DRIVER ALCOHOL DETECTION SYSTEM FOR SAFETY (DADSS) – RISK BASED APPROACH TO ALCOHOL SENSING OUTCOMES MODELING

Timothy Allen, Maura Campbell, Kelly Ozdemir, Kianna Pirooz, Michael Willis, Abdullatif K. Zaouk
KEA Technologies, Inc.
U.S.A.

Susan Ferguson
U.S.A.

George Bahouth, Rebecca Spicer
Impact Research
U.S.A.

Scott E Lukas
Behavioral Psychopharmacology Research Laboratory (BPRL), McLean Imaging Center, McLean Hospital
McLean Hospital, Belmont, MA, 02478,
U.S.A.

Robert Strassburger
Automotive Coalition for Traffic Safety
U.S.A.

Paper Number 23-0291

ABSTRACT

A large number of fatal crashes every year in the United States are caused by alcohol-impaired drivers The Automotive Coalition for Traffic Safety and the National Highway Traffic Safety Administration entered into a research agreement to explore the feasibility of developing a passive in-vehicle alcohol detection system, known as the Driver Alcohol Detection System for Safety, with the goal of significantly reducing the incidence of drunk driving. This paper presents an analysis of the net benefit that could be achieved by installing such technology in the passenger vehicle fleet, using a risk-based approach to model potential outcomes. This outcomes model will calculate the net benefit of, and the public policy challenges associated with, more widespread use of non-invasive technology. Such an approach can be beneficial in determining the merits of the new technology and could be used to help guide public policy with respect to implementation. Furthermore, the technical data can be used to further refine the Driver Alcohol Detection System for Safety performance specifications.

INTRODUCTION

Motor vehicle crashes involving alcohol-impaired drivers result in a large number of deaths and injuries in the United States every year. In 2020, there were an estimated 38,824 motor vehicle crash deaths, of which almost 30 percent involved fatally injured drivers with blood or breath alcohol concentrations at or above 0.08 g/dL, the legal limit in all but one U.S. state (National Highway Traffic Safety Administration (NHTSA), 2022, Insurance Institute for Highway Safety (IIHS), 2022). To address this continuing problem, a Cooperative Research Agreement was instigated in 2008 between the Automotive Coalition for Traffic Safety (ACTS) and NHTSA, to explore the feasibility of developing passive in-vehicle technology that will detect driver breath (BrAC) and blood alcohol concentration (BAC) and ultimately prevent drunk driving when drivers exceed a preset limit. This is known as the Driver Alcohol Detection System for Safety (DADSS) program (Ferguson et al., 2011, Zaouk et al., 2019). The goal is to prevent the vehicle from being driven if the driver's BrAC or BAC is at or above 0.08 g/dL, the legal limit in all but one U.S. state, although other limits can be adopted.

Before such systems can be implemented, policy makers must define performance criteria for these devices. It clearly would be preferable to prevent all incidences of drunk driving. However, there is a concern that trying to prevent the vast number of drunk drivers from driving may result in incorrectly identifying a large number of drivers who are below the limit, thus inconveniencing them. The question is, to what degree would such an outcome affect acceptance of the technology.

Many other vehicle technologies that are currently in use have been shown to decrease crashes and save lives, yet it is acknowledged that they do not perform perfectly. For example, frontal airbags do not always deploy as expected. They occasionally fail to deploy in higher speed crashes or deploy at lower speeds than anticipated, resulting in injuries and deaths that would otherwise not be expected (Ferguson, 1996, 1998). Nonetheless, many studies have confirmed their life saving benefits (Ferguson, et al., 1995, Lund and Ferguson, 1995). NHTSA estimates that as of 2017, 50,457 lives have been saved by frontal airbags (National Center for Statistics and Analysis, 2020). Automatic Emergency Breaking (AEB) systems also do not perform perfectly. Studies that tested the rear-end crash performance of AEB systems when encountering a stationary vehicle at speeds of 30 and 40 mph, have shown that the technology is only about 85% effective in preventing collisions with a stationary vehicles at 30 mph and about 30% effective at 40 mph (AAA, 2022). However, research using real-world crash data, have shown that AEB systems reduce front-to-rear crash rates by 43% and front-to-rear injury crash rates by 45% (Cicchino, 2017). Thus, vehicle safety technologies do not have to be perfect to reduce crashes, injuries, and deaths. Nevertheless, both air bags and AEB technologies still receive public support.

There is one example of a new technology that did not have a successful introduction in the vehicle fleet, because the potential negative impact of the technology on drivers was not sufficiently taken into consideration ahead of its introduction. When driver seatbelt interlocks were mandated in the 1970s there was a public outcry about driver inconvenience. As a result, Congress eliminated the requirement, and the technology was removed from vehicles (New York Times, 1974).

Thus, policy makers must balance the two competing requirements such that alcohol detection systems should have a high rate of success in preventing alcohol impaired driving, but should not inconvenience drivers to the extent that there is an unwillingness to adopt the new technology. If the performance requirements are set too high, manufacturers may have difficulty meeting the requirements or the costs may be prohibitively high. On the other hand, if the requirements are set too low, ineffective devices may be implemented that may be less effective in preventing deaths. Each of these could result in low levels of public acceptance of the device so that their use will be inhibited. Determining what requirements to set that are effective enough, but not so demanding that it prevents manufacturers from developing and implementing the technology, is fundamentally a policy decision. This paper discusses the tools that can be used by policy makers to guide their decision-making process.

RISK-BASED ANALYSIS OF OUTCOMES

Risk-based analysis of outcomes is a method for calculating outcomes based on the intended use of a device, a model of the population of users that will use it, and is based on the sensitivity and the specificity of the device. Sensitivity of the alcohol detection device determines the likelihood that impaired drivers will not be allowed to drive (thus potentially saving lives – known as the benefit). On the other hand, specificity determines the likelihood that sober drivers (or drivers with BrACs/BACs under the set limit) will not be allowed to drive, thus inconveniencing the driver (referred to as harm for the purposes of the calculation). For a given sensitivity and specificity the net value of the use of that device (Benefit - Harm) can be calculated. In addition, general performance acceptance criteria can be set which would give device manufactures, automobile makers, government rule makers, and other interested parties guidance on the minimum performance requirements for a device, such that the device would have a positive impact for users.

As noted above, sensitivity and specificity are indicators of the success in the detection of whatever it is that the device is intended to detect. In the case of a breath-based alcohol detection system the sensor is designed to measure whether a driver's BrAC is above a predetermined limit. If the device reports the driver's BrAC is above the limit, the result is referred to as positive. If the device reports the driver's BrAC to be below the limit, the result is referred to as negative. If the drivers true BrAC is actually known through the use of a reference test device, it is possible to calculate the success of the device that is being evaluated. If the driver's BrAC is above the limit and there is a positive result this is referred to as a True Positive (TP). If on the other hand, the driver's BrAC is above the limit but there is a negative result this is referred to as a False Negative (FN). Sensitivity is the term used to describe the overall success rate with subjects that are impaired. This is calculated as the fraction of all impaired subjects who were tested that gave a True Positive result (Sensitivity = TP/(TP+FN)).

Conversely, if the driver has a BrAC below the limit and the device being tested gives a negative result this is a True Negative (TN). Similarly, if the driver has a BrAC below the limit, but the test device gives a positive result this is referred to as a False Positive (FP). Specificity is the term used for the overall success rate with test subjects that are unimpaired. This is calculated as the fraction of all unimpaired subjects that were tested which gave a True Negative result (Specificity = TN/(TN+FP)). Figure 1 provides an illustration of these concepts.

	Unimpaired/Allowed to operate the vehicle	Impaired/Not allowed to operate the vehicle
Test Result = Positive	False Positive (FP)	True Positive (TP)
Test Result = Negative	True Negative (TN)	False Negative (FN)
	Specificity = TN/(TN+FP)	Sensitivity = TP/(TP+FN)

Figure 1. Table of possible outcomes from the use of a device

CALCULATION OF THE BENEFIT AND THE HARM FROM THE USE OF A DEVICE

In a risk-based analysis of outcomes the value of a device is calculated as the benefit derived from using the device minus the harm done by using the device. Value = Benefit – Harm. As noted above, there are four possible outcomes: True Positive, True Negative, False Positive, and False Negative. Looking at each of these in turn it can be determined if these outcomes are beneficial, harmful, or neutral. In the case of the DADSS sensor a True Positive result is when a driver with a BrAC above the limit receives a positive test. Depending on the intended use of the device, a range of outcomes might occur; the car could be prevented from being put in gear, could be limited to lower top speed using a "limp mode," or the driver could simply be given a warning. All of these outcomes would be positive as they would likely result in a reduction in the number of impaired drivers. In contrast, in the case of a False Positive, resulting from a driver with a BrAC below the limit receiving a positive test, the outcome would result in inconveniencing the driver. As noted above, the amount of the inconvenience would depend on the implementation. A warning could potentially be merely an annoyance, but if the car is prevented from moving, the inconvenience would be much more significant. A True Negative, when the unimpaired driver receives a negative result is a neutral result, there is no benefit in this case, but also there is no inconvenience to the driver. Similarly, in the case of False Negatives the result is considered neutral. Currently, there is no in-vehicle system preventing impaired drivers from driving, so if an impaired driver receives a negative result from the sensor, the situation would be no different than the current situation. For risk-based analyses, in cases where a device is being proposed to replace an existing device, False Negatives have to be taken in to account because the test subject might have had a positive result which is no longer detected with the new device. However, since the DADSS system would be a new technology that is not replacing an existing system, this outcome is considered neutral for the purposes of this analysis. In summary, for a DADSS system, a True Positive is considered a beneficial outcome, a False Positive is a harmful outcome, and both True Negatives and False Negatives are considered neutral.

The overall benefit is the value of the benefit multiplied by the number of times that benefit is obtained, which is the total number of True Positive results for a given population. The sensitivity (i.e., the rate of True Positive results for the positive population, TP/(TP+FN)) is multiplied by the total number of the positive population. Thus, as shown in equation1, the benefit of using a DADSS system is Benefit = Value of Benefit * Sensitivity * Population of Drivers over the limit. Similarly, the overall harm of using the DADSS system is the cost of the harm multiplied by the number of times that the harm occurs. The total number of times the harm occurs is the total number of False Positive results in the population, times the size of the population of drives with drivers under the limit. Because Specificity (i.e., TN/(TN+FP) is the rate of True Negative results, the rate of harm is the complement of specificity, Rate of Harm = 1-Specificity. The resulting harm of using a DADSS sensor is thus Harm = Value of harm *(1-Specificity)* number of drives with drivers under the limit. The total value of using the device is shown in equation 1 below.

Value = Value of Benefit * Sensitivity * Population of impaired Drives –

Value of harm *(1-Specificity)* population of unimpaired Drives

Equation 1. The Value equation for an Alcohol Detection system using risk-based analysis of outcomes.

HOW SENSITIVITY AND SPECIFICITY RELATE TO ACCURACY AND PRECISION OF A DEVICE.

The risk-based analysis of outcomes value equation uses sensitivity and specificity to calculate the net benefit of a device. In practice, it is very burdensome to test individual devices for sensitivity and specificity because these are population-based statistics. Thus, in order to accurately measure the sensitivity and specificity of a device, large scale studies with large number of subjects would be required. As a result, policy makers will need to define other requirements in a device model specification, such as accuracy and precision, which can be measured in a laboratory setting. Therefore, it is important to understand how accuracy and precision relate to sensitivity and specificity.

Sensitivity and specificity are measures of the success rate of a device with respect to the entire population that are tested. This success depends on the accuracy and precision of the device. Accuracy is how close a given set of measurements are to their true value, while precision is how close the measurements are to each other, in other words the standard deviation. Because we expect the error of a device to be random, the entire population of the results for a specific test device will be normally distributed and the expected results can be graphed with a Gaussian, or bell curve. The curve will be wider or narrower depending on the precision and accuracy of the device. A more accurate and precise device will have a narrower distribution of results, while a less accurate and less precise device will have a wider distribution of results. Because devices are calibrated, it would be expected that, on average, the bias is zero, and that any inaccuracies in the individual devices would result in a general broadening of the curve for the whole population.

The assumption of the normal distribution of the test results is that the variability is due to the normal random error in measurements. This does not account for erroneous results that may occur if a device is not functioning as intended, for example, if the device breaks. It also does not account for non-random errors (for cause errors) that occur, for example, from an interfering substance. If results from either of these sources of error are a significant contribution to the total results, then the calculated sensitivity and specificity will not match the sensitivity and specificity actually measured in practice, undermining the validity of the model calculations. If, on the other hand, these non-random errors are not a significant contribution to the total number of results, for example because the device is very robust or has a mechanism for detecting device failures, and it does not report such results, then the non-random errors are not a major contributor to model and can be ignored. The DADSS program has done extensive testing of breath-based alcohol sensors in both laboratory and field settings. It has been found that the sensitivity and specificity as measured in natural settings with a wide range of subjects and in a variety of environmental conditions, matches what is predicted from the calculated accuracy and precision.

How accuracy and precision are related to sensitivity and specificity can be shown graphically. For example, in Figure 2, a curve of results that would be expected from repeated testing of a population of drivers with a true BrAC of 0.07%, (less than the 0.08% limit) is shown. The width of the curve depends on the precision of the device; that is, a more precise device will have a narrower distribution of results, while a less precise device will have a wider distribution of results. Since these drivers all have a BrAC of 0.07% they should be allowed to drive if they are registering a True Negative. If the result is found to be 0.08% and greater, the result is a False positive. Because specificity is the percent of the results that are True Negatives, the percent of the area under the curve to the left of 0.08% is the expected specificity of the device (90%) that has the precision and accuracy that resulted in the Gaussian curve in the graph.

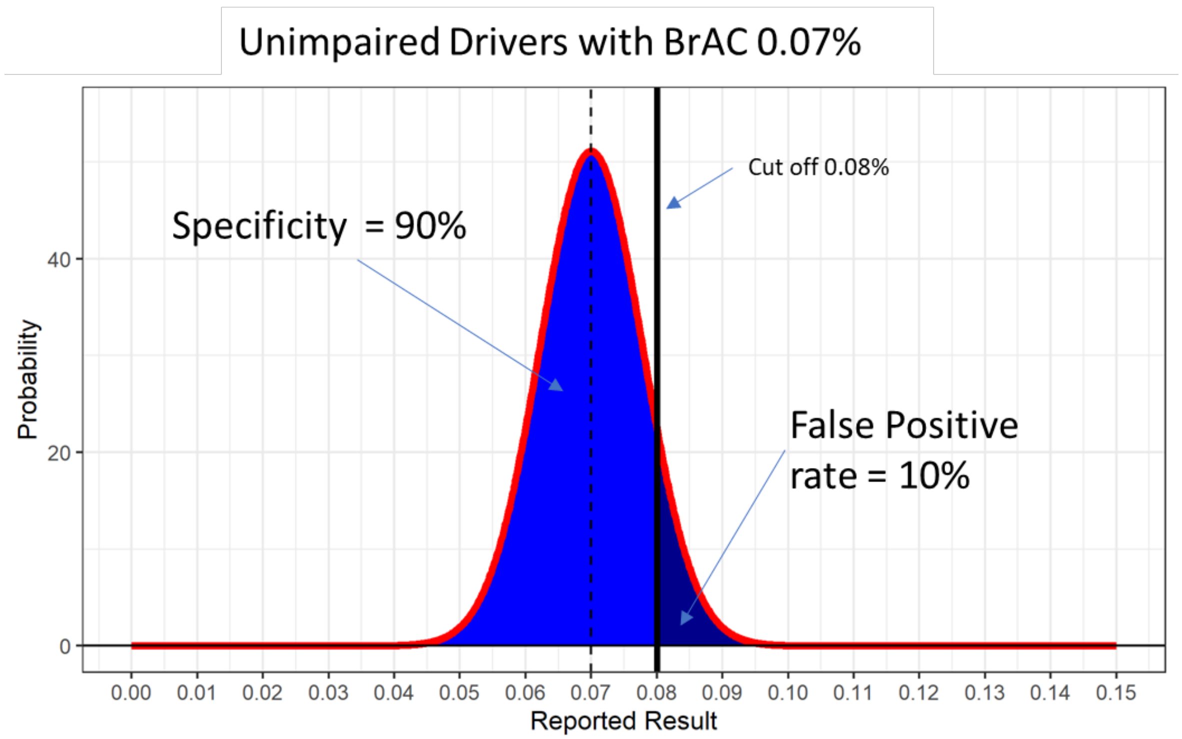

Figure 2. Unimpaired drivers with BrACs of 0.07%

In Figure 3, a curve of results that would be expected from repeated testing of a population of drivers with a true BrAC of 0.095% is shown. Because all these drivers have a BrAC of 0.095% they should not be allowed to drive. If the device reports a result of greater than 0.08% then the result is a True Positive, but if the result is less than 0.08% the result is a False Negative. Because Sensitivity is the percent of the results that are True Positives the percent of the area under the curve to the right of 0.08% (95%) is the expected sensitivity of a device that has the precision and accuracy that resulted in the Gaussian curve in the graph.

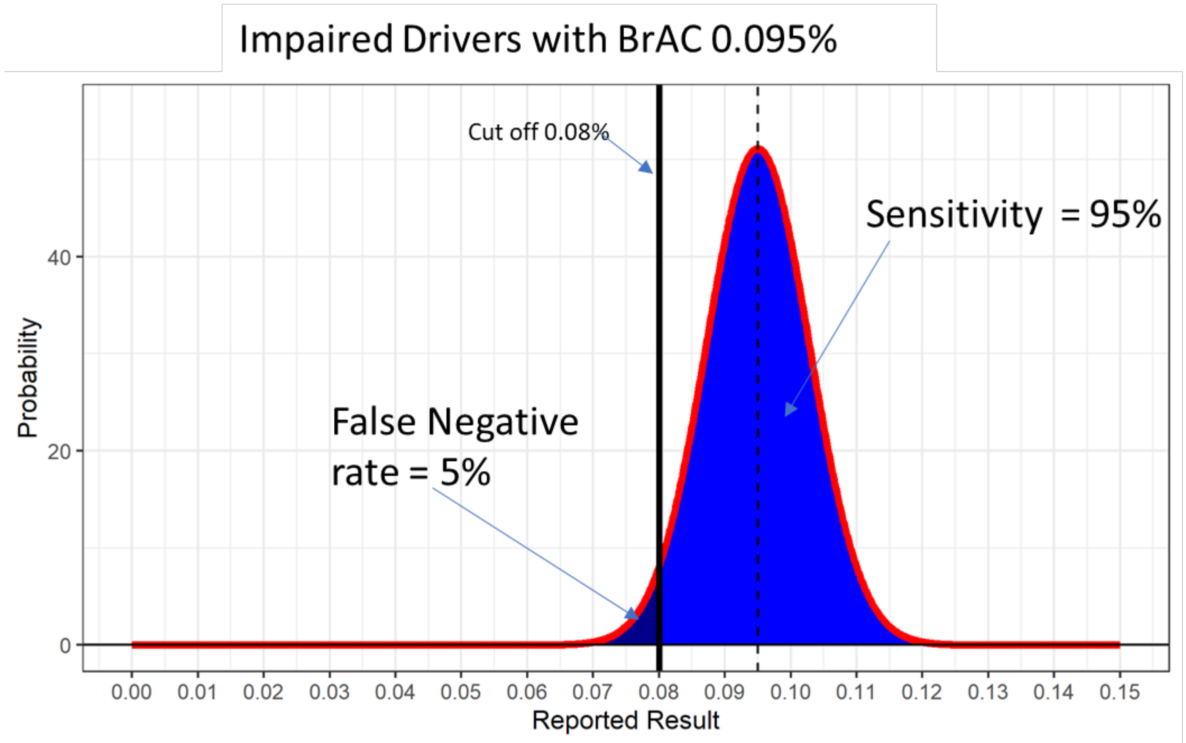

Figure 3. Impaired drivers with BrACs of 0.095%

Sensitivity and specificity depend enormously on the population that is being tested. For example, in Figure 2, the device has a specificity of 90% when testing a population of drivers with BrACs of 0.07% - a result very close to the cut off of 0.08% used to designate drivers as impaired. Using a device that had exactly the same accuracy and precision but instead testing a population of drivers that had BrACs of 0.05% (see Figure 4), should result in a much higher specificity because a much lower percentage of the results would be across the cutoff of 0.08%. Such a device would have a specificity of 99.99% with drivers that had a BrACs of 0.05%. As shown in Figure 5, with sober drivers (BrAC of 0.00%) the device would have a specificity of > 99.99999%.

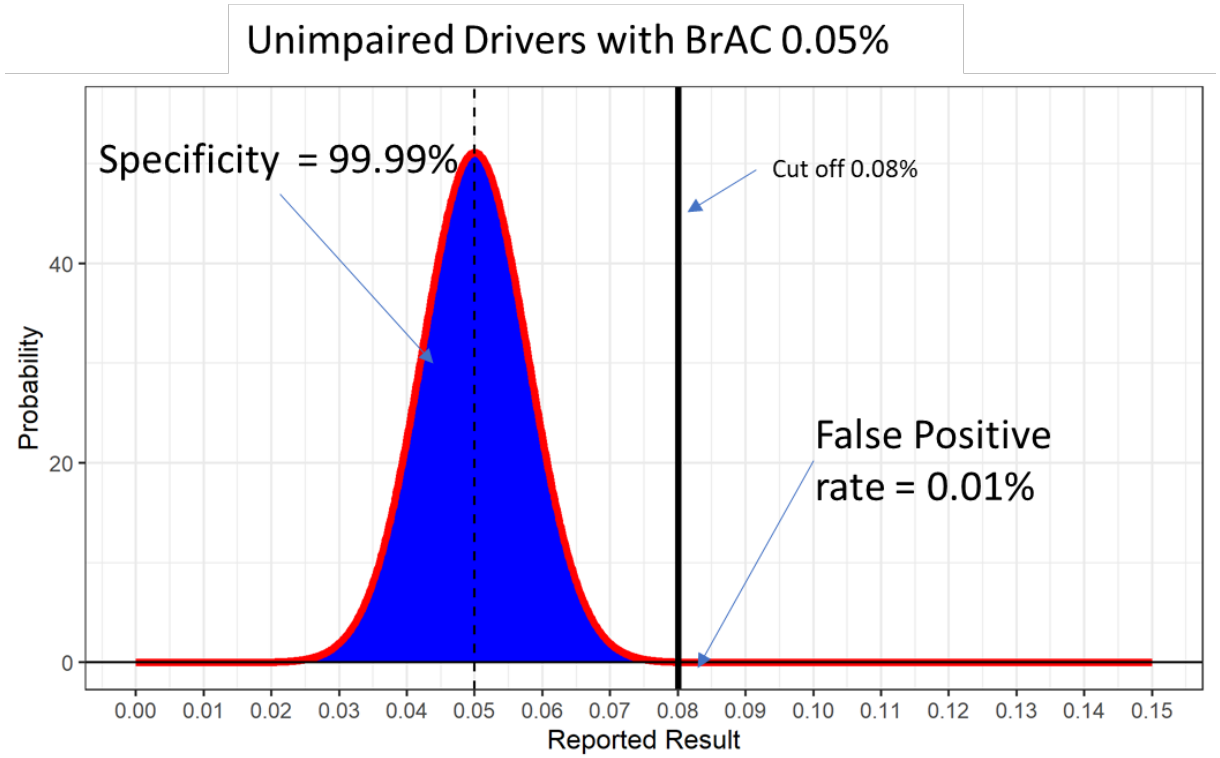

Figure 4. Unimpaired drivers with BrACs of 0.05%

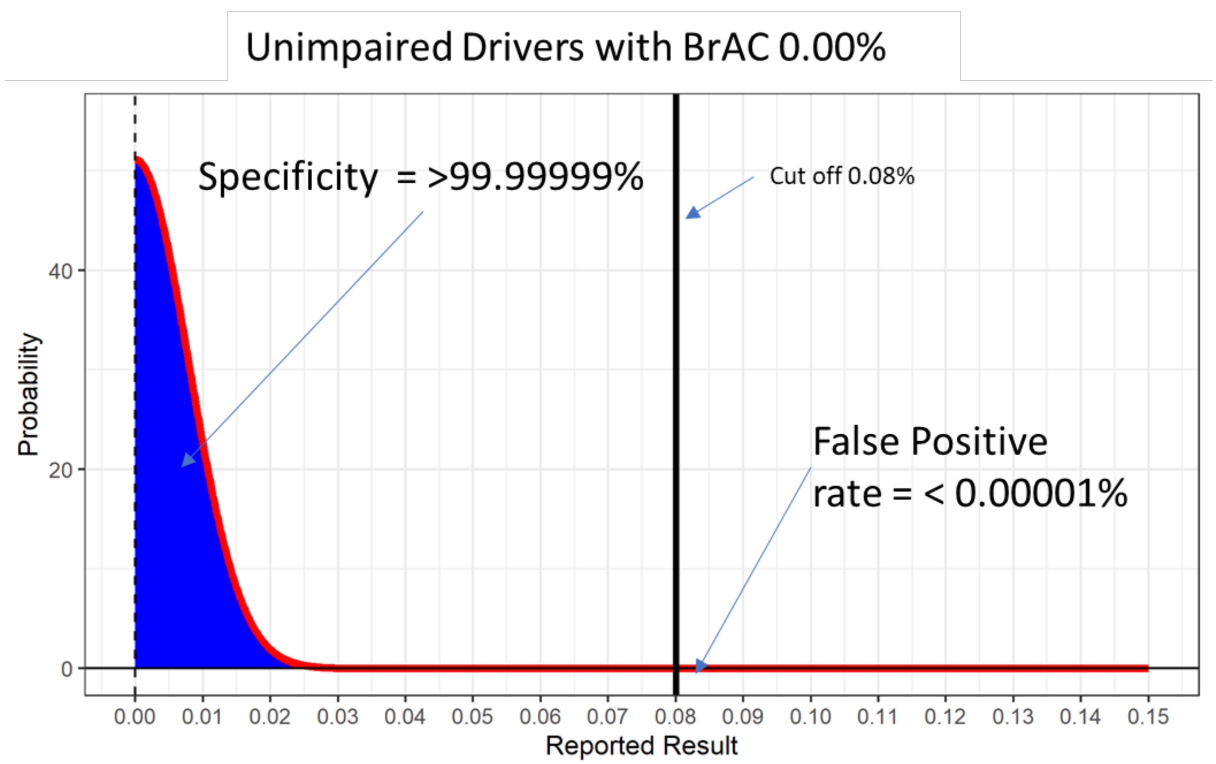

Figure 5. Unimpaired drivers with BrACs of 0.00%

An important concept of risk-based analysis of outcomes is that the sensitivity and specificity used to calculate the value to the population must be the sensitivity and specificity of the device for the whole population of people that will use the device. Thus, results used to calculate sensitivity and specificity must be weighted to account for the actual population of users. Sensitivity is additive by population if there are different populations, for example, drivers with different BACs. For example, if a device had a sensitivity of 50% with a given population A and 75% with a different population B, and if the A population was 10% of the total population and the B population was the remaining 90% the device sensitivity for the whole population is

50% (sensitivity of pop A) x 10% (fraction of total pop that is pop A)

+

75% (sensitivity of pop B) x 90% (fraction of total pop that is pop B)

=

72.5% overall sensitivity

Analyses of data from the 2013-2014 National Roadside Survey (Ramirez et al, 2016) combined with data from the 2017 National Household Travel Survey (FHWA, 2017), shows that there are approximately 600,000,000 drives each day in the U.S. Of these drives, there are an estimated 97.89% of trips that are driven sober. Of the drives by drivers who have some level of blood alcohol:

- 59% (1.25% of all drivers) are between 0.001 and 0.049 % BrAC,
- 15% (0.31% of all drivers) are between 0.05 and 0.079 % BrAC, and
- 26% (0.55% of all drivers) are over 0.08 %BrAC.

In the case of the example device which has a specificity of 99.99999% for sober drivers, 99.99% for drivers with 0.05% BrACs and 90% for drivers with a 0.07% BrACs, the overall specificity of the device can be calculated as follows:

Sober drivers are 98% of all drivers and will have a specificity of 99.99999%, therefore:

Specificity = 99.99999* # of Sober Drives = 97.89%

If we assume that the 1.25% of all drivers that are between 0.001 and 0.049 % BrAC have a specificity of 99.99%:

Specificity = 99.99 * # of all drives between 0.001 and 0.049% = 1.25%

and that the 0.31% of all drivers that are between 0.05 and 0.079 % BrAC have a specificity of 90%:

Specificity = 90 * # of all drives between 0.05 and 0.79% = 0.31%

Combining the results together:

(99.99999*97.89) + (99.99 * 1.25) + (90*0.31))/(97.89+1.25+0.31) = 99.97.

Thus, the device would result in an overall sensitivity of approximately 99.97%.

In reality, it would be expected that the specificity would be higher than this because the 1.25% of drivers who have BrACs of between 0.001% and 0.049 % will mostly be made up of drivers that have a BrAC that is significantly less than 0.05. However, since the exact distribution of drivers in this range is not available from the data, it is only possible to conservatively estimate the sensitivity as 99.99%. Similarly, many of the drivers in the 0.05% - 0.08% BrAC range likely will have specificity higher than 90% (the 0.07% BrAC specificity), because drivers with a BrAC of 0.05 will have a specificity of 99.99%. Since the actual distribution of BrACs of drivers in this range is unknown, the specificity could be estimated at 90%.

IMPLICATIONS OF THE RISK-BASED ANALYSIS OF OUTCOMES MODEL FOR DADSS SYSTEMS.

For an alcohol detection system to be effective it would need to be installed in a large number of vehicles. For example, policy makers could require their installation in all passenger cars. Based on this assumption, it is useful to look at the implication of the risk-based model for DADSS if it were to be implemented nationally. As can be seen from equation 1 on page 4, the overall value from the use of a device depends on the value of the benefit from a True Positive and the value of the harm from a False Positive. This depends on the intended use of the device, which is to say, what action is taken when a positive result is obtained. As described earlier, potential actions could include preventing motive power, a warning indicator, or a limp mode. For instance, an impaired driver could be prevented from putting the car into motion. This would maximize the value of the benefit because the impaired driver would be compelled to find alternative means of transportation and would be unlikely to get into crashes resulting in property damage, injuries, deaths, or other outcomes. However, this also means that the inconvenience to an unimpaired driver with a false positive would be greater, as the driver would be prevented from driving.

Since it is not possible to distinguish between a true positive and a false positive in terms of the device measurement, all positive results are treated the same. So drivers with false positives would also be prevented from driving, resulting in a major inconvenience. At the other extreme, an impaired driver could simply be given a warning. The overall value of the benefit from this system would be decreased because some fraction of impaired drivers would choose to drive in spite of the warning. However, the overall value of the harm would be significantly decreased because false positive drivers would not be inconvenienced to the same degree. It is possible to imagine other possibilities between these extremes, for example a car might be put into limp mode limiting its speed, or a system might send a message to a designated driver. The actual value of the benefit and the harm are subjective and will depend on the values of the person or group doing the evaluation, but in general when analyzing the value equation for a device, the question that should be asked is, does the benefit from using the system outweigh the harm? If the answer is yes, then the system should be put into use, and if no, other alternatives should be considered.

By way of an example, the net value can be calculated of an alcohol detection system using a risk-based model of outcomes assuming a device prevents impaired drivers from driving. A significant benefit from preventing impaired drivers from driving is deaths prevented. There are approximately 36,000 traffic fatalities in the U.S. annually, of which one third involve an alcohol impaired driver. Thus, if the device prevents impaired driving, the potential benefit if the device is implemented in all cars is 12,000 deaths prevented annually. The harm from this implementation comes from drivers inconvenienced (i.e., drivers inconvenienced per million drives.

Because of the high value of preventing deaths even a modest sensitivity can result in large benefits. For example, if an alcohol detection system prevented drivers that had a positive reading at or above a BrAC of 08% from driving, and that device only had a sensitivity of 50% it would still potentially prevent 6,000 deaths:

0.50 sensitivity * 12,000 deaths annually = 6,000 deaths prevented annually.

However, based on the enormous number of drives each year by sober drivers, a high specificity would still result in a large amount of harm. If the system prevented drivers with lower BrACs from driving if it detected a BrAC at or above 0.08% and the sensitivity was 99.9%, this would result in 1,000 drivers inconvenienced per million drives:

Inconvenienced Drivers per million Drives *(1-99.9%) * 1,000,000 drives = 1,000 drivers

In the U.S. there are an estimated 600,000,000 drives per day of which approximately 98% are made by sober drivers. With a rate of 1,000 inconvenienced drivers per million drives there would be 588,000 unimpaired drivers prevented from driving each day:

Prevented from driving * 1,000/1,000,000 * sober drives per day (98% * 600,000,000) = 588,000 drivers prevented from driving per day.

Thus, for the net value of the alcohol detection system to be positive, the specificity must be high enough to limit the inconvenience to drivers.

POTENTIAL WAYS TO IMPROVE SENSITIVITY OR SPECIFICITY BY TREATING TEST RESULTS DIFFERENTLY

Using the risk-based analysis of outcomes, it is possible to determine sensitivity and specificity thresholds at which the net benefit is positive. This in turn can be used by policy makers to set performance criteria for alcohol detection systems. Once those performance criteria are established, system manufacturers will need to ensure that their systems meet those thresholds. Understanding what these goals are, system manufactures can design devices that are optimized for the intended use. Obviously, manufacturers will want to make systems that are as accurate and precise as practical, but in addition, the way in which the devices are implemented can improve either the sensitivity or specificity of the device. For a device with a given accuracy and precision it would be expected that the specificity and sensitivity of that device would be affected for the population of all drivers by treating the test results differently. One option is that for a given definition of impaired, the threshold at which a positive result is reported can be modified. Another option is to combine more than one test result into a single outcome decision, which also can affect the sensitivity and specificity.

Increasing the BrAC threshold

The specificity of the device can be improved by increasing the threshold at which drivers are reported as positive. For example, setting the threshold for a positive report to 0.10% BrAC would dramatically increase the specificity. Considering drivers with a BrAC of 0.07% as shown in Figure 2, the device with the precision shown has a specificity of 90% with a threshold of 0.08%. However, moving the threshold to 0.10% changes the specificity for drivers with a 0.07% BrAC to 99.99, and drivers with a BrAC of 0.05% and 0.0% would have a specificity of > 99.999999. This would improve the specificity for the entire population of drivers from 99.97% to 99.9995%.

However, the change in the threshold comes at the cost of sensitivity. If the threshold for reporting a positive result is set at 0.10%, impaired drivers that had a BrAC between 0.08% and 0.10% would be reported positive at a very low rate. For example only 10% of drivers with a BrAC of 0.09% would be reported as positive. Drivers with a BrAC greater than 0.10% make up the majority of alcohol impaired drivers involved in fatal crashes so this might be an acceptable trade off. It can be estimated that the sensitivity of a device that has the accuracy and precision that have been used in our examples would decrease from 96% to 71% if the threshold for reporting a positive result is changed from 0.08% to 0.10%. Using risk-based analysis of outcomes the net benefit of these sorts of changes can be calculated. Using equation 1, our example device would drop from a benefit of 96% *12,000 deaths prevented, that is 11,520 deaths prevented, to only 71% *12,000 deaths prevented, that is 8,520 deaths prevented. However, the harm would be reduced from (1- 99.97%), that is, 300 drivers inconvenienced per million drives, to (1-99.9995%), or 5 drivers inconvenienced per million drives (see Figure 6).

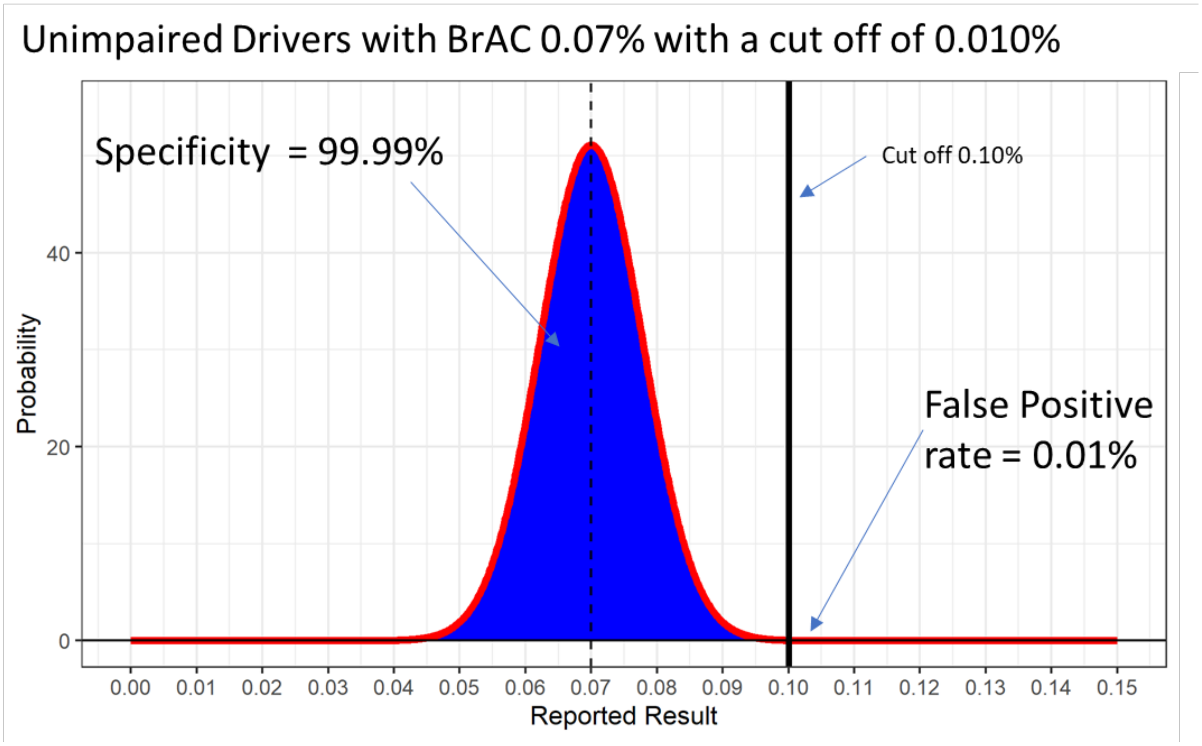

Figure 6. Unimpaired drivers with a BrAC 0.07% with a cut off of 0.01%

Decreasing the BrAC threshold

If sensitivity is more important than specificity, lowering the threshold has the effect of increasing sensitivity at the cost of specificity. For example, using the same process as described above, if the threshold is lowered from 0.08% BrAC to 0.06% BrAC the sensitivity of our example device increases from 96% to 99.98% increasing the benefit from 11,520 deaths prevented to 11,998 deaths prevented. However, specificity is lowered from 99.97% to 99.78% increasing inconvenienced drivers from 300 per million drives to 2134 per million drives.

Multiple breath tests

The second technique examined is the effect of combining more than one breath test into a single outcome. Averaging the results of multiple samples reduces the variability of the reported results. In general, the variability between outcomes (the standard deviation) is reduced by the square root of the number of samples averaged together. This reduction in the variability improves both the sensitivity and the specificity of a device but it comes at the cost of having to do multiple independent tests. In the case of a breath alcohol device, measuring multiple breaths could potentially increase inconvenience for the driver. At a minimum it would mean a somewhat longer time before a result could be delivered, potentially increasing the time before a car could be started. However, if the focus is on improving either sensitivity or specificity, retests could be undertaken only in certain conditions with the result that a much smaller number of drivers are inconvenienced. As we have demonstrated, because of the enormous number of drives every day it is very important to have very high specificity to avoid having the harm of the device outweigh the benefit. For example, if a device has a specificity of 99.9%, one in a thousand drivers would be inconvenienced. If a device asked for a retest sample for all drivers that were above 0.08% on initial analysis and only returned a positive result if both test results had a measured BrAC above 0.08%, then the sensitivity would be dramatically improved.

This effect is multiplicative, so the rate of harm (1-specificity) would improve from (1-99.9%) = 0.1% to (1-99.9%) * (1-99.9%) = 0.0001% or 99.9999% specificity (see Figure7). The rate of inconvenience would improve from 1 in 1000 drivers inconvenienced to 1 in 1,000,000 drivers inconvenienced. This does come at the cost of sensitivity. The impact on sensitivity is also multiplicative, so if the device had a 96% sensitivity the sensitivity would be reduced to

(96%) * (96%) = 92%. This is because some drivers who had a true BrAC greater than 0.08% could have a second measured result lower than 0.08% which would be considered a negative result.

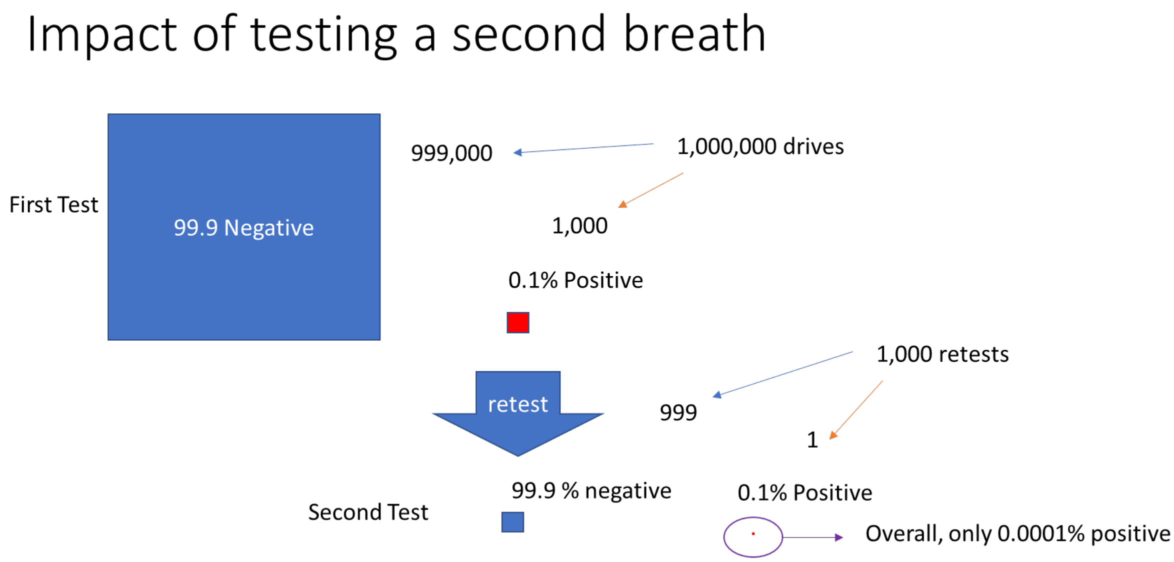

Figure 7. Impact of testing a second breath compared to the first.

ANALYSIS OF POTENTIAL OUTCOMES WITH THE CURRENT GENERATION 3.3 BREATH BASED DADSS SYSTEM

For the DADSS prototype generation 3.3 directed breath alcohol detection system, the sensitivity and specificity have been measured in multiple human subjects driving tests across a wide range of conditions and different alcohol levels. When using a limit of 0.08% as the definition of impaired, and weighting to account for the different levels of BrAC expected in the US population, the calculated sensitivity is 91.6% and the specificity is 99.93%. If we analyze the outcomes of using this device nationwide to prevent impaired driving we calculate a potential benefit of 10,992 deaths prevented (12,000 x .916 = 10,992) with a potential harm of 700 drivers inconvenienced per million drives ((1-0.9993) * 1,000,000). Using the method of retesting positive results as outlined above, there is a potential benefit of 10,069 deaths prevented (12,000 x 0.916 x 0.916 = 10,069) with potential harm reduced to only 0.5 drivers inconvenienced per million drives ((1-0.9993) * (1-0.9993) * 1,000,000).

CONCLUSIONS

The DADSS system, currently under development, has the potential to save thousands of lives a year by preventing drunk drivers from driving their vehicle. It is important to set device specifications that not only save as many lives as possible, but limit driver inconvenience to the extent possible. To calculate the appropriate specifications, a model for using a risk-based analysis of outcomes to determine the net value of implementation of a DADSS type system has been described. This should be a valuable tool that can be used by policy makers to guide their decision-making process in setting performance criteria for such devices. Specifically the models show that even modest sensitivity can have a significant positive benefit to the population because of the large number of deaths as a result of alcohol impaired driving. However, a very high specificity is required to prevent the negative impact from outweighing the benefit because of the enormous number of sober drives that take place every day.

That being said, there are many life-saving technologies currently in use in the vehicle fleet, such as frontal air bags, that also have potential negative effects on the driving population, and yet they are still acceptable to drivers. Initially, frontal air bags were developed to provide protection in frontal crashes for unbelted drivers and as a result were more powerful. However, these air bags were found to be overpowered resulting in injuries and deaths to occupants who were too close to them when they deployed. Frontal crash tests were modified to allow air bags to deploy with less force, resulting in dramatic reductions in air bag injuries and deaths, but no loss of protection for front seat passengers (Ferguson and Schneider, 2008).

In summary, it is important to balance the positive and negative effects of new technology so that the maximum lives can be saved without potentially risking a backlash among the driving population. This paper has suggested ways in which these two can be balanced appropriately using a risk-based analysis of outcomes to determine the potential sensitivity and specificity, and hence accuracy and precision, of the device.

REFERENCES

American Automobile Association. 2022. Automatic Emergency Braking Performance in the Context of Common Crash Scenarios 2022, American Automobile Association Inc. https://newsroom.aaa.com/wp-content/uploads/2022/09/Research-Report-2022-AEB-Evaluation-FINAL-9-26-22.pdf

Cicchino J.B. 2017. Effectiveness of forward collision warning and autonomous emergency braking systems in reducing front-to-rear crash rates. Accident Analysis and Prevention, 99:142–152.

Federal Highway Administration. 2017. 2017 National Household Travel Survey, U.S. Department of Transportation, Washington, DC. Available online: https://nhts.ornl.gov.

Ferguson, S.A. 1996. Update on air bag performance in the United States: Benefits and problems. Airbag 2000+: Third Fourth International Symposium on Sophisticated Car Occupant Systems, 7:7-17, Fraunhofer- Institut Fur Chemische Technologie (ICT) Karlsruhe, Germany.

Ferguson, S.A. 1998. An update on the real-world experience of passenger air bags in the United States. Fourth International Symposium on Sophisticated Car Occupant Systems, 2-1:2-18, Fraunhofer- Institut Fur Chemische Technologie (. ICT) Karlsruhe, Germany.

Ferguson, S.A., Greene, M., Lund, A.K. 1995. Driver fatalities in 1985-1994 air bag cars. Insurance Institute for Highway Safety, Arlington, VA.

Ferguson S A, Traube E, Zaouk A, Strassburger R. 2011. Driver Alcohol Detection System For Safety (DADSS) – Phase I Prototype Testing And Finding. Paper Number 11-0230. Proceedings of the 22nd International Technical Conference on the Enhanced Safety of Vehicles.

Ferguson, S.A. & Schneider, L.W. 2008. An overview of frontal air bag performance with changes in frontal crash-test requirements: Findings of the Blue Ribbon Panel for the Evaluation of Advanced Technology Air Bags. Traffic Injury Prevention, 9:421-431.

Gross, A. 2022. Braking Bad: Automatic Emergency Braking Absent When You Need It Most. AAA. https://newsroom.aaa.com/2022/09/braking-bad-automatic-emergency-braking-absent-when-you-need-it-most/

Insurance Institute for Highway Safety. 2022. Alcohol and Drugs. https://www.iihs.org/topics/alcohol-and-drugs#by-the-numbers. Arlingon, VA.

Lund, A.K., Ferguson, S.A. 1995. Driver fatalities in 1985-1993 cars with airbags. Journal of Trauma, Vol. 38, 469-475.

National Center for Statistics and Analysis. 2018. Traffic Safety Facts, 2018 data. Occupant protection. DOT HS 812 967. National Highway Traffic Safety Administration. Washington DC.

National Highway Traffic Safety Administration. 2022. NHTSA Releases 2020 Traffic Crash Data. https://www.nhtsa.gov/press-releases/2020-traffic-crash-data-fatalities. Washington, DC

New York Times. 1974. Congress Clears Auto Safety Measure Eliminating Seat Belt Interlock System. https://www.nytimes.com/1974/10/16/archives/congress-clears-auto-safety-measure-eliminating-seat-belt-interlock.html

Ramirez A, Berning A, Kelley-Baker T, Lacey JH, Yao J, Tippetts AS … Compton R. 2016. 2013-2014 National Roadside Study of Alcohol and Drug Use by Drivers: Alcohol Results (Report No. DOT HS 812 362). Washington, DC: National Highway Traffic Safety Administration.

Zaouk A K, Willis M, Traube E, Strassburger R. 2019. Driver Alcohol Detection System for Safety (DADSS) – A non-regulatory approach in the research and development of vehicle safety technology to reduce alcohol-impaired driving - Status Update. Paper Number 19-0260. Proceedings of the 26th International Technical Conference on the Enhanced Safety of Vehicles.