MassDOT IMPACT Phase II - Identification of Risk Factors for SHSP Emphasis Areas

Motorcycle Crashes

PREPARED FOR

massDOT
Massachusetts Department of Transportation

PREPARED BY

vhb

REPORT DATE: AUGUST 2021

Purpose & Background

The Massachusetts Department of Transportation (MassDOT) was awarded a grant by the United States Department of Transportation (USDOT) under its Safety Data Initiative (SDI) competition. MassDOT's work under this grant includes the creation of a Safety Analysis Module in their online IMPACT tool. One feature in this module will be a mapping component which will include crash-based and systemic network screening maps. As part of this work, MassDOT is identifying focus crash types, facility types, and risk factors for their Strategic Highway Safety Plan (SHSP) Emphasis Areas. This report was developed under the SDI project and summarizes the risk factor analysis performed for intersection crashes. It also describes a method to identify risk factors using binary logistic regression, which is one potential method to identify risk factors under the SDI grant. Reports for other emphasis areas describe different methods used to adapt to the needs of those areas.

Focus Crash Types and Focus Facility Types

After averaging 49 motorcycle-related fatalities between 2012 and 2016, MassDOT identified those crashes as an emphasis area in the 2018 Strategic Highway Safety Plan (SHSP)[1]. VHB and MassDOT used the crash data field "Vehicle Configuration (All Vehicles)" to identify motorcycle crashes. A crash was included for consideration if any of its vehicle configuration contained the text motorcycle.

After querying the crash data in the MassDOT IMPACT tool, VHB identified 1,588 fatal and serious injury (KA) crashes related to motorcycles between 2013 and 2017. Crashes in the city of Boston were excluded due to underreporting issues that can bias the models[2]. Figure 1 shows a summary of the crash tree that VHB used to identify focus crash types and facility types for motorcycle-related crashes.

[1] https://www.mass.gov/doc/massachusetts-shsp-2018/download

[2] This does not prohibit the use of the risk factors on Boston roads.

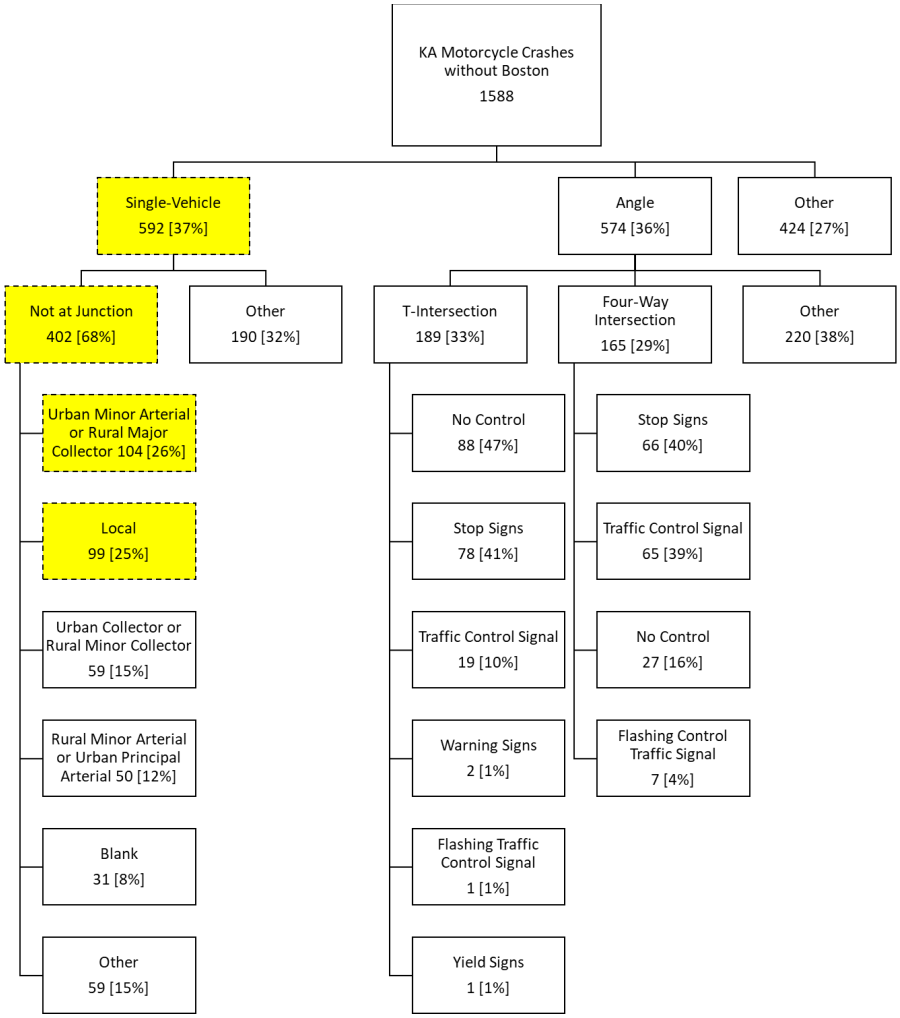

Figure 1. Crash tree to identify focus crash types and facility types for KA motorcycle crashes.

VHB used the data fields "Manner of Collision," "Roadway Junction Type," and "Functional Class" to build the crash tree. The crash tree results showed that single-vehicle crashes that were not at a junction accounted for 25.3 percent of the KA motorcycle crashes. Within single-vehicle not at a junction crashes, 26 percent occurred on urban minor arterial or rural major collectors and 25 percent occurred on local roads. These two functional classes captured 203 KA motorcycle crashes over the study period. Altogether, VHB and MassDOT used these combinations to identify the focus facility type:

- Single-Vehicle
    - Not at a junction
        - Urban Minor Arterial or Rural Major Collector
        - Local

MassDOT plans to evaluate angle motorcycle crashes at intersections in the future when intersection data are available.

Risk Factor Analysis

After identifying the focus crash type and focus facility types, VHB proceeded with the risk factor analysis. The following sections describe the methodology, data, and results of this analysis.

Methodology

Based on discussions with MassDOT, VHB used a modeling approach, similar to other risk factor analyses, to identify risk factors for intersection crashes. Due to the binary nature of the crash severity outcome of interest, the project team used binary logistic regression. This probabilistic modeling technique assesses the probability that an event has occurred (i.e., a KA motorcycle crash) at a given segment based on the model inputs. Agresti (2007) provides more background information on this method.[3] In this context, odds ratios for variables greater than 1.0 indicate the independent variable increases the probability of a KA crash on the segment, while odds ratios less than 1.0 indicate a decrease in probability. VHB developed one model for the two functional classes due to sample size.

When modeling, VHB added variables one at a time, monitoring the coefficients to ensure the inclusion of a variable did not result in large changes in the magnitude of odds ratios for the other variables. Additionally, VHB included variables with p-values upwards of 0.30 assuming the magnitude of the results made sense. VHB did not select a strict level of significance, as Hauer noted this could lead to misunderstanding or outright disregard for potentially noteworthy results.[4] Additionally, VHB monitored the variables for serial correlation. The maximum correlation between any two variables included in a model was 0.67, but most variable pairs had correlation values less than 0.10.

Data

VHB used ArcGIS to manage and integrate data for this analysis. MassDOT provided VHB with various sources of data. As stated in the methodology section, the binary logit model was developed at the road segment level, so VHB tied crashes and other data to each road segment for modeling.

[3] Agresti, A. (2007). An Introduction to Categorical Data Analysis. Second Edition. John Wiley & Sons, Inc., New York.

[4] Hauer, E. (2004). The harm done by tests of significance. *Accident Analysis & Prevention*, *36*(3), 495-500.

Motorcycle Data

MassDOT provided motorcycle-related Excel data and the following parameters were computed at the town-level:

- Average number of motorcycle licenses by town
    - The number of motorcycle licenses by town was provided for the years between 2015 and 2018. The average number of motorcycle licenses by town was computed using data from 2015 and 2017 in order to be consistent with the crash analysis period.
- Average number of motorcycle registrations by town
    - The number of motorcycle garaging (or registration) by town was provided for the years between 2015 and 2019. The average number of motorcycle registrations by town was computed using data from 2015 and 2017 in order to be consistent with the crash analysis period.
- Average vehicle purchase time before crash by town
    - The average vehicle purchase time before crash by town was computed using three datasets. The first dataset contained the vehicle purchase date and vehicle identification number (VIN). The second dataset contained the KA motorcycle crashes (excluding Boston) by vehicle identification number, crash number, and crash town. The third dataset contained the crash data (described below) containing the crash number and crash date. The VIN was first used to link the vehicle purchase date and the corresponding crash number and crash town, and then the crash number was used to link the corresponding crash date. Altogether, the processed data was used to compute a vehicle purchase time (or difference between crash date and purchase date) for each VIN. This result was then aggregated up to the town-level to determine the average vehicle purchase time before crash by town.
- Average license time before crash by town
    - The average license time before crash by town was computed using two datasets. The first dataset contained the license number, license credential, and license issue date. The second dataset contained the license number, crash town, and crash date. The license number was first used to link the license issue date and the corresponding crash date and crash town. Since some licensees held multiple driving credentials, the resulting data was filtered to only motorcycle-related credentials (e.g., JODM, JOM, MCYM, NCLDM, and PERM). For the license time before crash, this was computed using the difference between the crash date and the earliest non-PERM license issue date. The average permit time before crash by town was analyzed separately since some licensees only held permits. This was done to minimize skewing of the license results by inexperienced permit motorcyclists. The license time before crash per license holder was aggregated up to the town-level to determine the average license time before crash by town.
- Average permit time before crash by town (permit holders only)
    - The average permit time before crash by town was computed using the same processed dataset as the average license time before crash by town. The key difference is that the permit time before crash was computed for licensees with only PERM credentials. Therefore, the permit time before crash is the difference between the crash date and the

earliest permit time per permit-only licensee. These results were then aggregated up to the town-level to the determine the average permit time before crash by town. Correspondence with the Massachusetts Registry of Motor Vehicles indicated that motorcycle permits only last two years, and drivers can renew a permit with a written test, whereas they have to pass a road test to receive their license.

Note that considerations of the risk factor scoring implementation led to the decision to aggregate some of these parameters up to the town-level. In addition, the average purchase time before crash, average license time before crash, and average permit time before crash are determined using data where motorcycle crashes have occurred. Lastly, the town-level motorcycle results were joined to the corresponding segment data in the same town.

Crash Data

MassDOT provided statewide geolocated crash data for the years 2013 through 2017. VHB used ArcGIS Pro to compute a KA crash occurrence field (1 = a crash occurred; 0 = no crash occurred) for the applicable "Vehicle Configuration (All Vehicles)" (contains motorcycle), "Manner of Collision" (single-vehicle), "Roadway Junction Type" (not at a junction), and "Functional Class" (urban minor arterial/rural major collector, local). The city of Boston is excluded separately in the binary logistic regression model. VHB then spatially joined the crash occurrence field to the segment data within a buffer area.

Curve Data

MassDOT provided statewide curve data. VHB used ArcGIS Pro to compute a weighted degree of curvature using the curve radius and length of curve for each intersecting road segment. The resulting weighted degree of curvature was joined to the segment data described below using the segment ID.

Segment Data

MassDOT provided statewide roadway segment data. VHB used ArcGIS Pro was used to first filter the roadway segment data to the identified functional classes (e.g., urban minor arterial/rural major collector and local). The motorcycle data, crash data, and curve data were then joined to the segment data by town, buffer area, and segment ID, respectively. Note that some segments contained default AADT values. These default AADT values were set to 0 prior to the binary logistic regression modeling.

Results

This section describes the results of the risk factor models. The models were run as described in the methodology. As part of the modeling efforts, VHB used correlation matrices to verify low correlation between variables and used summary tables to make sure that there were sufficient observations for inclusion in the model. The dependent variable for the models was the occurrence of a KA motorcycle crash at the identified focus facility type: 1 if a crash occurred between 2013 and 2017; 0 otherwise.

Table 1 summarizes the binary logit regression model for KA motorcycle crashes. All variables are binary – meaning the variable is equal to 1 if the condition is true for the segment and 0 otherwise. This model excludes segments in Boston (due to known under-reporting issues with the crash data) and excludes segments length less than 0.05 miles.

Table 1. Binary logit regression model for KA motorcycle crashes.

Variable	Odds Ratio	Standard Error	z-value	P>|z|	95% Confidence Interval	
Segment in Northern Middlesex Council of Governments (NMCOG) MPO	1.74	0.57	1.71	0.09	0.92	3.29
Segment in Old Colony Planning Council (OCPC) MPO	2.14	0.58	2.83	0.01	1.26	3.64
Segment in Southeastern Regional Planning and Economic Development District (SPREDD) MPO	1.73	0.41	2.34	0.02	1.09	2.74
Average vehicle purchase time before crash by town less than or equal to 10 years	1.78	0.53	1.94	0.05	0.99	3.19
Segment containing curb of right side only	1.79	0.66	1.57	0.12	0.86	3.70
Segment containing median width between 11 and 50 feet	4.73	2.87	2.56	0.01	1.44	15.52
Segment in urban type urban-large area	2.36	0.94	2.16	0.03	1.08	5.14
Segment in urban type rural	3.28	1.55	2.52	0.01	1.30	8.28
Segment in MassDOT jurisdiction	2.97	1.66	1.94	0.05	0.99	8.91
Segment in City or Town Accepted Road jurisdiction	2.88	1.22	2.50	0.01	1.26	6.59
Average number of motorcycle registration by town greater than 250	1.84	0.55	2.05	0.04	1.03	3.30
Segment with AADT greater than 1000	3.93	0.71	7.61	0.00	2.76	5.60
Average license time before crash by town less than or equal to 10 years	1.23	0.21	1.21	0.23	0.88	1.72
Average permit time before crash by town greater than 3 years	1.30	0.24	1.41	0.16	0.90	1.87

Note: Number of observations = 184,386; Log likelihood = -1131.8287; Pseudo R2 = 0.0808; LR chi2(15) = 192.89; Prob > chi2 = 0.0000.

The risk factors identified in the binary logit model in Table 1 generally falls into three categories: spatial location risk factors, road condition risk factors, and motorcycle-specific risk factors. In terms of spatial location risk factors, the binary logit model showed an increased probability of a KA motorcycle crash when the segment falls within the NMCOG, OCPC, and SPREDD MPOs; in either urban-large or rural areas; and in MassDOT or City or Town Accepted Road jurisdictions. In terms of road condition risk factors, the model showed an increased probability for segments with right curbs, median widths between 11 and 50 feet, and AADTs greater than 1000. Lastly, in terms of motorcycle-specific risk factors, the model showed an increased probability for the average number of motorcycle registrations by town greater than 250, the average license time before crash by town less than or equal to 10 years, the average purchase time before crash by town less than or equal to 10 years, and the average permit time before crash by town greater than 3 years. The average license time and purchase time before crash by town less than 10 years suggest that more inexperienced drivers have a higher probability of being involved in a KA crash. The

average permit time before crash by town greater than 3 years suggests that long-term permit-only holders have a higher probability of being involved in a KA crash than those who become licensed.

Conclusions and Recommendations

The purpose of this report is to summarize the systemic analysis of motorcycle crashes in Massachusetts. VHB and MassDOT identified KA crashes involving single-vehicle, not at a junction, and on urban minor arterial/rural major collector or local roads as the focus facility type.

The identification of risk factors followed a similar approach to the other emphasis areas. MassDOT and VHB considered motorcycle, crash, curve, and segment data to develop the binary logit model and to determine the likely motorcycle risk factors. Table 2 summarizes the risk factors identified in this analysis and their recommended scoring.

VHB recommends MassDOT disregard the Odds Ratio results from the binary logit models and assign risk scores using the recommended schema in Table 2. Table 3 provides an example of how to calculate a risk score using the model results. In this example, the segment has a total score of 6.18 out of a total possible score of 10.

MassDOT can normalize the score using the total number of potential risk factors, for example assigning a segment a risk score of 100 percent if all risk factors for the facility type are present. Under this approach, the risk score for the example segment in Table 3 is 61.8%.

Table 2. Summary of risk factors for KA motorcycle crashes

Motorcycle Risk Factors	Motorcycle Risk Factor Scoring
MPO	• Risk = 1 if MPO is OCPC • Risk = 0.75 if NMCOG or SRPEDD • Risk = 0 otherwise
Jurisdiction	• Risk = 1 if jurisdiction is in MassDOT or City or Town Accepted Road • Risk = 0 otherwise
Urban Type	• Risk = 1 if urban type is rural • Risk = 0.5 if urban type is urban-large urbanized area or rural • Risk = 0 otherwise
AADT	• Risk = 0 if AADT is less than or equal to 1000 • Risk = $1.02*10^{-5}$*AADT + 0.490 if AADT is greater than 1000 and less than or equal to 50000 • Risk = 1 if greater than 50000
Median Width	• Risk = 1 if median width is between 11 and 50 feet • Risk = 0 otherwise
Curb	• Risk = 1 if curb on right side only • Risk = 0 otherwise
Average Number of Motorcycle Registration by Town	• Risk = 0 if registration number is less than or equal to 250 • Risk = $3.081*10^{-4}$*RN + 0.423 if RN or registration number is greater than 250 and less than or equal to 1873 • Risk = 1 if greater than 1873
Average License Time Before Crash by Town	• Risk = -0.05*LT + 1 if LT or license time is between 0 and 10 years • Risk = 0 if greater than 10 years
Average Permit Time Before Crash by Town	• Risk = 0 if permit time is less than or equal to 3 years • Risk = 1 if greater than 3 years
Average Vehicle Purchase Time Before Crash by Town	• Risk = -0.05*PT + 1 if PT or purchase time is between 0 and 10 years • Risk = 0 if greater than 10 years

Table 3. Example risk score calculations for KA motorcycle crashes.

Variable	Characteristic	Risk Factor Scoring	Risk Score
MPO	OCPC	• Risk = 1 if MPO is OCPC • Risk = 0.75 if NMCOG or SRPEDD • Risk = 0 otherwise	1
Jurisdiction	City or Town Accepted Road	• Risk = 1 if jurisdiction is in MassDOT or City or Town Accepted Road • Risk = 0 otherwise	1
Urban Type	Rural	• Risk = 1 if urban type is rural • Risk = 0.5 if urban type is urban-large urbanized area or rural • Risk = 0 otherwise	1
AADT	500	• Risk = 0 if AADT is less than or equal to 1000 • Risk = $1.02*10^{-5}$*AADT + 0.490 if AADT is greater than 1000 and less than or equal to 50000 • Risk = 1 if greater than 50000	0
Median Width	0	• Risk = 1 if median width is between 11 and 50 feet • Risk = 0 otherwise	0
Curb	None	• Risk = 1 if curb on right side only • Risk = 0 otherwise	0
Average Number of Motorcycle Registration by Town	350	• Risk = 0 if registration number is less than or equal to 250 • Risk = $3.081*10^{-4}$*RN + 0.423 if RN or registration number is greater than 250 and less than or equal to 1873 • Risk = 1 if greater than 1873	0.53
Average License Time Before Crash by Town	5 years	• Risk = -0.05*LT + 1 if LT or license time is between 0 and 10 years • Risk = 0 if greater than 10 years	0.75
Average Permit Time Before Crash by Town	3.5 years	• Risk = 0 if permit time is less than or equal to 3 years • Risk = 1 if greater than 3 years	1
Average Vehicle Purchase Time Before Crash by Town	2 years	• Risk = -0.05*PT + 1 if PT or purchase time is between 0 and 10 years • Risk = 0 if greater than 10 years	0.9
		Total Risk Score (out of 10):	6.18
		Risk Percent Score:	61.8%

In order to finalize the data, MassDOT dissolved the road inventory based on the risk factor inputs to generate uniform corridors. These corridors can be used to identify targeted safety improvement projects. Additionally, MassDOT identified the closest address geospatially to the beginning and end of each corridor as reference points. The addresses include the street number, street name, and town of the address. Note these are the closest addresses geospatially, so the reference address may not be on the same street as the corridor itself, and the beginning and end reference address may be the same. MassDOT continues to provide mileposts for MassDOT routes and encourages users to use both mileposts and address points as references.

The segments were then ranked at both the Statewide and MPO levels using the normalized risk score and the percentile of score ranking (rank kind equal to weak) function in ArcGIS. For each normalized risk score, a percentile rank for the given score was computed relative to all the normalized risk scores. If there are repeated occurrences of the same normalized risk score, then the percentile rank corresponds to values that are less than or equal to the given score. The advantage of the weak ranking approach is that it guarantees that the highest normalized score will receive a percentile rank of 100%. The risk categories were then determined using the computed ranks. For example, sites ranked in the top 5 percentile (95 through 100) were categorized as "Primary Risk Site," sites ranked in the next 10 percentile (85 through 95) were categorized as "Secondary Risk Site," and the remaining sites were not categorized. In instances where there are large repeated occurrences of the same normalized risk score, the percentage of segments computed for top 5% or next 10% may not be equal to 5 or 10%. This is a byproduct of the weak ranking approach used. Table 4 and Table 5 show the distribution of focus facility type segments (urban minor arterial/rural major collect, local) with the normalized risk score (presented as percentages) across these categories for Statewide and MPO rankings, respectively.

Table 4. Statewide risk categories.

State	Risk Category	Minimum Normalized Risk Score Percentage	Maximum Normalized Risk Score Percentage	Number of Segments	Percent of Scored State Segments
MA	Primary Risk Site	51.30%	84.97%	13,582	5.00%
	Secondary Risk Site	44.03%	51.30%	27,165	10.00%

Table 5. MPO risk categories.

MPO	Risk Category	Minimum Normalized Risk Score Percentage	Maximum Normalized Risk Score Percentage	Number of Segments	Percent of Scored MPO Segments
Berkshire Regional Planning Commission	Primary Risk Site	39.90%	53.83%	618	5.74%
	Secondary Risk Site	29.90%	39.60%	1,156	10.74%
Boston Region MPO	Primary Risk Site	46.39%	67.41%	4,471	5.00%
	Secondary Risk Site	39.89%	46.38%	9,146	10.23%
Cape Cod Commission	Primary Risk Site	40.32%	60.32%	1,686	6.41%
	Secondary Risk Site	34.96%	40.26%	2,515	9.55%
Central Massachusetts Regional Planning Commission	Primary Risk Site	50.02%	75.52%	1,289	5.07%
	Secondary Risk Site	45.18%	50.01%	2,699	10.63%
Franklin Regional Council of Governments	Primary Risk Site	40.04%	59.84%	336	5.36%
	Secondary Risk Site	34.44%	39.94%	879	14.03%
Martha's Vineyard Commission	Primary Risk Site	25.04%	40.74%	396	11.98%
	Secondary Risk Site	19.75%	25.00%	240	7.26%
Merrimack Valley Planning Commission	Primary Risk Site	48.97%	59.97%	862	6.51%
	Secondary Risk Site	43.17%	48.87%	1,726	13.04%
Montachusett Regional Planning Commission	Primary Risk Site	45.28%	64.82%	674	5.03%
	Secondary Risk Site	39.02%	45.18%	1,348	10.05%
Nantucket Planning and Economic Development Commission	Primary Risk Site	25.80%	41.80%	196	8.27%
	Secondary Risk Site	15.80%	22.00%	1,243	52.45%
Northern Middlesex Council of Governments	Primary Risk Site	61.68%	75.36%	558	5.01%
	Secondary Risk Site	56.98%	61.66%	1,399	12.56%
Pioneer Valley Planning Commission	Primary Risk Site	58.43%	75.03%	1,912	11.52%
	Secondary Risk Site	55.10%	58.38%	582	3.51%
Old Colony Planning Council	Primary Risk Site	46.14%	63.53%	1,241	5.01%
	Secondary Risk Site	39.24%	46.11%	3,619	14.62%
Southeastern Regional Planning and Economic Development District	Primary Risk Site	57.87%	84.97%	1,964	6.87%
	Secondary Risk Site	50.20%	57.75%	2,328	8.14%