.009) Limited Flexibility (0.025) Severe Arthritis (0.003)	0.321/ 0.257	Medical Data (All Variables)	Random Forest	Training: 82.1 % Validation: 86.8 %
Day/Night (0.001) Clear/Inclement Weather (0.001) SD of Acceleration (0.001) Nervous System/Sleep Condition (0.011)	0.499/ 0.452	Medical, Behavioral, Video & Vehicle Data (15 Variables)	Support Vector Machine	Training: 83.7 % Validation: 94.3 %

* Significant Variables, and R-Square and Adjusted R-Square values are extracted from Ordinary Least Squares models and are independent of Machine Learning Models.

Table 5
Machine learning and OLS models for predicting older drivers' risk of left-turn crashes at signalized intersections.

Significant Variables* (P-values)	R-Squared */Adjusted R-Squared*	Input	Model	Performance (Accuracy)
Heart Conditions (0.007) Nervous System/Sleep Conditions (0.001) Severe Arthritis (0.006) Impaired Visual Search (0.004) Multiple Med. Conditions (0.05)	0.749/ 0.652	Medical Data (All Variables)	Support Vector Machine	Training: 91.9 % Validation: 70.6 %
Nervous System/Sleep Conditions (0.001) Severe Arthritis (0.049) Impaired Field of View (0.038)	0.730/ 0.669	Medical, Behavioral, Video & Vehicle Data (10 Variables)	Logistic Regression	Training: 83.8 % Validation: 76.5 %
Gap Decision (0.037) Nervous System/Sleep Conditions (0.003) Severe Arthritis (0.052)	0.803/ 0.728	Medical, Behavioral, Video & Vehicle Data (15 Variables)	Random Forest	Training: 99.0 % Validation: 83.0 %

* Significant Variables, and R-Square and Adjusted R-Square values are extracted from OLS models and are independent of Machine Learning Models.

These values presented in Fig. 1 differed from those who did no experience crashes. Overall, the percentage of drivers with certain visual, cognitive, and/or medical challenges was higher for those drivers who experienced a crash than those who did not. This is demonstrated in Table 3.

3.2. Model results

Several models were tested with different variables to first predict drivers' crash risk during left turns at signalized intersections for both age cohorts, or all drivers in this study, and then for the older cohort only. For both of these two groups, all four modeling methods were used. Each of these models used variables available from the vehicle data, video scoring data, medical screening data, and behavioral questionnaire data to predict the occurrence of a crash (1) or no crash event (0). The machine learning models that performed the strongest at predicting drivers' left-turn crash risk are presented in Table 4. For example, the first model was a logistic regression model trained with video scoring and vehicle data. Four primary significant variables of this model included the time of day, weather, the presence of an opposing lane, and the standard deviation of acceleration. The R-squared values presented represent the variance of the same variables run from using Ordinary Least Squares modeling; in the case of this model, the adjusted R-squared value was 0.411. This logistic trained model achieved a training accuracy of 90 % and a validation accuracy of 83 %, meaning that the model was able to accurately predict 90 % of crash cases from the training data, and 83 % of cases from the set-aside validation data. The overall best model for predicting crash risk was the Support Vector Machine Model with the best 15 variables selected using the univariate feature selection method from the medical, behavioral, video scoring, and vehicle data. This model achieved a validation accuracy of 94 %. The Decision Tree model details are not included in the results as they did not perform as well as the other methods.

The second round of model testing focused on identifying the best model at predicting left-turn crash risk for older drivers specifically. These results are summarized in Table 5. The best model for predicting this crash risk was found to be the Random Forest Model, with the best 15 variables identified using the univariate feature selection method from the medical, behavioral, video scoring, and vehicle time series data. This model had a validation accuracy of 83 %.

As the results presented in Tables 4 and 5 show, Ordinary Least Squares (OLS) models were developed and used to identify the statistically significant variables in predicting left-turn crashes from the Machine Learning models. This model type was chosen given the ability to integrate in the coding process and as it has been shown to perform at the same level as logistic regression for classifying dependent variable outcomes with binary criterion (Pohlmann and Leitner, 2003). For the older drivers, these models revealed that most of the statistical

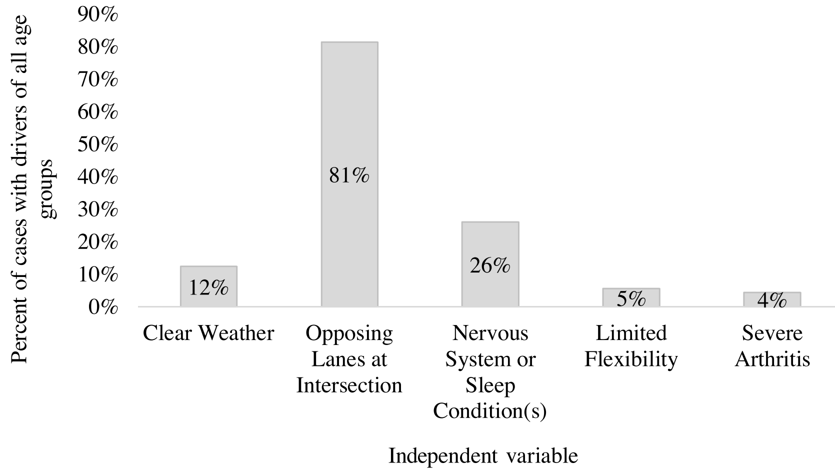

Fig. 2. Summary of distribution of significant model variables for all crash and no crash cases of all driver age groups in study.

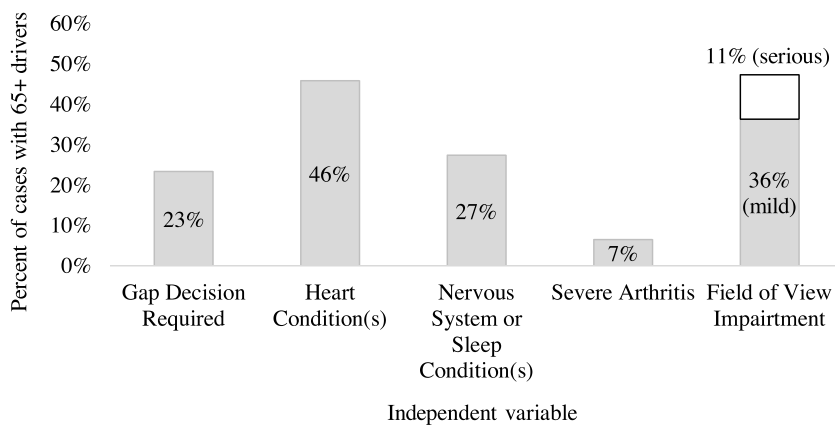

Fig. 3. Summary of distribution of significant model variables for all crash and no crash cases of drivers 65 years and older.

significant variables were health variables, such as the presence of arthritis or nervous system or sleep ailments, and visual and cognitive factors, such as having to monitor oncoming traffic, make good gap decisions, and accurately gauge the edges of the roadway while turning. Variables from the vehicle data and video scoring were less significant in these models. The summary distributions of some of the primary significant variables of the models are in Figs. 2 and 3 for additional context.

4. Conclusions

The SHRP2 NDS offers a rich dataset of driver performance and behaviors across a diverse age range, including older drivers. Given the documented challenges with the older drivers safely making left turns at signalized intersections for older drivers, the primary goal this study was to assess the usefulness of the SHRP2 NDS data to investigate infrastructure factors that contribute to these crashes. The study results did suggest that the most common type of crash for older drivers making a left turn at a signalized intersection involve them hitting a curb or leaving the roadway; this event accounted for 71 % of the older driver left-turn crashes at signalized intersections. However, this finding was not significant within the modelling results, likely due to the small number of crashes in the dataset. Thus, definitive findings and recommendations with regards to infrastructure changes were unable to be identified using this data.

While not the primary focus of the study, the results did suggest several health-related factors are correlated with crash risk for older drivers during left turn movements at signalized intersections. Specifically, the analysis found that statistically significant factors for such crashes included whether a driver has arthritis, a nervous system condition (e.g. epilepsy, Parkinson's, multiple sclerosis), and/or a sleep condition (e.g. insomnia, sleep apnea), or visual or cognitive challenges. This finding aligns with previous studies on older drivers, including research showing that older drivers have more confusion than younger drivers with the differences between permissive and protected left-turn signalizations at intersections (Knoblauch et al., 1995) and research showing the importance of understanding increased visual challenges faced by older drivers at intersections (Staplin et al., 1989).

Overall, the results of this study suggest that the use of SHRP2 NDS data alone will not lead to definitive findings or recommendations for infrastructure changes to increase safety for older drivers at signalized intersections during left turns. Thus, this study demonstrates the need for researchers and practitioners to consider using other data sources and/or data collection methods other than NDS-based datasets when investigating the key gap in literature on older driver safety and the impacts of different infrastructure and intersection configurations.

5. Limitations

This study included only a relatively small number of events and trips with older drivers turning left at signalized intersections and a small number of crashes. Thus, while the study's results provide an insight into the factors contributing to left-turn crashes, they are not fully generalizable. Further, the SHRP2 NDS database was not designed to be a comprehensive crash database but rather to complement existing crash databases by providing a wealth of information on driver attributes and behaviors.

6. Future research

Future research should continue to explore the use of available datasets to investigate the role of infrastructure in the crash risk for older driving during left turns at signalized intersections and potential methodologies for collecting detailed older driver crash information on a larger scale. Additionally, given the results of the study, future studies should investigate further the health, cognitive, and behavioural factors that impact older driver safety during intersection left-turns.

Funding details

This work was supported in part by the New England Transportation Consortium (NETC) under NETC Grant 15-2.

CRediT authorship contribution statement

Tracy Zafian: Conceptualization, Methodology, Resources, Investigation, Data curation, Visualization, Writing - original draft, Writing - review & editing. **Alyssa Ryan:** Conceptualization, Investigation, Writing - original draft, Writing - review & editing. **Ravi Agrawal:** Methodology, Investigation, Formal analysis, Resources, Data curation. **Siby Samuel:** Conceptualization, Methodology, Supervision, Funding acquisition, Writing - original draft, Writing - review & editing. **Michael Knodler:** Conceptualization, Project administration, Supervision, Funding acquisition, Writing - review & editing.

Declaration of Competing Interest

No authors have any conflicts of interest. There is no financial/personal interest or belief by any authors that could affect their objectivity.

Acknowledgements

The research team would like to thank the New England Transportation Consortium (NETC) for its funding of this project and the NETC Technical Advisory Committee for its feedback throughout the study. The team would also like to acknowledge Dr. Miguel Perez and the other VTTI staff who provided the SHRP2 NDS data for this study and data support.

References

Antin, J., Stulce, K., Eichelberger, L., Hankey, J., 2015. Naturalistic Driving Study: Descriptive Comparison of the Study Sample With National Data (No. S2-S31- RW-1). Washington, D.C.. https://doi.org/10.17226/22196.

Ball, K., 1990. Clinical visual perimetry underestimates peripheral field problems in older adults. Clin. Vis. Sci. 5 (2), 113–125.

Bergstra, J., Yamins, D., Cox, D.D., 2013. Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures. In: 30th International Conference on Machine Learning. Atlanta, GA.

Blatt, A., Pierowicz, J., Flanigan, M., Lin, P.-S., Kourtellis, A., Jovanis, P., Jenness, J., Wilaby, M., Campbell, J., Richard, C., Good, D., Czar, N., Hoover, M., 2014. Naturalistic Driving Study: Field Data Collection (No. S2-S07- RW-1), Naturalistic Driving Study: Field Data Collection. Washington, D.C.. https://doi.org/10.17226/22367.

Braitman, K.A., Kirley, B.B., Ferguson, S., Chaudhary, N.K., 2007. Factors leading to older drivers' intersection crashes. Traffic Inj. Prev. 8 (3), 267–274. https://doi.org/10.1080/15389580701272346.

Brewer, M., Debbie, M., Pate, A., 2014. Handbook for Designing Roadways for the Aging Population, 3rd ed. Washington, D.C.

Caird, J.K., Edwards, C.J., Creaser, J.I., Horrey, W.J., 2005. Older driver failures of attention at intersections: using change blindness methods to assess turn decision accuracy. Hum. Factors 47 (2), 235–249.

Chen, Y.T., Mazer, B., Myers, A., Vrkljan, B., Koppel, S., Charlton, J.L., Marshall, S.C., Gélinas, I., 2020. Changes in older drivers' self-awareness of driving ability over one year. Accid. Anal. Prev. 144 https://doi.org/10.1016/j.aap.2020.105552.

Chin, H.C., Zhou, M., 2018. A study of at-fault older drivers in light-vehicle crashes in Singapore. Accid. Anal. Prev. 112, 50–55. https://doi.org/10.1016/j.aap.2017.12.024.

Colcombe, A.M., Kramer, A.F., Irwin, D.E., Colcombe, S., Peterson, M.S., Hahn, S., 2003. Age-related effects of attentional and oculomotor capture by onsets and color singletons as a function of experience. Acta Psychol. (Amst.) 113 (2), 205–225.

Dinakar, S., Muttart, J., 2019. Driver behavior in left turn across path from opposite direction crash and near crash events from SHRP2 naturalistic driving. SAE Tech. Pap. https://doi.org/10.4271/2019-01-0414.

Eby, D.W., Trombley, D.A., Molnar, L.J., Shope, J.T., 1998. The Assessment of Older Drivers' Capabilities: A Review of the Literature (No. UMTRI-98-24). The University of Michigan, Transportation Research Institute. The University of Michigan Transportation Research Institute, Warren, MI.

Eck, R.W., Winn, G., 2002. Older-driver perception of problems at unsignalized intersections on divided highways. Transp. Res. Rec. 1818, 70–77. https://doi.org/10.3141/1818-11.

Freund, Y., Mason, L., 1999. The alternating decision tree algorithm. In: 16th International Conference on Machine Learning. Bled, Slovenia. https://doi.org/10.1093/jxb/ern164.

Garber, N.J., Srinivasan, R., 1991. Characteristics of accidents involving elderly drivers at intersections. Transp. Res. Rec. 1325, 8–16.

Hankey, J.M., Perez, M.A., McClafferty, J.A., 2016. Description of the SHPR2 Naturalistic Database and the Crash, Near-Crash, and Baseline Data Sets, The Strategic Highway Research Program 2. Washington, D.C.. https://doi.org/10.5151/cidi2017-060.

Hearst, M.A., 1998. Support vector machines. IEEE Intell. Syst. Appl. 13 (4), 18–28.

Herrell, F.E., 2015. Ordinal logistic regression. Regression Modeling Strategies. Springer, Cham, New York, NY, pp. 311–325.

Huang, T., Wang, S., Sharma, A., 2020. Highway crash detection and risk estimation using deep learning. Accid. Anal. Prev. 135 https://doi.org/10.1016/j.aap.2019.105392.

Hutton, J.M., Bauer, K.M., Fees, C.A., Smiley, A., 2015. Analysis of Naturalistic Driving Study Data: Offset Left-Turn Lanes (No. S2-S08B- RW-1), Analysis of Naturalistic Driving Study Data: Offset Left-Turn Lanes. Washington, D.C.. https://doi.org/10.17226/22315.

Insurance Institute for Highway Safety, 2019. Fatality Facts 2018: Older People [WWW Document]. URL https://www.iihs.org/topics/fatality-statistics/detail/older-people (Accessed 9.24.20).

Isabelle, M.P., Simon, M., 2020. Comparison between elderly and young drivers' performances on a driving simulator and self-assessment of their driving attitudes and mastery. Accid. Anal. Prev. 135 https://doi.org/10.1016/j.aap.2019.105317.

Isler, R.B., Parsonson, B.S., Hansson, G.J., 1997. Age related effects of restricted head movements on the useful field of view of drivers. Accid. Anal. Prev. 29 (6), 793–801.

Janke, M.K., 1994. Age-Related Disabilities That May Impair Driving and Their Assessment (No. 156). California Department of Motor Vehicles, Sacremento, CA doi:x.

Janke, M.K., Eberhard, J.W., 1998. Assessing medically impaired older drivers in a licensing agency setting. Accid. Anal. Prev. 30 (3), 347–361.

Jensen, M.S., Yao, R., Street, W.N., Simons, D.J., 2011. Change Blindness and Inattentional Blindness. Cogn. Sci. 2 (5), 529–546. https://doi.org/10.1002/wcs.130.

Johnson, M.M.S., 1990. Age differences in decision making: a process methodology for examining strategic information processing. J. Gerontol. 45 (2), 75–78. https://doi.org/10.1093/geronj/45.2.P75.

Knoblauch, R., Nitzberg, M., Reinfurt, D., Council, F., Zegeer, C., Popkin, C., 1995. Traffic Operations Control for Older Drivers (No. FHWA-RD-94-119). Washington, D.C..

Kramer, A.F., Hahn, S., Gopher, D., 1999a. Task coordination and aging: explorations of executive control processes in the task switching paradigm. Acta Psychol. (Amst.) 101 (2–3), 339–378.

Kramer, A.F., Hahn, S., Irwin, D.E., Theeuwes, J., 1999b. Attentional capture and aging: Implications for visual search performance and oculomotor control. Psychol. Aging 14 (1), 135–154.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R News 2 (3), 18–22.

Madden, D.J., Whiting, W.L., 2004. Age-related changes in visual attention. In: Costa, P. T., Siegler, I.C. (Eds.), Advances in Cell Aging and Gerontology. Elsevier, Amsterdam, pp. 41–88. https://doi.org/10.1016/S1566-3124(03)15003-1.

Mayhew, D.R., Simpson, H.M., Ferguson, S.A., 2006. Collisions involving senior drivers: high-risk conditions and locations. Traffic Inj. Prev. 7 (2), 117–124.

McKnight, A.J., McKnight, A.S., 1999. Multivariate analysis of age-related driver ability and performance deficits. Accid. Anal. Prev. 31 (5), 363–370. https://doi.org/10.1016/S0001-4575(98)00082-7.

McPherson, K., Ostrow, A., Shaffron, P., Yeater, R., 1988. Physical Fitness and the Aging Driver: Phase I (No. HS-040 558). AAA Foundation for Traffic Safety, Washington D. C.

McPherson, K., Ostrow, A., Shaffron, P., Yeater, R., 1989. Physical Fitness and the Aging Driver: Phase II. AAA Foundation for Traffic Safety, Washington D.C.

National Highway Traffic Safety Administration, 2020. Traffic Safety Facts: Older Population (No. DOT HS 812 928). Washington, D.C.

Owsley, C., Ball, K., Sloane, M.E., Roenker, D.L., Bruna, J.R., 1991. Visual/cognitive correlates of vehicle accidents in older drivers. Psychol. Aging 6 (3), 403–415. https://doi.org/10.1037/0882-7974.6.3.403.

Owsley, C., Ball, K., McGwin Jr., G., 1998. Visual processing impairment and risk of motor vehicle crash among older adults. J. Am. Med. Assoc. 279 (14), 1083–1088.

Pohlmann, J.T., Leitner, D.W., 2003. A comparison of ordinary least squares and logistic regression. Ohio J. Sci. 103 (5), 118–125.

Preusser, D.F., Williams, A.F., Ferguson, S.A., Ulmer, R.G., Weinstein, H.B., 1998. Fatal crash risk for older drivers at intersections. Accid. Anal. Prev. 30 (2), 151–159. https://doi.org/10.1016/S0001-4575(97)00090-0.

Rizzo, M., Sparks, J., McEvoy, S., Viamonte, S., Kellison, I., Vecera, S.P., 2009. Change blindness, aging and cognition. J. Clin. Exp. Neuropsychol. 31 (2), 245–256.

Samuel, S., Yamani, Y., Fisher, D.L., 2016. Large reductions are possible in older driver crashes at intersections. Clin. Exp. Optom. https://doi.org/10.1111/cxo.12443.

Somasundaram, R.S., Nedunchezhian, R., 2011. Evaluation of three simple imputation methods for enhancing preprocessing of data with missing values. Int. J. Comput. Appl. 21 (10), 14–19. https://doi.org/10.5120/2619-3544.

Staplin, L., Lococo, K., Sim, J., Drapcho, M., 1989. Age differences in a visual information processing capability underlying traffic control device usage. Transp. Res. Rec. 1244, 63–72.

Staplin, L., Lococo, K.H., Stewart, J., Decina, L.E., 1999. Safe Mobility for Older People Notebook (No. DOT HS 808 853). Washington, D.C.. https://doi.org/10.1017/S0959259800000290.

Staplin, L., Lococo, K., Byington, S., Harkey, D., 2001. Highway Design Handbook for Older Drivers and Pedestrians (No. FHWA-RD-01-103). Federal Highway Administration, Washington, D.C.

Staplin, L., Lococo, K.H., Martell, C., Stutts, J., 2012. Taxonomy of Older Driver Behaviors and Crash Risk (No. DOT HS 811 468A). Washington, D.C..

Victor, T., Dozza, M., Bärgman, J., Boda, C.N., Engström, J., Flannagan, C., Markkula, G., 2015. Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and Crash Risk (No. S2-S08A- RW-1). Washington, D.C.

Virginia Tech Transportation Institute, n.d. Insight Data Access Website SHRP2 Naturalistic Driving Study [WWW Document]. URL https://insight.shrp2nds.us/ (Accessed 8.20.07).

Walker, N., Fain, W.B., Fisk, A.D., McGuire, C.L., 1997. Aging and decision making: driving-related problem solving. Hum. Factors 39 (3), 438–444. https://doi.org/10.1518/001872097778827188.

Zacks, R.T., Hasher, L., Li, K.Z., 2000. Human memory. The Handbook of Aging and Cognition. Lawrence Erlbaum Associates Publishers, pp. 292–357. https://doi.org/10.4324/9781315807195.