Integrating Information from Prior Research into a Before-After Road Safety Evaluation through Bayesian Approach and Data Sampling

Ravi Kumar Mangi, T Prakash Lazarus and Dr. Divakar C

Abstract: Before-after road safety evaluation (B/A) to measure safety treatment effect is a key mission in road safety management, and has fueled considerable research. However, previous research in this area has been overwhelmingly dedicated to safety model estimation with less emphasis on other methodological issues. As a result, there continues to be uncertainty in the validity of treatment effect estimates. This study seeks, with innovative paradigms, a systematic solution by solidifying methodologies for every essential step of a thorough B/A process to secure its ultimate validity. Methodologies of data sampling and processing, and before and after model development, both vital procedures that have been sampling approach to select reference groups is established in the context of B/A application. A post-assignment propensity score matching method is developed in order to further eliminate statistical bias while the treatment effect indicator – collision reduction ratio (CRR) – is being estimated. Rather than focus on single safety model development as is common in traffic safety research, this study seeks all viable knowledge by employing various safety measures including collision and safety surrogates, by embedding several adaptable random distributions, by fitting models through both "Frequentist" and "Bayesian" approaches, and by exploring a variety of model forms and components. Accordingly, the output of this study is not a "best" single model, but rather an amalgamation of diversified models. The diversity is shown to be attractive in terms of information conveyed, especially for the B/A process. Finally, this study succeeds in finding a methodology to integrate all of the diverse knowledge sources. The Bayesian Model Averaging (BMA) method is investigated and developed to integrate a variety of statistical significant models without exclusion, in forging a unified model. All methodologies explored and developed in this study are essential to secure the validity of the B/A process. As important, they are substantially connected to each other. Should one method be deficient, the remaining steps cannot guarantee validity of B/A process. As a whole, these methodologies, if properly developed and applied, constitute a logical chain to estimate treatment effect with minimal errors and high validity.historically neglected, are investigated. A pre-test data.

I. INTRODUCTION

The implementation of before after road safety evaluation (before-after evaluation or B/A) to measure the effects of safety remedies is one of the two key missions in road safety analysis. Another is network screening, which identifies sites with potential for safety treatment (Persaud et al., 2010a). As one of two pivots in this domain, B/A has fueled considerable research work, with a large body of published literature. The empirical Bayes (EB) method (Hauer, 1985; Persaud et al., 2010a), for instance, is one of the most well established approaches in before-after evaluations to date. Recently, the full Bayesian (FB) approach has also generated many efforts as a viable option for the conduct of before-after evaluations (Persaud et al., 2010a; Lan et al., 2009; Lan, 2010; Yanmaz-Tuzel and Ozbay, 2010; El-Basyouny and Sayed, 2010). A B/A process utilizes four datasets: treated group, reference population, reference group and comparison group. The treated group is the group receiving a certain treatment. The reference population, in traffic safety practice, is the total collection of intersections or roadway segments of a jurisdiction with same features as the treated group before treatment. For example, for the treatment of protected left turn provision, the reference population is the total signalized intersections in a city without protected left turn, or, for the treatment of median barrier, the reference population is the entire highway network in a state/province without median barrier. Reference groups are the legitimate samples selected from reference population. The comparison group is a subset of the same type sites as treated groups and is usually used to compare observed before and after period collisions between the comparison and treated groups to enhance treatment effect estimation. Figure 1-1 illustrates the relationship among these different datasets.

The working mechanism of B/A utilizes the above four datasets in different ways. Among them, the treated group (TG) is the core of B/A processing, since the final result of B/A, namely the treatment effect, is estimated by comparison for the treated group itself between observed collisions after treatment and the "postulated" collisions without treatment (Hauer, 1985; Persaud et al., 2012a; Persaud et al., 2012b).

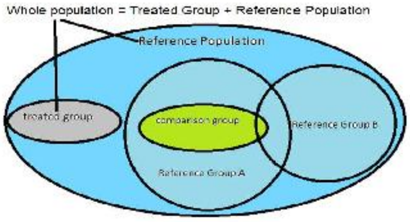


Figure: Different Data Groups Involved in Before-after Evaluation for Protected Left Turn Control

Reference population and reference groups are utilized to develop safety models, which are introduced to offer referential information for before-after evaluations. Reference Group A and B at Figure 1-1 represent different samples obtained from one reference population. Both the EB and FB methods, the best- established approaches, require a reference population (RP) or reference group (RG).The comparison group is not necessarily present for all

B/A approaches. It is only for comparison group (C/G) method (Hauer, 1985) that a comparison group is a necessity. For the currently applied EB and FB methods, the conventional practice does not require the physical presence of a comparison group; instead, postulated collisions from safety models replace the role of observed collisions in the comparison group in the B/A process.

II. MOTIVATION AND RESEARCH OBJECTIVES

The general goal of this dissertation research is to achieve a higher GOC for B/A processes, or, as expressed in conventional statistical terms, to minimize the bias of before-after evaluations by securing the internal and external validity of before-after evaluations. The external validity stems from the appropriate selection of a reference group and a referential knowledge base while the internal validity is achieved by appropriate assignment of treated groups (Dattalo, 2010). Since the assignment of treated groups, or network screening, is not a topic of this dissertation, the internal validity is instead pursued through a post-assignment matching process on the reference group. In realizing this goal, the above-mentioned five issues as outlined in Figure 1-2 are to be addressed. To supplement the previous conceptual descriptions, this section will accordingly re-investigate these five issues from a statistical perspective, so as to focus on the dissertation research objectives.

A. Pre-Test Data Sampling on Select Local Reference Group

A statistical perspective, the B/A process is a "test" (Dattalo, 2010). This means that the before period is the "pre-test" stage, so that selection of an reference group is statistically a "pre-test" data sampling procedure.The local reference group has multiple roles. First, any model needs to be calibrated before being used in the local context, as recommended by the Highway Safety Manual (HSM) (AASHTO, 2010). Hence, local data collection will at least require a calibration database. However, one should always consider developing a local model whenever possible since this is pertinent to local traffic system characteristics. That is to say, local data can be used as the basis for local model development as well.

B. Identification of Multi-level Structure for Locally Developed Models

The current HSM framework for collision prediction is actually not directly derived from a fully specified SPF equation. Instead, it is built on a base SPF and several collision modification factors (CMFs) as follows:

$$N_{predicted} \square N_{spf\,x} \square (CMF_{1x} \square CMF_{2\,x} \square \ldots\ldots \square CMF_{yx}) \square C_x$$

where

NN $spf\,x$

$N_{predicted}$ = predicted average crash frequency for a specific year for site type x,= predicted average crash frequency determined for the base conditions of the SPF developed for site type x,CMF_{yx}=crash modification actors specific to an SPF for site type x, and C_x = calibration factor that adjusts the SPF to local conditions for site type x.

C. Utilizing Knowledge from Surrogates towards Collision Measures

The intuitive statistical solution for utilizing knowledge from surrogates towards collision measures is a regression model that associates the former with the latter, which will be explored in this dissertation. Moreover, an alternative solution in the event that no statistical model is available will be investigated as well. Generally, this is a rank-based algorithm that conveys the ranking of the surrogates to the ranking of collision measures and finally estimates the relevant collisions in accordance with their rankings.

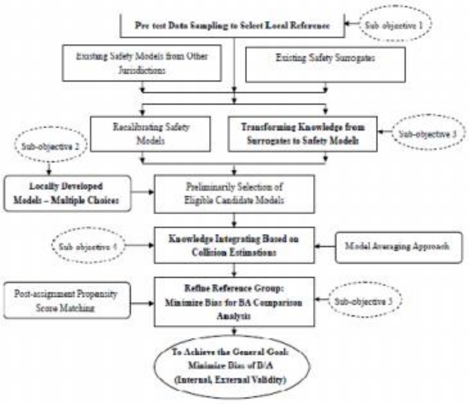


Figure: Research Flowchart of Dissertation – Objectives and Statistical Approaches

III. ORGANIZATION OF DISSERTATION

This dissertation is composed of 9 chapters. A literature review and sample data summaries are provided in the relevant chapters. All processes are developed by the use of the SAS statistical package (SAS Institute Inc., 2012).pre-test data sampling to select a local reference group. This is a two-stage procedure. The first stage is to determine an approximate sample size controlled by a given Type I error rate (probability of incorrectly identifying a statistically significant effect, denoted as "α"), and a given model power error (probability of not identifying a statistically significant effect when one exists, denoted as "B"; Accordingly, (1- B) is called as "model power level", usually simplified as "power") (Dattalo, 2010). The result of the first stage provides the basis for the second stage, which comprises data sampling to select a local reference group. In this dissertation, sequential stratified sampling is applied to achieve objective such as being representative of the entire population, coverage of all sub-categories of the population, and adequacy for significance of local safety model development.local safety performance function (SPF) development with a variety of different models. Chapter 3 develops SPFs via standard approach while Chapter 4 explores diversified SPF development, which is aimed at including a variety of local SPFs in order to capitalize on the diversity of knowledge sources. This part of research will explore and favor multi-level SPFs but retains other types of SPFs, including calibrated HSM models and single-level full local models, as optional choices. The preferable multilevel local models has first level average daily traffic (ADT)-only model with shape parameters and intercepts which are all functions of sub-hierarchical models with other covariates, including items with a local context.

The utilization of knowledge from safety surrogates as a substitute for collision measures. There are two scenarios: with or without statistical models. The ideal scenario is the former when there is adequate data to support surrogate-based safety model estimation. This dissertation investigates the speed of modern roundabouts as the sample surrogate and speed-based roundabout safety models are developed. Considering data inadequacy, an indirect approach is investigated for using the ranking of safety surrogates as a substitute for ranking of collisions, and for quantitatively estimating the collisions according to these rankings.

A model averaging algorithm to integrate all collision estimations from different sources and approaches. In condideration of the heterogeneous nature of candidate models or estimations, a Bayesian model averaging (BMA) algorithm is applied. This algorithm seeks a unified model in which the coefficients are respectively estimated from equivalent coefficients of all candidate models based on a weighted averaging mechanism for which loglikehoods are employed as weights. The multi-level model structure introduced in Chapter 3, facilitates the averaging of calibrated and locally developed models.Finally, all knowledge sources are merged together and a unique integrative model is formed. : the post-assignment statistics to refine the efficiency of reference groups in B/A processes. It is retrospective to the procedure ibut moves forward, with a post-assignment statistical process on the reference groups to further lower the comparison bias of the B/A. In this dissertation, an algorithm known as propensity score matching is applied. The principle is to measure the heterogeneity of the reference group versus the treated group and then apply calibrations accordingly for referential estimation. Finalizes the last step of B/A process - treatment effect estimation - by an application example and then provides some brief discussion, comparing the pros and cons of the dissertation methodologies investigated versus more conventional approaches, from both conceptual and statistical perspectives.

IV. SAMPLE FACILITY, MEASURE AND DATA

For a B/A process, the central feature is a safety treatment applied to a certain type of facility. Considering data availability and rationality, this research has selected urban 4-legged signalized (4SG) intersection as the sample facility, and introduced left turn protection of signalized intersection (also called exclusive left turn signal) as the sample treatment. Hence the B/A process in this dissertation study has the following key characteristics:

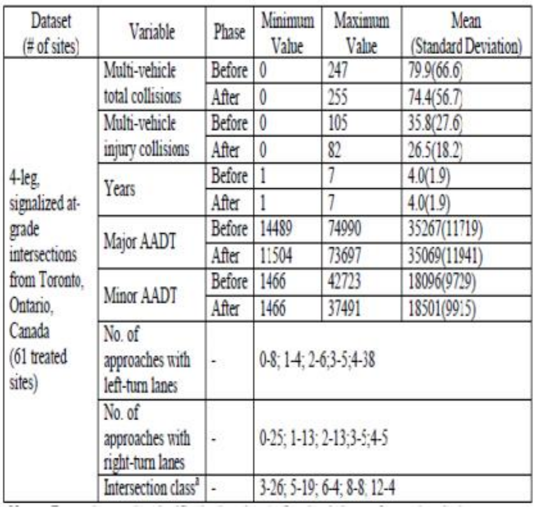
1.Phasing before - permitted left turn control,

2.Phasing after – protected, protected/permitted left turn control,

3.TG – intersections originally with permitted left turn, then converted to protected, protected/permitted left turn control,

4.RP - Reference population, all other 4-legged signalized intersections except for treated group in a city or region, and

5.RG – a sample extracted from a reference population, applied as representative of the population.

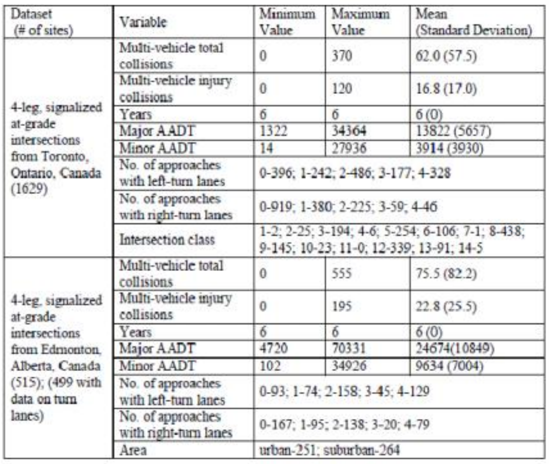


Dataset (# of sites)	Variable	Phase	Minimum Value	Maximum Value	Mean (Standard Deviation)
4-leg, signalized at-grade intersections from Toronto, Ontario, Canada (61 treated sites)	Multi-vehicle total collisions	Before	0	247	79.9(66.6)
		After	0	255	74.4(56.7)
	Multi-vehicle injury collisions	Before	0	105	35.8(27.6)
		After	0	82	26.5(18.2)
	Years	Before	1	7	4.0(1.9)
		After	1	7	4.0(1.9)
	Major AADT	Before	14489	74990	35267(11719)
		After	11504	73697	35069(11941)
	Minor AADT	Before	1466	42723	18096(9729)
		After	1466	37491	18501(9915)
	No. of approaches with left-turn lanes	-	0-8; 1-4; 2-6;3-5;4-38		
	No. of approaches with right-turn lanes	-	0-25; 1-13; 2-13;3-5;4-5		
	Intersection class[a]	-	3-26; 5-19; 6-4; 8-8; 12-4		

Table: Summarized Statistics of Treated Group Data

A. Summary Statistics of Reference Population Data

The reference population is the entire collection of 4SG intersections except for the 61 treated sites in Toronto. This group comprises 1629 sites. In addition, the entire collection of 4SG intersections in Edmonton, Alberta, Canada was also selected as a supplemental reference population per the requirements of the methodological aspects of the research. Table 2 provides the summary statistics of these data.



Dataset (# of sites)	Variable	Minimum Value	Maximum Value	Mean (Standard Deviation)
4-leg, signalized at-grade intersections from Toronto, Ontario, Canada (1629)	Multi-vehicle total collisions	0	370	62.0 (57.5)
	Multi-vehicle injury collisions	0	120	16.8 (17.0)
	Years	6	6	6 (0)
	Major AADT	1322	34364	13822 (5657)
	Minor AADT	14	27936	3914 (3930)
	No. of approaches with left-turn lanes	0-396; 1-242; 2-486; 3-177; 4-328		
	No. of approaches with right-turn lanes	0-919; 1-380; 2-225; 3-59; 4-46		
	Intersection class	1-2; 2-25; 3-194; 4-6; 5-254; 6-106; 7-1; 8-438; 9-145; 10-23; 11-0; 12-339; 13-91; 14-5		
4-leg, signalized at-grade intersections from Edmonton, Alberta, Canada (515); (499 with data on turn lanes)	Multi-vehicle total collisions	0	555	75.5 (82.2)
	Multi-vehicle injury collisions	0	195	22.8 (25.5)
	Years	6	6	6 (0)
	Major AADT	4720	70331	24674(10849)
	Minor AADT	102	34926	9634 (7004)
	No. of approaches with left-turn lanes	0-93; 1-74; 2-158; 3-45; 4-129		
	No. of approaches with right-turn lanes	0-167; 1-95; 2-138; 3-20; 4-79		
	Area	urban-251; suburban-264		

Table: PRE-Test Data Sampling To Select

V. LOCAL REFERENCE GROUPS

An often neglected, but essential step in traffic safety analysis is data sampling. Researchers in the traffic safety field tend to pour their efforts into model development or comparison analysis, with less focus on the procedure for data sampling. In before-after evaluation (B/A) procedures, reference groups are used to develop SPFs. While there is plenty of research on SPF development, very few studies have examined the actual selection of reference groups. Practitioners and researchers tend to include the entire reference population or to arbitrarily select any available data sources without carrying out statistical sampling. The drawbacks of this ad-hoc approach are clear: on the one hand, given that the whole reference population is applied, the data items are not always available or worthwhile to collect, especially when many items need field surveying or manual inputs; on the other hand, if arbitrary selection is applied, the reference group may not be sufficient enough to conduct the next step in modeling, or may not be consistently representative of the reference population.

This chapter aims to address these drawbacks by exploring and establishing a data sampling and a data assignment mechanism that are specifically designed to work for B/A processes. This investigation is comprised of three steps.The first step, described in Section 4.1, is to review data sampling and data assignment related literature and then to recommend the methods most suitable for the data sampling for this dissertation. The second step, described in Section 4.2, is to estimate an appropriate sample size for the reference group by controlling the modeling power error level. The third step, described in Section 4.3, is based on the outcome from the previous two steps and conducts random data sampling procedures to select the reference group to meet two goals: the reference group will have sufficient samples to develop models with the controlled power level; and the reference group will be a legitimate representative of the reference population so that the models developed from the reference group are identical to the models developed from reference population.

After these three steps are conducted, Section 4.4 will examine the data sampling effects by comparing variables of reference group vs. reference population, in order to prove that the reference groups have consistent statistical features with relevant reference population, and therefore, are legitimate representatives of the reference population.

VI. CONCEPTS, METHODOLOGY AND NECESSITY

This section constructs the theoretical and methodological foundation for all following analysis applications. Sub-section A is the general introduction for random sampling (RS) and random assignment (RA) strategies. Sub-section B describes the selection of appropriate RS and alternative RA approaches which will be applied for this dissertation studies. Sub-section C emphasizes the rationality of data sampling in the context of B/A process.

A. Basic Concepts

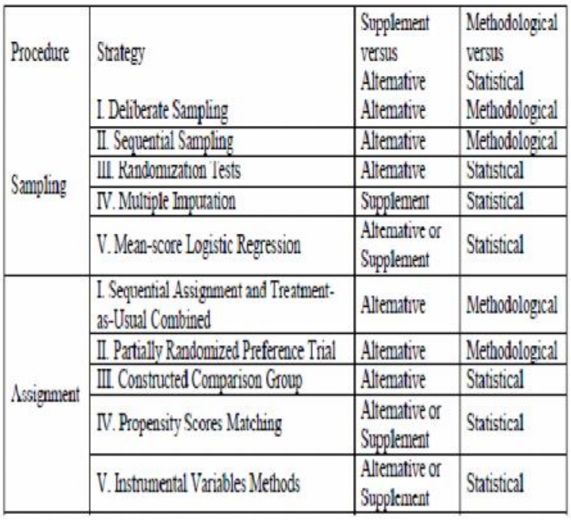
This sub-section introduces the basic concept of RS, RA and their relation with B/A validity or bias. Dattalo (2010) systematically described B/A oriented random sampling (RS) and random assignment (RA) strategies, among which there are three key concepts: selection bias, external validity and internal validity.

Selection bias is the introduction of errors due to systematic differences in the characteristics of participants and nonparticipants in a study (reference groups and treated groups in B/A processes). Two types of selection biases can be distinguished: sampling and assignment. In sampling bias, error results from failure to ensure that all members of a reference population have a known chance of being selected for inclusion in a sample. In assignment bias, error results from systematic differences in the characteristics of those allocated to an intervention (treated) group versus a control group in an experimental study (Dattalo, 2010). (Note: a "control" group is used for experimental studies; however, traffic safety B/A processing is an observational study. In such observational cases, a reference group is used instead).

B. Methodology Selection

Both RS and RA have a variety of approaches and methods, this subsection aims to compare the characters and utilities of all those optional RS and RS approaches and then to recommend the one most appropriate for this dissertation investigationIn practice, there are adjustments and/or substitutions for RS and RA, which are called "alternatives", while others used as compensation on top of RS and RA are called "supplements"(Dattalo, 2010). Meanwhile, strategies used before or during RS and RA are categorized as "methodological" while others used as adjustments after RS and RA are classified as "statistical". Dattalo (2010) listed the available strategies based on different combinations, as shown in Table. Table implies that sampling strategies do not have to be "random"; as a result, the terminology is generalized thereafter in this dissertation: "data sampling" replaces random sampling (RS) while "data assignment" replaces random assignment (RA).

The dissertation research pertains to specific sampling and assignment approaches by taking into account two considerations. The first consideration is the stage when sampling is conducted. At the pre-testing stage, the dissertation needs a "before and during" sampling approach so it has to be methodological rather than statistical. Another consideration is the sample size. Sampling strategies include fixed-sample design, in which sample size is set in advance, or sequential sampling in which sample size is eventually determined (Stephens, 2001). The dissertation research will not fix the sample size before sampling, so it has to be sequential. As a result, the appropriate sampling strategy for this dissertation is a "sequential sampling methodology" per Table.



Procedure	Strategy	Supplement versus Alternative	Methodological versus Statistical
Sampling	I. Deliberate Sampling	Alternative	Methodological
	II. Sequential Sampling	Alternative	Methodological
	III. Randomization Tests	Alternative	Statistical
	IV. Multiple Imputation	Supplement	Statistical
	V. Mean-score Logistic Regression	Alternative or Supplement	Statistical
Assignment	I. Sequential Assignment and Treatment-as-Usual Combined	Alternative	Methodological
	II. Partially Randomized Preference Trial	Alternative	Methodological
	III. Constructed Comparison Group	Alternative	Statistical
	IV. Propensity Scores Matching	Alternative or Supplement	Statistical
	V. Instrumental Variables Methods	Alternative or Supplement	Statistical

Table:sequential sampling methodology

C. *Necessity of Data Sampling*

The fundamental purpose of this dissertation is to determine a better B/A methodology for safety treatment effect evaluation. As for local practices, not all data items are easily accessible. Some documented or conventional inventory data items, such as collisions and traffic volume, might be obtained with ease while other observed data items, such as turn lanes of intersections, usually require field surveys or labor-intensive manual means and can only be processed site by site, which can be costly.

VII. ACCOMPLISHMENTS,CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER STUDY:

A. Accomplishments

B/A methodologies are not a new topic in the traffic safety domain. The basic framework for processing before-after evaluations has been well-established and widely applied in practices in the real world for a lengthy amount of time. Both traditional EB and the more recent FB approaches have fueled a substantial amount of research. Nevertheless, there are still some methodological issues that are causing uncertainty and bias in current treatment effect estimations.

In particular, previous research may have concentrated on safety model development itself while paying insufficient attention to the stages before and after model development. This dissertation has sought to balance the research in all of the stages for a thorough processing of the before-after evaluation in order to address all major methodological issues in current applications. In so doing, the following accomplishments have resulted from this dissertation study.

1. Conducting pre-test data sampling to select appropriate local reference groups

If treatment effect estimation is seen as one test, pre-test data sampling is an important and inevitable step to assemble adequate, sufficient and appropriate reference groups for the next step in model development. Regardless of its importance, few researchers in the road safety field have carried out data sampling prior to modeling. Most research immediately commences from the modeling procedure itself, which is based on existing, arbitrarily selected or readily available reference group, without applying a statistical data sampling process.

This dissertation has challenged this status quo by investigating a data sampling approach to select appropriate reference groups through two consecutive steps: Step 1, which estimates the appropriate sample size to seek a specific level of modeling power (1-B, see details in Chapter 2); and Step 2, in which, in accordance with the pre-estimated sample size, stratified sequential probability proportional to size (PPS) data sampling is conducted to select appropriate reference groups that are adequate for model development, while maintaining statistical consistency with the whole reference population.

2. Development of local safety models with multiple hierarchies, various random distributions and with different approaches

Unlike most safety modeling efforts that only concentrate on one single model, and thus neglecting other choices, this dissertation has developed a variety of local safety models. First of all, model structures have been identified with multiple hierarchies, including single-level full models and multi-level (hierarchical) structures. While multi-level safety models have key merits, such as addressing local specifics while maintaining structural consistency, this dissertation research did not reject the full model. Traditionally, SPF developments favored NB distribution, e.g., a special case of Poisson-gamma distribution. This dissertation also took NB distribution into consideration, but at the same time, included others from the mixed Poisson family, such as Poisson-lognormal and Poisson-Weibull distributions. Moreover, this dissertation has applied both the “Frequentist” and “Bayesian” approaches to develop a local model. The former uses the MLE process and yielded fixed model parameters. The latter uses a simulation process and treated parameters as random variables.

The advantage of development via multiple models is clear: any statistically significant (SS_ model would have useful information and different SS models provide different knowledge sources from different perspectives. To keep all of these SS models, means that no useful information and knowledge sources are excluded. The dissertation research has contributed to advancing this philosophy.

3. Converting knowledge from safety surrogates into collision measures

Collision measures are the most favored indicator for safety performance. However, the estimation of collisions relies on adequate historical collision data and these are not always available and sufficient for such estimation. If this is the case, indirect safety measures are to be applied, i.e., safety surrogates. This dissertation has selected the predicted speed of modern roundabouts as the sample and proved its connection with both collision measures and design features, which confirmed that predicted speed can be used as a safety surrogate in the event that collision measures are absent. In so doing, the dissertation has contributed to knowledge on the validity of using safety surrogates.

4. Exploration of Bayesian model averaging to integrate different knowledge sources

In the B/A context, safety models are applied to calculate “postulated collisions without treatments”, i.e., referential information. So the most important characteristic of these models is that they must have wide representation. Single models, regardless of their positive attributes, have difficulties in providing widespread enough information.

This dissertation has investigated an innovative approach to integrating all eligible models together without exclusions. This is achieved by the use of the BMA, an approach that takes many eligible models and merges them into one, by averaging their parameters which are weighted by their posterior model probabilities.

This is one of the most important achievements of this dissertation study. It has developed a viable option for traditional model comparison and selection in traffic safety practices that usually end with a single recommended model that and may be inappropriate for before-after evaluations, as was demonstrated. The final BMA models developed in this dissertation were tested and found to display better application performance in addition to conceptual superiority.

5. Refining validity of treatment effect estimation by propensity score matching and applying comparison groups to adjusted BMA models

Due to the conventional assignment method of treated groups, they are usually created with very high heterogeneity compared to the reference group. This has led to the observation from the dissertation data that, when BMA modeling is directly applied to compute the "postulated collisions without treatment", there is still the tendency to exaggerate the treatment effect.

In order to solve this problem, this dissertation has investigated a propensity score matching approach that is carried out post-assignment, to generate comparison groups that are relatively more similar to the treated groups. Then, these comparison groups are applied into the calibration process to further adjust the BMA models and secure higher validity for the final treatment effect estimation.

VIII. CONCLUSIONS

After this series of systematic investigations on the B/A methodologies, the following conclusions can be drawn. It is important to address, as this dissertation has done, not one, but five different B/A methodologies that comprise the whole process. The fundamental reason is that all of these methodologies are essential for a valid before-after treatment effect analysis and none of them are already well established through previous research. With any one of the methodologies deficient, the treatment effect would be distorted. Also, all five methodologies are sequentially followed until the treatment effect estimation is optimized when comparison bias is minimal and internal and external validities are maximized. In addition, none of the datasets developed in this dissertation study are redundant. Although the final treatment effect was estimated through comparisons of the treated group itself, and with and without treatments, other datasets including reference population, reference groups and comparison groups were also used to model the "postulated collisions without treatment" of the treated group so they all played their own roles in the treatment effect analysis. The sake of a valid analysis on the treatment effect, to determine all eligible models through different channels, and then finally combine them all together, is more promising than recommending a single model. To serve this purpose, this dissertation has developed models by many means and finally applied the BMA approach to integrate all significant models together. The way that they were developed is not important here; they could be either imported from an external source or locally developed. Pre-phase data sampling and post-stage adjustment are as equally important as the model development itself. From a practical perspective, data sampling before modeling means that the data are more accessible and post-assignment sample matching and model adjustments enhance the internal validity when assignment of a treated group is already carried out beforehand and beyond control. From a theoretical perspective, both pre-phase data sampling and post-assignment sample matching in this dissertation require advanced statistical methods and considerable analytical processing. They are all worthy research topics, but were neglected in previous road safety studies. This dissertation has thus enhanced these two aspects.

Most significantly, this dissertation has explored and applied alternative model selection and an averaging mechanism that are unique in comparison to conventional model recommendation practices. Rather than recommending one candidate while neglecting all others, this dissertation integrates all eligible models by means of the BMA without exclusion. In the before-after evaluation scenario, this has contributed to less comparison bias and higher external validity.

A. Future Studies

This dissertation research can be further enhanced and extended as follows.

1. A breakthrough strategy can be explored for the assignment of a treated group which can achieve lower "innate" systematic heterogeneity between treated and reference groups.
2. Innovative safety model paradigms can be further explored, e.g., multi-level model applied in the dissertation can include more dynamic safety attributes in traffic operations and better address local specifics with more flexible parameters.
3. The mechanism to integrate knowledge sources from safety surrogates and collision prediction models can be further investigated. Within the scope of this part of the dissertation investigation, the studied safety surrogate is still an independent referential knowledge that is not merged into the final model. In the future, this should be carried out and knowledge sources from safety surrogates and collision prediction models should be integrated.
4. Further research on other before-after evaluation control factors are beyond scope of this dissertation. One example is, where appropriate data are available, investigation on how the expected treatment effect, or expected collision reduction, could be taken into consideration for reference group size determination based on the new approach recommended by this dissertation. One example is the development of applicative tools, e.g., spreadsheet macros or other software that can automatically conduct the before-after evaluation procedures based on methodologies established by this dissertation. Another example is the construction of CMFs or CM-Functions more accessible to general practitioners. The dissertation focused on models, so development of new CMFs or CM-Functions can follow in future studies, based on components developed in this dissertation.

REFERENCES

1. Aguero-Valverde, J. (2013). Full Bayes Poisson gamma, Poisson lognormal, and zero inflated random effects models: Comparing the precision of crash frequency estimates. Accident Analysis and Prevention 50, pp. 289-297.
2. Ando, T. and Tsayb, R. (2010). Predictive Likelihood for Bayesian Model Selection and Averaging. International Journal of Forecasting 26, pp. 744–763.
3. Akaike, H. (1977). On entropy maximization principle. In Applications of Statistics (Proceedings of Symposium, Wright State University, Dayton, Ohio, 1976), pages 27–41. North-Holland, Amsterdam. Akaike, H. (1978). A new look at the Bayes procedure. Biometrika, 65:53–59.
4. American Association of State Highway and Transportation Officials (AASHTO) (2010). Highway Safety Manual. Washington, DC, United States of America.
5. Amundsen, F. and Hyden, C. (1977). Proceedings of First Workshop on Traffic Conflicts, Oslo, Institute of Transport Economics.
6. Ando, T. and Tsay, R. (2010). Predictive Likelihood for

Bayesian Model Selection and Averaging. International Journal of Forecasting 26, pp. 744-763.

7. Chen, Y., Persaud, B., and Sacchi, E. (2012) Improving transferability of safety performance functions by Bayesian model averaging. Journal of Transportation Research Record (TRR, accepted and In-Press).
8. Chin, H. C. and Huang, H. (2008). Modeling Multilevel Data in Traffic Safety: A Bayesian Hierarchical Approach. Transportation Accident Analysis and Prevention. Nova Science Publishers, Inc., New York, pp. 53-106.
9. Chromy, J. R. (1979). Sequential Sample Selection Methods. Proceedings of the American Statistical Association, Survey Research Methods Section, 401–406.
10. Claeskens, G. and Hjort, N. L. (2009). Model Selection and Model Averaging. Cambridge University Press.
11. Coca-Perraillon, M. (2007). Local and Global Optimal Propensity Score Matching. Paper 185-2007, Statistics and Data Analysis, Proceedings of SAS Global Forum 2007, Orlando, Florida, U.S.A.
12. Congdon, P. (2001). Bayesian statistical modeling. Wiley, Chichester. Wiley series in probability and statistics.
13. Cook, T.D., Campbell, D.T. (1979). Quasi-Experimentation: Design and Analysis for Field Settings. Boston: Houghton Mifflin.
14. Dmitrienko, A., Chuang-Stein, C. and D'Agostino, R. B (2007). Pharmaceutical Statistics Using SAS: A Practical Guide (Chapter 14: Decision Analysis in Drug Development). SAS Institute Inc., Cary, NC, USA.
15. Dougherty, C. (2002). Introduction to Econometrics: Statistical Tables, Second edition, Oxford University Press, Oxford, United Kingdom.
16. El-Basyouny, K. and Sayed, T. (2010). Full Bayes Approach to Before-and-After Safety Evaluation with Matched Comparisons: Case Study of Stop-Sign In-Fill Program, Transportation Research Record: Journal of the Transportation Research Board, No. 2148, Transportation Research Board of the National Academies, Washington, D.C., USA, pp. 1–8.

AUTHOR'S INFORMATION

Ravi Kumar Mangi Student of Masters of Engineering in Civil Engineering Department in Pydah College of Engineering&Technology, Gambheeram, Viskhapatnam.

TPrakashLazarus, M.Tech, Assoc.Professor& HOD,Pydah College of Engineering & Technology Gambheeram, Viskhapatnam

Dr. Divakar C M.Tech Ph.D, Principal, Pydah College of Engineering&Technology Gambheeram, Viskhapatnam.