BTSCRP Project BTS-17

Appendices A and B

The Behavioral Traffic Safety Cooperative Research Program (BTSCRP) is sponsored by the Governors Highway Safety Association and funded by the National Highway Traffic Safety Administration. BTSCRP is administered by the Transportation Research Board (TRB), part of the National Academies of Sciences, Engineering, and Medicine. Any opinions and conclusions expressed or implied in resulting research products are those of the individuals and organizations who performed the research and are not necessarily those of TRB; the National Academies of Sciences, Engineering, and Medicine; or BTSCRP sponsors.

Appendix A. Survey

This survey is to gain insight into how agencies conduct High Visibility Enforcement (HVE) campaigns, such as how behaviors to target are selected, what data are collected, and how HVE are evaluated. The project is conducted as part of BTS-17: Determining the Effectiveness of Combined High Visibility Enforcement through the Behavioral Traffic Safety Cooperative Research Program.

This survey is designed to take 5-10 minutes. If you have any further questions feel free to reach out to the project PI Shauna Hallmark at shallmar@iastate.edu.

We thank you for your time in taking this survey as part of this effort.

Contact Information

We may have additional questions and wish to follow up. If this is something you'd agree to, please fill in your name and e-mail below.

Name (optional)

Agency

E-mail (optional)

Phone (optional)

How frequently does your agency conduct or participate in high visibility enforcement (HVE) campaigns?

- ◯ Infrequently or not at all
- ◯ Have conducted 1 or more in the last 3 years
- ◯ Conduct one or more annually

How are road user behaviors selected to target in HVE campaigns? (select all that apply)

- ☐ Priority goals in Statewide Highway Safety Plans (SHSPs)
- ☐ Other formal safety plans
- ☐ Analysis of crash data
- ☐ Other (Describe)

What is the primary focus of your HVE campaigns? (select all that apply)

- ☐ Driving under the influence
- ☐ Speeding
- ☐ Pedestrian safety

- [ ] Seat belt use
- [ ] Distracted driving
- [ ] Other (Describe)

Do you conduct HVE campaigns that include more than one focus (combined)?

- ○ Yes, typically combine
- ○ No

Which data are collected for evaluation or reporting purposes for HVE campaigns? And when are these data collected? (select all that apply)

	Before campaign	During campaign	After campaign	Not collected
Officers involved, hours, or other resources used for campaign	☐	☐	☐	☐
Crashes	☐	☐	☐	☐
Number of vehicles stopped	☐	☐	☐	☐
Citations	☐	☐	☐	☐
Arrests	☐	☐	☐	☐
Public surveys to assess visibility of HVE or changes in driver impression of risk	☐	☐	☐	☐
Number of paid media spots (i.e., radio ad)	☐	☐	☐	☐
Number of unpaid media (i.e., mentions on morning radio shows)	☐	☐	☐	☐

	Before campaign	During campaign	After campaign	Not collected
Observed metrics such as seat belt or cell phone use	☐	☐	☐	☐
Speed	☐	☐	☐	☐
Other (Describe)	☐	☐	☐	☐

What is the rationale for collection of this data? (select all that apply)

- ☐ Best data to show impact of HVE campaign
- ☐ Agency resources (i.e., cost, time)
- ☐ This is what has been collected historically
- ☐ Required (by agency, laws)
- ☐ Difficulty in collecting data
- ☐ Data cannot be collected due to agency/legal reasons

What additional data would you collect if resources/other were not a constraint? (select all that apply)

- ☐ Number of officers involved, hours, or other resources used for campaign
- ☐ Crashes
- ☐ Number of contacts with drivers
- ☐ Number of citations
- ☐ Number of arrests
- ☐ Public surveys to assess visibility of HVE or changes in driver impression of risk
- ☐ Instances of paid instances of media (i.e., number of radio spots)

- [ ] Instances of spontaneous (unpaid) media (i.e., mentions on morning radio shows)
- [ ] Observed metrics such as seat belt or cell phone use
- [ ] Other (Describe)

How is the effectiveness of HVE evaluated? (select all that apply)

- [ ] Only raw numbers are provided (i.e., number of arrests)
- [ ] Information is shown in charts or graphs
- [ ] Simple statistics (i.e., before and after comparison, t-test)
- [ ] More advanced statistical methods
- [ ] Other (Describe)

Why are the current evaluation methods used? (select all that apply)

- [ ] Best way to share data
- [ ] Method is most cost efficient or other methods would be too resource intensive
- [ ] Limited by data that can be collected
- [ ] Analysis is limited by agency expertise
- [ ] Other (Describe)

Appendix B. Common Data Collected and Metrics Summary

This appendix summarizes common data collected and metrics used to implement and evaluate HVE campaigns. Methods that may be appropriate to evaluate the effectiveness of HVE campaigns are summarized in the following sections. The methods are presented in terms of their technical, data, and computational requirements; their strengths; their limitations (e.g., potential biases); the considerations involved in their deployment; example outcomes; and some basic (noncomprehensive) method details.

B.1 Descriptive Statistics

B.1.1 Technical, Data, and Computational Requirements

This method has few requirements. Once measures of interest are identified and collected, simple summaries such as mean, standard deviation, minimum, maximum, count, etc. are computed. The computations can easily be performed by hand or by using a calculator, spreadsheet, etc. Thus, the time and costs associated with descriptive statistics are low compared to other methods.

B.1.2 Strengths

Few resources are required beyond the data collected on the measures of interest. Additionally, the results are easy to communicate and interpret.

B.1.3 Limitations

This method is purely descriptive. It does not provide detailed insights into effectiveness or the relationships between the intervention and outcomes of interest.

B.1.4 Considerations

Descriptive statistics are useful for describing basic items such as the number of tickets issued, the mean age of distracted drivers cited, etc. These statistics can provide context for additional in-depth analysis. They can also be used to describe the scope of the HVE implementation.

B.1.5 Example Outcomes for Analysis

- Number of citations issued
- Number of commercials aired per day
- Mean age and genders of drivers not wearing seat belts
- Self-reported frequency of cell phone use while driving (from surveys)

B.1.6 Method Details

Descriptive statistics are methodologically simple. Data are collected on the measures of interest, and then basic statistics, including mean, standard deviation, minimum, maximum, count, range, etc., can be computed and summarized. For example, the mean value of a variable is typically computed as follows:

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

where μ = the mean value, N = the number of observations, and x_i = the individual observation values for the variable of interest.

Additionally, the standard deviation of a sample is computed as follows:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N - 1}}$$

Where σ = the standard deviation. The other variables are as previously defined.

B.1.6.1 Utility for HVE Evaluation

Descriptive statistics are useful for displaying information about HVE campaigns such as the number of citations issued or a comparison of cell phone use before and after the campaign. They are most useful for reporting purposes.

Simple statistics are not able to assess the actual effectiveness of HVE campaigns. They are also not likely to detect differences due to combined HVE campaigns.

B.2 z-Test or t-Test

B.2.1 Technical, Data, and Computational Requirements

The z-test and t-test are simple statistical methods for hypothesis testing. They require data for the measures of interest (assuming they have continuous values and are normally distributed or are measures of proportions), a statistical hypothesis to test (both null and alternative hypotheses), and statistical tables or software to compute p-values from the z- or t-statistics.

B.2.2 Strengths

The methods are simple and easy to implement. They can be used to compare the outcome with a specified value (e.g., to determine whether seat belt use was greater than 90%) or to compare two samples (e.g., seat belt use rates in similar areas that implemented/did not implement a campaign related to seat belt use).

B.2.3 Limitations

These methods are only valid for data that are continuous and approximately normally distributed or for comparing proportions, and a null hypothesis and an alternative hypothesis must be specified. These methods do not account for common sources of bias such as selection bias, aggregation bias, regression to the mean, endogeneity, etc. These methods are often used on outcomes that do not meet the basic method assumptions underlying the methods.

B.2.4 Considerations

It should be determined whether the descriptive statistics resulting from this method will provide useful information for decision-makers or other stakeholders. This could include providing context for additional analysis performed.

B.2.5 Example Outcomes for Analysis

- Percent seat belt use (a documented case of this is in Retting et al. 2018)
- Comparison of the mean speeds before and after an HVE campaign related to speeding

B.2.6 Method Details

There are multiple versions of these tests, including single-sample tests (where the data are tested against a specified value), two-sample tests (where the data are collected in two areas and/or time periods and compared), and paired samples (where the data are collected for two conditions for the sample people/locations).

The single-sample t-test evaluates whether the mean value of a sample is statistically different from a specific value. The t-test can be used with both small and large samples and converges to the z-test when the sample size approaches infinity. The null hypothesis (H_0) for the t-test states that the mean value is equal to some value (μ_0). The alternative hypothesis (H_1) states that the mean value does not equal/is greater than/is less than μ_0. The t-statistic standardizes the difference between μ_1 and μ_0. The degrees of freedom for the single-sample t-test are (df) = n-1. The equation for computing the single-sample t-test is as follows:

$$t = \frac{\mu_1 - \mu_0}{\frac{\sigma}{\sqrt{N}}}$$

After calculating the t-statistic, the p-value is computed using the t-statistic, degrees of freedom, and alternative hypothesis (which is used to determine whether it is a single-tail or two-tail p-value).

The key difference between the t-test and z-test is that the z-test does not require using degrees of freedom, which makes computing the p-values slightly easier. Common statistical practice is that if the sample has at least 30 observations, a z-test is used.

The t-test and z-test can also be modified for use in comparing independent proportions. For instance, the measure of interest may be the percentage of occupants that used seat belts. If the proportions are measured before and after an HVE campaign targeting seat belt use, the two-sample z-test for comparison of proportions would be computed using the following:

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{N_1} + \frac{p_2(1-p_2)}{N_2}}}$$

where p_1 = the proportion using a seat belt after the campaign, p_2 = the proportion using a seat belt before the campaign, N_1 = the number of data points for the period after the campaign, and N_2 = the number of data points for the period before the campaign.

B.2.7 Utility for HVE Evaluation

The z-test and t-test can be used to compare differences between metrics collected before, during, and after HVE campaigns and are able to show whether changes are statistically significant. They are best suited to metrics such as speed, seat belt use, or changes in the number of violations.

The z-test and t-test are not able to account for issues such as regression to the mean or other factors that may have influenced changes in metrics. For instance, adverse weather conditions during the campaign may have impacted speed independently of the campaign. They are also not likely to detect differences due to combined HVE campaigns.

B.3 Naïve Before-and-After Comparisons

B.3.1 Technical, Data, and Computational Requirements

This method is considered to be a simple statistical method. It requires basic data (the measures of interest) from before and after implementation of the intervention. Then, basic comparisons are made for the data before and after implementation. Thus, the requirements for data, technical knowledge, and computational equipment/software are relatively low.

B.3.2 Strengths

This method is easy to compute, can show benefits (e.g., reductions in crashes), and has low data needs.

B.3.3 Limitations

This method does not control for omitted variables, regression to the mean, or other potential sources of bias. It also requires that data for the measures of interest be collected/available from both before and after the intervention. Additionally, it is possible that the measure of interest, such as crash counts, can have no observations in one of the periods (before or after), which causes problems for evaluating the effectiveness of the intervention.

B.3.4 Considerations

It should be determined whether the use of a naïve before-and-after method gives adequate results for decisions that need to be made or whether additional biases need to be accounted for.

B.3.5 Example Outcomes for Analysis

- Comparison of the percent of traffic traveling more than 5 mph over the speed limit
- Comparison of crash frequencies
- Comparison of seat belt use rates
- Comparison of self-reported drowsy driving
- Comparison of the percent of vehicles properly yielding to pedestrians

B.3.6 Method Details

The comparison of data from before and after implementation of an intervention (i.e., an HVE campaign) utilizes appropriate statistical methods. For instance, a simple z-test can be used to compare the percent of traffic traveling 5 mph or more over the speed limit before and after an enforcement campaign. Examples of statistical methods that can be used for naïve before-and-after analyses based on the type of outcome (continuous, proportions, counts, etc.) are provided in Table B-1.

Table B-1. Example statistical methods for naïve before-and-after analysis based on outcome type.

Outcome Type	Example	Statistical Methods (Noncomprehensive)
Continuous	Comparison of mean speeds	t-test, z-test
Proportions (binomial)	Percent seat belt use	t-test, z-test, McNemar’s chi-squared test, RR, odds ratio
Count	Crash frequency	RR, odds ratio
Multinomial or ordinal	Crash severity	Chi-squared test, Mann-Whitney U test, Wilcoxon signed rank sum test, Fisher’s exact test

B.3.7 Utility for HVE Evaluation

Naïve before-and-after comparisons include and are similar to z- and t-tests. They can be used to compare differences between metrics collected before, during, and after HVE campaigns and are able to show whether changes are statistically significant. They are best suited to metrics such as speed, seat belt use, or changes in the number of violations.

Naïve before-and-after comparisons are not able to account for issues such as regression to the mean or other factors that may have influenced changes in metrics. For instance, adverse weather conditions during the campaign may have impacted speed independently of the campaign. They are also not likely to detect differences due to combined HVE campaigns.

B.4 Before-and-After Comparisons with Comparison Group

B.4.1 Technical, Data, and Computational Requirements

The requirements for this method are similar to those of the naïve before-and-after method but with the addition of a comparison group that did not receive the intervention (i.e., the HVE campaign). The need for data from the comparison group increases the data requirements for this method relative to simpler methods. This analysis method is also slightly more difficult to implement due to the need to account for the comparison group.

B.4.2 Strengths

The use of a comparison group reduces the likelihood of bias due to factors correlated with the intervention (e.g., increased traffic in the after period). This method can also allow adjustments for time-based differences and patterns that are not possible with the naïve before-and-after method (e.g., changes in driver behaviors based on seasons of the year). Additionally, this method can include additional variables (e.g., traffic volumes, weather) in the evaluation to adjust for their impacts on the estimates.

B.4.3 Limitations

This method does not account for some biases such as regression to the mean and selection bias. It also requires data from both before and after the intervention.

B.4.4 Considerations

It should be determined whether the use of a before-and-after study with a comparison group will give adequate results for the decisions that need to be made or whether there's a need to account for additional biases. Additionally, the variables that should be adjusted for (beyond the key measures of interest or outcomes) need to be identified.

B.4.5 Example Outcomes for Analysis

- Comparison of the percent of traffic traveling more than 5 mph over the speed limit
- Comparison of crash frequencies
- Comparison of seat belt use rates
- Comparison of self-reported drowsy driving
- Comparison of the percent of vehicles properly yielding to pedestrians

B.4.6 Method Details

The before-and-after method with a comparison group uses the comparison group as a baseline for changes between the before and after periods that would have occurred had the intervention not taken place. Using this baseline, the effect of the intervention is then evaluated. This can range from use of the baseline for simple adjustments to detailed methods that use predictive models developed using the comparison group and then applied to the data that received the intervention. Details on the use of this method for evaluating crash counts are readily available in the American Association of State Highway and Transportation Officials *Highway Safety*

Manual and other sources (e.g., the U.S. DOT publication *A Guide to Developing Quality Crash Modification Factors*).

Another application for this method would be a comparison of the percentage of drivers that properly yield to pedestrians. A simple analytical method for evaluating this outcome using a before-and-after analysis with a comparison group would be to use *RR*. The equations for this are as follows:

$$R_{after} = \frac{N_{proper,after}}{N_{total,after}}$$

$$R_{before} = \frac{N_{proper,before}}{N_{total,before}}$$

$$R_{after,comp} = \frac{N_{proper,after,comp}}{N_{total,after,comp}}$$

$$R_{before,comp} = \frac{N_{proper,before,comp}}{N_{total,before,comp}}$$

$$RR = \frac{\frac{R_{after}}{R_{before}}}{\frac{R_{after,comp}}{R_{before,comp}}} = \frac{R_{after} \cdot R_{before,comp}}{R_{before} \cdot R_{after,comp}}$$

where:

- $N_{proper,after}$ = the number of proper yields to pedestrians in the intervention area in the after period
- $N_{total,after}$ = the total number of yields to pedestrians in the intervention area in the after period
- $N_{proper,after,comp}$ = the number of proper yields to pedestrians in the comparison area in the after period
- $N_{total,after,comp}$ = the total number of yields to pedestrians in the comparison area in the after period
- $N_{proper,before}$ = the number of proper yields to pedestrians in the intervention area in the before period
- $N_{total,before}$ = the total number of yields to pedestrians in the intervention area in the before period
- $N_{proper,before,comp}$ = the number of proper yields to pedestrians in the comparison area in the before period
- $N_{total,before,comp}$ = the total number of yields to pedestrians in the comparison area in the before period

The standard error (*SE*) for the log-transformed RR is computed as follows:

$$SE = \sqrt{\frac{R_{after}}{R_{before}} + \frac{R_{after,comp}}{R_{before,comp}}}$$

The 95% confidence interval ($CI_{95\%}$) for the RR is then computed as follows:

$$CI_{95\%} = \exp(\ln(RR) \pm 1.96SE)$$

If the 95% confidence interval includes the value 1, then the change associated with the intervention is not statistically significant at the 95% confidence level (i.e., the p-value is greater than 0.05).

B.4.7 Utility for HVE Evaluation

The before-and-after method with a comparison group is similar to a naïve before-and-after comparison but is able to account for factors not related to the intervention that may have impacted the campaign metrics. For instance, if adverse weather conditions were present during a campaign, changes in the control group would account for those conditions. If, for example, speeds in the campaign area decreased by 5 mph and speeds in the control areas decreased by 2 mph, approximately a decrease of 3 mph could be attributed to the campaign. The before-and-after method with a comparison group can be used to compare differences between metrics collected before, during, and after HVE campaigns and is able to show whether changes are statistically significant. This method is best suited to metrics such as speed, seat belt use, or changes in the number of violations.

The before-and-after method with a comparison group is not able to account for specific factors that may have impacted metrics. It is also not likely to detect differences due to combined HVE campaigns.

B.5 Empirical/Full Bayes Before-After Comparison

B.5.1 Technical, Data, and Computational Requirements

This method builds on the before-after method with a comparison group and requires, in some cases, additional data. This method requires additional variables not associated with the intervention in order to develop predictive models using the comparison data. The predictions are then used on the intervention data to conduct a before-after analysis while making additional adjustments. Thus, this method entails additional data, technical, and computational requirements compared to other methods.

B.5.2 Strengths

This method adjusts for multiple sources of bias, including selection bias and regression to the mean.

B.5.3 Limitations

Statistical software is required to estimate the prediction models, and advanced statistical expertise is required to perform the evaluation. The results may be difficult for stakeholders to understand if not presented appropriately.

B.5.4 Considerations

It should be determined whether the use of a simpler method will produce adequate results for the intervention based on the needs of the analysis. The availability of adequate resources (including technical expertise and software) should also be considered.

B.5.5 Example Outcomes for Analysis

- Comparison of the percent of traffic traveling more than 5 mph over the speed limit
- Comparison of crash frequencies
- Comparison of seat belt use rates
- Comparison of self-reported drowsy driving
- Comparison of the percent of vehicles properly yielding to pedestrians

B.5.6 Method Details

As with the other before-after methods, the specific computations used depend on the type of outcome being evaluated. For the purposes of the report, the method described in this section outlines the traditional process used for an empirical Bayes before-after analysis for crash frequency. The method involves the following steps:

Step 1. Estimate a SPF using a comparable reference group. SPFs are estimated using count regression (e.g., negative binomial regression).

Step 2. Estimate the expected number of crashes for each year in the before period for each treated entity using the SPF.

Step 3. Compute the sum of SPF predictions for each treated entity ($N_{pred,before}$) in the before period.

Step 4. Estimate the expected number of crashes in the before period for each treated entity ($N_{before,EB,i}$) and the associated variance for the expected number using the empirical Bayes adjustments shown in the following equations:

$$N_{before,EB,i} = w_i N_{pred,before} + (1 - w_i) N_{observed,before}$$

where $N_{observed,before}$ = the sum of the reported crashes on treated entity i in the before period, $N_{pred,before}$ = the predicted number of crashes on treated entity i in the before period (using the SPF), and w_i = a weighting factor estimated using the following:

$$w_i = \frac{1}{1 + \alpha N_{pred,before}}$$

where $\alpha =$ the overdispersion parameter from the SPF.

Also, compute the variance of the empirical Bayes prediction as follows:

$$Var\left(N_{before,EB,i}\right) = (1 - w_i)N_{before,EB,i}$$

Step 5. Estimate the number of crashes for each year in the after period for each treated entity using the SPF. Calculate the sum of the SPF-predicted crashes for the after period ($N_{pred,after}$).

Step 6. Calculate the ratio of SPF-predicted crashes from the before ($N_{pred,before}$) and after ($N_{pred,after}$) periods for each treated entity using the following:

$$R_i = \frac{N_{pred,after}}{N_{pred,before}}$$

Step 7. Estimate the number of crashes that would have occurred had the treatment not been implemented ($N_{expected,after}$) and the variance of the estimate of $N_{expected,after}$ using the following:

$$N_{expected,after} = \sum R_i N_{before,EB,i}$$

$$Var(N_{expected,after}) = \sum R_i^2 Var(N_{before,EB,i})$$

Step 8. Estimate the CMF (denoted as θ) and the variance of the treatment effect using the following:

$$\theta = \frac{N_{observed,after}}{N_{expected,after}\left(1 + \frac{1}{N_{expected,after}}\right)}$$

$$Var(\theta) = \frac{\theta^2\left(\frac{1}{N_{observed,after}} + \frac{Var\left(N_{observed,after}\right)}{\left(N_{observed,after}\right)^2}\right)}{\left(1 + \frac{Var\left(N_{observed,after}\right)}{\left(N_{observed,after}\right)^2}\right)^2}$$

where $N_{observed,after} =$ the sum of reported crashes on treated entity *i* in the after period.

The 95% confidence interval for CMFs developed using this method are commonly computed using the following:

$$CI_{95\%} = \theta \pm 1.96\sqrt{Var(\theta)}$$

As with the RR, if the 95% confidence interval contains the value 1, the result is considered not statistically significant at the 95% confidence level. Values smaller than 1 indicate reductions in crashes, while values larger than 1 indicate increases in crashes associated with the intervention.

B.5.7 Utility for HVE Evaluation

Empirical/full Bayes before-after comparison is able to account for factors such as regression to the mean and other factors that may have influenced changes in the metrics. This method is also more likely than other methods to be able to detect the impact of combined HVE campaigns.

B.6 Logistic Regression

B.6.1 Technical, Data, and Computational Requirements

Logistic regression is a traditional statistical method that predicts probabilities of a dichotomous outcome (i.e., categorical outcome with two possible values). The outcome variable has values of 0 or 1, indicating whether the outcome is observed or not (i.e., 0 = false, 1 = true). As a rule, the outcome with the fewest number of observations (0 or 1) should have at least 10 observations for each variable included in the model. For example, a data set with an outcome that is dichotomous where there are 5,000 observations but only 50 of those have an outcome value of "1" can, as a rule, include up to five predictor variables in the regression model.

Logistic regression also is difficult to compute without statistical software. While there are many statistical packages that can be used to estimate these models, interpretation of the results is not always simple. Thus, it is important that the results be carefully described and interpreted by someone with the appropriate training and experience to ensure the results are not misused.

B.6.2 Strengths

This method can be used to evaluate the effects of interventions on binary, or dichotomous, outcomes while accounting for additional variables (reducing the likelihood of omitted variable bias). It can also be used for supporting models based on utility theory (i.e., a type of decision theory).

B.6.3 Limitations

The results of these models are not always easily interpreted. Odds ratios are typically used, although marginal effects can also be used. Additionally, logistic regression is sensitive to outliers.

B.6.4 Considerations

Logistic regression can be considered whenever the measure of interest being evaluated has binary values and there are other variables (i.e., predictors) that can or should be adjusted/accounted for in the analysis.

B.6.5 Example Outcomes for Analysis

- Seat belt use
- Whether the vehicle is insured
- Speeding versus not speeding
- Cell phone use

B.6.6 Method Details

Logistic regression is a probabilistic method that predicts the probability of a binary outcome using the following equation:

$$P(y = 1|\beta, X) = \frac{\exp(\beta X)}{1 + \exp(\beta X)}$$

where β = a vector of estimated coefficients, X = a vector of predictor variables, and $P(y = 1|\beta, X)$ = the probability the outcome has a value 1, given β and X.

The probability of the outcome having a value 0 ($P(y = 0|\beta, X)$) is as follows:

$$P(y = 0|\beta, X) = 1 - P(y = 1|\beta, X) = 1 - \frac{\exp(\beta X)}{1 + \exp(\beta X)}$$

The coefficients are estimated using maximum likelihood methods. The quality of the model is evaluated using one or more of many existing methods and measures including the following:

- Accuracy
- Precision
- Recall
- Specificity
- The area under the curve from a receiver operating characteristic plot
- Pseudo R^2
- AIC
- BIC
- Validation methods (such as cross-validation and k-fold cross-validation)

B.6.7 Utility for HVE Evaluation

Logistic regression is well suited for evaluation of survey results. This method is also more likely than other methods to be able to detect differences in surveys with multiple barriers.

B.7 Quasi-Induced Exposure

B.7.1 Technical, Data, and Computational Requirements

This method is based on police-reported crash data involving multiple vehicles. However, it can also be extended to single-vehicle crashes for cases where the intervention of interest only impacts specific crash types and can be reasonably argued that it does not impact other crash types (i.e., independence). Once the crash data are available, the computations and process range from simple calculations to the use of logistic regression.

B.7.2 Strengths

The method accounts for exposure, even when exposure metrics (such as vehicle miles traveled) are not available.

B.7.3 Limitations

The method requires identifying “at-fault” drivers in multi-vehicle collisions. This may be difficult, depending on the intervention of interest.

B.7.4 Considerations

If the exposure metrics of interest are available, then this method may not be necessary or appropriate. However, it is useful for many applications where exposure is not available or is known to have significant error in the exposure data.

B.7.5 Example Outcomes for Analysis

- Seat belt use
- Fatigued driving
- Distracted driving
- Cell phone use

B.7.6 Method Details

The quasi-induced exposure method is based on the concept that within police-reported crash databases an analyst can use subpopulations that are statistically independent of an intervention, characteristic of interest, etc., to adjust for or estimate exposure metrics. It is well documented that failing to account for exposure (e.g., VMT) results in biased estimates (AASHTO 2010). However, it is often difficult, infeasible, or impossible to directly measure exposure for the group of interest in an analysis, such as pedestrian crossing counts, VMT with/without seat belt usage, etc. Thus, this method provides an analytical solution to account for and estimate these measures of exposure. For details of variations of the quasi-induced exposure method, along with example applications, see Carr (1969), Lyles et al. (1991), Stamatiadis and Deacon (1997), Yan et al. (2005), Keall and Newstead (2009), Jiang and Lyles (2010), and Sharmin et al. (2020).

The method relies on either: (1) identifying drivers at-fault and not-at-fault in multi-vehicle crashes or (2) identifying crash types associated with (i.e., statistically dependent) the intervention and other crash types that are statistically independent (i.e., not influenced) by the intervention. Given that enforcement agencies are intimately familiar with how the crash data is collected and reported, the identification of these conditions is likely to be the most accurate when they are implementing the method or are involved in it.

After identification of the different conditions (i.e., at-fault/not-at-fault or crash types dependent/independent of the intervention), the statistical analysis often involves simple computations of RR or IRR, binary logistic regression, or including inferred exposure measures into other statistical models (e.g., negative binomial regression).

B.7.7 Utility for HVE Evaluation

Quasi-induced exposure is best suited to crash analyses. It is most useful when exposure values are difficult to obtain. It is less likely to provide a method to identify differences in combined HVE methods.

B.8 Regression with Spatial and/or Temporal Adjustments

B.8.1 Technical, Data, and Computational Requirements

These models require spatial and/or temporal information for inclusion in the regression methods. The spatial correlations are often handled using Bayesian or other, similar, methods. The temporal aspects can be handled using multiple approaches, although one of the most common is using time series approaches. Thus, there are significant computational and technical requirements for these models. Software specific to these models (such as the R packages: Surveillance, McGLM, hhh4adon, hhh4contacts, Carst, BUGS, R-INLA, and CARBayesST) and hardware that can support the computations (adequate processors, memory, etc.) often require specialized training and expertise.

B.8.2 Strengths

The method adjusts for spatial and temporal correlations that are not captured in other methodological frameworks.

B.8.3 Limitations

The method can be resource intensive and require data that are not readily available. Additionally, it may be difficult for most decision-makers to interpret and understand.

B.8.4 Considerations

It should be considered whether there are likely to be significant temporal trends and/or significant spatial correlations that need to be captured.

B.8.5 Example Outcomes for Analysis

- Speeding-related crashes per road segment or intersection (spatial and temporal)
- Seat belt use (spatial and temporal)
- Distracted driving (spatial and temporal)
- Speeding by time of day, day of week, etc. (temporal)
- Number of citations issued (spatial and temporal)

B.8.6 Method Details

These regression models allow for greater use of information when available. For example, the outcome one week may be a good predictor for the outcome in the following week. Thus, when spatial or temporal data are available, and may improve understanding of the intervention of interest, spatiotemporal models can be used to leverage the additional information. For detailed examples of several spatiotemporal regression models applied to count outcomes, see González-Pérez et al. (2021).

B.8.7 Utility for HVE Evaluation

Regression with spatial and/or temporal adjustments is able to account for temporal and spatial variations. As a result, it is well suited to evaluate impacts of different methodologies.