Accident Analysis and Prevention 152 (2021) 106003

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Quantifying the automated vehicle safety performance: A scoping review of the literature, evaluation of methods, and directions for future research

Soheil Sohrabi [a,b,*], Ali Khodadadi [a], Seyedeh Maryam Mousavi [a,b], Bahar Dadashova [b], Dominique Lord [a]

[a] *Zachry Department of Civil & Environmental Engineering, Texas A&M University, Texas, USA*
[b] *Texas A&M Transportation Institute (TTI), Texas A&M University, Texas, USA*

ARTICLE INFO

Keywords:
Automated vehicle
Safety evaluation
Safety implications
Driving simulator
Traffic simulation
Road test data
Target crash population
Safety effectiveness
Failure risk
Evaluation metric

ABSTRACT

Vehicle automation safety must be evaluated not only for market success but also for more informed decision-making about Automated Vehicles' (AVs) deployment and supporting policies and regulations to govern AVs' unintended consequences. This study is designed to identify the AV safety quantification studies, evaluate the quantification approaches used in the literature, and uncover the gaps and challenges in AV safety evaluation. We employed a scoping review methodology to identify the approaches used in the literature to quantify AV safety. After screening and reviewing the literature, six approaches were identified: target crash population, traffic simulation, driving simulator, road test data analysis, system failure risk assessment, and safety effectiveness estimation. We ran two evaluations on the identified approaches. First, we investigated each approach in terms of its input (required data, assumptions, etc.), output (safety evaluation metrics), and application (to estimate AVs' safety implications at the vehicle, transportation system, and society levels). Second, we qualitatively compared them in terms of three criteria: availability of input data, suitability for evaluating different automation levels, and reliability of estimations. This review identifies four challenges in AV safety evaluation: (a) shortcomings in AV safety evaluation approaches, (b) uncertainties in AV implementations and their impacts on AV safety, (c) potential riskier behavior of AV passengers as well as other road users, and (d) emerging safety issues related to AV implementations. This review is expected to help researchers and rulemakers to choose the most appropriate quantification method based on their goals and study limitations. Future research is required to address the identified challenges in AV safety evaluation.

1. Introduction

Vehicle automation have the potential to improve traffic safety profoundly, mainly by eliminating driver error. According to the National Highway Traffic Safety Administration (NHTSA), human error contributes to 94 % of crashes, and Automated Vehicles (AVs) are optimistically expected to prevent those crashes (NHTSA, 2018). Nevertheless, more accurate AV safety evaluations are required before deploying AVs. Particularly, the intent to use AVs and their market success are contingent upon the safety evaluation of AVs (Sener et al., 2019). In addition, not only can manufacturers and the automotive industry benefit from the accurate safety evaluations of AVs, but legislative and executive agencies require such information to advocate with industry stakeholders and society (Junietz et al., 2018). Evaluating the safety implications of AVs is necessary for formulating regulations and policies to alleviate the unintended consequences of AV implementations and increase their benefits, as outlined by the United States Department of Transportation (US DOT, 2018) and the US Congressional Research Service (Canis, 2020).

Given the significant role of safety evaluation in the successful and efficient implementation of AVs, quantification studies are needed to evaluate AV safety. This study was designed to synthesize the lessons learned from existing studies that quantified the safety of AVs. To this end, we implemented a scoping review methodology to identify, screen, and review the existing literature about AV safety quantification. After identifying quantification methods, we ran two evaluations to compare the quantification methods. First, we investigated each approach in terms of its input (required data, assumptions, etc.), output (evaluation metrics), and application (to estimate AV safety implications at the vehicle, transportation system, and society levels). Second, we

* Corresponding author at: 3127 TAMU #303E, College Station, TX, 77843-3127, USA.
E-mail address: sohrabi.s@tamu.edu (S. Sohrabi).

https://doi.org/10.1016/j.aap.2021.106003
Received 7 October 2020; Received in revised form 18 December 2020; Accepted 16 January 2021
Available online 8 February 2021
0001-4575/

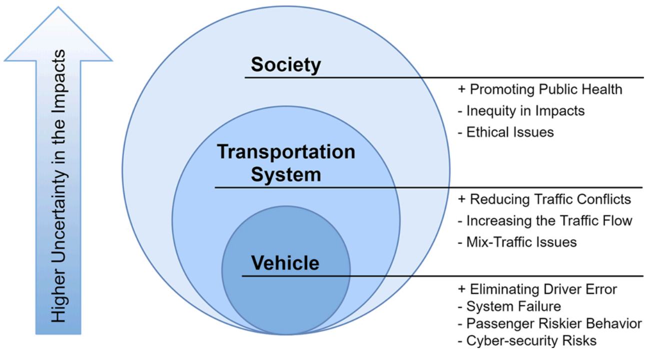

Fig. 1. AV Levels of safety implications.

compared the identified approaches based on three criteria: availability of input data, suitability for evaluating different levels of automation, and reliability of estimations. Then, we identified and discussed the gaps in the literature and AV safety evaluation challenges that need to be addressed in the future.

We expect this study to serve as a stop knowledge point and introduce future research avenues to contribute to AV safety evaluation. The results of evaluating AV safety quantification methods can help researchers, policy makers, and practitioners to choose an appropriate evaluation method based on their objective. This study did not examine nominal safety or perceived safety but rather targeted the substantive safety of AVs. Nominal safety refers to whether or not a vehicle is fulfilling all standards and laws that apply to the vehicle. The perceived safety indicates users' perception of the vehicle's safety. Moreover, this study does not intend to synthesize the AV safety quantifications results but rather to explore the methodologies.

The remainder of this paper is structured as follows. In Section 2, we briefly provide some background information about AVs' potential safety implications at different automation levels and discuss different levels of AV safety implications. Section 3 introduces the review methodology, including the review question, study identification process, and study inclusion criteria. In Section 4, we report the total number of studies included in this review, a descriptive analysis of the selected studies, and a literature review of the identified AV safety evaluation approaches. In Section 5, we discuss the findings of the literature review. This section consists of in-depth evaluation of the identified approaches and diccussion about their limitations and gaps.. A potential list of future research avenues is suggested in this section. Finally, in the Summary and Conclusions section, we summarize the key findings and conclude the paper.

2. Background information

2.1. Safety implications at different automation levels

AV safety, and the related safety implication complexity, vary in terms of driving automation levels, as defined by the Society of Automobile Engineers (SAE, 2018). In the lower levels of automation (Levels 1 and 2), the driver is responsible for dynamic driving tasks (DDTs), and advanced driver assistance systems (ADASs) on the vehicle can sometimes assist the human driver with steering or/and braking/accelerating (SAE, 2018). ADASs have the potential to prevent or mitigate crashes by partially eliminating driver error. In higher levels of automation, the automated driving system (ADS) performs the entire DDT while engaged. In Level 3, the DDT fallback-ready user needs to intervene when requested (SAE, 2018). On the other hand, Levels 4 and 5 of automation do not require a DDT fallback-ready user, and Level 5 has an unlimited operation design domain (ODD), unlike Levels 3 and 4. An ADS is expected to entirely eliminate driver error in its ODD; however, disengagement from ADSs in Level 3 of automation can be challenging.

2.2. Levels of safety implications

Vehicle automation impacts on safety can be investigated at three levels: vehicle, transportation system, and society (Fig. 1). At the vehicle level, AVs can be examined in terms of how they contribute to the critical driver-related reasons for crashes, such as inattention; internal and external distractions; inadequate surveillance; decision error caused by false assumptions and perceptions; performance error (i.e., execution of improper driver response); and nonperformance error mainly due to impairment, drowsiness, and fatigue (NHTSA, 2018). At the transportation system level, AV has the potential to reduce traffic conflicts and, consequently, reduce crashes. However, AVs' implementation carries higher levels of uncertainty at the transportation system level. At the society level, crashes pose a public health crisis, and the health impacts of AVs can be investigated based on the changes in motor vehicle crashes (Sohrabi et al., 2020). Previously, the role of motor vehicle crashes in public health has been measured in the form of premature mortalities from fatality crashes (Sohrabi and Khreis, 2020) and the disability-adjusted life year from injury crashes (Tainio, 2015).

Although traffic crashes caused by driver error are expected to be eliminated after AVs' deployment, other safety issues may compromise the positive impacts (Kockelman et al., 2016; Litman, 2017; Yang et al., 2017). System operation failure (Koopman and Wagner, 2016), cybersecurity (Lee, 2017), and AV users' riskier behaviors related to feeling overly safe while using AVs are some examples of potential safety concerns at the vehicle level. At the transportation system level, AVs may experience safety issues related to the interaction between human drivers and AVs in mixed traffic (Virdi et al., 2019; Taeihagh and Lim, 2018), as well as AVs' potential to increase traffic flow and, consequently, exposure to crashes as a result of induced demand, increased mobility, and changes in land use (Milakis et al., 2017). Moreover, due to the high cost of AVs, only wealthy consumers might be able to afford AVs as personal vehicles (Raj et al., 2019; Cohen and Shirazi, 2017) and, therefore, the disproportionate deployment of AVs may lead to health inequities that challenge AV safety impacts at the society level. The controversial discussion about how AVs should react during an unavoidable crash is another example of AV safety challenges at the society level.

3. Review methodology

We followed a scoping review methodology proposed by Arksey and O'Malley (2005) to (a) examine the nature of previous studies on quantifying the impacts of AVs on traffic safety, (b) summarize and document the quantification methodologies in previous research, and (c) identify research gaps in the existing literature. In this study, we chose to conduct a scoping review since we aimed to identify previous studies answering a general question and then review the evidence from previous quantifications on AVs' impact on traffic safety (Munn et al., 2018). In this context, the findings are not aggregated, nor is the quality of evidence assessed (Arksey and O'Malley, 2005). In the subsequent sections, we discuss the scoping review methodology..

3.1. Review question

The first step in a scoping review is to identify a research question to be answered (Arksey and O'Malley, 2005). The research question for this review was the following: "What are the methodologies and the gaps in the existing research on quantifying the potential impacts of AVs on traffic safety?" We were specifically interested in research that quantified the impacts of AVs rather than studies of a speculative nature.

3.2. Identifying relevant studies

A search strategy was developed to retrieve relevant research evidence from four electronic research databases—Scopus, Institute of

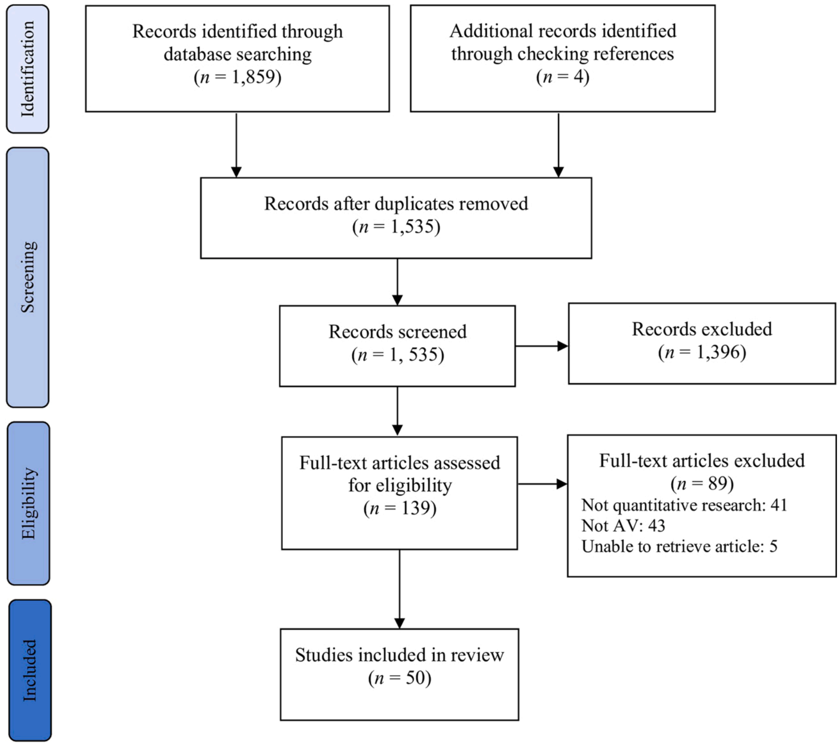

Fig. 2. Study identification and selection mechanism of the implemented scoping review.

Electrical and Electronics Engineers digital library (IEEE Xplore), Web of Science, and Transport Research International Documentation (TRID)—as well as reference lists of the retrieved publications. IEEE Xplore is a research database that covers more than 5 million journal articles, conference proceedings, standards, and related materials on multiple disciplines, including but not limited to computer science, electrical engineering and electronics, and allied fields.[1] IEEE Xplore is sponsored by the IEEE and other partner publishers. Scopus is Elsevier's research database, which covers more than 75 million records from 50,000 publishers in four core areas: life sciences, social sciences, physical sciences, and health science.[2] The Web of Science, sponsored by the Institute of Scientific Information, is a publisher-independent research database that covers more than 79 million records from several areas, such as life sciences, biomedical sciences, engineering, social sciences, arts and humanities, natural sciences, health sciences, engineering, computer science, and materials sciences.[3] We also explored the TRID database, a research database that combines the records from the Transportation Research Board's Transportation Research Information Services, which is solely focused on transportation research and provides access to more than 1.25 million records.[4]

The databases were searched to identify published articles, letters, reports, book chapters, and books using any combination of two sets of keywords in their title, abstract, and keywords: ["vehicle automation" or "automated vehicle" or "automated vehicle" or "autonomous vehicle" or "autonomous car" or "self-driving car" or "driverless car" or "automated driving"] and ["crashes" or "accidents" or "collision" or "safety"]. We included only published material written in English due to the burdensome translating process. All material considered in the review was published as of October 2020.

3.3. Study selection

To ensure consistency in selecting studies that answers the review question and excluded irrelevant studies, we defined a set of inclusion and exclusion criteria. The included studies had to meet the following established criteria:

1. Must explicitly quantify AVs' impacts on traffic safety rather than merely offer speculations and qualitative assessments.
2. Must evaluate the AV as a vehicle for ground transportation, such as automated cars, buses, shuttles, trucks, and the like.
3. Must investigate the safety of different levels of vehicle automation rather than individual AV technologies (e.g., ADASs, sensors, and algorithms).

Based on the inclusion criteria, connected vehicles' safety evaluations did not fall within the scope of this study. However, we elected to include the literature on connected and automated vehicles (CAVs) in our review, with a focus on the safety evaluation of automation components of CAVs. The selection process was divided into two stages. First, the titles and abstracts of the identified publications were assessed, and potentially relevant publications were selected. Second, the full text of the potentially relevant publications was retrieved and reviewed against the inclusion criteria, and studies that did not meet all inclusion criteria were excluded. The reference lists of included publications were also reviewed to find any relevant articles that were not identified through the developed search strategy.

4. Results

4.1. Search results

The implemented scoping review process is shown in Fig. 2. As of October 2020, a total of 1,859 publications were identified using the developed search strategy. After checking for duplicates, screening the identified articles, and reviewing articles' full text, we excluded 1,809

[1] Source: https://innovate.ieee.org/about-the-ieee-xplore-digital-library/
[2] Source: https://www.elsevier.com/solutions/scopus/why-choose-scopus
[3] Source: https://clarivate.libguides.com/webofscienceplatform/coverage
[4] Source: http://www.trb.org/InformationServices/AboutTRID.aspx

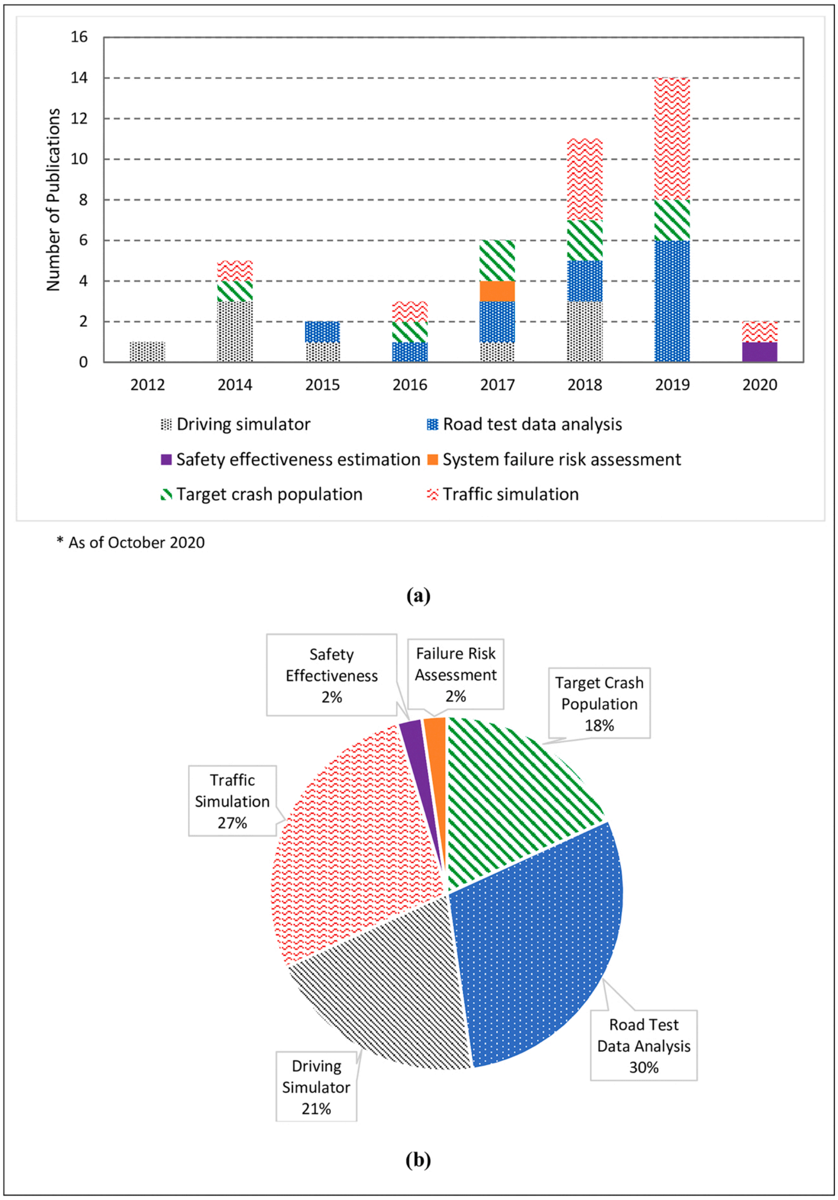

Fig. 3. (a) Publication date of the studies included in this review, and (b) distribution of the identified AV safety quantification approaches.

articles: 324 duplicates, 1,396 after screening, and 89 after full-text review. Ultimately, 50 articles met the inclusion criteria and were included in this review.

4.2. *Characteristics of included studies*

The number of publications increased significantly beginning in 2012, although in 2019, only 14 articles were published on quantifying AVs' safety implications (Fig. 3a). The AV safety quantification approaches can be classified into six groups: target crash population, traffic simulation, driving simulator, road test data analysis, system failure risk assessment, and safety effectiveness estimation. Fig. 3b shows the distribution of quantification approaches. Road test data analysis and simulation studies were more commonly used in the literature, followed by the driving simulator and target crash population approaches. Failure risk assessment and safety effectiveness quantification received the least attention. A time-series analysis of publications indicated that traffic simulation and road test data analysis methods began receiving more attention over time. Increases in availability of road test data may be one of the reasons behind this change.

4.3. *Identified approaches*

In this section, we discuss the AV safety quantification approaches identified in the literature.

4.3.1. *Target crash population*

The target crash population approach quantifies the number of preventable crashes after AV implementation. The quantification process in the examined studies followed three steps (Rau et al., 2015; Yanagisawa et al., 2017):

1 Identify AVs' ADS and ADAS functionality.
2 Match AV functionality with the target crash type.
3 Explore the crash datasets and identify preventable crashes.

In the first step, AV functions were investigated on the basis of (a) levels of automation (Lubbe et al., 2018; Agriesti et al., 2019), and (b) individual or combined ADS and ADAS functions (Combs et al., 2019; Detwiller and Gabler, 2017; Hendrickson and Harper, 2018; Li and Kockelman, 2016; Kusano and Gabler, 2014).

In the second step, the AV functionality was matched with corresponding crash characteristics. Previous studies assessed AV technology to mitigate either specific crash types (e.g., rear-end collision, pedestrian crashes) (Combs et al., 2019; Detwiller and Gabler, 2017; Hendrickson and Harper, 2018), specific crash-contributing factors (e.g., distracted driving, speeding, etc.), or critical pre-crash events (e.g., running a red light, vehicle failure) (Yanagisawa et al., 2017; Lubbe et al., 2018; Li and Kockelman, 2016; Kusano and Gabler, 2014). In addition, some AV functions are programmed to operate under a certain ODD to activate and achieve the maximum desired effectiveness; therefore, the crash dataset had to be filtered out to mirror those conditions properly. Lighting condition (day/night) (Yanagisawa et al., 2017; Agriesti et al., 2019), weather condition (clear/adverse) (Yanagisawa et al., 2017; Agriesti et al., 2019), road surface condition (wet/dry) (Yanagisawa et al., 2017; Agriesti et al., 2019), travel speed range (Yanagisawa et al., 2017; Agriesti et al., 2019; Hendrickson and Harper, 2018), visual obstruction (Lubbe et al., 2018; Combs et al., 2019), pedestrian crossing condition (Lubbe et al., 2018; Detwiller and Gabler, 2017), lane marking condition (Lubbe et al., 2018; Agriesti et al., 2019), and stable vehicle condition (Lubbe et al., 2018) are conditions under which AV safety was examined in the literature. The AV safety implications were explored for various road facilities and areas (Detwiller and Gabler, 2017; Hendrickson and Harper, 2018) as well. However, in some studies, facility type was automatically filtered out by selecting possible crash scenarios (e.g., running a red light, which is specific to intersections only) and beneficial safety equipment specific to that facility (e.g., cooperative intersection collision avoidance systems, which are applicable in intersections only) (Li and Kockelman, 2016; Kusano and Gabler, 2014). The safety effectiveness of AV technology was widely presumed to be 100 % in the literature (Yanagisawa et al., 2017; Agriesti et al., 2019; Detwiller and Gabler, 2017; Hendrickson and Harper, 2018; Kusano and Gabler, 2014); however, some studies accounted for the shortcomings in the safety implications of AVs by considering the effectiveness of AV technology (Lubbe et al., 2018; Combs et al., 2019; Li and Kockelman, 2016). AV safety effectiveness was either extracted from simulation studies (Combs et al., 2019) or indirectly through defining different sets of rules (Lubbe et al., 2018; Li and Kockelman, 2016). Each set consisted of assumptions regarding weather, road condition, vehicle condition, speed range, and so forth, through which both maximum effectiveness and lower effectiveness due to adverse conditions could be taken into account. Moreover, different rule sets provided a lower and upper bound for the expected number of preventable crashes instead of a constant value for effectiveness. Most of the literature assumed a 100 % market penetration rate (MPR); indeed, only two studies considered the MPR in their analysis (Agriesti et al., 2019; Li and Kockelman, 2016).

In the third step, the crash datasets were explored, and the crash characteristics were extracted. Next, the safety benefits of AVs were quantified in terms of the number of preventable crashes (Yanagisawa et al., 2017; Lubbe et al., 2018; Agriesti et al., 2019; Combs et al., 2019; Detwiller and Gabler, 2017; Hendrickson and Harper, 2018; Kusano and Gabler, 2014) and/or reduced cost of crashes (Yanagisawa et al., 2017; Hendrickson and Harper, 2018; Li and Kockelman, 2016). As a result, AV safety was attributed to ADSs (Yanagisawa et al., 2017; Lubbe et al., 2018; Agriesti et al., 2019; Combs et al., 2019; Detwiller and Gabler, 2017; Hendrickson and Harper, 2018; Kusano and Gabler, 2014; Li and Kockelman, 2016) and ADASs (Combs et al., 2019; Hendrickson and Harper, 2018; Li and Kockelman, 2016; Kusano and Gabler, 2014). The total number of preventable crashes was estimated in the target crash population approach, and some studies stratified crashes based on severity level (Detwiller and Gabler, 2017; Hendrickson and Harper, 2018; Li and Kockelman, 2016; Kusano and Gabler, 2014). Table A1 in the appendix summarizes the target population studies.

4.3.2. *Road test data analysis*

Analyzing AV road tests is one of the approaches used in the literature to evaluate AV safety. AV incident data were sourced from the California Department of Motor Vehicles (CA DMV) (Schoettle and Sivak, 2015; Teoh and Kidd, 2017; Favarò et al., 2017; Matysiak and Razin, 2018; Banerjee et al., 2018; Xu et al., 2019; Wang and Li, 2019; Petrović et al., 2020; Boggs et al., 2020; Das et al., 2020), US National Transportation Safety Board (NTSB) (Wang and Li, 2019), or AV manufacturers' self-reports (Schoettle and Sivak, 2015). CA DMV mandates that all manufacturers testing AVs on public roads file two different types of reports: (a) a report of a collision involving an AV within ten days after the collision; and (b) an annual report summarizing the disengagements.

Three types of analyses were found in the literature. First, the rate of AV incidents was compared to conventional car crashes as a benchmark (Schoettle and Sivak, 2015, Teoh and Kidd, 2017; Matysiak and Razin, 2018; Banerjee et al., 2018; Favarò et al., 2017). The AV incident rate was estimated as either number of crashes per number of AV vehicle miles traveled (VMT) (Schoettle and Sivak, 2015, Teoh and Kidd, 2017; Favarò et al., 2017) or the number of disengagements per VMT (Matysiak and Razin, 2018; Banerjee et al., 2018). The AV incident rates were then compared to either conventional vehicle crash rates (Schoettle and Sivak, 2015, Teoh and Kidd, 2017; Favarò et al., 2017; Banerjee et al., 2018) or injury and fatality crash rates (Matysiak and Razin, 2018). Unlike AV crashes, where the auto manufacturers report every single incident involving AVs, conventional vehicle crashes are reported by police based on the dollar amount of the property damage and therefore are significantly underreported. To have a fair comparison between AVs and conventional vehicle crash rates, Teoh and Kidd (2017) used AV police-reportable crashes, and Schoettle and Sivak (2015) adjusted the conventional vehicle crash rates for underreporting. Given the disparities in the equivalence between AV and conventional vehicle crash rates, mixed conclusions were drawn in the literature regarding AV safety in terms of crash rates.

Second, some studies investigated the characteristics of AV crashes in terms of collision type, crash location, speed, and causes of the crash. The majority of the literature ran a descriptive analysis of AV characteristics (Schoettle and Sivak, 2015; Favarò et al., 2017; Xu et al., 2019; Petrović et al., 2020), whereas some compared AV crash characteristics to conventional vehicle crashes (Schoettle and Sivak, 2015; Favarò et al., 2017; Petrović et al., 2020). Researchers found that the rate of rear-end crashes is higher in AV crashes (Schoettle and Sivak, 2015; Favarò et al., 2017; Petrović et al., 2020), while the severity of crashes is lower (Schoettle and Sivak, 2015). More rigorous statistical analyses, in the form of logistic regression (Wang and Li, 2019; Xu et al., 2019), a decision tree (Wang and Li, 2019), a Bayesian latent class model (Das et al., 2020), and logit discrete choice models (Boggs et al., 2020) were used to uncover the factors contributing to AV crash risk (Boggs et al., 2020), collision type (Xu et al., 2019; Wang and Li, 2019), and severity (Xu et al., 2019; Wang and Li, 2019). Driving speed, on-street parking, speed limit, and collision location—highway, arterial and collector, streetlights, and intersections—were shown to be associated with AV crash risk. The number of lanes marked with a centerline and clear weather conditions were shown to reduce the likelihood of AV crashes. AV driving mode (AV mode or conventional driver), collision location, roadside parking, rear-end collision, and one-way road were the main factors found to contribute to the severity level of AV-involved crashes. AV driving mode, AV stopped or not, vehicle turning movement, and whether crashes were associated with yielding to pedestrians/cyclists were the factors found to affect the collision type of AV crashes. The cause of AV disengagement was investigated by Banerjee et al. (2018), who found that 64 % of disengagements were the result of problems in, or untimely decisions made by, the machine learning system.

Third, the safety reliability of AVs was examined by comparing (a) the AV failure rate to other safety-critical autonomous systems (Banerjee et al., 2018); (b) the number of miles driven by AVs until a crash to the number of miles driven by conventional cars until a crash (Favarò et al., 2017); (c) the number of failure-free miles AVs should drive to reach conventional cars' failure rates (Kalra and Paddock, 2016; Li and Zhai, 2019); (d) the total number of miles driven to evaluate AV failure rate (Kalra and Paddock, 2016; Li and Zhai, 2019); and (e) the total number of miles AVs need to drive to demonstrate their failure rate is statistically lower than that of conventional cars (Kalra and Paddock, 2016). Banerjee et al. (2018) compared AV reliability with other safety-critical autonomous systems in terms of reliability per mission and demonstrated that AVs are 4.22 times worse than airplanes and 2.5 times better than surgical robots. Favarò et al. (2017) estimated that AVs drive 500, 000 miles before a crash, which shows AVs' reliability versus conventional vehicles. However, estimations regarding the number of failure-free miles AVs should drive to reach conventional vehicles' failure rate resulted in higher thresholds of 1.6 million miles (Kalra and Paddock, 2016) and 140 million miles (Li and Zhai, 2019). Kalra and Puddok (2016) showed that AVs need to be driven 51 and 61 million miles to be able to test their failure rate and statistically examine their failure rate, respectively. However, much higher numbers (71 billion miles) have been estimated for AV testing requirements to be able to properly investigate AV safety (Kalra and Paddock, 2016). Table A2 summarizes the studies that used AV road test data to evaluate their safety.

4.3.3. *Traffic simulations*

During the last decade, traffic simulation models have been implemented frequently to replicate conventional vehicles' driving characteristics in a fleet (Young et al., 2014). Research studies have employed traffic simulation models to assess AVs' safety effects and the assumption, methodologies, and limitations behind them (see Table A3 for a summary of related literature).

In the identified traffic simulation studies, various traffic microsimulation computer software was used, such as VISSIM (Kockelman et al., 2016; Katrakazas et al., 2019; Morando et al., 2018; Deluka Tibljaš et al., 2018; Rahman et al., 2019; Arvin et al., 2020; Mousavi et al., 2020), MATLAB, SUMO, VENTOS, and PELOPS (Bahram et al., 2014; Arvin et al., 2018, 2019; Qin and Wang, 2019). Depending on the study purpose, safety was evaluated at roadway segments (Katrakazas et al., 2019; Bahram et al., 2014; Ye and Yamamoto, 2019; Virdi et al., 2019; Qin and Wang, 2019; Zhang et al., 2015; Sinha et al., 2020), intersections (Kockelman et al., 2016; Arvin et al., 2018, 2019; Morando et al., 2018; Virdi et al., 2019; Rahman et al., 2019; Arvin et al., 2020; Mousavi et al., 2020), roundabouts (Morando et al., 2018; Deluka Tibljaš et al., 2018), or on/off-ramps (Kockelman et al., 2016).

For developing the simulation scenarios, different car-following models were utilized for conventional vehicles and AVs. Various car-following models were implemented to replicate conventional vehicles' driving behavior, such as Wiedemann 74 (Arvin et al., 2018; Deluka Tibljaš et al., 2018; Virdi et al., 2019; Arvin et al., 2020; Mousavi et al., 2020), Wiedemann 99 (Katrakazas et al., 2019; Morando et al., 2018; Zhang et al., 2015; Sinha et al., 2020), and user-defined models (Ye and Yamamoto, 2019). For AVs, car following was in the form of modified built-in models, including modified Wiedemann models (Kockelman et al., 2016; Arvin et al., 2018; Morando et al., 2018; Deluka Tibljaš et al., 2018; Arvin et al., 2020; Mousavi et al., 2020) or automated vehicle-specific models using external coding interfaces to either adjust a variable, introduce a new following strategy, or test various models (Bahram et al., 2014; Arvin et al., 2018; Ye and Yamamoto, 2019; Papadoulis et al., 2019; Virdi et al., 2019; Sinha et al., 2020). In general, Wiedemann characterizes the car-following behavior by look-ahead distance, look-back distance, and average standstill distance, while modified Wiedemann 99 also considers headway time (PTV, 2018).

Based on driving behaviors, various scenarios were developed to evaluate the impact of AVs on safety. The majority of the studies explored different AV MPRs as the main variable (Katrakazas et al., 2019; Bahram et al., 2014; Rahman et al., 2019; Arvin et al., 2018, 2019; Morando et al., 2018; Deluka Tibljaš et al., 2018; Ye and Yamamoto, 2019; Papadoulis et al., 2019; Qin and Wang, 2019; Arvin et al., 2020; Sinha et al., 2020). Depending on the study, each simulation scenario was run multiple times to obtain reliable outputs for evaluating traffic safety. Since simulations do not lead to any crash, near-miss events were used instead to assess safety.

Surrogate safety measures (SSMs) were used to determine the number of near-miss events and, consequently, the associated level of safety. The most commonly used SSMs in AV safety evaluation studies were time-to-collision (TTC) and post-encroachment time (PET) (Kockelman et al., 2016; Katrakazas et al., 2019; Bahram et al., 2014; Arvin et al., 2018, 2019; Morando et al., 2018; Deluka Tibljaš et al., 2018; Ye and Yamamoto, 2019; Papadoulis et al., 2019; Mousavi et al., 2020; Sinha et al., 2020). Acceleration rate and velocity difference (Ye and Yamamoto, 2019; Sinha et al., 2020), time-exposed time-to-collision (TET) (Bahram et al.), time-integrated time-to-collision (TIT) (Bahram et al., 2014; Qin and Wang, 2019; Zhang et al., 2015; Rahman et al., 2019), time-exposed rear-end crash risk index (TERCRI) (Zhang et al., 2015; Rahman et al., 2019), number of critical jerks (NCJ) (Rahman et al., 2019), and lane-change conflicts (Zhang et al., 2015) were the other types of SSMs used in these studies.

Most of the studies concluded that by increasing the AV MPR, the number of near-miss events decreased on road segments (Bahram et al., 2014; Morando et al., 2018; Ye and Yamamoto, 2019; Qin and Wang, 2019; Sinha et al., 2020), at intersections (Kockelman et al., 2016; Arvin et al., 2018, 2019; Morando et al., 2018; Rahman et al., 2019; Arvin et al., 2020; Mousavi et al., 2020), at priority intersections (Virdi et al., 2019), in bottlenecks, at on/off-ramps (Kockelman et al., 2016), and in roundabouts (Morando et al., 2018; Virdi et al., 2019). However, Deluka et al. (2018) indicated that an increase in the AV MPR in roundabouts led to an increase in the number of conflicts. Moreover, Kockelman et al. (2016) showed an increase in conflicts by increasing the AV MPR at intersections. On the other hand, other studies showed that low AV MPRs were associated with a higher number of conflicts compared to zero MPR, yet, the number of conflicts decreased at intersections (Arvin et al., 2018; Virdi et al., 2019) and diverse diamond interchange (DDI) intersections (Virdi et al., 2019) by increasing the MPR in the simulation environment. Katrakazas (2019) also proposed a method to enable AVs to determine their trajectories to enhance safety in emergency situations. Study results indicated that the proposed method is capable of improving safety.

4.3.4. *Driving simulators*

Probable challenges in human-vehicle interaction in the AV domain can take place in either the AV driver and AV interface stage (e.g., taking-over process) or the interaction between conventional vehicles and AVs (e.g., conventional vehicles entering the platoon of AVs). At different levels of automation, the AV driver needs to monitor or even intervene in the automation system to some extent in order to compensate for automation biases. On the other hand, AVs, at any MPR, will interact with conventional vehicles before they entirely dominate the future transportation system. In both cases, detailed knowledge of human driving behavior and reactions is necessary to evaluate AV safety. All the safety-related scenarios in reviewed studies could be categorized as (a) vehicle-human interaction (take-over situations in different driving states, such as drunk driving, drowsy driving, distracted driving, unplanned disengagement from the ADS, planned disengagement, etc.) (Strand et al., 2014; Kundinger et al., 2018; Berthelon and Gineyt, 2014; Gold et al., 2018; Happee et al., 2017; Blommer et al., 2015; Yun and Yang, 2020; Lee et al., 2020), or (b) vehicle-vehicle interaction (joining a conventional vehicle to a platoon of AVs) (Gouy et al., 2012; Lee et al., 2018). In both categories, a hazard scenario must be designed to determine the driver's performance in the evasive situation of interest. A hazard scenario is a situation that triggers the driver to make a maneuver and might be (a) a sudden blocked lane by another vehicle(s) or an obstacle (Gold et al., 2018; Happee et al., 2017; Blommer et al., 2015; Yun and Yang, 2020; Lee et al., 2020), a sudden drift toward the edge of the road (Desmond et al., 1998), or a deceleration failure (Strand et al., 2014); or (b) safety challenges faced during driving, such as entering a platoon environment (Gouy et al., 2012; Lee et al., 2018) or controlling the vehicle while drowsy or drunk (Kundinger et al., 2018; Berthelon and Gineyt, 2014). The simulator experiments included three aspects—participants, experimental variables, and safety measurements—that had to be designed before the main experiment.

Different characteristics of participants used in designing simulator experiments included the following: age (Berthelon and Gineyt, 2014; Gold et al., 2018; Happee et al., 2017; Blommer et al., 2015; Gouy et al., 2012; Strand et al., 2014; Lee et al., 2018; Kundinger et al., 2018; Desmond et al., 1998; Yun and Yang, 2020; Lee et al., 2020), gender (Happee et al., 2017; Gold et al., 2018; Blommer et al., 2015; Gouy et al., 2012; Strand et al., 2014; Lee et al., 2018; Kundinger et al., 2018; Berthelon and Gineyt, 2014; Desmond et al., 1998; Yun and Yang, 2020; Lee et al., 2020), annual mileage driven (Strand et al., 2014), driving experience (Strand et al., 2014; Gouy et al., 2012; Berthelon and Gineyt, 2014; Yun and Yang, 2020; Lee et al., 2020), previous experience with automated driving (Strand et al., 2014; Blommer et al., 2015), prior experience with a driving simulator (Gouy et al., 2012; Gold et al., 2018; Happee et al., 2017), and mental/physical health condition (Kundinger et al., 2018; Berthelon and Gineyt, 2014; Lee et al., 2020). Each experiment took place in a controlled ODD and was based on a predefined procedure. Predesigned factors, such as (a) traffic density (Gold et al., 2018; Happee et al., 2017; Blommer et al., 2015; Strand et al., 2014; Gouy et al., 2012; Lee et al., 2018; Kundinger et al., 2018; Berthelon and Gineyt, 2014), (b) MPR (Lee et al., 2018), (c) facility type (Gold et al., 2018; Happee et al., 2017; Blommer et al., 2015; Strand et al., 2014; Gouy et al., 2012; Lee et al., 2018; Kundinger et al., 2018; Berthelon and Gineyt, 2014; Yun and Yang, 2020), and (d) repetition of experiment (Happee et al., 2017; Gold et al., 2018; Strand et al., 2014; Gouy et al., 2012; Desmond et al., 1998; Yun and Yang, 2020) and controlled factors—including the facility geometry design characteristics (Gold et al., 2018; Happee et al., 2017; Blommer et al., 2015; Gouy et al., 2012; Lee et al., 2018; Berthelon and Gineyt, 2014) and speed (Gold et al., 2018; Happee et al., 2017; Blommer et al., 2015; Strand et al., 2014; Gouy et al., 2012; Lee et al., 2018; Kundinger et al., 2018; Berthelon and Gineyt, 2014; Desmond et al., 1998; Yun and Yang, 2020)—were common experimental characteristics found in simulator studies. Some studies conducted only one experiment per participant to avoid learning effect bias (Blommer et al., 2015; Kundinger et al., 2018; Lee et al., 2018); others repeated the experiment to extract the maximum information from the available resources and tried to mitigate the learning effect bias by incorporating it as a variable in the model. However, almost all studies conducted a trial run before the main experiment to familiarize the participants with the simulator environment.

A metric is required to measure AVs' performance and quantify the risks and benefits of AVs using simulator studies. To this end, SSMs were widely used as the response variable to quantify safety risks and benefits of AVs, namely average/maximum/minimum speed (Berthelon and Gineyt, 2014; Lee et al., 2020), time headway (Strand et al., 2014; Gouy et al., 2012), take-over time (TOT) (Gold et al., 2018), TTC (Gold et al., 2018; Happee et al., 2017; Strand et al., 2014; Lee et al., 2020), distance to collision (DTC) (Lee et al., 2020), time to lane change (TTL) (Yun and Yang, 2020), brake application (Gold et al., 2018), crash/crash probability (Gold et al., 2018; Berthelon and Gineyt, 2014), steering response time (Happee et al., 2017; Lee et al., 2018), response time (Blommer et al., 2015; Strand et al., 2014; Yun and Yang, 2020), percent of the time with eyes on the road (Blommer et al., 2015), clearance toward the obstacle (Happee et al., 2017), road clearance metric (Happee et al., 2017), steering magnitude (Lee et al., 2018), lateral/longitudinal control (e.g., longitudinal/lateral deceleration) (Desmond et al., 1998; Lee et al., 2020), standard deviation of lane position (SDLP) (Yun and Yang, 2020; Lee et al., 2020), steering wheel reversed (SWR) (Yun and Yang, 2020), Karolinska Sleepiness Scale (Kundinger et al., 2018), physical and perceptual fatigue (Desmond et al., 1998), skin conductance response time (SCR) (Yun and Yang, 2020), and average heart rate (AHR) (Yun and Yang, 2020). The point of modeling different SSMs relates to the difference in their ability to capture near-crash events and critical maneuvers.

Finally, the SSMs were used to (a) find contributing factors to safety risk and benefits of AVs in different settings (Gold et al., 2018; Happee et al., 2017; Blommer et al., 2015; Strand et al., 2014; Gouy et al., 2012; Lee et al., 2018; Berthelon and Gineyt, 2014; Yun and Yang, 2020; Lee et al., 2020), and (b) compare AV safety with conventional vehicle safety (Happee et al., 2017; Kundinger et al., 2018; Desmond et al., 1998). Linear regression (Gold et al., 2018), logistic regression (Lee et al., 2018), univariate/multivariate analysis of variance (ANOVA) (Blommer et al., 2015; Strand et al., 2014; Gouy et al., 2012; Lee et al., 2018; Berthelon and Gineyt, 2014; Yun and Yang, 2020; Lee et al., 2020), Fisher's exact test (Strand et al., 2014), analysis of covariance (ANCOVA) (Strand et al., 2014), and Cochran's Q test (Strand et al., 2014) were used to identify significant variables that influenced AV safety. Besides the participant characteristics and experiment characteristics (or elements) mentioned before, other variables—such as time budget (Gold et al., 2018; Happee et al., 2017), lanes driven (Gold et al., 2018; Happee et al., 2017), type of secondary tasks (Gold et al., 2018; Happee et al., 2017; Blommer et al., 2015; Lee et al., 2020), automation level (Strand et al., 2014), disengagement scenarios (planned/unplanned) (Yun and Yang, 2020), types of take-over warnings (Yun and Yang, 2020), extent of hazard scenario and challenges (e.g., moderate/severe/complete deceleration failure, or different time headway within the platoon) (Strand et al., 2014; Gouy et al., 2012), platoon size (Lee et al., 2018), and alcohol concentration (Berthelon and Gineyt, 2014)—were considered. Results showed that take-over scenarios, traffic density, experiment repetition, and defined time budget were highly influential factors affecting SSMs (Gold et al., 2018). In addition, scheduled disengagement (Blommer et al., 2015), lower automation levels, lower extent of hazard scenarios (Strand et al., 2014), engaging in non-driving-related tasks with less cognitive load (Lee et al., 2020), and use of multimodal take-over warning systems (Yun and Yang, 2020) led to better performance of drivers during the take-over situation. Drunk

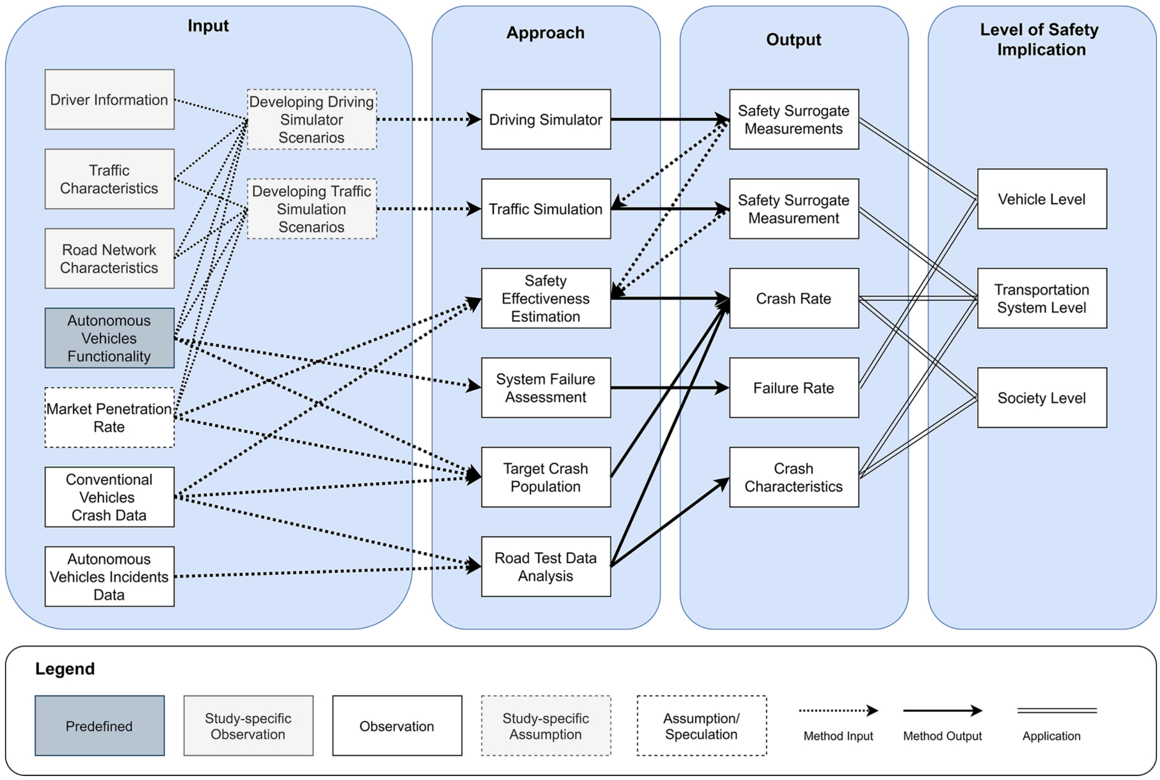

Fig. 4. Summary of the inputs, outputs, and potential application of AV safety quantification methodologies.

driving affected the longitudinal and lateral control of the vehicle and driver reaction to evasive maneuver, especially in lower automation levels (Berthelon and Gineyt, 2014). Moreover, in the platoon environment, the higher MPR (Lee et al., 2018) and lower time headway of AVs resulted in more aggressive driving behavior from conventional vehicles joining the platoon. To compare conventional vehicles and AVs in terms of safety risks and benefits, researchers mostly used ANOVA (Kundinger et al., 2018; Desmond et al., 1998) and Fisher's exact test (Happee et al., 2017). Results showed that automated driving would negatively affect a take-over scenario in response to a risk while the vehicle is disengaged from the ADS (Happee et al., 2017; Desmond et al., 1998) and increase driver drowsiness (Kundinger et al., 2018) compared to manual driving.

More details on the reviewed driving simulator studies can be found in Table A4.

4.3.5. System failure risk assessment

System operation failure is one probable risk that AVs encounter (Koopman and Wagner, 2016). Malfunctioning sensors in detecting objects (pedestrians, bikes, vehicles, obstacles, etc.), misinterpretation of data, and poorly executed responses can jeopardize AVs' reliability and have serious safety consequences in an automated environment (Bila et al., 2017). The failure rate of each component of AVs was synthesized by Bhavsar et al. (2017). To this end, each component of the ADS and ADAS was examined individually, and the failure rate was determined for each component based on the evidence from the existing literature. The researchers developed a hierarchical model to synthesize AV failure risks associated with the vehicle and infrastructure. The communication system's failure risks, hardware system (sensor and integration platform failure), and software system were ranked the highest, with 9.5 %, 4.2 %, and 1.0 % failure probability, respectively. The failure probability of an AV involved in a crash with a non-AV was also calculated by multiplying the risk of failure of AVs and the crash probability of conventional vehicles.

4.3.6. AV safety effectiveness

AV safety effectiveness can be defined using AV SSMs and crash rates. For example, the safety effectiveness of AVs can be estimated as follows:

$$\text{Safety Effectiveness} = 1 - \frac{\text{AVs' crash rate}}{\text{Conventional vehicles' crash rate}} \tag{1}$$

However, decisions about AV safety effectiveness or AV safety validity cannot be based on the results of a single study because results typically vary from one study to the next (see Sections 4.3.3 and 4.3.4 for more details). Rather, a mechanism is needed to synthesize data across studies. Wang et al. (2020) synthesized the results of previous simulation and field experiments that estimated safety effectiveness by performing a meta-analysis of 89 studies. They estimated the safety effectiveness of nine ADASs, in descending order: intersection movement assists, pedestrian collision and mitigate (PCAM), lane-departure warning (LDW), lane-change warning (LCW), forward collision warning (FCW), electronic stability control (ESC), blind-spot warning, automated emergency braking (AEB), and adaptive cruise control (ACC).

Wang et al. (2020) further designed a target crash population study to implement the estimated ADASs' safety effectiveness rates and quantify the potential impacts of CVs and AVs on different crash types. The results of their analyses showed that 3.4 million crashes could be prevented between 2012–2016; this figure represented a significant reduction in crashes in India (54.24 %), Australia (51.55 %), the United States (48.07 %), New Zealand (45.36 %), Canada (44.71 %), and the UK (40.95 %).

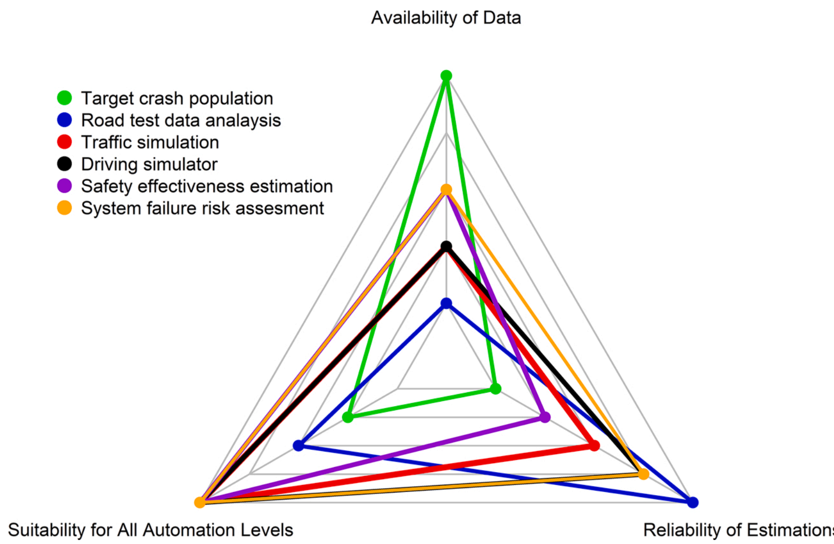

Fig. 5. Trade-offs between relative availability of data, suitability for evaluating levels of automation, and reliability of estimations.

5. Discussion

5.1. Evaluating AV safety quantification approaches

Six AV safety quantification approaches were identified in this review. The identified approaches were investigated in terms of their input, output, and level of safety implications they address. Theii identified approaches' inputs included predefined information on AVs' functionality, conventional vehicle crashes, AV road test crashes and errors, study-specific observations, and assumptions and speculations regarding AV implementation. This review showed that the *target crash population* approach can be used to estimate the number of preventable crashes for evaluating AV safety at the transportation system and society-level. *Road test data analysis* was mainly used to investigate the characteristics of AV crashes. In addition, AV safety was evaluated using this approach by comparing system failure and crash frequencies of AVs with conventional vehicles. *Road test data analysis* approach can be used for evaluating AV safety at the transportation system and society levels. *Driving simulator* and *traffic simulation* studies can be used for evaluating AV safety in terms of SSMs under different implementation scenarios. While driving simulators investigate AV safety and its potential operational challenges (e.g., disengagement from ADS) at the vehicle level, traffic simulation studies consider AVs' performance and their interactions with other vehicles in a fleet at the transportation system level. The driving simulator studies also unveil some information regarding the user's behavior, such as car-following behavior, that is later used as an input in the traffic simulation studies. AVs' safety effectiveness is estimated as a result of synthesizing the *drivingsimulator* and traffic simulation studies and statistically analyzing their outputs. Although *safety effectivenessassesement approach* was was developed to evaluated ADASs safety, this method can evaluate the safety of ADS as well. The estimated safety effectiveness can provide insights into AV safety at both the transportation system and society levels. The *system failure assessment* approach evaluates AV safety at the vehicle level based on the system components' failure rates. Fig. 4 summarizes the inputs, outputs, and potential application of the identified approaches.

The AV safety quantification approaches vary in terms of (a) availability of input data, (b) suitability for evaluating different levels of automation, and (c) reliability of estimations. Fig. 5 shows the trade-off between AV safety quantification methods based on their relative capabilities in terms of these three criteria[5] . The road test data analysis approach is able to evaluate the safety of higher levels of automation with minimal uncertainty, however, with the cost of extensive and reliable AV crash data. Since higher levels of automation have not yet been tested on roads, this approach can not provide insights into the safety implications of Levels 4 and 5 of automation. The target crash population approach needs relatively fewer input data and can estimate the safety benefits of lower automation levels (given the uncertainties in safety implications of higher levels of automation); nevertheless, considerable uncertainty exists in the estimates. Since driving simulator and traffic simulation approaches are mainly based on assumptions and speculations regarding AV implementation and performance, AV safety evaluations using these approaches would carry considerable uncertainties. However, they can be used to evaluate all levels of automation. Safety effectiveness estimations inherit the limitation of driving simulator and traffic simulation studies. Still, synthesizing driving simulator and traffic simulation results using statistical methods would reduce the reliability of safety effectiveness estimations results. Although system failure risk assessment can evaluate different levels of automation, its reliability significantly depends on the accuracy and availability of the system components' failure rates. Accessing verified information from manufacturers is challenging, questioning the reliability of AV safety evaluations using the system failure risk assessment approach.

5.2. AV safety validation challenges

We identified four challenges in AV safety evaluation:

1 limitations in the existing quantification methodologies,
2 uncertainties in AV implementations and their impacts on AV safety,
3 potential riskier behaviors of AV passengers as well as other road users, and
4 new safety issues related to AV implementations.

5.2.1. Limitation in the existing quantification methodologies

Certain limitations in existing AV safety quantification

[5] This qualitative analysis is based on a comprehensive review of the literature and a detailed evaluation of each approach's capabilities rather than quantitative analyses.

methodologies can jeopardize the safety evaluation of this new technology. Target crash population studies did not account for the risky scenarios that AVs might cause (e.g., disengagement or system failure) and totally disregarded probable new crashes. Also, the mixed traffic safety issues (interaction of AVs and conventional vehicles) and the way an AV driver reacts to hazards were not considered in the target crash population approach. Thus, this method is expected to represent a theoretical upper bound (or optimistic estimations) of AVs' potential safety benefits, as opposed to their expected actual impacts.

Driving simulator studies were designed to evaluate AVs' potential safety challenges. Traffic simulation studies can also be used to account for both AV and conventional vehicles' driving behaviors and mixed traffic safety issues. Nevertheless, driving simulator and traffic simulation studies have certain limitations. They are subject to biases from a variety of sources, such as participants (e.g., driving behavior and fatigue), simulator and simulation environment (e.g., physical fidelity and functional fidelity), and SSM selection. Employing different SSMs to evaluate AV safety in driving simulators and traffic simulations makes it almost impossible to directly compare the literature, although a general comparison in terms of the overall safety trend of AVs could be conducted using SSMs. Another challenge in driving simulator and traffic simulation studies is the limitations in calibration and validation of experiment results since AV road test data—which is the ground truth data—are limited. Because safety effectiveness estimations are based on the results of *driving simulator* and *traffic simulation* studies, they carry remarkable uncertainty as well.

The system failure risk assessment approach was used to quantify the crash risks associate with the failure probability of ADASs/ADSs technologies. However, looking at the system failure rates individually can result in overestimating AV failures, given that other components can compensate for the failure of the deficient components. For example, in the event of an AV radar malfunction, the camera vision can help to activate the collision prevention system and avoid a collision. Moreover, system failure risk assessment relies on system failure rates from private companies. Collecting accurate system failure rates is challenging since this information should be collected from the manufacturer and might be underreported.

The road test data analysis was purported to be the most reliable method for evaluating AV safety. However, existing road tests are limited, and more data are required to draw reliable conclusions on AV safety. Accounting for AVs' safety implications at different market penetration levels is another limitation of road test data, given that higher market penetrations cannot be expected in the near future. Also, a decisive comparison between AV and conventional vehicle crashes is subject to accurate and reliable information about the AV testing environment (ODD and fallback-ready user) as well as conventional vehicle crashes (non-reportable crashes). Increases in AV road test data analysis studies in recent years (Fig. 3) can be associated with larger and more reliable road test datasets. Nevertheless, validating AV safety with road test data has been criticized because they expose road users to road hazards (Kalra, 2017).

5.2.2. Uncertainties in AV implementations and their impacts on AV safety

AV impacts on transportation go beyond safety impacts. By offering a safer, cheaper, and more comfortable travel option to individuals, including those with disabilities, AVs may induce additional transportation demand and encourage longer trips. AVs can also encourage shifting from public transit and active transportation (walking and cycling) to private cars (Fagnant and Kockelman, 2015). Transportation and land use are tightly linked in urban areas (Rodrigue et al., 2016); consequently, changes in transportation can ultimately result in urban sprawl (i.e., migrating to areas with lower density and consequently spreading a city's boundaries). Urban sprawl increases total VMT (Childress et al., 2015) and negatively influences accessibility in an urban area (Milakis et al., 2017). In addition, the uncertainties in AVs' intention of use and disproportionate ownerships will affect transportation systems, travel patterns, and urban design.

Changes in VMT and modal shifts along with the level of market penetration are factors that can impact traffic safety at the transportation system and society levels. Therefore, these changes need to be considered in AV safety evaluations to attain accurate insights into AV safety implications. Full-chain assessment of AV safety—including AV adoption modeling, urban growth modeling, travel demand modeling, and safety analysis—can be a potential avenue to address the uncertainties associated with AV implementations in the transportation system, travel patterns, and urban design.

Table 1
Suggested Future Studies.

Research area	Study subtopics	Level of safety impact
Address the limitations of existing AV safety quantification methods	Consider mix-traffic issues, system failure, and the risk associated with fallback-ready user reaction at the time of AV disengagement from ADS in target crash population approach	Transportation system + Society
	Evaluate traffic simulation and driving simulator results using AV road test data	Vehicle + Transportation system
	Collect and analyze reliable system failure rates	Vehicle
	Perform reliable statistical analysis on a larger AV road test dataset	Society
Perform full-chain assessment of AVs' safety implications	Account for AV market penetration and its influence on urban areas, trip patterns, and transportation systems	Society
Investigate the potential risky behavior of AV users	Examine the risk homeostasis hypothesis	Transportation system + Society
Study the emerging safety issues associated with AV implementations	Address AVs' cybersecurity issues	Vehicle
	Preprogram AVs to follow the best course of action during unavoidable crashes	Vehicle

5.2.3. The potential riskier behavior of AV passengers and other road users

Changes in AV and conventional vehicle users' behavior need to be considered in AV safety evaluation. Based on research conducted by AAA Foundation, a substantial minority of early adopters of braking assistance systems reported having had a crash or near-crash while driving a vehicle without this technology, supposedly because of incorrect expectations from the unequipped vehicle to provide warnings (Jenness et al., 2007). Gouy et al. (2012) ran a driving simulator experiment and showed that the conventional vehicles would be driven more aggressively if joining a platoon of AVs. The riskier behavior of drivers during interaction with AVs can be explained by the risk homeostasis hypothesis (Wilde, 1998). Based on this hypothesis, every person has an acceptable amount of risk that they find tolerable. According to Wilde (1998), "If the perceived level of risk in one part of a person's life changes, they will compensate by either reducing or increasing the risks they take–all in order to maintain an equilibrium of perceived risk."

5.2.4. New safety issues related to AV implementations

Cybersecurity is another potential concern related to AV operation because hacking and vehicle misuse can result in catastrophic crashes (Lee, 2017; Taeihagh and Lim, 2018; Cui et al., 2019). A car hacking experiment conducted by (Jafarnejad et al., 2015) demonstrated that electric vehicles could be easily controlled remotely by mobile applications that forced the vehicles to go forward or backward, limited their speed, etc. In addition, the ethical dilemma associated with AV reactions during unavoidable situations introduces another challenge in AV operation (Goodall, 2014; Awad et al., 2018) that requires further attention. Although AVs' ethical issues cannot directly impact AV safety evaluation, they concern the liability of AVs in crashes, which requires judiciary attention.

5.3. Limitations of review methodology

This review has certain limitations. First, this study focuses on AV safety quantification methods; therefore, we did not include the literature that evaluated ADAS safety implications or proposed frameworks and conceptual models for AV safety evaluation rather than quantifying the impacts. Both ADAS safety evaluation methods and proposed frameworks for AV safety evaluation might have the potential to address some of the limitations of the existing quantification methods. Second, we examined the AV safety evaluation methodologies qualitatively and relatively. Future research provide a more accurate comparison between the methods by conducting quantitative analyses. Third, we focused on methodologies that quantified AVs' substantive safety rather than the nominal safety and perceived safety. However, the safety of vehicles should be validated based on three definitions of safety. For example, even though AV safety can be comparable to that of conventional vehicles, users' degrading perceptions of AV safety may hinder the adoption of this new technology. Future research is required to review the literature and examine the methodologies used for evaluating AV nominal safety and perceived safety for more accurate evaluation and understanding of AV safety.

5.4. Gaps in the literature and future studies

In light of this comprehensive review of the literature, we specify four areas of research that would add to the AV safety evaluations. First, identified approaches have some shortcomings and limitations that need to be addressed. Mix-traffic issues, system failure, and fallback-ready user errors have not been considered in target crash population studies. Accounting for these factors can potentially lead to more accurate estimations of AV safety implications. Driving simulator and traffic simulation studies can benefit from ground truth data (e.g., AV road test data) to verify their assesments and study findings . The AV system failure risk assesment should be revisited using more reliable data on AV system failure rates. Running statistical analyses on a large amount of AV road test data in future studies can provide more reliable conclusions regarding AV safety. Second, AV safety studies do not generally account for uncertainties in AV implementations—i.e., AV market penetration and its role in urban design, trip patterns, and transportation systems. Future research can address this limitation in order to assess the safety impact of AVs at the society level. Third, since we expect riskier behavior of AV passengers as well as other road users when interacting with AVs, further investigations of the risk homeostasis hypothesis are needed to measure the potential safety implications. Fourth, emerging safety issues related to AVs, including cybersecurity and AVs' reactions during unavoidable crashes, should be studied further. In addition, future studies should address the limitations of this review, namely (a) defining a broader review question, (b) evaluating the identified methodology quantitatively, and (c) investigating AVs' nominal and perceived safety implications.

Table 1 shows the list of future research directions, the study topics, and the level of safety implications these studies can address.

6. Summary and conclusions

This study implemented a scoping review methodology to synthesize the AV safety quantification approaches. We evaluated the identified quantification approaches and uncovered the gaps and challenges in AV safety evaluation. The AV safety quantification methods were categorized into six groups: target crash population, road test data analysis, traffic simulation, driving simulator, safety effectiveness estimation, and system failure assessment. We compared and relatively evaluated the identified approaches. Results can be used as a guideline for future research to select the appropriate AV safety evaluation method based on the study objective and limitations. This review showed that existing methodologies for AV safety evaluation carry certain shortcomings, and further investigations are required for a reliable evaluation of AV safety. In addition, we discussed major challenges in AV safety evaluations—uncertainties in AV implementations and their impacts on AV safety, potential riskier behavior of AV passengers as well as other road users, and emerging safety issues related to AV implementations. Future research is required to better understand and evaluate AV safety while addressing the gaps in the existing methods and the challenges in AV safety evaluation.

Funding

The work conducted in this study was partly funded by the Robert Wood Johnson Foundation, Grant ID 76489, Quantifying the Harms and Benefits of Connected and Automated Vehicle Technologies on Public Health and Equity, and the University Transportation Center Safety through Disruption (Safe-D) Project, ID TTI-Student-006. This research was partly funded by the A.P. and Florence Wiley Faculty Fellow provided by the Texas A&M University College of Engineering and the Zachry Department of Civil and Environmental Engineering to Dr. Dominique Lord.

CRediT authorship contribution statement

Soheil Sohrabi: Conceptualization, Methodology, Investigation, Visualization, Writing - original draft, Writing - review & editing, Supervision. **Ali Khodadadi:** Investigation, Writing - original draft. **Seyedeh Maryam Mousavi:** Investigation, Writing - original draft. **Bahar Dadashova:** Validation, Writing - review & editing, Supervision, Funding acquisition. **Dominique Lord:** Validation, Writing - review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that there is no conflict of interest.

Appendix A

Table A1
Summary of target crash population studies.

Author	ADS/ADAS	Target Crashes	ODD						Significant Results
			Road Type	Road Surface Condition	Weather Condition	Lighting Condition	Speed	Effectiveness	
Kusano and Gabler (2014)	FCW, PCAM, LDW	18 pre-crash scenarios	NA	NA	NA	NA	NA	NA	• Safety systems can mitigate 20 % and 26 % of serious injury and fatal crashes, respectively.
Lee and Kockelman (2016)	CACC, LKA, ESC	37 pre-crash scenarios	NA	NA	NA	NA	NA	NA	• Reduction of crash costs by 126 million annually. • Reduction of functional human-years lost by nearly 2 million (per year).
Detwiller and Gabler (2017)	AEB	Transportation-related pedestrian crashes	Urban area	NA	NA	NA	✓	100 %	• Employing two different sets of rules resulted in a reduction or mitigation of 40 % and 95 % of crashes, respectively.
Yanagisawa and Rau (2017)	Level 2 to Level 4 ADS	37 pre-crash scenarios	Intersection, ramp, highway, work zone	✓	✓	✓	✓	100 %	• L2 to L4 can address 35–250 billion dollars in comprehensive costs and 1100–11,000 fatal crashes annually.
Hendrickson and Harper (2018)	BSM, LDW, and FCW	Lane-change crashes, lane-departure crashes, and rear-end collision	NA	NA	NA	NA	✓	100 %	• All technologies together can mitigate 1.3 million crashes annually, including 133,000 injury and 10,000 fatal crashes.
Lubbe et al. (2018)	AEB, LCW, LKA*, ESC	30 pre-crash scenarios	NA	✓	✓	NA	✓	100 %	• Fatality reduction from 12 to 13% (using passive safety systems only) to 45–63% (using advanced ADAS and assuming cautious driving). • Reduction of vulnerable road user fatalities by 33–41 %.
Agriesti et al. (2019)	Level 3 ADS	Distracted driving, insufficient safety distance, speeding, skidding, road departure	Highways	✓	✓	✓	✓	100 %	• 66 % of crashes involving AVs and 6.6 % of crashes involving conventional vehicles (considering 10 % MPR) could be avoided.
Combs et al. (2019)	Pedestrian detection	Transportation-related pedestrian crashes	Urban/rural, intersection/not intersection, freeway/not freeway	NA	NA	NA	✓	100 % except for adverse condition (20 %)	• Different combinations of sensors can lead to a 30%–90% reduction of fatal pedestrian crashes.

Note: NA = Not Applicable.
* LKA: Lane Keeping Assistant.

Table A2
Summary of AV road test data analysis studies.

Study	Type of Analysis	Data Source	Approach	Significant Results
Schoettle and Sivak (2015)	Frequency; characteristics of the incident	CA DMV (2014–2015) and Google self-report (2012–2014) (11 crashes)	• Comparing AV and conventional cars' crash rates after adjusting for underreporting. • Descriptive analysis of crash characteristics (vehicle motion at the time of the crash, crash type, and crash severity) and comparison to conventional vehicles.	• Most of the crashes happened while the AV's speed was less than 5 mph. • The rate of rear-end crashes in AVs is higher than conventional cars. • The severity of AV crashes is lower than conventional cars. • The rate of AV crashes is 8 times higher than conventional vehicles.
Kalra and Paddok (2016)	Reliability	Accident rates in the US (2013)	• Estimating number of failure-free miles AVs should drive to reach conventional cars' failure rate using survival analysis. • Estimating the required total number of miles driven to evaluate AVs' failure rate. • Estimating the total number of miles AVs need to drive to demonstrate their failure rate is statistically lower than conventional cars.	• AVs need to drive 1.6 million miles failure-free to be as safe as conventional cars. • AVs need to drive 51 and 61 million miles to be able to test their failure rate and statistically examine if their failure rate is lower than conventional cars, respectively.
(Teoh and Kidd, 2017)	Frequency	CA DMV (2009–2015)*	• Comparing AV (police-reportable) crash rate to conventional cars' crash rate.	• Google self-driving cars are safer than conventional human-driven passenger vehicles (2.19 vs. 6.06 per million VMT).
Favarò et al. (2017)	Frequency; characteristics of the incident; reliability	CA DMV (September 2014 to March 2017) (5326 disengagements and 26 accidents)	• Descriptive analyses of crashes by collision type, location, and manufacturer. • Comparing AV crash rate and number of miles driving until an accident to conventional cars' crash rate and number of miles driving until an accident.	• The rate of crashes was lower for AVs than conventional cars, and AVs will drive longer before an accident (~42,000 vs. 500,000 miles). • Most of the AV crashes happened at intersections. • Rear-end crashes are higher for AVs than for conventional cars.
Matysiak and Razin (2018)	Frequency	CA DMV (2015–2017)	• Comparing AVs' disengagement data to injury and fatal crashes in Europe and US.	• AVs' crash rate is 2–3 times higher than conventional cars. • AVs should drive more than 442 million km fatal-free to be considered safer than human-driven cars.
Banerjee et al. (2018)	Frequency; characteristics of the incident; reliability	CA DMV (September 2015 to November 2017)	• Comparing AVs' disengagement rate to conventional cars' accident rates. • Analyzing the cause of disengagement from manufacturer report (after excluding unknown causes). • Comparing to other safety-critical autonomous systems.	• Conventional vehicles were 15–4000 times less likely (depending on the AV manufacturer) than AVs to have an accident. • 64 % of disengagements were the result of problems in, or untimely decisions made by, the machine learning system. • In terms of reliability per mission, AVs are 4.22 times worse than airplanes and 2.5 times better than surgical robots.
Wang and Li (2019)	Characteristics of the incident	CA DMV (2017–2018) (107 crashes) NTSB (2017–2018) (6 crashes)	• Investigating the factors contributing to AV crash collision types and severity using logistic regression and decision tree.	• The highway and automated driving mode were identified as the location where severe injuries are likely to happen due to high travel speed. • Collision types of AV-related crashes depend upon the driving mode, location, and whether crashes are associated with yielding to pedestrians/cyclists. • Both ordinal logistic regression and the decision tree models show consistent results.
Xu et al. (2019)	Characteristics of the incident	CA DMV (January 2015 and June 2018) (72 crashes)	• Descriptive statistics analysis to investigate the characteristics of AV-involved crashes. • Binary logistic regressions were developed to investigate the factors contributing to the collision type and severity of AV-involved crashes.	• AV driving mode, collision location, roadside parking, rear-end collision, and one-way road are the main factors contributing to the severity level of AV-involved crashes. • AV driving mode, AV stopped or not, AV turning or not, normal vehicle turning or not, and normal vehicle overtaking or not are the factors affecting the collision type of AV-involved crashes.
Li and Zhai (2019)	Reliability	The accident rate on China highways (2008–2015)	• Finding the minimum fault-free distance of AVs to be as safe as conventional cars by inferring the overall distribution from the sample distribution and calculating how much sample size is needed at minimum.	• With a 95 % confidence interval, AVs need to drive fault-free for ~226 million km and should be tested for 115,972 million km to be considered as safe as conventional cars.
Petrović et al. (2020)	Characteristics of the incident	CA DMV (2015–2017) (53 accidents)	• Analyzing the type of collision frequencies using descriptive statistics of crash data.	• The rear-end type of collision is statistically more significantly frequent in traffic accidents with AVs.
Boggs et al. (2020)	Characteristics of the incident	CA DMV (2014–2018) (113 crashes)	• Frequentist and Bayesian binary logit model to examine the factors contributing to the AV crashes.	• Speed of conventional vehicle, missing speed, on-street parking, speed limit, driving through arterial and collector, and intersections were positively associated with AV crash assurance. • The number of lanes marked with a centerline and clear weather conditions increase the risk of crashes.
Das et al. (2020)	Characteristics of the incident	CA DMV (2014–2019) (151 crashes)	• Bayesian latent class model to classify AV crashes and to examine the factors. contributing to each class of crashes. • Text mining of AV crash narratives.	• Six classes of AV crashes were identified and associated with turning, multivehicle collisions, dark lighting conditions with streetlights, and sideswipe. • More detailed collision narratives are required to draw reliable conclusions.

* Only Google self-driving car crashes.

Table A3
Summary of traffic simulation studies.

Authors	Simulation Information					Driving Behavior Model		SSM	Results
	Facility Type	Length	Software	Technology	MPR	Conventional Vehicle	AV		
Bahram et al. (2014)	Four-lane highway	6000 m	PELOPS	Highly automated vehicles (HAVs)	0%, 50 %, and 100 %		The model of HAV controller developed in Simulink; the model is coupled via Xface2 to the interface in PELOPS	• TTC • TET (lower values represent safer situations) • TIT (lower values are associated with higher level of safety)	• TTC =3.0 s: 1440, 729, and 16 conflicts for MPRs of 0%, 50%, and 100%, respectively. • At 50 % MPR of HAV, the critical situation < 1.5 s increased remarkably compared to the base scenario. • MPRs of 0%, 50 %, and 100 % are associated with the TET of 144.1 s to 72.9 s and 1.6 s, respectively. • By increasing the MPR from 0% to 50 % and 100 %, the TIT changed from 66 to 76.29 and 1.10 s^2, respectively. • MPR of 50 % is not as safe of the other cases since AVs tend to follow other vehicles closely.
Zhang et al. (2015)	Four-lane freeway	7 km	VISSIM	CAV	0%, 10 %, 20 %, and 30 %	Wiedemann 99	Car-following and lateral lane-change decisions coded in C++	• TET • TIT • TERCRI • LCC	• Compared to the base scenario: • Providing 1 or 2 exclusive lanes led from −1.8 % to −87.1 % and −2.1 % to −85.3 % of lateral conflicts. • Installing 1 or 2 exclusive lanes resulted in +42.4 % to −52.90 % and +45.7 % to −55.2 % of longitudinal risk. • Only MPRs of 10 % and demands < 6000 veh/h providing exclusive lanes had mainly adverse effects on longitudinal conflicts ranging from 1.8 to −40.4, but for other scenarios with different MPRs and traffic demands, the overall safety improved.
Kockelman et al. (2016)	Intersection Freeway on/off-ramp	NA	VISSIM	AV	25 %, 50 %, 75 %, and 100 %	NA	NA	TTC	• Bottleneck: 40–88 % reduction in the number of conflicts by increasing the AV MPR from 0%–100%. • 4-leg intersection: 4% reduction in the number of conflicts by increasing the AV MPR from 0%–100%. • 77 % and 31 % reduction in the number of conflicts for two other intersections. • 17 % increase for another intersection. • Freeway on-ramps/off-ramps: 49 % reduction in the conflicts by increasing the MPR from 0% to 100.
Deluka et al. (2018)	Roundabout	NA	VISSIM	AV	0%, 10 %, 25 %, and 50 %	Wiedemann 74	Calibrated Wiedemann 74	TTC and PET	• By increasing the AV MPR from 0% to 50 %: • Omisalj roundabout: number of conflicts increased from 0 to 45; the majority of them were rear-end conflicts. • Malinska roundabout: the conflicts increased from 2 to 5, with all the conflicts being rear-end.
Morando et al. (2018)	Signalized intersection Roundabout	NA	VISSIM	AV Level 4	0%, 25 %, 50 %, 75 %, and 100 %	Wiedemann 99 car-following model with default parameters	Modified Wiedemann 99	TTC	• Intersection: AVs reduced the number of conflicts by 20%–65%, with an AV MPR of between 50% and 100%.

(*continued on next page*)

Table A3 (*continued*)

Authors	Simulation Information					Driving Behavior Model		SSM	Results
	Facility Type	Length	Software	Technology	MPR	Conventional Vehicle	AV		
									• Roundabout: the number of conflicts was reduced by 29%–64% with 100% AV penetration rate.
Arvin et al. (2018)	Intersection	NA	SUMO	AV Levels 3 and 5	0%, 7%, 15 %, 40 %, 60 %, 80 %, and 100 % (for MPR 100, different combinations of AV Level 3 and AV Level 5 were used)	Wiedemann 74	Modified Wiedemann 74	TTC	• Cases with human-driven vehicles, Level 3 and Level 5 AVs: the average crashes decreased from 9 to 0 by increasing the MPR from 0%–100%. • Cases with AV Level 5 and human-driven vehicles: at low AV MPR (below 40 %), the number of crashes increased from 9 to 10. • Cases with AV Level 5 and human-driven vehicles: by increasing the AV MPR (over 40 %), the number of crashes reduced from 10 to 0.
Papadoulis et al. (2019)	Three-lane motorway section	4.27 km	PTV VISSIM 9.0 and API	CAV	0%, 25 %, 50 %, 75 %, and 100 %	Wiedmann 99	External CAV driver model API written in C++	• TTC • PET	• Reduction in conflicts by 12–47 %, 50–80 %, 82–92 %, and 90–94 % for MPRs of 25 %, 50 %, 75 %, and 100 %, respectively.
Arvin et al. (2019)	Intersection	NA	VENTOS	HAVs and low-level AVs (LAVs)	Various combinations of conventional vehicles, LAVs, and HAVs	ACC model	Wiedemann	• TTC • Driving volatility	• For AV MPR of 0%, an average of 9.43 conflicts was observed. • At AV MPR of 100 %, there was a 90.1 % improvement compared to the baseline. • Where all the vehicles were HAVs: the intersection became conflict-free. • By increasing the MPR of LAVs and HAVs, the volatility decreased from 8.5 to 5.5 for acceleration. • By increasing the MPR of LAVs and HAVs, the speed volatility decreased from 6.9 to 3.8.
Katrakazas et al. (2019)	A section of highway	4.52 km	VISSIM	AV	NA	Wiedemann 99	NA	• TTC	• The artificial and the real-world datasets indicated that: • If the network-level, real-time collision risk indicates a situation as conflict-prone traffic, the probability of detecting if a vehicle poses a threat to an AV increases by 10 %. • When traffic conditions were marked as safe, the prediction did not improve the probability of a road user being a threat for the ego-vehicle. • By using disaggregated traffic data (i.e., 30 s), the probability of a traffic participant posing a threat to the ego-vehicle was enhanced by about 6%. • The proposed method allows AVs to change their trajectory, reduce their speeds, or even prompt a passenger to take the controls to ensure safety even when other sensor systems fail since network-level predictions utilize data at a higher temporal interval than the sampling frequency.

(*continued on next page*)

Table A3 (*continued*)

Authors	Simulation Information					Driving Behavior Model		SSM	Results
	Facility Type	Length	Software	Technology	MPR	Conventional Vehicle	AV		
Virdi et al. (2019)	Intersection	NA	VISSIM	CAV	0% to 100 % (10 % incremental)	Wiedemann 74 and Wiedemann 99	Virdi CAV control protocol algorithm	• TTC • PET	• The first 20 % MPR of CAVs resulted in: • +22 % change in conflicts at signalized intersections. • 87 % reduction in conflicts at priority intersections. • −62 % change in conflicts at roundabouts. • 33 % increase in conflicts at DDI intersections. • At high CAV MPR, a global reduction in conflicts occurred such that the 90 % CAV MPR was accompanied by: • −48 % change in conflicts at signalized intersections. • 100 % reduction in near-miss events at priority intersections. • −98 % change in near-crash events at roundabouts. • 81 % reduction in conflicts at DDI intersection.
Rahman et al. (2019)	Arterial segment Intersection	3.8 miles	VISSIM	CV and CV lower-level automation (CVLLA) (two automated features such as automated braking and lane-keeping assistance)	0%, 40 %, 60 %, 80 %, and 100 %	Wiedemann	C++ programming	• TTC • TET • TIT • TERCRI • LCC • NCJ	• Segment: by increasing the MPR from 0% to 100 %: • TET decreases from approximately 1750 to 1450 and 1370 for CV and CVLLA, respectively. • TIT decreases from 445 to 345 and 310 for CV and CVLLA, respectively. • TERCRI reduces from 390 to 308 and 265 for CV and CVLLA, respectively. • LCC decreases from 520 to 455 and 405 for CV and CVLLA, respectively. • Intersection: for different evaluated values of TTC and PET thresholds: • Total number of conflicts were decreased by 21–24 % for CV technologies compared to base scenario. • Total number of conflicts were reduced by 31–34 % for CVLLA compared with that of base condition.
Ye and Yamamoto (2019)	Two-lane road segment	10 km	NA	CAV	10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, and 90 %	User-defined	User-defined	• TTC • Acceleration rate • Velocity difference	• Reduction in the number of dangerous situations by increasing the MPR depends on traffic density and TTC. • By increasing the MPR from 0% to 100 %, the reduction in the dangerous situations falls within 0% and 97 %.
Qin and Wang (2019)	Freeway	20 km	MATLAB	CAV	Different MPRs	NA	NA	• TET • TIT	• Average reduction of 75%–95% depending on the number of feedback links by increasing the CAV MPR. • By increasing the feedback links from 1 to 2, average reduction in collision risks changes from 75%–95%. • There is not a significant reduction in the number of conflicts between 2, 3, and 4 links.

(*continued on next page*)

Table A3 (*continued*)

Authors	Simulation Information					Driving Behavior Model		SSM	Results
	Facility Type	Length	Software	Technology	MPR	Conventional Vehicle	AV		
Mousavi et al. (2020)	Urban unsignalized intersections	NA	VISSIM	AV	0% and 100 %	Wiedemann 74	Modified Wiedemann 74	• TTC	• Overall, regardless of the traffic LOS, AVs are capable of decreasing the total number of conflicts by 3.16. • The higher the traffic congestion, the better the performance of AVs compared to conventional vehicles.
Sinha et al. (2020)	Freeway	NA	VISSIM	CAV	0%, 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, and 100 %	Wiedemann 99	User-defined driving behavior	• TTC, PET, relative speed	• Manual vehicle–manual vehicle crash rate decreased from 0.9 to 0.0 by increasing the CAV MPR from 0%–100%. • CAV–manual vehicle crash rate started escalating to 0.3 by increasing the CAV MPR to 90 %. • The overall crash rate dropped from 0.9 to 0.0 by increasing the CAV MPR from 0%–100%.
Arvin et al. (2020)	Intersection	NA	VISSIM	AV CAV	0%, 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, and 100 %	Wiedemann 74	AVs: ACC and cooperative ACC (CACC) models CAVs: modified Wiedemann	• Number of longitudinal conflicts • Driving volatility	• Number of conflicts: • Implementing only AVs resulted in a reduction in the number of conflicts from 10 to zero by increasing the MPR from 0%–100%. • At AV MPR of 10 %, the number of conflicts increased compared to the AV MPR of 0%. • By adding coordination into AVs, the number of conflicts decreased steadily from 10 to zero for CAV MPRs of 0% and 100 %. • Overall, from MPR of 10%–90%, the CAV scenarios had fewer number of conflicts than the AV scenarios. • Speed volatility: • For the AV environment, the speed volatility experienced two peaks at MPRs of 40 % and 80 %. • AV MPRs of 0% and 100 % experienced seven and zero conflicts. • For the CAV environment, the number of conflicts decreased constantly from MPR of 0%–100%. • Speed volatilities in the CAV environments were lower than the AV environments.

Note: NA = Not Applicable.

Table A4
Summary of driving simulator studies.

Author	Participants' Information			Experiment Factors				AV Challenge	Scenario Parameters	Statistical Tool	Response Variable	Significant Results
	Age	Annual mileage	Driving experience	Facility	Speed	Traffic	Repetition					
Desmond et al. (1998)	18–27	×	2 to 8 years of driving experience	Not considered	80 km/h	×	✓	Fatigue	Perturbing events	ANOVA	• Physical fatigue items, perceptual fatigue items, boredom/apathy • Lateral control such as heading error, deviation of the vehicle	• A similar level of workload. • Better performance recovery in manual driving. • Automated driving results in undermobilizing driver's effort.
Gouy et al. (2012)	20–63	2000–56,000 km	Experience with a driving simulator, at least 1 year of driving experience	Three-lane highway	90 km/h	✓	✓	Platoon environment	Time headway within the platoons	ANOVA	Time headway	• Smaller average and minimum time headway when driving adjacent to AV platoons with short time headway.
Berthelon and Gineyt (2014)	21–29	×	At least 2 years of driving experience	Three-lane highway	Highway: 110 km/h Urban scenario: 70–90 km/h	With/without	×	Drunk driving	Driving environment (urban area, car following, highway), different alcohol concentration	ANOVA	Number of collisions, mean speed	• Lateral and longitudinal control of the AV is more likely to be impaired compared to strategies adopted in evasive situation.
Strand et al. (2014)	24–65	>10,000 km	No automated driving experience & > 5 yr driving experience	Two-lane undivided rural road	70 km/h	✓	✓	System failures	Automation level, extent of system failure (moderate/severe/completely)	ANOVA, ANCOVA, Fisher's exact tests	Minimum TTC, minimum time headway, response time, point-of-no-return, number of collisions	• Further automation leads to lower performance of driver. • Drivers performed better at controlling the lower extent of system failure.
Blommer et al. (2015)	40 (24 < 45 yr and 16 > 45 yr)	×	No experience of automated driving	Four-lane undivided roadway	50–70 mph	Light traffic	×	Disengagement	Continuous and scheduled automated driving, secondary tasks	ANOVA	Response time, eye glance behavior, percent eyes-on-road time	• Radio listeners responded significantly faster. • The scheduled driver engagement strategy performed better when visual distraction was used.
Happee et al. (2017)	33.5 (SD = 9)	×	Familiarity with the driving simulator	Three-lane highways	120 km/h	With (30 veh/km)/Without traffic	✓	Disengagement	Time budget, lane driven, traffic density, secondary tasks	Linear regression, Fisher's exact tests	In total, 19 performance metrics in terms of risk, braking, and steering, such as TTC, clearance toward the obstacle and the roadside, peak accelerations, overshoot, etc.	• AV can cause delayed initial steering and braking, lower TTC, and stronger braking or steering. • No difference between cognitive and visual distraction. • The precision of maneuver remained unaffected.
Gold et al. (2018)*	19–79	×	At least 1 year of driving experience	Three-lane highways	120 km/h	0, 10, 20, 30 veh/km	✓	Disengagement	Time budget, lane driven, traffic density, secondary	Generalized linear regression	TOT, TTC, crash, brake application	• Traffic density (negatively), repetition (positively), and time

(*continued on next page*)

Table A4 (*continued*)

Author	Participants' Information			Experiment Factors				AV Challenge	Scenario Parameters	Statistical Tool	Response Variable	Significant Results
	Age	Annual mileage	Driving experience	Facility	Speed	Traffic	Repetition					
									tasks, repetition of the experiment			budget were highly influential. • TOT, TTC, and crash probability showed reliable results.
Lee et al. (2018)	23 below and 7 above 50 years of age	×	Not considered	Three-lane highway	100 km/h	✓	×	Platoon environment	Platoon size, different MPR	ANOVA, logistic regression	Steering magnitude, steering velocity, lane-change duration, lane-change (success/failure)	• Smaller average and minimum time headway when driving adjacent to AV platoons with short time headway.
Kundinger et al. (2018)	18–64	×	Not considered	Three-lane highway	MV: 120 km/h AV: 110 km/h	Light traffic	×	Drowsiness	Age group, different time of the day, different sleepiness category	ANOVA	Karolinska Sleepiness Scale	• Time and driving mode have a significant effect on the development of drowsiness.
Yun and Yang (2020)	22–33	×	At least 6 months of driving experience	Four-lane highway	100 km/h	×	✓	Disengagement	Diverse warning combinations (visual, auditory, haptic), disengagement scenarios (planned/unplanned)	MANOVA	• Human behavior metrics: response time, TTL • Vehicle control metrics: SDLP, SRR • Psychological metrics: SCR, AHR	• The multimodal warning method showed superiority over unimodal warnings. • Each modality is preferred in a specific situation (e.g., haptic and auditory modality elicits a more immediate and stable warning, respectively). • Response time in unplanned disengagement is faster than planned events.
Lee et al. (2020)	25–39	×	More than 1 year of driving experience		×	×	×	Disengagement	Different secondary tasks with different physical/visual/cognitive loads	Non-parametric ANOVA	• Mean longitudinal/lateral acceleration, maximal longitudinal/lateral acceleration, maximum speed, minimum speed, DTC, TTC, SDLP	• Resource allocation associated with each of the non-driving-related tasks did not significantly affect the take-over quality. • The cognitive load of the non-driving-related tasks more effectively affect the longitudinal and lateral control than their physical and visual attributes.

[*] This study used a series of driving simulator experiments with the same design.

References

Agriesti, S., Studer, L., Gandini, P., Marchionni, G., Ponti, M., Visintainer, F., 2019. Safety on the Italian Highways: Impacts of the Highway Chauffeur System. Smart Transportation Systems 2019. Springer.
Arksey, H., O'Malley, L., 2005. Scoping studies: towards a methodological framework. Int. J. Soc. Res. Methodol. 8, 19–32.
Arvin, R., Kamrani, M., Khattak, A., Rios-Torres, J., 2018. Safety impacts of automated vehicles in mixed traffic. In: Transportation Research Board Annual Meeting, 2018. Washington D.C., District of Columbia, United States of America.
Arvin, R., Khattak, A., Rios Torres, J., 2019. Evaluating safety with automated vehicles at signalized intersections: application of adaptive cruise control in mixed traffic. In: Transportation Research Board Annual Meeting, 2019. Washington D.C., District of Columbia, United States of America.
Arvin, R., Khattak, A.J., Kamrani, M., Rio-Torres, J., 2020. Safety evaluation of connected and automated vehicles in mixed traffic with conventional vehicles at intersections. J. Intell. Transp. Syst. Technol. Plan. Oper. 1–18.
Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., Rahwan, I., 2018. The moral machine experiment. Nature 563, 59.
Bahram, M., Ghandeharioun, Z., Zahn, P., Baur, M., Huber, W., Busch, F., 2014. Microscopic traffic simulation based evaluation of highly automated driving on highways. In: 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2014/11//. Qingdao, China. IEEE, pp. 1752–1757.
Banerjee, S.S., Jha, S., Cyriac, J., Kalbarczyk, Z.T., Iyer, R.K., 2018. Hands off the wheel in autonomous vehicles?: A systems perspective on over a million miles of field data. In: 2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN). IEEE, pp. 586–597.
Berthelon, C., Gineyt, G., 2014. Effects of alcohol on automated and controlled driving performances. Psychopharmacology 231, 2087–2095.
Bhavsar, P., Das, P., Paugh, M., Dey, K., Chowdhury, M., 2017. Risk analysis of autonomous vehicles in mixed traffic streams. Transp. Res. Rec. 2625, 51–61.
Bila, C., Sivrikaya, F., Khan, M.A., Albayrak, S., 2017. Vehicles of the future: a survey of research on safety issues. IEEE Trans. Intell. Transp. Syst. 18, 1046–1065.
Blommer, M., Curry, R., Kochhar, D., Swaminathan, R., Talamonti, W., Tijerina, L., 2015. The effects of a scheduled driver engagement strategy in automated driving. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications Sage CA: Los Angeles, CA, pp. 1681–1685.
Boggs, A.M., Wali, B., Khattak, A.J., 2020. Exploratory analysis of automated vehicle crashes in California: a text analytics & hierarchical Bayesian heterogeneity-based approach. Accid. Anal. Prev. 135, 105354.
Canis, B., 2020. Issues in Autonomous Vehicle Testing and Deployment. Congressional Research Service. Available: https://fas.org/sgp/crs/misc/R45985.pdf (April 2020).
Childress, S., Nichols, B., Charlton, B., Coe, S., 2015. Using an activity-based model to explore the potential impacts of automated vehicles. J. Trans. Res.Board 2493, 99–106.
Cohen, S., Shirazi, S., 2017. Can We Advance Social Equity With Shared, Autonomous and Electric Vehicles? Institute of Transportation Studies at the University of California, Davis. Available: <https://www.transformca.org/sites/default/files/3R.Equity.Indesign.Final_.pdf> (January 2020).
Combs, T.S., Sandt, L.S., Clamann, M.P., McDonald, N.C., 2019. Automated vehicles and pedestrian safety: exploring the promise and limits of pedestrian detection. Am. J. Prev. Med. 56, 1–7.
Cui, J., Liew, L.S., Sabaliauskaite, G., Zhou, F., 2019. A review on safety failures, security attacks, and available countermeasures for autonomous vehicles. Ad Hoc Netw. 90, 101823.
Das, S., Dutta, A., Tsapakis, I., 2020. Automated vehicle collisions in California: applying Bayesian latent class model. IATSS Res.
Deluka Tibljaš, A., Giuffrè, T., Surdonja, S., Trubia, S., 2018. Introduction of autonomous vehicles: roundabouts design and safety performance evaluation. Sustainability 10, 1060.
Desmond, P.A., Hancock, P.A., Monette, J.L., 1998. Fatigue and automation-induced impairments in simulated driving performance. Transp. Res. Rec. 1628, 8–14.
Detwiller, M., Gabler, H.C., 2017. Potential reduction in pedestrian collisions with an autonomous vehicle. In: The 25th ESV Conference Proceedings. NHTSA, pp. 1–8.
Fagnant, D.J., Kockelman, K., 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. Transp. Res. Part A Policy Pract. 77, 167–181.
Favarò, F.M., Nader, N., Eurich, S.O., Tripp, M., Varadaraju, N., 2017. Examining accident reports involving autonomous vehicles in California. PLoS One 12, e0184952.
Gold, C., Happee, R., Bengler, K., 2018. Modeling take-over performance in level 3 conditionally automated vehicles. Accid. Anal. Prev. 116, 3–13.
Goodall, N.J., 2014. Ethical decision making during automated vehicle crashes. Transp. Res. Rec. 2424, 58–65.
Gouy, M., Diels, C., Reed1, N., Stevens1, A., Burnett, G., 2012. The Effects of Short Time Headways Within Automated Vehicle Platoons on Other Drivers.
Happee, R., Gold, C., Radlmayr, J., Hergeth, S., Bengler, K., 2017. Take-over performance in evasive manoeuvres. Accid. Anal. Prev. 106, 211–222.
Hendrickson, C.T., Harper, C., 2018. Safety and Cost Assessment of Connected and Automated Vehicles.
Jafarnejad, S., Codeca, L., Bronzi, W., Frank, R., Engel, T., 2015. A car hacking experiment: when connectivity meets vulnerability. In: 2015 IEEE Globecom Workshops (GC Wkshps). IEEE, pp. 1–6.
Junietz, P., Wachenfeld, W., Klonecki, K., Winner, H., 2018. Evaluation of different approaches to address safety validation of automated driving. In: 2018 21st International Conference on Intelligent Transportation Systems (ITSC). IEEE, pp. 491–496.
Kalra, N., 2017. Challenges and Approaches to Realizing Autonomous Vehicle Safety. RAND Corporation, Santa Monica, CA.
Kalra, N., Paddock, S.M., 2016. Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? Transp. Res. Part A Policy Pract. 94, 182–193.
Katrakazas, C., Quddus, M., Chen, W.H., 2019. A new integrated collision risk assessment methodology for autonomous vehicles. Accid. Anal. Prev. 127, 61–79.
Kockelman, K., Avery, P., Bansal, P., Boyles, S.D., Bujanovic, P., Choudhary, T., Clements, L., Domnenko, G., Fagnant, D., Helsel, J., 2016. Implications of Connected and Automated Vehicles on the Safety and Operations of Roadway Networks: a Final Report. Center for Transportation Research, University of Texas at Austin. FHWA/TX-16/0-6849-1, Available: <https://library.ctr.utexas.edu/ctr-publications/0-6849-1.pdf> (January 2019).
Koopman, P., Wagner, M., 2016. Challenges in autonomous vehicle testing and validation. SAE Int. J. Transp. Saf. 4, 15–24.
Kundinger, T., Riener, A., Sofra, N., Weigl, K., 2018. Drowsiness detection and warning in manual and automated driving: results from subjective evaluation. Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications 229–236.
Kusano, K.D., Gabler, H.C., 2014. Comprehensive target populations for current active safety systems using national crash databases. Traffic Inj. Prev. 15, 753–761.
Lee, C., 2017. Grabbing the wheel early: moving forward on cybersecurity and privacy protections for driverless cars. Federal Commun. Law J. 69, 25.
Lee, S., Oh, C., Hong, S., 2018. Exploring lane change safety issues for manually driven vehicles in vehicle platooning environments. IET Intell. Transp. Syst. 12, 1142–1147.
Lee, S.C., Yoon, S.H., Ji, Y.G., 2020. Effects of non-driving-Related task attributes on takeover quality in automated vehicles. Int. J. Hum. Interact. 1–9.
Li, T., Kockelman, K.M., 2016. Valuing the safety benefits of connected and automated vehicle technologies. Transportation Research Board 95th Annual Meeting.
Li, R., Zhai, R., 2019. Estimation and analysis of minimum traveling distance in self-driving vehicle to prove their safety on road test. J. Phys. Conf. Ser., 032101. IOP Publishing.
Litman, T., 2017. Autonomous Vehicle Implementation Predictions. Victoria Transport Policy Institute Victoria, Canada. Available: <https://www.vtpi.org/avip.pdf> (January 2019).
Lubbe, L., Jeppsson, H., Ranjbar, A., Fredriksson, J., Bärgman, J., Östling, M., 2018. Predicted road traffic fatalities in Germany: the potential and limitations of vehicle safety technologies from passive safety to highly automated driving. In: Proceedings of IRCOBI Conference. Athena, Greece.
Matysiak, A., Razin, P., 2018. Analysis of the advancements in real-life performance of highly automated vehicles' with regard to the road traffic safety. MATEC Web of Conferences.
Milakis, D., Van Arem, B., Van Wee, B., 2017. Policy and society related implications of automated driving: a review of literature and directions for future research. J. Intell. Transp. Syst. Technol. Plan. Oper. 21, 324–348.
Morando, M.M., Tian, Q., Truong, L.T., Vu, H.L., 2018. Studying the safety impact of autonomous vehicles using simulation-based surrogate safety measures. J. Adv. Transp.
Mousavi, S.M., Lord, D., Dadashova, B., Reza Mousavi, S., 2020. Can Autonomous vehicles enhance traffic safety at unsignalized intersections?. In: International Conference on Transportation and Development 2020. American Society of Civil Engineers Reston, VA, pp. 194–206.
Munn, Z., Peters, M.D., Stern, C., Tufanaru, C., McArthur, A., Aromataris, E., 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med. Res. Methodol. 18, 143.
NHTSA, 2018. Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey. U.S. Department of Transportation. Available: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812506> (March 2020).
Papadoulis, A., Quddus, M., Imprialou, M., 2019. Evaluating the safety impact of connected and autonomous vehicles on motorways. Accid. Anal. Prev. 124, 12–22.
Petrović, Đ., Mijailović, R., Pešić, D., 2020. Traffic accidents with autonomous vehicles: type of collisions, manoeuvres and errors of conventional vehicles' drivers. Transp. Res. Procedia 45, 161–168.
PTV, 2018. PTV VISSIM Manual.
Qin, Y., Wang, H., 2019. Influence of the feedback links of connected and automated vehicle on rear-end collision risks with vehicle-to-vehicle communication. Traffic Inj. Prev. 20, 79–83.
Rahman, M.S., Abdel-Aty, M., Lee, J., Rahman, M.H., 2019. Safety benefits of arterials' crash risk under connected and automated vehicles. Transp. Res. Part C Emerg. Technol. 100, 354–371.
Raj, A., Kumar, J.A., Bansal, P., 2019. A Multicriteria Decision Making Approach to Study the Barriers to the Adoption of Autonomous Vehicles. arXiv Preprint arXiv: .12051.
Rau, P., Yanagisawa, M., Najm, W.G., 2015. Target Crash Population of Automated Vehicles. In 24th International Technical Conference on the Enhanced Safety of Vehicles (ESV), pp. 1–11.
Rodrigue, J.-P., Comtois, C., Slack, B., 2016. The Geography of Transport Systems. Routledge.
SAE, 2018. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. Available: <https://www.sae.org/standards/content/j3016_201806/> (November 2019).
Schoettle, B., Sivak, M., 2015. A Preliminary Analysis of Real-world Crashes Involving Self-driving Vehicles. The University of Michigan, Transportation Research Institute. UMTRI-2015-34.

Sener, I.N., Zmud, J., Williams, T., 2019. Measures of baseline intent to use automated vehicles: a case study of Texas cities. Transp. Res. Part F Traffic Psychol. Behav. 62, 66–77.
Sinha, A., Chand, S., Wijayaratna, K.P., Virdi, N., Dixit, V., 2020. Crash severity and rate evaluation of conventional vehicles in mixed fleets with connected and automated vehicles. Procedia Comput. Sci. 170, 688–695.
Sohrabi, S., Khreis, H., 2020. Burden of disease from transportation noise and motor vehicle crashes: Analysis of data from Houston, Texas. Environ. Int., 105520
Sohrabi, S., Khreis, H., Lord, D., 2020. Impacts of autonomous vehicles on public health: a conceptual model and policy recommendations. Sustain. Cities Soc. 63, 102457.
Strand, N., Nilsson, J., Karlsson, I.M., Nilsson, L., 2014. Semi-automated versus highly automated driving in critical situations caused by automation failures. Transp. Res. Part F Traffic Psychol. Behav. 27, 218–228.
Taeihagh, A., Lim, H.S.M., 2018. Governing autonomous vehicles: emerging responses for safety, liability, privacy, cybersecurity, and industry risks. Transp. Rev. 39, 103–128.
Tainio, M., 2015. Burden of disease caused by local transport in Warsaw, Poland. J. Transp. Health 2, 423–433.
Teoh, E.R., Kidd, D.G., 2017. Rage against the machine? Google's self-driving cars versus human drivers. J. Safety Res. 63, 57–60.
US DOT, 2018. Preparing for the Future of Transportation: Automated Vehicles 3.0.
Virdi, N., Grzybowska, H., Waller, S.T., Dixit, V., 2019. A safety assessment of mixed fleets with connected and Autonomous Vehicles using the Surrogate Safety Assessment Module. Accid. Anal. Prev. 131, 95–111.
Wang, S., Li, Z., 2019. Exploring the mechanism of crashes with automated vehicles using statistical modeling approaches. PLoS One 14.
Wang, L., Zhong, H., Ma, W., Abdel-Aty, M., Park, J., 2020. How many crashes can connected vehicle and automated vehicle technologies prevent: a meta-analysis. Accid. Anal. Prev. 136, 105299.
Wilde, G.J., 1998. Risk homeostasis theory: an overview. Inj. Prev. 4, 89–91.
Xu, C., Ding, Z., Wang, C., Li, Z., 2019. Statistical analysis of the patterns and characteristics of connected and autonomous vehicle involved crashes. J. Safety Res. 71, 41–47.
Yanagisawa, M., Najm, W.G., Rau, P., 2017. Preliminary estimates of target crash populations for concept automated vehicle functions. 25th International Technical Conference on the Enhanced Safety of Vehicles (ESV).
Yang, J., Ward, M., Akhtar, J., 2017. The Development of Safety Cases for an Autonomous Vehicle: a Comparative Study on Different Methods. SAE Technical Paper, 0148-7191, Available: <https://saemobilus.sae.org/content/2017-01-2010> (January 2019).
Ye, L., Yamamoto, T., 2019. Evaluating the impact of connected and autonomous vehicles on traffic safety. Phys. A Stat. Mech. Its Appl. 526, 121009.
Young, W., Sobhani, A., Lenné, M.G., Sarvi, M., 2014. Simulation of safety: a review of the state of the art in road safety simulation modelling. Accid. Anal. Prev. 66, 89–103.
Yun, H., Yang, J.H., 2020. Multimodal warning design for take-over request in conditionally automated driving. Eur. Transp. Res. Rev. 12, 1–11.
Zhang, W., Guhathakurta, S., Fang, J., Zhang, G., 2015. Exploring the impact of shared autonomous vehicles on urban parking demand: an agent-based simulation approach. Sustain. Cities Soc. 19, 34–45.