Public health principles to inform testing and build trust in automated vehicles

Johnathon P Ehsani 1, Jeffrey Michael, 1 Takeru Igusa 2

ABSTRACT
Highly publicised crashes involving self-driving or autonomous vehicles (AVs) have raised questions about safety and eroded public trust in the technology. In this State of the Art Review, we draw on previous successes in injury prevention and public health to focus attention on three strategies to reduce risk and build public confidence as AVs are being tested on public roads. Data pooling, a graduated approach to risk exposure, and harm reduction principles each offer practical lessons for AV testing. The review points out how the eventual deployment of AV technology could have a substantial impact on public health. In this regard, inclusive testing, public education and smart policy could extend the social value of AVs by improving access to mobility and by directing deployments towards scenarios with the greatest population health impact. The application of these strategies does not imply slowing down progress; rather, their implementation could accelerate adoption and result in realising the benefits of AVs more quickly and comprehensively while minimising risks.

INTRODUCTION
The development of vehicles with automated driving systems is rapidly advancing, with estimates of cumulative industry investment approaching $100 billion or more 3 and initial deployments already occurring. 4 5 While the promise that highly automated (driverless or autonomous) vehicles (AVs) will reduce the number of traffic injuries and fatalities 6 7 is appealing, public trust and confidence in this technology have wavered over the past years, partly due to media coverage about several serious crashes involving AVs. 8

In the absence of specific federal safety standards for AVs, public confidence in AV technology—and interest in adopting this potentially life-saving technology—will come from other sources. These include the extent to which local officials (city and state) create a safe testing environment and the industry provides adequate safety assurance during testing. In addition, the public is likely to be influenced by the viewpoints expressed by trusted sources, such as law enforcement officers and public safety officials.

The emerging literature on public acceptance and trust in AVs has examined the degree to which individuals intend to use the technology according to a range of factors such as demographic and contextual characteristics. 9 10 While younger, urban dwelling and more affluent individuals report greater acceptance and interest in using AVs, concerns regarding safety, privacy and cybersecurity are consistently reported in both national and international samples. 9 10 Access to existing mobility options and experience using advanced vehicle technologies are also factors influencing acceptance of AVs, as well as perceived societal benefits such as safety or enhanced mobility for vulnerable populations. 4 11

Part of the challenge in addressing public confidence is that conventional approaches to regulatory safety assurance, such as barrier crash testing, are not adequate for assessing the risks inherent in self-driving technologies. New metrics are needed to demonstrate safety. For example, the notion that self-driving vehicles could be tested on public roadways until they demonstrate a certain benchmark of safety has been challenged by an analysis that suggests hundreds of millions or billions of miles of testing would be needed. 12 The current requirement for companies to report the number of times a human driver needed to over-ride the automated driving system could be counterproductive, as testing could then be skewed to low-demand environments. 13

A number of alternative safety assurance approaches have been discussed in the literature, and new paradigms have been proposed for optimising safety potential. 14 For example, Ryerson and colleagues 15 present an innovative extension of the widely used Haddon Matrix. 16 Their model proposes that the software and the hardware components of both the vehicle and the physical environment should be considered when designing and testing the safety of AVs.

In this State of the Art Review, we focus attention on steps that could be taken to bolster public confidence in the near term when AVs have passed in silico and closed-course track tests and are being tested on public roads. We use a public health lens to examine critical issues associated with real-life testing, identify gaps and future directions in research, and provide recommendations that will inform AV stakeholders on pathways towards safe and favourable deployment. This review goes on to examine how the eventual deployment of AV technology could have a substantial impact on public health. Mobility is a social determinant of health 17 and AVs—if strategically deployed—could extend and enhance mobility, especially among those populations that are currently, and have historically been, underserved by mobility options. While such deployment may be far in the future, this review will demonstrate the importance of the manner of testing in the near term, as well as the influential role that public opinion leaders can play in building trust for subsequent deployment.
CRITICAL ISSUES

Successes in injury prevention and public health can be used to guide a thoughtful approach to development and deployment of AVs. The application of these principles does not imply slowing down progress; rather, the strategies we describe in the following section could potentially accelerate testing and result in realising the benefits of AVs more quickly and comprehensively while minimising the risk of a public backlash.

Six critical issues discussed further address actions that could be taken by entities that test AVs on public roads and governments in order to reduce risk and increase public confidence in AV technology. Three address operational measures: data pooling, graduated approach to risk exposure and harm reduction. Three others—inclusive testing, public education and smart policy—focus on extending the social value of AVs by directing testing towards applications that could result in the highest public health impacts.

Strategies to reduce risk and build public trust

Data pooling

Scientist, innovator and entrepreneur Sebastian Thrun, a pioneer in AV development, describes the importance of AV safety data: ‘When one of the self-driving cars makes an error, all of the self-driving cars learn from it. In fact, new self-driving cars are “born” with the complete skill set of their ancestors. So collectively, these cars can learn faster than people’.18 This underlines the unique potential of AVs to learn from one another, where data associated with a safety issue experienced by one AV could be analysed and disseminated so that lessons learnt could be adopted across the entire fleet.19 20

The value of data pooling has been demonstrated in the field of reliability engineering, in which reliability is defined as the ability of a product or system to perform its functions without failure or safety incidents. When data are systematically collected and pooled, lessons learnt from past operational performance can be applied to future designs to address reliability issues. This process, known as reliability growth, follows the general principles of the learning curve and results in continuous improvement.

Reliability growth is not restricted to products from a single manufacturer; if data are shared between manufacturers, the learning rate increases and the entire industry benefits. The aeronautics industry, with its strong emphasis on safety and reliability, has successfully pursued data pooling efforts through two ongoing efforts, the Commercial Airline Safety Team (CAST) consortium and the Government-Industry Data Exchange Program.21 22 These programmes improve the total quality and reliability of operations, systems and components by pooling technical information from multiple manufacturers and vendors.

Data pooling programmes are credited with a number of critical safety enhancements for commercial aviation, including refinements to warning systems that alert pilots that they are flying too close to mountains.23 The refinements corrected a situation where the systems had been excessively conservative and pilots were desensitised to the warnings. Data pooling can be effective, but the programme need to be designed with incentives for participation. Corporations may be reluctant to contribute data that reveal proprietary information. Such issues would need to be addressed before a system could be implemented across the AV industry. However, the aviation example shows that such an achievement is feasible.

Graduated approach to risk exposure

The effectiveness of graduated risk exposure in reducing crashes among novice teenage drivers has been well established. Graduated licensing programmes have been adopted in some form by all US states and these policies have demonstrated fatality reductions of up to 20% among 16-year-old drivers.24 Human drivers learn most effectively and safely when they are initially exposed to relatively simple environmental conditions, such as lower speeds, less dense traffic, good lighting, fewer distractions and better weather. Later, as their skills increase and they demonstrate competence, they are exposed to progressively more complex and challenging conditions.

The same approach could be extended to the testing of AVs. Similar to human drivers, there are several dimensions of risk that could be adjusted during the initial testing phase for AVs. For example, suburban arterials with heavy, fast traffic and frequent vehicle interactions have relatively high crash risk per mile, while limited access divided highways have much lower risk.25 Geospatial crash data could reveal differences in crash risks across road segments within a city and inform route choice for AVs in a graduated testing approach. Visible roadway lane markings and the predictability of road user behaviour also vary considerably among roadway types and locations.26 An inventory of appropriate infrastructure and an analysis of the makeup of road users within a city could provide additional dimensions of risk for gradual exposure to AVs.

Several jurisdictions and companies have already taken this type of phased-risk approach for AV testing, requiring demonstrated competence in low-risk circumstance before permitting use in more complex driving environments.27 Waymo’s decision to deploy their vehicles in Chandler, Arizona, where the streets are flat and wide, there is little rain or fog, and detailed mapping data are available, was a design choice to reduce risk.28 More could be done in managing risk while accumulating operational experience with AV testing, and this caution could help prevent future crashes involving AVs and avoid the subsequent erosion in public trust.

This analogy between driver training for teenagers and AVs could be extended to include the requirement for a supervisor to be present during the learner period. During initial on-road testing, AV companies include a ‘safety driver’ in the vehicle. This is necessary to ensure that a ‘back-up’ safety mechanism remains in place as a computer system learns how to safely handle the wide range of environments and road user situations encountered on public roads. Some companies have extended this concept to allow any employee the authority to ground the self-driving fleet when safety issues are detected in the product development and testing phase.29

Harm reduction

A concept that has been used alongside graduated licensing is harm reduction. When applied to teenage drivers, this approach attempts to compensate for higher crash risk by providing them with the safest possible vehicle (rather than the smaller, older and inherently less-safe vehicles that are too often given to young drivers).30 If applied to AVs, the same approach would have AV companies voluntarily using vehicles with the highest safety ratings (which include pedestrian protections), at least until a predetermined level of demonstrated safety competence had been achieved. Similarly, AVs that have been custom built (rather than conventional vehicles retrofitted with instrumentation) could, for example, be voluntarily crash tested at higher speeds than mandated for passenger cars as a further measure to...
build public trust. Harm reduction principles would also suggest that initial testing environments for AVs should meet minimum safety standards (eg, adequate sidewalks for pedestrians, well-illuminated roads and the presence of protected bicycle lanes).

Harm reduction principles differ from graduated learning in that rather than attempting to lower crash risk, they acknowledge that some crashes may occur and seek to compensate by providing a safer environment. While AVs that have been carefully developed are likely to present a lower crash risk than human-driven vehicles and therefore may not justify such measures, harm reduction approaches could be worthwhile because they may bolster public confidence by demonstrating an abundance of caution.

Maximising the public health impact of AVs

In this section, we examine strategies for extending the social value of AVs by directing testing toward applications that could significantly impact public health.

Inclusion

Strategic early deployment of AVs in rideshare programmes or other innovative configurations promises to extend mobility by making transportation services cheaper and more convenient. This could mean that mobility services will be economically feasible in underserved areas and that health disparities could be reduced. Enhanced mobility can impact disparities by providing access to healthy foods, healthcare, economic opportunities, and recreation and exercise, which influence quality of life. However, since the roadway environment—both physical infrastructure and road user characteristics—can be quite different in these underserved areas than in more affluent neighbourhoods, it is important that AVs develop safety competencies in such areas as quickly as possible. The critical issue here is developing testing strategies that will facilitate AV deployment in underserved neighbourhoods.

The underlying hypothesis is that AV deployment will be quicker and safer in areas that mimic the real-world setting where testing took place. There is a parallel in clinical trials, in which a medical product is tested for efficacy and safety in a targeted population and subsequently approved for the same population. Use of this medical product on a substantially different population (eg, prescribing a drug to children, when the drug was only tested on adults) is unproven and therefore discouraged until clinical tests can be conducted with this population.

We acknowledge that the need to test in underserved areas may sometimes conflict with the need to begin testing in a low-risk environment. However, because of historical mobility deficits, underserved areas should be included in geospatial analyses of crash risk and given priority as testing locations when consistent with safety needs. Moreover, testing should continue until such areas are addressed by the graduated risk approach described previously. This will help ensure that AVs fully deliver their potential public health benefits.

Educating the public

Most people do not think about crash risk in a rational way and education may improve understanding of relative safety risks. Many drivers have an optimistic and inflated sense of their ability to drive safely, and generally place a premium on being able to control a situation, rather than on delegating it to another entity (even if the other entity is more capable). By implication, many people will demand a much higher standard of safety from AVs than they would expect from other drivers, including themselves. The ease with which one can bring to mind examples of an event (such as highly publicised AV fatalities) can also influence perceptions of risk and the likelihood of an event happening to themselves. This natural human bias, which psychologists termed the ‘availability heuristic’, might explain some of the uneasiness among the public with the thought of riding in an AV or sharing a road with them. Public education initiatives, such as the Partners for Automated Vehicle Education, represent steps towards addressing misconceptions about the safety risks posed by AVs.

In addition to driver education and campaigns, public information comes from media attention concerning the safety of AVs. Statements from trusted local sources, including highway safety officials and law enforcement officers, either through their own information outlets or through media interviews, are key in determining public attitudes. A recent white paper from the Governors Highway Safety Association outlines information needs, operational requirements and other prerequisites that highway safety officials and law enforcement agencies will need in order to effectively perform their functions when AVs are operating in their jurisdiction. For example, highway safety officials need information they can use in advising drivers of conventional vehicles concerning how to interact with AVs on the road. Law enforcement officials need a uniform way to identify AVs and procedures to interact with them at crash scenes. Organisations that are testing AVs should respond to the needs of these local officials in order to maintain their trust and confidence.

Smart policy

Regulation is a potentially powerful tool but needs to be used carefully, especially in a dynamic area such as the development of automated driving technology. Regulation of conventional motor vehicles and safety devices has been extremely effective. An evaluation by the National Highway Traffic Safety Administration of 21 federally mandated safety technologies concluded that, between 1960 and 2012, these technologies—from energy-absorbing steering columns implemented in the 1960s to electronic stability control mandated in all cars in 2012—saved a total of 613,501 lives.

However, the pace of innovation and development of automated driving systems is incredibly rapid compared with these conventional technologies. Serious questions have been raised about the ability of traditional regulatory approaches to provide safety assurance without stifling innovation and safety progress. Many questions remain regarding the most appropriate substitute for traditional motor vehicle safety standards for assuring the safety of highly automated vehicles. Alternatives have been considered based on experience in other fields, including aviation and drug safety, but there is no consensus yet on the best method to assess the safety performance of these technologies or on the most reasonable threshold for acceptability.

The history of automotive regulation suggests that collaboration may be more efficient than traditional adversarial relationships among regulators, safety advocates and industry. A recent example is the collaboration between National Highway Traffic Safety Administration and the Insurance Institute for Highway Safety’s initiative to work with 20 automakers representing more than 99% of the US auto market to make automatic emergency braking a standard feature on all new cars. Another relevant example can be found in the history of airbags, where 40 years passed from the time federal regulators first pursued the goal of equipping cars with air bags to their
The introduction of automated vehicle technology has the potential to transform mobility, reduce the societal burden of motor vehicle crashes and improve health disparities. However, their rapid development does not guarantee successful adoption and widespread use, especially among populations that might benefit most from their use. The history of previous transformative safety interventions such as the airbag provides sobering lessons of how the deployment of a life-saving technology can stall for decades as competing interests are resolved, technical and safety issues are recognised and public expectations mature.

The principles of data pooling, graduated risk exposure and harm reduction offer lessons for minimising risks associated with testing and early deployments and nurturing public confidence in new technologies. Just as provisions of graduated driver licensing are lifted when teenagers have accumulated sufficient driving experience, additional safety precautions could be temporary accommodations while the characteristic risks of AVs are established. Inclusive testing, public education, coordination with local public officials and smart policy could further strengthen public confidence and accelerate successful deployment and widespread adoption of this transformative technology.

The principles of data pooling, graduated risk exposure and harm reduction offer lessons for minimising risks associated with testing and early deployments and nurturing public confidence in new technologies. Just as provisions of graduated driver licensing are lifted when teenagers have accumulated sufficient driving experience, additional safety precautions could be temporary accommodations while the characteristic risks of AVs are established. Inclusive testing, public education, coordination with local public officials and smart policy could further strengthen public confidence and accelerate successful deployment and widespread adoption of this transformative technology.

**CONCLUSIONS**

Adoption of automated mobility could have far-reaching effects that not only improve transportation safety but also substantially benefit the health and well-being of the nation. The safety potential of AVs demands a re-examination of how we view and design mobility policy. A public health approach focuses on the well-being of the entire population, with an emphasis on the greatest benefit for the largest number of people while also recognizing the need to maximise health equity.

It is indisputable that AVs have the potential to transform mobility, reduce the societal burden of motor vehicle crashes and improve health disparities. However, their rapid development does not guarantee successful adoption and widespread use, especially among populations that might benefit most from their use. The history of previous transformative safety interventions such as the airbag provides sobering lessons of how the deployment of a life-saving technology can stall for decades as competing interests are resolved, technical and safety issues are recognised and public expectations mature.

The principles of data pooling, graduated risk exposure and harm reduction offer lessons for minimising risks associated with testing and early deployments and nurturing public confidence in new technologies. Just as provisions of graduated driver licensing are lifted when teenagers have accumulated sufficient driving experience, additional safety precautions could be temporary accommodations while the characteristic risks of AVs are established. Inclusive testing, public education, coordination with local public officials and smart policy could further strengthen public confidence and accelerate successful deployment and widespread adoption of this transformative technology.

**Research using simulation**

While simulation is not likely to ever replace the need for realistic testing, it is an essential tool for assessing automated driving systems prior to physical testing and could be used to address the critical issues outlined in the previous section.

Single-vehicle simulation studies could address the benefits of harm reduction approaches, while fleet simulations may be necessary to examine the graduated approach to risk exposure. Simulation studies can assess safety impacts of the interactions of AVs and conventional vehicles and could even be used to predict the benefits of AV deployments in areas now underserved by other mobility options. An analysis of data pooling may require simulation of multiple fleets of AVs from competing companies. Such a study could inform decisions as to whether a CAST or similar data pooling effort could be effective in improving the overall safety of AVs.

**Research using innovative survey and focus group methods**

Surveys and focus groups have long been instrumental tools for planners seeking to develop transportation networks, particularly in predicting demand for transportation services and assessing acceptance or resistance to proposed projects. Such tools will be needed particularly in the study of critical issues associated with public education and smart policy. Given the unique nature of AVs, the types of surveys necessary for certain questions may be quite different from conventional practice. Specifically, the need for new ideas for graduated and inclusive testing and for AV use cases that can alleviate equity issues in low-income neighbourhoods may require survey methods that are designed specifically for ideation, such as the wiki survey developed by sociologists at Princeton University.

**Formulation of standards and policy approaches**

Just as the development of graduated licensing systems was guided by high-quality crash data, formulation of new AV testing protocols will require uniform safety data collection systems. This may involve the development of a taxonomy that includes characteristics of testing areas, near misses, crashes and other critical events. Candidate systems will be needed in order to evaluate relative merits and to refine a uniform system. Existing conventions for data standardisation, such as the protocol used in the Fatality Analysis Reporting System, could be used as models. Coalitions of automakers, tech companies and policy groups are already working together on these issues, and an effective standard-setting process will emerge from those efforts. Finally, a deeper look at relevant regulatory history, such as the evolution of air bag rules, could be very useful for making decisions about future AV regulatory paths.

**REFERENCES**


State of the art review