TECHBRIEF

U.S. Department of Transportation
Federal Highway Administration

Research, Development, and Technology

Turner-Fairbank Highway Research Center

6300 Georgetown Pike

McLean, VA 22101-2296

www.fhwa.dot.gov/research

Safety Evaluation of Access Management Policies and Techniques

FHWA Publication No.: FHWA-HRT-15-038

FHWA Contact: Wei Zhang, HRDS-10, (202) 493-3317, wei.zhang@fhwa.dot.gov

This document is a technical summary of the Federal Highway Administration report, *Safety Evaluation of Access Management Policies and Techniques*, FHWA-HRT-14-057.

Introduction

Access management is the process that provides (or manages) access to land development while simultaneously preserving the flow of traffic on the surrounding road network for safety, capacity, and speed. Access management provides important benefits to the transportation system. These benefits have been increasingly recognized at all levels of government, and a growing number of States, cities, counties, and planning regions are managing access by requiring driveway permit applications and establishing where new access should be allowed. They are also closing, consolidating, or improving driveways, median openings, and intersections as part of their access management implementation strategy. However, these decisions are often challenged for various reasons.

Additional information is needed to help guide decisions related to access management. This information will help agencies better explain the safety and operational benefits of their policies and practices. Previous studies and empirical evidence have shown positive operational and safety benefits associated with good access management practices. While the operational effects of access management have been investigated quantitatively through different modeling and analysis approaches, there have been few scientifically rigorous evaluations to quantify safety effectiveness, particularly for corridor access management. The Federal Highway Administration initiated this study to help fill some of the research gaps, namely quantifying the safety impacts of corridor access management decisions.

Study Objective and Scope

The objective of this research was to develop corridor-level crash prediction models to evaluate the potential safety effects of access management strategies. Agencies can apply the algorithms to assess the safety impacts of their decisions related to access management.

The intent of this study was to focus on corridors based on functional classification, area type, and land use. All corridors included in this study are functionally classified as arterials and fall under one of nine area type/land use scenarios. Table 1 identifies the nine area type/land use categories and provides a definition for each area type.

Residential and commercial areas are characterized by the type of development but are also differentiated by the type and distribution of vehicle types accessing the areas. Residential areas serve mainly passenger cars, and commercial areas serve a larger proportion of heavy vehicles. Commercial areas are generally defined as those areas with office buildings and other businesses that operate primarily during normal business hours on weekdays. Commercial areas, as defined in this study, do not include large shopping centers (e.g., malls) that have a larger percentage of trips on the weekends. Mixed-use area types are defined as those areas with a balanced mix of both commercial and residential establishments and access points. Figure 1 and figure 2 provide examples of corridors included in the study.

Overview of Strategies

Table 2 identifies the access management strategies considered in this research project and notes the related access management safety principle.

Methodology

An observational cross-sectional study design was employed for this project. The safety impact of a given feature can be derived from a cross-sectional study by comparing the safety of a group of sites with that feature to the safety of a group of sites without that feature. Multiple variable regression models were used to estimate the effects of one feature while controlling for other characteristics that vary among the sites. These cross-sectional models are also called *crash prediction models*, which are mathematical equations that relate crash frequency to site characteristics. While cross-sectional models provide a means to estimate the safety impacts of access management strategies, potential issues need to be addressed. The following potential sources of bias were identified in this study with an explanation of how they were addressed or dismissed:

- *Selection of appropriate functional form*: Generalized linear modeling (GLM) techniques were applied to develop corridor-level crash prediction models. A log-linear relationship was specified using a negative binomial error structure, following the state-of-the-art in modeling crash data. The appropriate model form for each variable was determined following a review of the data.

Table 1. Area type and land use categories.

Area Type	Land Use
Urban: Metropolitan area with population of at least 250,000.	Residential
	Commercial
	Mixed Use
Suburban: Nearby areas with population of 50,000 to 250,000.	Residential
	Commercial
	Mixed Use
Urbanizing: Areas with build out-plans to reach or exceed population of 50,000.	Residential
	Commercial
	Mixed Use

Figure 1. Example of an urban arterial in a residential area.

Figure 2. Example of a suburban arterial in a commercial area.

Table 2. Strategies/policies in relation to access management safety principles[1]			
Access Management Strategy/Policy	**Limit Conflicts**	**Separate Conflicts**	**Reduce Conflicts**
Access Spacing			
Unsignalized access spacing		•	
Traffic signal spacing		•	
Interchange crossroad spacing		•	
Corner clearance		•	
Roadway Cross-Section			
Median type: TWLTL			•
Median type: Non-traversable median	•		
Median type: Replace TWLTL with non-traversable median	•	•	
Directional median opening	•		
Median opening spacing		•	
Property Access			
Frontage/backage roads	•	•	
Internal cross-connectivity	•	•	

TWLTL = Two-way left-turn lane.

- *Accounting for State-to-State differences*: Data from four regions were used to develop the crash prediction models, and indicator variables were included in the models to identify the respective region for each corridor.
- *Correlation among independent variables*: The correlation matrix of the estimated parameters was examined to determine the extent of correlation among independent variables. Several access management strategies were highly correlated. As such, it was necessary to develop multiple models with subsets of the independent variables rather than one single model with all variables.
- *Over-fitting of prediction models*: Relatively few parameters were included in the final models due to the correlation issue. As such, over-fitting was dismissed as a potential issue.
- *Low sample mean and sample size*: Corridor-level models help overcome issues related to low sample mean because each site (i.e., corridor) typically experiences multiple crashes per year. Sample size was addressed during the early planning stages of the study, and more than 600 mi of data were obtained to provide a large database for analysis. When possible, 3 years of crash data were obtained to further increase the sample size.

- *Aggregation, averaging, or incompleteness in data*: Data were obtained from various sources and supplemented with field measurements to ensure a relatively complete and accurate dataset. While multiple years of crash data were used in the analysis, the number of years was also included as an independent variable to account for the multiple years of data (i.e., the model prediction results in crashes per year). A maximum of 3 years of data were used for any site.

- *Temporal and spatial correlation*: Temporal correlation may arise if multiple observations are used for the same entity. In this case, multiple years of data were used for each corridor, but these data were aggregated into a single observation, since the maximum was limited to 3 years, and the number of years was included as an independent variable. Spatial correlation is a potential issue, but the corridors were selected from four regions and relatively dispersed within each of the regions. Indicator variables were included to account for similarities within regions.

- *Endogenous independent variables*: Endogeneity arises when one or more of the independent variables depend on the dependent variable. For example, left-turn lanes may be installed due to the frequency of left-turn crashes at an intersection. A cross-sectional model that predicts crash frequency based on the presence of left-turn lanes and other factors may conclude that left-turn lanes increase crashes. Similar examples could be drawn for other access management strategies. In this study, endogeneity was not considered to be a substantial threat because data were aggregated at the corridor level.

- *Omitted variable bias*: It is difficult to completely account for the potential effects of omitted variable bias in an observational cross-sectional study. In this case, omitted variable bias was addressed to the extent possible by carefully considering the roadway and traffic characteristics to be included in the models. Detailed data were collected for each corridor, and numerous variables were tested for suitability in the models. There was the potential for omitted variable bias due to other factors such as weather, driver population, and vehicle fleet, but a regional indicator variable was included in the models to help to account for differences among the regions.

Data Collection

Detailed data were collected for more than 600 mi of corridors across four regions of the United States. The regions included North Carolina (Raleigh, Cary, and Wake Forest), Minnesota (St. Paul and Minneapolis), Northern California (Oakland, Sacramento, San Francisco, and San Jose), and Southern California (Los Angeles and San Diego). This section identifies the procedures for selecting corridors, collecting and verifying data, and merging the various sources of data for analysis.

Identifying Candidate Corridors

State and local agencies were contacted to solicit candidate corridors for inclusion in the study. Guidance was provided on what constituted suitable corridors to assist the State and local agencies with this process. The critical factors for corridor selection included the following:

- No major construction activity during the study period.
- Availability of crash, traffic volume, and roadway inventory data.
- Arterial functional classification (e.g., principal arterial, minor arterial).
- At least one of the target access management strategies is displayed.
- Area type of urban, suburban, or urbanizing.
- Land use of residential, commercial, or mixed use.

Collecting Highway Safety Information System (HSIS) Data

HSIS contains readily available crash, roadway, and traffic volume data for select States. By design, the three States included in this study are all members of HSIS. The HSIS guidebooks were examined, and any potentially useful variables were requested. At the time of the data request, the most recent year of available data was 2008. As such, the study period for this project was from 2006 to 2008.

Corridor Screening

Three rounds of screening were employed to ensure that there were no major construction activities or changes along the corridors during the study period. Initial screening was conducted by the participating agencies through a review of construction records. The study team performed a second phase of screening using HSIS data, comparing specific variables across years to detect changes that would indicate construction activity (e.g., number of lanes, lane width, shoulder width, median type, median width, and mileposts). The team performed a third round of screening using historical aerial imagery. They identified high-resolution aerial imagery for the identified corridors from the United States Geological Survey National Seamless Server. By comparing historical aerial images with current conditions, the team was able to identify where changes had taken place during the study period.

Supplemental Data Collection and Verification

The data obtained through HSIS were verified and augmented with additional data from video collected during field visits and aerial imagery. Detailed information such as lighting presence, visual clutter, and posted speed limit were obtained from the field videos. Aerial imagery was used to verify the land use, number of through lanes, and median type for each corridor. Aerial imagery was also used to collect information that was unavailable from HSIS such as the frontage type, presence of frontage or backage road, extent of internal cross-connectivity, condition of pavement markings, and access points (location, type, and density).

Setup of Geographic Information System (GIS) Database to Facilitate Data Collection

ArcGIS™ feature classes were created for signalized intersections, unsignalized intersections, driveways, and medians. This process allowed data collectors to insert symbols representing these objects on the aerial images of the corridors. Data fields were created for each object so its characteristics could be noted. The characteristics collected for each object are summarized in table 3.

Figure 3 provides an example of these objects for a 1-mi section of a study corridor (California Route 1). In total, the study corridors contained more than 1,500 signalized intersections, 3,500 unsignalized intersections, and 15,000 driveways.

Post-Processing

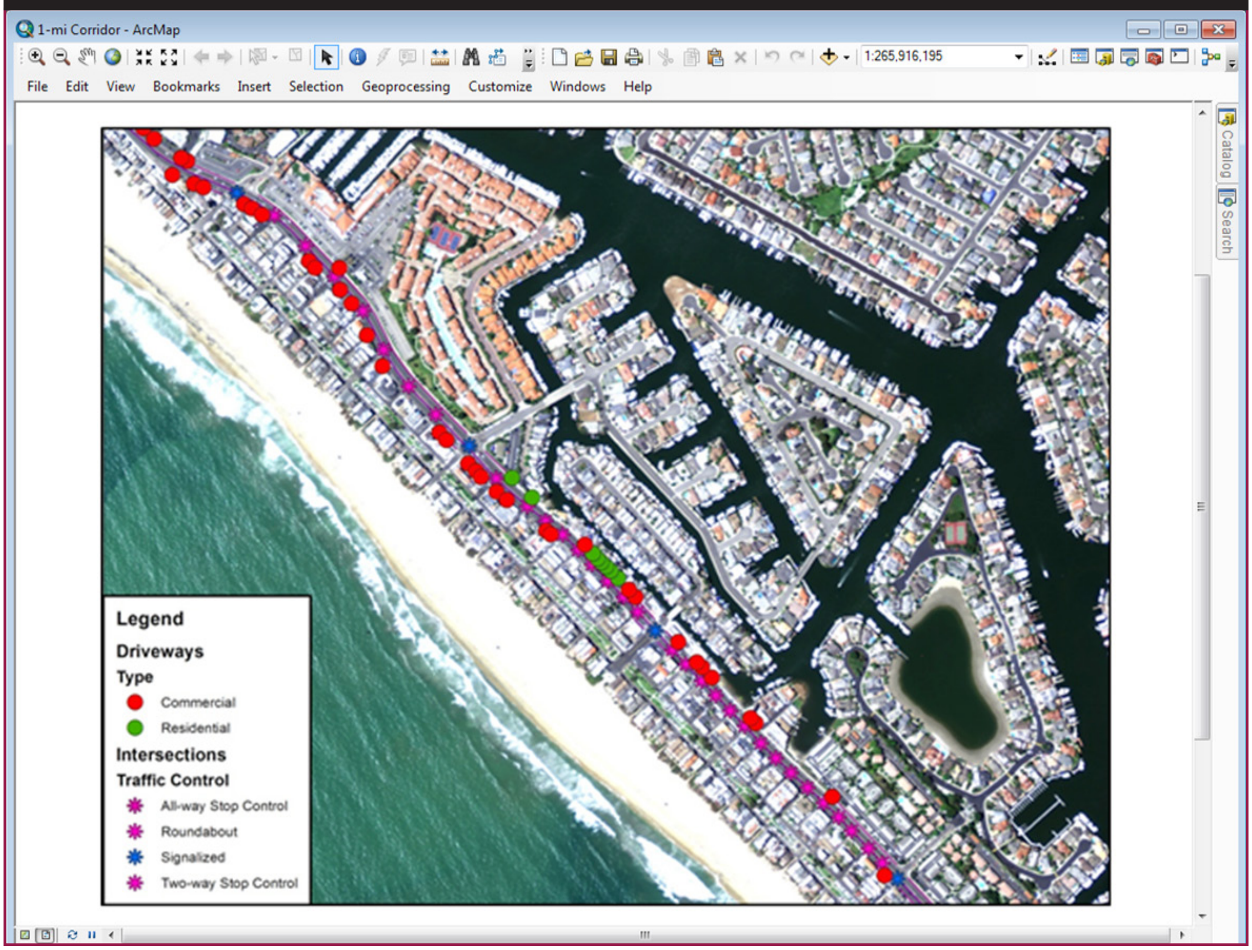
Data were obtained in various formats. Some information was provided in Microsoft® Word documents (e.g., area type, land use, and frontage type), specifying the beginning and ending points of each corridor. HSIS data were provided in Microsoft® Excel format. General corridor information and specific attributes for signalized intersections were identified in the video logs. Other information was identified and stored in the form of ArcGIS™ feature datasets (e.g., intersections and driveways). Due to the multiple formats, a post-processing step was required to transform these data sources into a well-integrated database to serve as the basis for statistical analysis. The following tasks were performed as part of the post-processing:

- *General corridor information*: A spreadsheet was developed and populated to identify the general corridor data, including a unique corridor ID, route number, beginning milepost, ending milepost, beginning cross-street, ending cross-street, area type, and land use for each corridor. The video logs and aerial images were then used to populate additional data fields for each corridor. HSIS data were then appended to identify crashes along each corridor.

Table 3. Objects and characteristics coded in ArcGIS™.

Object Type	Characteristics
Driveways	• Type (commercial or residential) • Movements permitted (limited movement or full movement)
Median openings and crossovers	Presence of left-turn lane
Unsignalized intersections	• Type (two-way stop control, all-way stop control, or roundabout • Presence of left-turn lane(s) on mainline • Presence of right-turn lane(s) on mainline • Presence of left-turn lane(s) on cross street • Presence of right-turn lane(s) on cross street • Movements permitted (right in right out, left from major only, or full) • Maximum number of through lanes on the cross street
Signalized intersections	• Number of approaches • Presence of left-turn lane(s) on mainline • Presence of right-turn lane(s) on mainline • Presence of left-turn lane(s) on cross street • Presence of right-turn lane(s) on cross street • Presence of non-traditional accomodation of left turns • Maximum number of through lanes on the cross street

Figure 3. Example of point objects for a 1-mi corridor.



- *Segmentation*: Within each corridor, new segment links were created whenever variables changed. The beginning and ending mileposts were identified for each individual link. The links were aggregated based on area type and land use to create the study corridors, and each study corridor was assigned a unique ID. Specifically, each study corridor was consistent with respect to area type and land use, but other variables were allowed to change.
- *Access density, spacing, and corner clearance*: The ArcGIS™ files were used to query access-related information for each study corridor. Several variables were created, including the total number of specific access points (driveways, stop-controlled intersections, and signalized intersections), minimum spacing, maximum spacing, number with/without right-turn lanes, number with/without left-turn lanes, and number with turning restrictions (full or limited movement).
- *Median openings and crossovers*: The ArcGIS™ files were used to query median opening information for each study corridor with a median. The total number of median openings was identified for each study corridor along with the number of median openings with and without left-turn lanes.
- *Interchange-related spacing*: The ArcGIS™ files were used to query interchange-related information for each study corridor. Specifically, the ramp location and type were identified for each interchange. For off-ramps, the distances to the first downstream driveway (on right), the first median opening allowing left turns, and the first major intersection were measured. For on-ramps, the distances from the last driveway (on right) and major intersection were measured.

The segmentation process required that links be combined into study corridors to achieve a reasonable length for analysis. In this way, some variables were summed over all links making up a study corridor (e.g., number of driveways). In other cases, new variables reflecting the percentage of the total length were created (e.g., number of lanes). The average annual daily traffic (AADT) and percentage truck variables were calculated as weighted averages, weighting by the lengths of the links within a corridor.

Summary Statistics

The result of the data collection and post-processing totalled 245 corridors representing over 600 mi and approximately 6,500 crashes. Table 4 presents corridor mileage by area type and land use. Table 5 presents a summary of the crashes that occurred along the study corridors from 2006 to 2008. The summary statistics are based on corridor-level data and represent the crash density (i.e., number of crashes per mile per year) for various crash types. The crash types are defined in the Results section.

Analysis

GLM techniques were applied to estimate the models. A negative binomial error structure was specified, following the state-of-the-art in modeling crash data. The negative binomial structure is now recognized as being more appropriate for crash counts than the normal distribution that is assumed in conventional regression modeling. Crash counts per year by crash type were used as estimates of the dependent variable, while corresponding roadway characteristics and traffic volume data were used as the independent variables.

Table 4. Mileage of corridors by area type and land use.

Scenario	Commercial	Mixed Use	Residential	Subtotal
Urban	79.1	92.4	48.7	220.2
Suburban	64.3	119.7	57.0	241.0
Urbanizing	63.8	31.9	62.4	158.3
Subtotal	207.4	244.1	168.2	619.5

Table 5. Crash density summary statistics (crashes per mile per year).

Crash Type	Region	Corridors	Minimum	Maximum	Mean	Standard Deviation
Total	North Carolina	74	0.84	195.31	28.57	25.91
	Northern California	61	0.18	64.52	20.16	15.27
	Southern California	51	1.14	108.99	23.29	19.66
	Minnesota	59	3.10	140.26	33.32	27.55
Injury	North Carolina	74	0.00	40.74	7.22	5.70
	Northern California	61	0.00	24.29	8.34	6.35
	Southern California	51	0.33	33.61	10.34	6.55
	Minnesota	59	1.19	52.39	11.07	10.80
Rear-end	North Carolina	74	0.00	69.63	12.91	11.76
	Northern California	61	0.00	25.57	8.85	7.07
	Southern California	51	0.00	43.60	9.60	9.27
	Minnesota	59	0.33	67.83	16.13	16.33
Right-angle	North Carolina	74	0.00	52.10	4.31	6.39
	Northern California	61	0.00	16.67	2.32	3.09
	Southern California	51	0.00	14.99	2.46	2.77
	Minnesota	59	0.36	27.64	5.72	6.10
Turning	North Carolina	74	0.00	24.44	4.94	4.73
	Northern California	61	0.00	17.18	5.23	4.23
	Southern California	51	0.33	35.88	7.19	6.29
	Minnesota	59	0.00	23.93	2.49	3.52

The first step in the analysis process was to develop a model using only the AADT as a predictor variable and both the number of years and corridor length as offset variables. The general form of this model is given by figure 4.

Figure 4. General form of crash prediction model.

$$\text{Crashes}=\text{years}*\text{segment length}*\alpha*\text{AADT}^{\beta}$$

The next step was to investigate additional variables. This investigation involved entering each variable one at a time such that only AADT and the new variable of interest were included. The estimated parameter and its standard error were examined to determine the following:

- If the direction and magnitude of effect are logical.
- If the estimate is close to being statistically significant.
- If the estimated dispersion parameter improved significantly. The properties of the dispersion parameter are such that lower values indicate a better fit.

Alternate model forms were explored using the procedure described by Hauer and Bamfo.[2] It was determined that the exponential model form is appropriate due to its flexibility, and this form was retained for development of the final models.

Pearson correlation statistics were computed for each independent variable and all crash types per mile-year. The correlation matrix was not the primary driver of model-building but helped to identify those variables most associated with the different crash types. The correlation matrix also helped identify independent variables that were highly correlated. High correlation between independent variables can be problematic in developing models. Specifically, the inclusion of highly correlated variables can lead to illogical results. While this issue can be avoided by omitting a highly correlated variable, this omission limits the practicality of the results when determining the safety impacts of the omitted variable. To overcome this challenge, a series of models were estimated with various combinations of variables, which helped address issues related to correlation and provide information for all variables of interest.

Following the development of preliminary models, feedback was requested from a steering committee as to which of the variables were most desired from a practical perspective. Not all variables could be included in the models due to sample size limitations and correlation between potential explanatory variables. As such, the steering committee was asked to identify the explanatory variables that would be most useful to practitioners. The following variables were indicated to be most important for practical use according to the feedback:

- Adjacent land use (i.e., no development, partial development, or full development).
- Driveway density.
- Median type (i.e., undivided, two-way left-turn-lane, or divided).
- Number of median openings.
- Signalized intersection density.
- Speed limit.

All of these variables were included in various models except for posted speed limit. Vehicle speed was related to the severity of a crash, but the posted speed limit was not included in these models because it was statistically insignificant after accounting for other variables. Posted speed tends to be highly correlated with other variables such as access density and frontage type, which is likely why it could not be included in the final models. It is also possible that posted speed is not providing an accurate representation of the actual speeds (i.e., operating speed may be a better alternative for capturing the impacts of speed).

Results

One or more models were successfully calibrated for each land use type and crash type combination. The three land use types are mixed use, commercial, and residential. The crash types include total, injury, turning, rear-end, and right-angle crashes. Note that individual models by crash type cannot be summed to estimate total crashes. Also, each State has specific crash codes, and, as such, the definitions vary slightly. The crash types are identified in table 6 with the associated definitions for each region.

In the modeling phase, the treatment of area type and regional variables required further resolution. Within each land use type (i.e., mixed use, commercial, and residential), each corridor was identified as being located within an urban, suburban, or urbanizing area. All area types were combined within the respective land use type in order to develop reliable models. A factor variable was included in each model to account for any differences due to area type, but the differences were minor and statistically insignificant. This is not to say that there was no difference in crash patterns by area type, but the data did not allow this relationship to be quantified. It is also likely that area type is better described by other variables in the model. For example, the traffic volume, number of lanes, access density, and frontage development can be used to describe the characteristics of a corridor and are more quantitative than defining a corridor as urban, suburban, or urbanizing. As such, area type was not included as an independent variable in the final models.

An indicator variable was included in each model to identify the region in which the corridor is located (i.e., North Carolina, Minnesota, Northern California, or Southern California). This variable accounts for differences between regions such as those related to crash reporting practices, driver demographics, weather, and other non-access-

Table 6. Crash type definitions.

Crash Type	Definition
Total	All regions: Defined as all crashes
Injury	All regions: Defined as KABC on KABCO scale
Turning	• California: Defined as any involved vehicle making a turn • Minnesota: Defined as left turn or right turn • North Carolina: Defined as rear-end turn, left-turn safe roadway, left-turn different roadway, right-turn same roadway, right-turn different roadway*
Rear-end	• California and Minnesota: Defined as rear-end • North Carolina: Defined as rear-end slow or stop and rear-end turn*
Right-angle	• California: Defined as broadside and no vehicle was turning • Minnesota: Defined as right angle • North Carolina: Defined as angle

KABCO = KABCO injury severity scale, where K = Fatal, A = Incapacitating injury, B = Non-incapacitating injury, C = Possible injury, and O = Property damage only.

*North Carolina crashes coded as rear-end turn crashes are included in both rear-end and turning crashes. Because the specific crash types cannot be summed to get total crashes, it was decided that double-colunting should not pose a problem for the crash type models.

related factors affecting reported crashes. The factors for Northern and Southern California were similar and sufficiently close to be considered as one region. Similarly, the factors for Minnesota and North Carolina were sufficiently similar to consider them as one region. The aggregate regions help to increase sample sizes within the models (i.e., two regions instead of four) and reflect the similarities between the aggregated regions. Note that the models presented in this section include a variable to identify the applicable region. Users should select an applicable region based on a comparison between the corridor of interest and the summary statistics in the full report, and not based on geographic proximity.

The models in this section are presented in one of two forms. In most cases, the model form is represented by figure 5. In these cases, the result is expressed as crashes per mile per year. In other cases, the traffic volume variable is statistically insignificant, indicating a linear relationship between traffic volume and crashes. In these limited cases, the model form is reduced to figure 6, and the result is expressed as crashes per million vehicle-miles.

Figure 5. Crash prediction model with regional calibration.

$$\text{Crashes/mi/year}=\exp^{(\text{intercept}+\text{Region})}*(\text{AADT})^{b}*\exp^{(c_1*X_1+...+c_n*X_n)}$$

Figure 6. Normalized crash prediction model with regional calibration.

$$\text{Crashes/million-vehicle-mi}=\exp^{(\text{intercept}+\text{Region})}*\exp^{(c_1*X_1+...+c_n*X_n)}$$

Where:

- Intercept is the coefficient estimated for the model to account for unobserved variables (see table 7 to table 9).
- Region is the coefficient estimated for the model when the applicable region is North Carolina or Minnesota; a value of 0 is used if the applicable region is Northern or Southern California (see table 7 to table 9).
- AADT is the annual average daily two-way traffic for the corridor.
- *b* is the coefficient estimated for the AADT term in the model (see table 7 to table 9).
- c_i is a vector of coefficients estimated for the other independent variables included in the model (see table 7 to table 9).
- X_i is a vector of other independent variables included in the model (i.e., the specific roadway attributes such as access density).
- ACCDENS is the number of driveways plus unsignalized intersections per mile.
- MEDOPDENS is the number of median openings per mile.
- PROPDIV is the proportion of corridor length with divided median.
- PROPFULLDEV is the proportion of corridor length with full roadside development.
- PROPLANE1 is the proportion of corridor length with two lanes.
- PROPNODEV is the proportion of corridor length with no roadside development.
- PROPVC is the proportion of corridor length with visual clutter.
- PROPTWLTL is the proportion of corridor length with two-way left-turn lanes.
- SIGDENS is the number of signalized intersections per mile.
- UNSIGDENS is the number of unsignalized intersections per mile.

Table 7. Model coefficients and standard error (in parentheses) by crash type for mixed use corridors.

Variables	Total			Injury		Turning			Rear-End	Right-Angle		
	1	2	3	1	2	1	2	3	1	1	2	3
Intercept	-3.1845 (1.9550)	-3.2905 (1.8743)	-0.8926 (0.5021)	-3.5700 (1.7816)	-1.7775 (0.5964)	-2.1083 (0.4338)	-2.0792 (0.3963)	-0.4146 (0.7632)	-3.3091 (0.6700)	-5.8048 (1.9472)	-5.2671 (2.1768)	-2.1485 (0.6851)
Region	1.1410 (0.2316)	1.0533 (0.2086)	0.6166 (0.1013)	0.5695 (0.1980)	0.2465 (0.0931)	0.9647 (0.2843)	0.8015 (0.2354)	-0.3163 (0.1301)	0.8113 (0.1136)	1.8390 (0.2616)	1.2134 (0.2457)	1.2344 (0.1377)
AADT	0.5187 (0.1819)	0.5266 (0.1738)	0.3766 (0.0468)	0.5010 (0.1659)	0.3880 (0.0558)			0.2179 (0.0729)	0.5015 (0.0618)	0.4656 (0.1856)	0.5678 (0.2103)	0.2433 (0.0648)
ACCDENS	0.0053 (0.0044)					0.0088 (0.0061)				0.0112 (0.0051)		
MEDOPDENS											0.1901 (0.0884)	
PROPDIV											-0.4710 (0.3461)	
PROPFULLDEV												0.6787 (0.1846)
PROPLANE1	-0.5185 (0.3789)	-0.6376 (0.3796)		-0.5814 (0.3582)	-0.6623 (0.1404)				-0.5548 (0.1713)			
PROPNODEV			-0.4252 (0.2268)		-0.3159 (0.2201)			-0.5890 (0.2827)				
PROPVC												
PROPTWLTL												
SIGDENS	0.1095 (0.0607)	0.0957 (0.0594)		0.1239 (0.0556)		0.1865 (0.0754)	0.1797 (0.0742)		0.0621 (0.0380)	0.2284 (0.0637)		
UNSIGDENS		0.0471 (0.0224)					0.0582 (0.0323)					
Dispersion (*k*)	0.5073	0.4897	0.5165	0.4248	0.4151	0.7920	0.7780	0.7791	0.6098	0.5585	0.6796	0.7674

Table 8. Model coefficients and standard error (in parentheses) by crash type for commercial corridors.											
Variables	Total		Injury				Turning		Rear-End	Right-Angle	
	1	2	1	2	3	4	1	2	1	1	2
Intercept	-0.7017 (0.6873)	-0.6854 (0.5010)	-2.0602 (0.7991)	-0.9792 (0.8386)	0.2127 (0.7288)	-1.9690 (0.5862)	-0.9816 (0.9366)	0.0085 (1.1277)	-3.0651 (0.6691)	-1.6746 (0.9312)	-1.9023 (0.6838)
Region	0.8353 (0.1883)	0.6166 (0.1013)	0.4672 (0.1815)	0.2383 (0.1497)	0.6769 (0.1559)	0.3056 (0.0923)		-0.2548 (0.2101)	0.8113 (0.1136)	1.4756 (0.2388)	1.2344 (0.1377)
AADT	0.3094 (0.0660)	0.3766 (0.0468)	0.3649 (0.0766)	0.3225 (0.0797)	0.2705 (0.0697)	0.3751 (0.0548)	0.1650 (0.0960)	0.1947 (0.1068)	0.5015 (0.0618)	0.1238 (0.0912)	0.2433 (0.0648)
ACCDENS	0.0069 (0.0048)		0.0085 (0.0047)				0.0110 (0.0052)			0.0165 (0.0064)	
MEDOPDENS											
PROPDIV											
PROPFULLDEV											0.6787 (0.1846)
PROPLANE1				-0.6047 (0.2631)	-0.6244 (0.2566)	-0.5245 (0.1430)		-0.7328 (0.3577)	-0.5548 (0.1713)		
PROPNODEV		-0.4252 (0.2268)		-0.6472 (0.3040)				-0.6967 (0.4150)			
PROPVC					0.5421 (0.1990)						
PROPTWLTL											
SIGDENS	0.1002 (0.0523)		0.0566 (0.0512)			0.1075 (0.0300)	0.1995 (0.0660)		0.0621 (0.0380)	0.1532 (0.0658)	
UNSIGDENS											
Dispersion (*k*)	0.4890	0.5165	0.4406	0.4228	0.4739	0.3951	0.7140	0.7802	0.6098	0.7288	0.7674

Table 9. Model coefficients and standard error (in parentheses) by crash type for residential corridors.

Variables	Total			Injury		Turning			Rear-End			Right-Angle	
	1	2	3	1	2	1	2	3	1	2	3	1	2
Intercept	-0.5615 (0.7076)	-1.3644 (0.4953)	-1.1048 (0.4876)	-2.7357 (0.8556)	-2.7379 (0.9147)	-1.1275 (1.1225)	-0.9528 (0.7286)	-0.7154 (0.7477)	-3.8941 (0.9816)	-2.6180 (1.0221)	-3.3056 (0.6549)	-1.4079 (1.0732)	-2.1173 (0.6540)
Region	0.4443 (0.1533)	0.6850 (0.1107)	0.6166 (0.1013)	0.1656 (0.1423)	0.2303 (0.1603)	-0.6520 (0.2073)	-0.1651 (0.1339)	-0.3163 (0.1301)	0.5803 (0.1984)	0.5406 (0.1865)	0.8113 (0.1136)	0.8858 (0.2180)	1.1970 (0.1314)
AADT	0.3094 (0.0673)	0.3883 (0.0463)	0.3766 (0.0468)	0.4189 (0.0820)	0.4615 (0.0867)	0.1826 (0.1059)	0.1759 (0.0708)	0.2179 (0.0729)	0.5392 (0.0945)	0.4782 (0.0967)	0.5015 (0.0618)	0.1332 (0.1051)	0.1768 (0.0639)
ACCDENS		0.0032 (0.0022)					0.0052 (0.0028)						0.0044 (0.0028)
MEDOPDENS													
PROPDIV													
PROPFULLDEV	0.3371 (0.2317)				0.3720 (0.2273)							0.4295 (0.3125)	
PROPLANE1	-0.5479 (0.1702)			-0.4040 (0.1669)	-0.6125 (0.1715)					-0.8174 (0.2078)	-0.5548 (0.1713)	-0.3633 (0.2383)	
PROPNODEV			-0.4252 (0.2268)					-0.5890 (0.2827)					
PROPVC													
PROPTWLTL										-0.5600 (0.2439)			
SIGDENS	0.1262 (0.0629)			0.2081 (0.0539)		0.2244 (0.0818)	0.1821 (0.0426)		0.1675 (0.0864)		0.0621 (0.0380)	0.2267 (0.0750)	0.2084 (0.0390)
UNSIGDENS						0.0635 (0.0283)							
Dispersion (*k*)	0.3277	0.5181	0.5165	0.2663	0.3220	0.5792	0.7030	0.7791	0.5541	0.4803	0.6098	0.5555	0.6790

In either case, the same general procedure is followed to select an appropriate model and compute the predicted crashes. Further discussion of model selection and related examples are provided in the full report. The model coefficients and dispersion parameters are provided in table 7 through table 9. The dispersion parameter (*k*) is provided to help select an appropriate model when multiple options are available. The following factors (in priority order) may be considered in the selection of an appropriate model if more than one option is available:

1. Availability of data to apply the model.
2. Statistical significance of the coefficients for the variables of interest. (The standard error is shown in parentheses under the respective coefficient.)
3. Dispersion parameter (*k*). (A smaller value indicates a better fitting model.)

Once a model is selected, the parameter estimates and the characteristics of the corridor of interest are used to compute the predicted crashes for the corridor. The following example provides sample results and a demonstration of how a given model can be used to compute the predicted crashes. Further details and examples are provided in the full report.

Example

Assume an analyst would like to predict the total number of crashes for a mixed-use corridor in North Carolina, and model 1 for total crashes is selected from table 7. The corridor is described by the following characteristics: AADT of 25,000 vehicles per day, 10 signalized intersections, 30 unsignalized intersections, and 80 driveways. The total corridor length is 2.5 mi, of which 0.625 mi is two lanes. Figure 5 presented the model form and sample estimates for the example problem.

The intercept and region are constants, and the region coefficient is included in this case since the corridor of interest is similar to the corridors in North Carolina. The intercept and region coefficients are summed and included in the first exponential term of the model. The traffic volume is identified for the scenario of interest and input as the AADT term in the model. The coefficient for traffic volume is identified from the table and input as the b term in the model. The appropriate values for relevant access management variables are then identified and input as the $X_1 - X_n$ terms in the model. Finally, the corresponding coefficients for access management variables are identified from the table and input as the c_1 to cn terms in the model. (Note that the coefficients c_1 to c_n correspond to predictor variables X_1 to X_n, respectively.)

In this case, the model coefficients from table 7, model 1, are intercept (-3.1845), region (1.1410), AADT (0.5187), ACCDENS (0.0053), SIGDENS (0.1095), and PROPLANE1 (-0.5185). The AADT is 25,000 vehicles per day. Predictor variables X1 to Xn are defined as follows:

- ACCDENS = (80 driveways + 30 unsignalized)/2.5 mi = 44.0 access points/mi.
- SIGDENS = 10 signals/2.5 mi = 4.0 signals/mi.
- PROPLANE1 = 0.625 mi of two-lane/2.5 mi = 0.25.

The predicted number of crashes per mile per year is computed as seen in figure 5. With the values used in model 1, crashes per mile per year would be computed as seen in figure 7:

Figure 7. Example of using crash prediction model 1 in table 7.

$$=\exp^{(\text{Intercept}+\text{Region})} * (\text{AADT})^{0.5187} * \exp^{(0.0053*\text{ACCDENS}+0.1095*\text{SIGDENS}-0.5185*\text{PROPLANE1})}$$
$$=\exp^{(-3.1845+1.1410)} * (25{,}000)^{0.5187} * \exp^{(0.0053*44.0+0.1095*4.0-0.5185*0.25)}$$
$$=42.6 \text{ crashes/mi/year}$$

In the previous example, a model was selected and applied to predict the number of total crashes for a corridor with specific characteristics. This example is just one potential use of the corridor crash prediction models. The following discussion identifies the two basic applications of the corridor crash prediction models from this research:

1. Assess the relative safety effects of one or more contemplated strategies (or combinations). This situation applies to both existing corridors and new construction. The following two scenarios explain how the models can be applied:

   - *Predicted crashes*: The models can be used to estimate the number of crashes for the corridor under both existing and proposed conditions. The results can then be compared to estimate the change in predicted crashes as various features change. This comparison is particularly useful if the expected AADT changes among the alternative conditions.

   - *Percent change in crashes*: The models can be used to compare the relative effects of various strategies without predicting crashes for a given scenario. In this case, there is no direct consideration of the costs of alternative strategies or strategy combinations. For this application, the individual coefficients from the selected model(s) are used to derive the effects of the variable(s) of interest. For each variable, the relative effect is derived as exp(coefficient). For example, consider the impacts of signal density (SIGDENS) on total crashes for a mixed-use corridor. The coefficient from table 7, model 1, is 0.1095. The relative effect of SIGDENS is exp(0.1095) = 1.12, which indicates that total crashes are expected to increase by 12 percent along this corridor for every unit increase in signal density (e.g., one to two signals per mi). For a change from one to three signals per mi, the expected change in crashes would be a 24 percent increase.

2. Compare the benefits and costs of two or more alternative strategies or strategy combinations in order to select the most appropriate alternative. This situation applies to existing corridors with an available crash history. The observed crash history is used in conjunction with the predicted crashes from the selected model(s) to obtain the expected number of crashes for the scenarios of interest. This number will be referred to as the Empirical Bayes (EB) estimate, which corrects for several potential sources of bias, including variables that are not in the model. A correction factor is computed as the ratio of the EB estimate and the model estimate, which can then be used to adjust model predictions for proposed scenarios.

The full report provides additional guidance on the selection and application of the most appropriate models depending on the intended application. Six typical scenarios are discussed in detail, and sample problems are provided in the full report to further illustrate the application of models in the six scenarios. The six scenarios include the following:

1. All variables of interest are available in one (and only one) land-use model for each crash type of interest.
2. All variables of interest appear in multiple models for the same land use and crash type of interest (i.e., choose the most appropriate model from multiple options).
3. Variables of interest appear in different models for the same land use and crash type of interest (i.e., using a combination of models to assess the impacts of multiple variables because some variables of interest are in one model while other variables of interest are in another model).
4. Variables of interest appear in models for different crash types (i.e., assessing the impacts of a variable over different crash types).

5. Variables of interest are available for a given crash type but not for the land-use type of interest (i.e., extrapolating the impacts of a variable on a given crash type from models related to a different land use).
6. Variables of interest do not appear in any models for any crash type or land use.

The following guiding principles are common to all scenarios.

- If necessary, models may be extrapolated with caution across land-use types for a given crash type. However, models for one crash type may not be extrapolated to another crash type.
- In some situations, it may not be possible to estimate the impacts of a strategy for all or some crash types.
- If the objective of the analysis is to compare the benefits and costs of alternatives, then the models should be calibrated for the local jurisdiction when possible using the EB method.

Recommendations

The results of this project will help users better understand the safety implications of their decisions related to access management. Specifically, users can apply the models to assess the relative safety effects of one or more contemplated strategies (or combinations), or they can compare the benefits and costs of two or more alternative strategies (or combinations). It is recommended that a safety evaluation software tool be developed to help users select and apply an appropriate model or set of models. Functional specifications were developed as part of this project to facilitate the development of such a tool. The specifications include a detailed description of the model selection process and identify the required and optional inputs as well as default values for the various scenarios included in this study.

Conclusions

This research was performed to develop corridor-level crash prediction models to estimate and analyze the safety effects of selected access management techniques for different area types, land uses, roadway variables, and traffic volumes. More than 600 mi of detailed corridor data were collected across four regions of the United States to facilitate the model estimation process. It was not possible to develop a single model for each crash-type and land-use scenario due to the strong correlations among many of the variables of interest. As a result, 41 crash prediction models were estimated for specific land-use and crash-type scenarios. In most cases, multiple models are presented for each land-use and crash-type scenario in which the alternate models contain subsets of access management strategies in an attempt to account for strong correlations among variables. A four-step process is provided in the full report to guide users through the model selection and application process, but it is envisioned that a basic software tool will be developed to simplify this process based on functional specifications. Sample problems are provided in the full report to illustrate the various uses of the models and to demonstrate the model selection and application process.

These models represent the first of their kind for evaluating the safety effects of access management strategies at the corridor level based on national data. While the results of this research will help advance the knowledge-base and state of the practice in access management, the crash prediction models are not without limitations. Specific limitations of the models include the following:

- *Omitted variables*: Ideally, a single crash prediction model would include all desired variables of interest. This model was not a preferred option in this study due to strong correlation among several of the independent variables. To overcome issues related to correlation, all variables could not be included in a single model. Other variables were omitted due to illogical effects and statistical insignificance. As a result, most models have few variables, and median type is not represented in most models.

- *Inability to quantify effects of turning restrictions*: Detailed data were collected to identify the type of access points (e.g., residential versus commercial driveway) and the associated turning restrictions (e.g., full movement, right in right out, and left from major only). Incorporating this information in the models proved difficult. Variables were created to represent these characteristics at the corridor level, but the results were statistically insignificant. While detailed data are available for each point, the models were not developed to assess the impacts of individual points (i.e., a specific driveway or intersection). As such, differences between full and limited movement access points and between three-legged and four-legged intersections are not clear from these models.

- *Lack of volumes on cross streets and driveways*: Traffic volume is a key variable in predicting crashes. The objective of this study was to develop corridor-level crash prediction models, so a weighted average of the traffic volume along the corridor was used to account for exposure. While the major road volume was included, the minor road volume and driveway volumes were not included. Variables such as land use, driveway density, and frontage development may serve as surrogates for traffic volume at minor roads and driveways.

- *Inability to quantify effects of interchange cross-road spacing*: Detailed data were collected to represent various characteristics of interchange crossroads (e.g., distance from ramp terminal to nearest turning opportunity); however, there were relatively few interchanges included in the dataset, and the results were statistically insignificant.

Based on the results of this research and lessons learned during the completion of the study, there are several opportunities for future research as follows:

- *Increase sample size and regional diversity*: There is an opportunity to increase the number of sites and years of data in the database. Increasing the sample size will likely improve the models and allow additional analysis of the variables of interest. Specifically, this increase could focus on resolving the shortcomings noted previously.

- *Corroborate results*: Cross-sectional methods are useful for developing crash prediction models, but there are several potential sources of bias as discussed in the methodology section. Rigorous before-after studies are preferred for estimating the effects of an individual strategy (e.g., access management characteristic). There is an opportunity to corroborate the results of these crash prediction models by collecting additional data to undertake before-after evaluations of each individual strategy.

- *Separate models for non-driveway and driveway crashes*: This study estimated models for a variety of crash types, including total, injury, turning, rear-end, and right-angle. It may be of interest to estimate additional models to explore the effects of specific access management strategies on driveway and non-driveway crashes. This estimation was not possible in this study due to the lack of specific information in the crash data (i.e., California does not indicate driveway-related crashes). Additional research could investigate the suitability of developing these separate models while considering the potential for extensive geographic diversity in how driveway and non-driveway crashes are defined.

- *Develop Highway Safety Manual (*HSM*)-type algorithms*: The HSM provides methods for estimating the expected number of crashes for individual intersections and homogeneous segments.[3] These estimates can be combined to estimate the crashes for a given corridor. The HSM uses a system of base models to predict crashes for an average scenario, and adjustment factors (i.e., crash modification factors) are used to adjust the base predictions to reflect actual conditions. The models developed in this study are corridor-level models, but there may be an opportunity to use these models as base models for average conditions and apply corridor-level adjustment factors to reflect actual corridor conditions. Additional research could investigate the suitability of using these models for this purpose.

References

1. Stover, V.G., *Access Management Techniques and Practices: A Toolbox for the Practitioner*, Teach America, 2007, revised 2009.
2. Hauer, E., and Bamfo, J.,"Two Tools for Finding What Function Links the Dependent Variable to the Explanatory Variables," *Proceedings of ICTCT 97 Conference*, Lund, Sweden, November 5–7, 1997.
3. American Association of State Highway and Transportation Officials, *Highway Safety Manual, 1st Ed.*, Washington, DC, 2010.

Researchers—This study was performed by Vanasse Hangen Brustlin, Inc. For more information about this research, contact Wei Zhang, FHWA Project Manager, HRDS-10 at (202) 493-3317, wei.zhang@dot.gov.

Distribution—This TechBrief is being distributed according to a standard distribution. Direct distribution is being made to the Divisions and Resource Center.

Availability—The report may be obtained from the FHWA Product Distribution Center by e-mail to report.center@fhwa.dot.gov, fax to (301) 577-1421, phone to (301) 577-0818, or online at http://www.tfhrc.gov/safety.

Key Words—Access management, safety analysis, crash prediction models.

AUGUST 2015

FHWA-HRT-15-038
HRDS-10/08-15(WEB)E