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

Rural and Urban Roadway Departures

PREPARED FOR

Massachusetts Department of Transportation

PREPARED BY

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 is a part of the SDI project and summarizes the risk factor analysis performed for rural and urban roadway departure (RwD) crashes. This report describes a method to identify risk factors using mileage as the exposure parameter, which is one potential method to identify risk factors under the SDI grant. Reports for the other emphasis areas describe different methods used to adapt to the needs of those areas.

The overall purpose of this report is to analyze and discuss the results of a systemic analysis of RwD crashes in rural and urban areas for MassDOT. VHB performed this analysis using crash, traffic, and roadway data in Massachusetts to identify focus facility types and risk factors for RwD crashes. VHB followed the tasks described in FHWA's *Systemic Safety Project Selection Tool*[1] for a systemic analysis:

1. Select Focus Crash Types
2. Select Focus Facilities
3. Identify and Evaluate Risk Factors

Systemic safety analysis is a complimentary approach to roadway safety management and is particularly useful for addressing crash types that tend to be spread-out across a network or facility (as opposed to crash types that tend to form hotspots). For example, the systemic approach is particularly well-suited to address the seemingly random locations of fatal and severe RwD crashes.[2] As such, agencies have used systemic analysis as a tool to reduce the likelihood of severe or fatal RwD crashes. With the systemic approach, agencies identify which crashes they will focus on (e.g., RwD crashes), which facilities they occur most frequently on, and what risk factors are associated with the focus crashes on the focus facilities. Risk factors are typically selected by identifying roadway characteristics which are "overrepresented" for a combination of fatal (K), serious/incapacitating (A), and sometimes non-incapacitating (B) injury crashes. Overrepresentation is identified by comparing the proportion of the high severity (KA or KAB) crashes with and without the characteristic(s) of interest to the proportion of the network with and without the characteristic(s) of interest. In this case, segment length is the measure of exposure for the proportion of the network with and without the characteristic(s) of interest. By using this approach, agencies can use crash and roadway data from events which occurred in the past to identify sites with a high risk of these events in the future. Agencies can then deploy low-cost and effective countermeasures to these high-risk sites in a proactive attempt to prevent these future crashes.

VHB compared the proportional distribution of the roadway characteristics in the high severity crashes against the segment length representing those same roadway characteristics on the focus facility type to identify elevated risk factors. Roadway characteristics that made up a statistically higher proportion of the

[1] Preston, H., Storm, R., Dowds, J. B., Wemple, B., Hill, C., & Systematics, C. (2013). *Systemic safety project selection tool* (No. FHWA-SA-13-019). United States. Federal Highway Administration. Office of Safety.

[2] https://www.fhwa.dot.gov/innovation/everydaycounts/edc_5/roadway_departures.cfm

high severity crashes than the measure of exposure were selected as risk factors. The selected roadway characteristics, or risk factors, include jurisdiction, annual average daily traffic (AADT), access control, median width, right shoulder width, right shoulder type, total number of lanes, curve radius, metropolitan planning organizations (MPOs), access control, and posted speed limit.

The analyses performed by VHB for rural and urban roadway departures are described in more detail in the following sections. More specifically, the discussion of results identify the risk factors for the selected focus facility type. MassDOT can use the results in this report to identify relevant countermeasures and potential installation locations to target RwD reductions on both rural and urban roadways. Additionally, this report includes three appendices: Appendix A includes the queries used to identify RwD crashes in rural and urban areas; Appendix B includes risk factor plots for the rural RwD analysis; and Appendix C includes the risk factor plots for the urban RwD analysis.

Focus Crash Types

In the most recent SHSP, MassDOT identified Lane Departure Crashes as an emphasis area.[3] MassDOT recommended non-intersection, RwD crashes as the focus crash type for this report. Based on discussions with MassDOT, VHB established two roadway departure focus crash types separated by rural and urban areas (see Table 1). Note that in the urban RwD analysis, crashes in the city of Boston were excluded per discussion with MassDOT.

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

Table 1. Focus crash types for roadway departure crashes.

Method of Identification	Focus Crash Type	
	Rural RwD	**Urban RwD**
Urban Type	Urban type is rural (Urban_Type = 5)	Urban type is not rural (Urban_Type <> 5) and city is not Boston (City <> 35)
Crash Severity	KAB	KA
Junction Type	Roadway Junction Type (RDWY_JNCT_TYPE_DESCR) is any of the following: • Not at junction • Driveway • Not reported • Unknown	
RwD Crashes	Vehicle sequence of events (VEHC_SEQ_EVENTS_CL) include any of the following: • Collision with curb • Collision with tree • Collision with utility pole • Collision with light pole or other post/support • Collision with guardrail • Collision with median barrier • Collision with ditch • Collision with embankment • Collision with highway traffic sign post • Collision with overhead sign support • Collision with fence • Collision with mail box • Collision with bridge • Collision with other fixed object (wall, building, tunnel, etc.) • Collision with unknown fixed object • Ran off road right • Ran off road left • Cross median or centerline	

Focus Facility Types

VHB identified the focus facility types for the rural RwD and urban RwD analysis using crash tree analysis on the MassDOT roadway inventory and crash data.

Rural RwD Crashes

VHB identified the rural RwD focus facility types using the processed crash and roadway inventory data. Due to the limited number of K and A severity crashes in the processed data, B injury crashes were included in the identification of the focus facility types. Both the crash and roadway inventory data have a functional class field, and the functional class from the roadway inventory was used in this crash tree analysis. Figure 1 summarizes the non-intersection, rural KAB RwD crashes as a crash tree. The crash tree identifies which facilities MassDOT should focus on for rural RwD efforts. The layer of the crash tree was built using the functional class field in the data. The rural local roads (functional class = 0) and rural major collector roads (functional class = 5) were identified as the top two focus facility types with the most KAB crashes. Based on these results, MassDOT chose to add minor collector roads (functional class = 6) to create a combined focus facility type – rural local and collector roads. These roads are eligible for High Risk Rural Road (HRRR) funding and will help MassDOT identify HRRR projects. Therefore, the combined HRRR focus facility types for rural RwD crashes include:

1. Rural local
2. Rural major collector
3. Rural minor collector

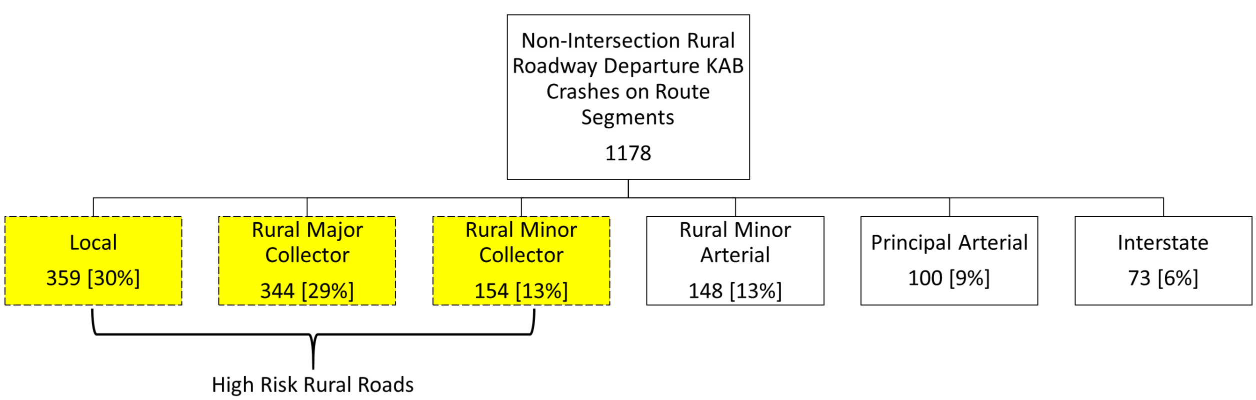

Figure 1. Crash tree to identify focus facility types for non-intersection rural RwD crashes.

Urban RwD Crashes

VHB identified the urban RwD focus facility types using the maximum crash severity and route-linked functional class from the crash data. Due to the high number of urban crashes, only the K and A severity crashes were used in the identification of the focus facility types. The functional class from the roadway inventory was not used in this crash tree analysis in order to minimize errors introduced in the spatial joining between the high number of urban crashes and dense road networks in urban areas. Note that the roadway inventory data is used later in the urban RwD risk factor analysis to identify geolocated crashes after the focus facility types have been determined.

Figure 2 summarizes the non-intersection, urban KA RwD crashes as a crash tree. The crash tree identifies which facilities MassDOT should focus on for urban RwD efforts. The layer of the crash tree was built using the functional class field in the crash data. The urban principal arterial (function class = 2 or 3) and urban minor arterial roads (functional class = 5) were identified as the top two focus facility types with the most KA crashes. Therefore, the combined focus facility type for urban RwD crashes include:

1. Urban principal arterial
2. Urban minor arterial

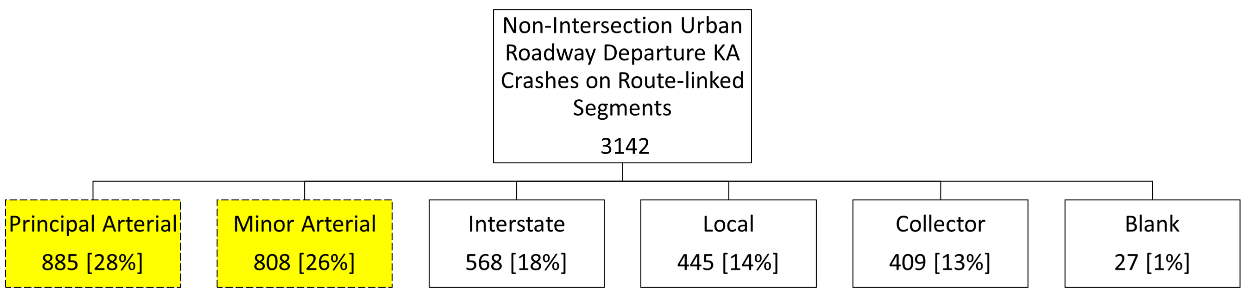

Figure 2. Crash tree to identify focus facility types for non-intersection urban RwD crashes.

Risk Factor Analysis

After identifying focus crash types and focus facility types, VHB proceeded with the risk factor analysis for the rural RwD and urban RwD analysis, respectively.

Rural RwD Crashes

The following sections describe the data, methodology, and results of the rural RwD analysis.

Data

MassDOT provided crash, curve, and roadway data to VHB for this analysis. The crash data included all crashes from 2013 to 2017; the curve data included the horizontal curve radius; and the roadway data included the roadway segment attributes, such as functional class, roadway geometry, physical conditions, traffic volumes, and more. All data were provided in a spatial format allowing VHB to integrate the data spatially using geoprocessing tools in ArcGIS Pro (version 2.5.2, ESRI Inc.).

Pre-processing steps were performed on both the crash and roadway inventory data prior to analysis (see Appendix A for more details on the queries used). First, the input crash data was filtered using the vehicle sequence of events-all vehicles (equal to multiple selections), roadway junction type (equal to multiple selections), and urban type (equal to rural) to identify the non-intersection, rural RwD crashes. Second, the input roadway inventory data was filtered using mileage counted (equal to 1), jurisdiction (not equal to null), facility type (less than 7), and urban type (equal to rural) to identify the unique, rural segments that were counted for the Highway Performance Monitoring System (HPMS). Filtering the roadway inventory in this way prevented potential double-counting of mileage for divided roads and roads with overlapping route numbers.

For the risk factor analysis, both the curve and filtered roadway inventory data were spatially joined to the crash data. This merged crash data was used to determine the proportion of KAB crashes. Similarly, the filtered roadway inventory was spatially joined to the curve data in order to determine the proportion of segment length for the curve segments. The curve length was computed using the calculate geometry tool in ArcGIS. Lastly, the filtered roadway inventory data was used directly to determine the proportion of segment length for the road segments. The segment length was provided in the input data and these values were consistent with the results obtained using the ArcGIS calculate geometry tool.

Methodology

Having defined the focus crash type and focus facility type, VHB reviewed the processed rural crash data, the rural roadway data, and the rural curve data to identify risk factors for the rural, KAB RwD crashes. The selected roadway characteristics, or potential risk factors, include jurisdiction, AADT, access control, median width, right shoulder width, right shoulder type, total number of lanes, curve radius, MPOs, and posted speed limit. Using the roadway characteristics, VHB compared the proportion of KAB crashes with these roadway characteristics against the proportion of segment length (in miles) with the same roadway characteristics. For each roadway characteristic (e.g., right shoulder width), a roadway attribute (e.g., right shoulder width = 0′,1′-4, 5′-8′, 9′-13′, 13′+, blank) was identified as a risk factor if the percentage of KAB crashes with that attribute was statistically higher (e.g., the KAB proportion including the error bar exceeds the proportion of the segment length). This statistical comparison was done using the margin of error[4] for a sample proportion at the 90% confidence interval. This error is defined as follows:

$$Margin\ of\ Error = z^* \sqrt{\frac{\hat{\rho}(1-\hat{\rho})}{n}}$$

where z^* is obtained from the standard normal distribution at the desired percentage confidence value, n is the sample size, and $\hat{\rho}$ is the sample proportion. For example, the margin of error for the KAB crashes in rural Local routes is computed using n = 359 for the KAB crashes, z^* = 1.645 for a percentage confidence of 90%, and $\hat{\rho}$ is the decimal proportion of each roadway attribute for that roadway characteristic. For the segment length proportions, error bars are not displayed because this error is mainly due to machine precision from the ArcGIS calculate geometry function.

Figure 3 shows an example of this risk factor comparison for right shoulder width on HRRR roads. In this example, a right shoulder width between 1′-4′ was identified as a risk factor because the KAB proportion (including the lower bound of its error) exceeded the segment length proportions.

For all the selected roadway characteristics, their attribute values were grouped as needed to better illustrate their proportions. Blanks and zero entries were often treated as separate attribute categories. A blank value means that either the input data was “null” or the processed crash data resulted in a “null” join. This information is shown in the risk factor plots because it may be helpful in identifying gaps in the data. If the blanks represented a significant percentage of the risk factor proportions, a follow-up risk factor plot was created without the blanks. This was done for the curve radius and speed risk factors.

[4] Moore, D. S. and McCabe G. P. Introduction to the Practice of Statistics. New York: W. H. Freeman, p. 443, 1999.

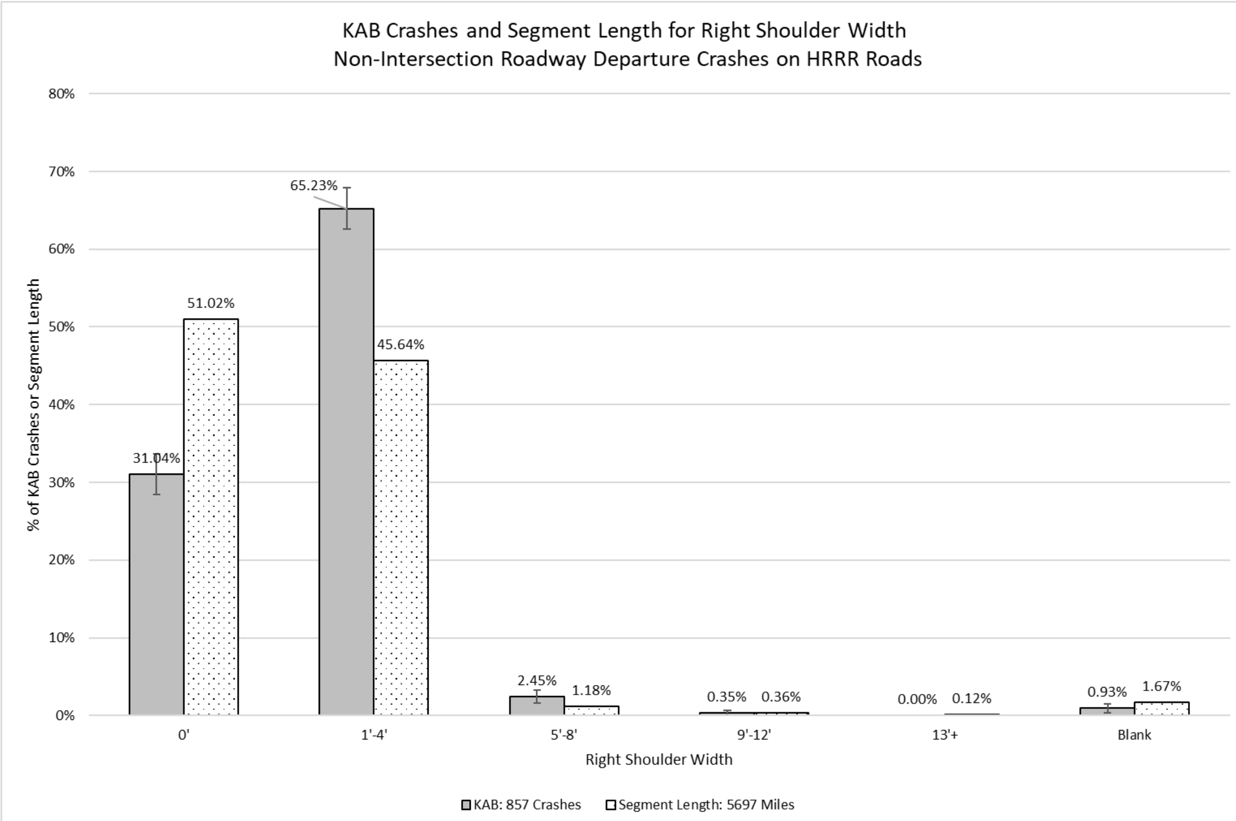

Figure 3. Sample risk factor analysis for right shoulder width on HRRR roads.

Results

Based on the above described method, the following risk factors were identified for non-intersection RwD crashes for HRRR roads (see Appendix B for all risk factor plots). When selecting a risk factor, VHB considered whether the characteristic made up a meaningful proportion of the roadway network.

Risk factors for HRRR Roads (Rural Major Collector, Rural Minor Collector, Rural Local):

1. MassDOT owned roadway.
2. AADT known and greater than or equal to 1,000 vehicles per day.
3. Right shoulder width between 1 and 4 feet.
4. Stable shoulder, unstable, or hardened shoulder.
5. Two-lane roadway.
6. Horizontal curve radius less than 500 feet.[5]
7. Central Massachusetts Regional Planning Commission (CMRPC), Martha's Vineyard Commission (MVC), Montachusett Regional Planning Commission (MRPC), Southeastern Regional Planning and Economic Development District (SRPEDD).

[5] Based on crashes only with known radius value.

8. Posted speed limit of 40 miles per hour.[6]

Note these risk factors are correlations and should not be interpreted as causation. As such, these risk factors are correlated with higher than expected KAB crash frequency compared to the distribution of mileage.

Urban RwD Crashes

The following sections describe the data, methodology, and results of the urban RwD analysis. The urban RwD analysis used the same input datasets and methodology as described in the rural RwD analysis. However, there were some differences due to the different focus crash types, focus facility types, and crash severities used. These differences are described below.

Data

In the urban RwD analysis, the urban type filter was updated to consider urban areas only (e.g., urban type not equal to rural) and the city filter was added to exclude Boston (e.g., city not equal to Boston) due to known underreporting issues with the City's crash data. Unless otherwise stated, all other filters remained the same (see Appendix A for more details on the queries used).

For the risk factor analysis, both the curve and filtered roadway inventory data were spatially joined to the crash data. Due to the density of the road networks and the high number of crashes in urban areas, the roadway inventory data was filtered to include only the focus facility types (function class equal to urban principal arterial or urban minor arterials) prior to the spatial joins. The final processed crash data was used to determine the proportion of KA crashes. The same filtered roadway inventory data was spatially joined to the curve data in order to determine the proportion of segment length for the curve segments. Lastly, the filtered roadway inventory data was used directly to determine the proportion of segment length for the road segments.

Methodology

In the urban RwD analysis, the proportion of KA crashes, instead of KAB crashes in the rural RwD analysis, was used to identify the risk factors in the urban RwD analysis. Note that the combined number KA crashes on urban principal arterial and urban minor arterial roads is 1963 using the functional classification from the roadway inventory data. This number is slightly higher than the number identified using the functional classification from the crash data in the crash tree analysis. This difference may be due to the following: 1) the difference in the functional class fields and filters used to developed these fields in the crash and roadway inventory; and 2) the spatial processing used to combine the crash and roadway inventory data for risk factor analysis may have captured additional segments in order to be conservative. In addition, for select roadway characteristics (for example, AADT), the groupings used on the observed values were updated to better reflect the risk factor proportions. Lastly, for the curve radius, two risk factor plots (with and without the blank entries) were created. This was necessary because the blank values for the curve radius curve represented a significant percentage of the risk factor proportions.

Figure 4 shows an example of this risk factor comparison for right shoulder width on urban principal arterials and urban minor arterials. In this example, a right shoulder width between 5' and 12' was identified as a risk factor because the KA proportion (including the lower bound of its error) exceeded the segment length proportions.

[6] Based on crashes only with known posted speed limit.

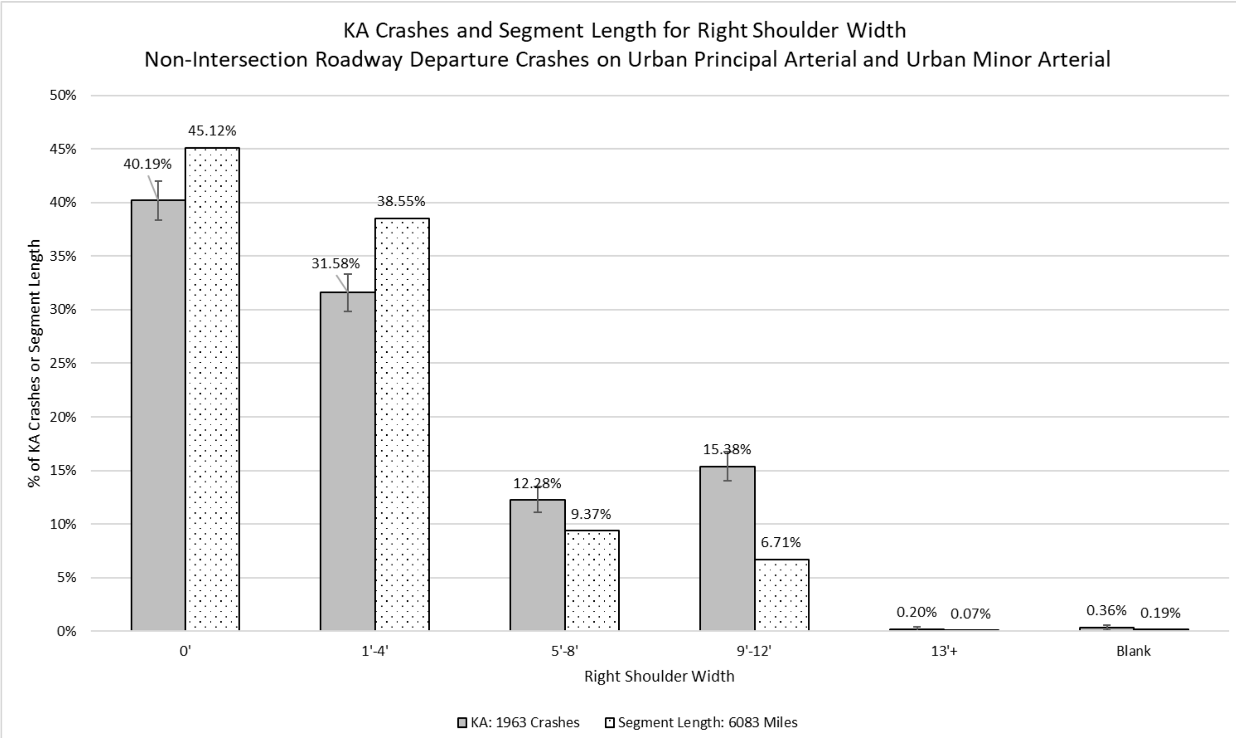

Figure 4. Sample risk factor analysis for right shoulder width on urban principal arterial and urban minor arterial roads.

Results

Based on the above described method, the following risk factors were identified for non-intersection RwD crashes for urban principal arterials and urban minor arterials (see Appendix C for all risk factor plots). When selecting a risk factor, VHB considered whether the characteristic made up a meaningful proportion of the roadway network.

Risk factors for Urban Principal Arterials and Urban Minor Arterials:

1. MassDOT owned roadway.
2. AADT known and greater than or equal to 20,000 vehicles per day.
3. Right shoulder width between 5 and 12 feet.
4. Hardened bituminous mix or penetration shoulder.
5. Four-lane or six-lane roadways.
6. Horizontal curve radius less than 500 feet.[7]
7. Old Colony Planning Council (OCPC) and Southeastern Regional Planning and Economic Development District (SRPEDD).
8. Posted speed limit of 50 to 65 miles per hour.

[7] Based on crashes only with known radius value.

9. Median width greater than or equal to 1 foot.
10. Partial or full access control.

Note these risk factors are correlations and should not be interpreted as causation. As such, these risk factors are correlated with higher than expected KA crash frequency compared to the distribution of mileage.

Conclusions and Recommendations

The purpose of this report is to summarize the systemic analysis of rural and urban RwD crashes on MassDOT highways. For rural RwD crashes, VHB and MassDOT focused on non-intersection crashes occurring on HRRR roads (rural major collector, rural minor collector, and rural local roads). For urban RwD crashes, VHB and MassDOT focused on non-intersection crashes occurring on urban principal arterials and urban minor arterials. To identify the risk factors, VHB used a comparison of the crash severity proportions (KAB or KA) against the mileage exposure proportions (segment or curve length).

Table 2 and Table 3 show the risk factors identified in this analysis for rural and urban roadway departures, respectively. As exemplified in these tables, MassDOT can calculate the risk score by assigning a risk point for every risk factor present for each segment. The risk factor scoring scheme was designed to range between 0 and 1 for each risk factor. Therefore, the total risk factor score is the sum of all the individual risk factor scores. In addition, these scores are normalized against the total number of risk factors in order to determine a normalized risk score percentage and to compute the primary and secondary risk sites to identify the segments with the most risk both statewide and within each MPO.

Rural RwD Crashes

The risk factor scheme for the rural RwD analysis is as follows. For the jurisdiction, AADT, right shoulder type, total number of lanes, and posted speed limit risk factors, a road segment received a score of 1 if that segment contained the risk factor value and 0 otherwise. For the right shoulder width, MPOs, and horizontal curve radius risk factor, the scoring scheme was modified to reflect a gradient of scores between 0 and 1:

- Right Shoulder Width Risk Factor Scoring Scheme:
    - Score = 1 if right shoulder width is between 1 and 2 feet;
    - Score = 0.5 if right shoulder width is between 3 and 4 feet; and
    - Score = 0 otherwise.
- Metropolitan Planning Organization Risk Factor Scoring Scheme:
    - Score = 1 for CMRPC;
    - Score = 0.75 for SRPEDD;
    - Score = 0.5 for MRPC;
    - Score = 0.25 for MVC; and
    - Score = 0 otherwise.
- Horizontal Curvature Risk Factor Scoring Scheme
    - Score = $\alpha * CR + \beta$

    - α = -0.001 or slope
    - β = 1.0354 or intercept
    - CR = curve radius between 1 and 499 feet
  - Score = 0 otherwise.

The slope and intercept of the horizontal curvature scoring scheme was designed such that the sharpest curves (the observed minimum curve radius was 33 feet) received a score of 1, the widest curves (the maximum curve radius equal to 499 feet) received a score of 0.5, and values in between received a linearly adjusted score between 0.5 and 1. The sum of the individual risk factor scores are then normalized by the total possible number of risk factors to obtain a normalized risk score percentage. Table 2 shows a normalized risk score percentage of 44% (3.5 out of 8).

Table 2. Example risk score calculations for rural RwD crashes on an HRRR segment.

Risk Factor	Value Constituting a Risk Factor	Value for Subject Segment	Risk Score
Jurisdiction	MassDOT	City or Town Accepted Road	0
AADT	Greater than or equal to 1,000 vehicles per day	1,350	1
Right Shoulder Width	1 foot to 4 feet	4 feet	0.5
Right Shoulder Type	Stable, Unstable, or Hardened	Stable	1
Number of Lanes	Two lanes	2	1
Horizontal Curvature	Curve radius less than 500 feet	850 feet	0
MPO	CMRPC, MVC, MRPC, SRPEDD	Berkshire Regional Planning Commission	0
Posted Speed Limit	40 miles per hour	50 miles per hour	0
		Risk Score =	**3.5**
		Normalized Risk Score Percentage =	**44%**

Urban RwD Crashes

The risk factor scheme for the urban RwD analysis is as follows. For the jurisdiction, right shoulder width, and right shoulder type risk factors, a road segment received a score of 1 if that segment contained the risk factor value and 0 otherwise. For the AADT, median width, MPOs, total number of lanes, posted speed limit, and access control, and horizontal curve radius risk factor, the scoring scheme was modified to reflect a gradient of scores between 0 and 1:

- AADT Risk Factor Scoring Scheme:
  - Score = $\alpha * AADT + \beta$
    - $\alpha = -1.7*10^{-5}$ or slope
    - β = 1.333 or intercept

    - *AADT* = AADT between 20,000 and 65,000
  - Score = 0.25 for AADT greater than 65,000; and
  - Score = 0 otherwise.
- Median Width Risk Factor Scoring Scheme:
  - Score = $\alpha * MW + \beta$
    - α = -6.3*10^{-4} or slope
    - β = 1.001 or intercept
    - *MW* = medium width greater than or equal to 1 feet
  - Score = 0 otherwise.
- Total Number of Lanes Risk Factor Scoring Scheme:
  - Score = 1 if total number of lanes is 4;
  - Score = 0.5 if total number of lanes is 6; and
  - Score = 0 otherwise.
- Metropolitan Planning Organization Risk Factor Scoring Scheme:
  - Score = 1 for SRPEDD;
  - Score = 0.5 for OCPC; and
  - Score = 0 otherwise.
- Posted Speed Limit Risk Factor Scheme:
  - Score = 1 if posted speed limit is between 60 and 65 mph;
  - Score = 0.75 if posted speed limit is 55 mph;
  - Score = 0.5 if posted speed limit is 50; and
  - Score = 0 otherwise.
- Access Control Risk Factor Scoring Scheme:
  - Score = 1 for full access control;
  - Score = 0.5 for partial access control; and
  - Score = 0 otherwise.
- Horizontal Curvature Risk Factor Scoring Scheme
  - Score = $\alpha * CR + \beta$
    - α = -0.001 or slope
    - β = 1.006 or intercept
    - *CR* = curve radius between 1 and 499 feet
  - Score = 0 otherwise.

The slope of intercept of the AADT scoring scheme was designed such that values in between 20,000 and 65,000 AADT received a linearly adjusted score between 1 and 0.25, values greater than 65,000 received a constant score of 0.25, and values outside of the risk factor range received a score of 0. The slope of intercept of the median width scoring scheme was designed such that the smallest non-zero median width (the observed minimum median width was 1 feet) received a score of 1, and the largest median width (the observed maximum median width was 800 feet) received a score of 0.5, and values in between received a linearly adjusted score between 0.5 and 1. Similarly, the slope and intercept of the horizontal curvature scoring scheme was designed such that the sharpest curves (the observed minimum curve radius was 6.4 feet) received a score of 1, the widest curves (the maximum curve radius equal to 499 feet) received a score of 0.5, and values in between received a linearly adjusted score between 0.5 and 1. The sum of the individual risk factor scores are then normalized against the total possible number of risk factors to obtain a normalized risk score percentage. Table 3 shows a normalized risk score percentage of 20% (2 out of 10).

Table 3. Example risk score calculations for urban RwD crashes on an urban principal arterial and urban minor arterial segments.

Risk Factor	**Value Constituting a Risk Factor**	**Value for Subject Segment**	**Risk Score**
Jurisdiction	MassDOT	City or Town Accepted Road	0
AADT	Greater than or equal to 20,000 vehicles per day	1,350	0
Right Shoulder Width	5 feet to 12 feet	4 feet	0
Right Shoulder Type	Hardened bituminous mix or penetration	Stable	0
Number of Lanes	Four lanes or six lanes	2	0
Horizontal Curvature	Curve radius less than 500 feet	850 feet	0
MPO	OCPC and SRPEDD	Berkshire Regional Planning Commission	0
Posted Speed Limit	50 to 65 miles per hour	50 miles per hour	0.5
Median Width	Median width greater than or equal to 1 foot	2 feet	1
Access Control	Partial or full access control	Partial access control	0.5
		Total Score =	**2**
		Normalized Risk Score Percentage =	**20%**

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 addresses issues with ranking when there are multiple sites with the same score. Ranking was done separately for Rural and Urban segments, where rural segments were ranked only against other rural segments and urban segments were ranked only against other urban segments. The segments were then combined – resulting in both ~5 percent of rural and ~5 percent of urban segments ranked between 95 and 100.

The risk categories were then determined using the computed ranks. For example, sites ranked in the top 5 percentile (95 through 100) for either urban or rural were categorized as "Primary Risk Site," sites ranked in the next 10 percentile (85 through 95) for either urban or rural 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 5 show the distribution of focus facility type segments 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	50.00%	93.90%	4475	6.94%
	Secondary Risk Site	37.50%	49.90%	5890	9.14%

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	50.00%	81.63%	369	9.54%
	Secondary Risk Site	37.50%	49.90%	613	15.85%
Boston Region MPO	Primary Risk Site	52.70%	82.50%	848	5.01%
	Secondary Risk Site	30.00%	52.50%	1744	10.31%
Cape Cod Commission	Primary Risk Site	44.70%	76.80%	212	5.09%
	Secondary Risk Site	30.00%	44.60%	442	10.60%
Central Massachusetts Regional Planning Commission	Primary Risk Site	58.50%	86.50%	380	5.05%
	Secondary Risk Site	43.75%	58.38%	823	10.94%
Franklin Regional Council of Governments	Primary Risk Site	50.00%	84.88%	307	8.18%
	Secondary Risk Site	37.50%	49.88%	503	13.41%
Martha’s Vineyard Commission	Primary Risk Site	59.38%	59.38%	13	6.81%
	Secondary Risk Site	34.38%	53.13%	50	26.18%
Merrimack Valley Planning Commission	Primary Risk Site	39.90%	71.80%	184	5.14%
	Secondary Risk Site	30.00%	39.80%	455	12.71%
Montachusett Regional Planning Commission	Primary Risk Site	56.25%	90.88%	275	6.29%
	Secondary Risk Site	43.75%	56.00%	440	10.06%
Nantucket Planning and Economic Development Commission	Primary Risk Site	37.50%	56.25%	20	8.23%
	Secondary Risk Site	31.25%	31.25%	45	18.52%
Northern Middlesex Council of Governments	Primary Risk Site	39.60%	75.60%	135	5.14%
	Secondary Risk Site	26.30%	39.50%	260	9.90%
Pioneer Valley Planning Commission	Primary Risk Site	45.00%	84.70%	187	5.17%
	Secondary Risk Site	34.20%	44.90%	359	9.93%
Old Colony Planning Council	Primary Risk Site	50.00%	79.00%	547	6.66%
	Secondary Risk Site	37.50%	49.90%	961	11.70%
Southeastern Regional Planning and Economic Development District	Primary Risk Site	59.38%	93.90%	347	6.45%
	Secondary Risk Site	45.10%	59.30%	463	8.61%

Appendix A – Data Queries

A1. The following queries were used to include the data of interest from the roadway inventory, crash, and curve datasets for the rural RwD analysis.

Roadway Inventory Query: The roadway inventory data was preprocessed using a Mile_Count, Facility, and Jurisdictn query.

Mile_Count = **1** And Facility < **7** And Jurisdictn IS NOT NULL

The processed road inventory data was filtered again to include only the rural segments.

Urban_Type = **5**

Crash Query: The crash data was preprocessed using a RDWY_JNCT_TYPE_DESCR and VEH_SEQ_EVENTS_CL query. The '%(Collision with bridge)%' attribute has to be filtered with parentheses due to another occurrence with 'Collision with bridge overhead structure' that was excluded.

(RDWY_JNCT_TYPE_DESCR = 'Not at junction' Or RDWY_JNCT_TYPE_DESCR = 'Driveway' Or RDWY_JNCT_TYPE_DESCR = 'Not reported' Or RDWY_JNCT_TYPE_DESCR = 'Unknown') And (VEHC_SEQ_EVENTS_CL LIKE '%Collision with curb%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with tree%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with utility pole%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with light pole or other post/support%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with guardrail%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with median barrier%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with ditch%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with embankment%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with highway traffic sign post%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with overhead sign support%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with fence%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with mail box%' Or VEHC_SEQ_EVENTS_CL LIKE '%(Collision with bridge)%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with other fixed object(wall, building, tunnel, etc.)%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with unknown fixed object%' Or VEHC_SEQ_EVENTS_CL LIKE '%Ran off road right%' Or VEHC_SEQ_EVENTS_CL LIKE '%Ran off road left%' Or VEHC_SEQ_EVENTS_CL LIKE '%Cross median or centerline%')

The processed crash data was filtered again to include only rural segments (Urban_Type_1 from the roadway inventory attributes) after joining with road inventory and curve data.

Urban_Type_1 = **5**

Curve Query: The curve data was spatially joined to the filtered roadway inventory data, and the merged curve data was filtered to include only rural segments (Urban_Type from the roadway inventory attributes). The mileage count, jurisdiction, and facility filters are consistent with the roadway inventory data due to the spatial join.

Urban_Type = **5**

A2. The following queries were used to include the data of interest from the roadway inventory, crash, and curve datasets for the urban RwD analysis.

Roadway Inventory Query: The roadway inventory data was preprocessed using a Mile_Count, Facility, and Jurisdictn query.

Mile_Count = **1** And Facility < **7** And Jurisdictn IS NOT NULL

The preprocessed road inventory data was filtered again to exclude rural segments, to exclude the city of Boston, and include only principal arterial and minor arterial segments

Urban_Type <> **5** And City <> **35** And (F_Class = **2** or F_Class = **3** or F_Class = **5)**

Crash Query: The crash data was preprocessed using a RDWY_JNCT_TYPE_DESCR and VEH_SEQ_EVENTS_CL query. The ‘%(Collision with bridge)%’ attribute has to be filtered with parentheses due to another occurrence with ‘Collision with bridge overhead structure.’

(RDWY_JNCT_TYPE_DESCR = 'Not at junction' Or RDWY_JNCT_TYPE_DESCR = 'Driveway' Or RDWY_JNCT_TYPE_DESCR = 'Not reported' Or RDWY_JNCT_TYPE_DESCR = 'Unknown') And (VEHC_SEQ_EVENTS_CL LIKE '%Collision with curb%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with tree%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with utility pole%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with light pole or other post/support%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with guardrail%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with median barrier%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with ditch%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with embankment%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with highway traffic sign post%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with overhead sign support%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with fence%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with mail box%' Or VEHC_SEQ_EVENTS_CL LIKE '%(Collision with bridge)%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with other fixed object(wall, building, tunnel, etc.)%' Or VEHC_SEQ_EVENTS_CL LIKE '%Collision with unknown fixed object%' Or VEHC_SEQ_EVENTS_CL LIKE '%Ran off road right%' Or VEHC_SEQ_EVENTS_CL LIKE '%Ran off road left%' Or VEHC_SEQ_EVENTS_CL LIKE '%Cross median or centerline%')

The preprocessed crash data was spatially joined with the filtered the roadway inventory data. The merged crash data filtered again to exclude rural segments (Urban_Type_1 from the roadway inventory). The city of Boston was automatically excluded using this filter due to how the roadway inventory data was processed.

Urban_Type_1 <> **5**

Curve Query: The filters used for the roadway inventory data was applied to the curve data after the two datasets were spatially joined.

Appendix B – Non-Intersection RwD Crashes in Rural Areas: Risk Factor Plots

The following plots were used to identify risk factors for non-intersection RwD crashes on HRRR roads.

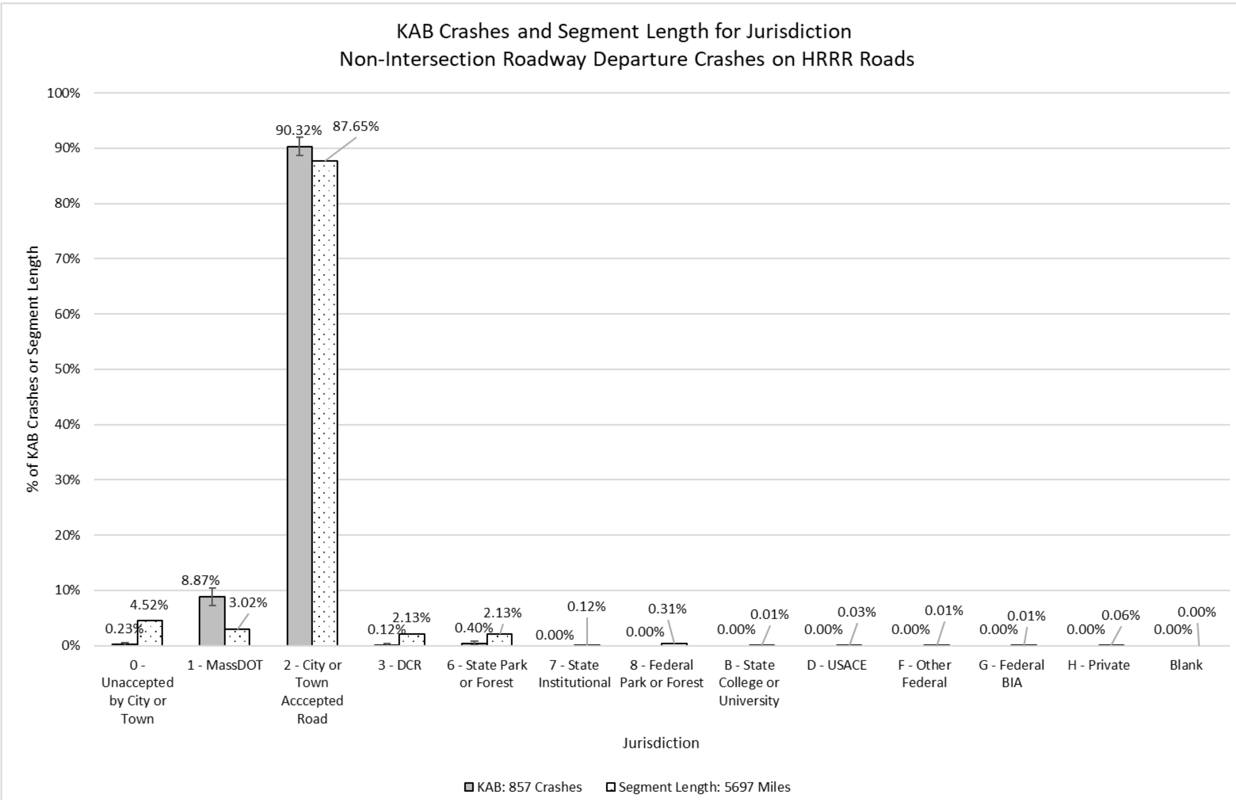

Figure B1. Risk factor analysis plot for Jurisdiction on HRRR roads.

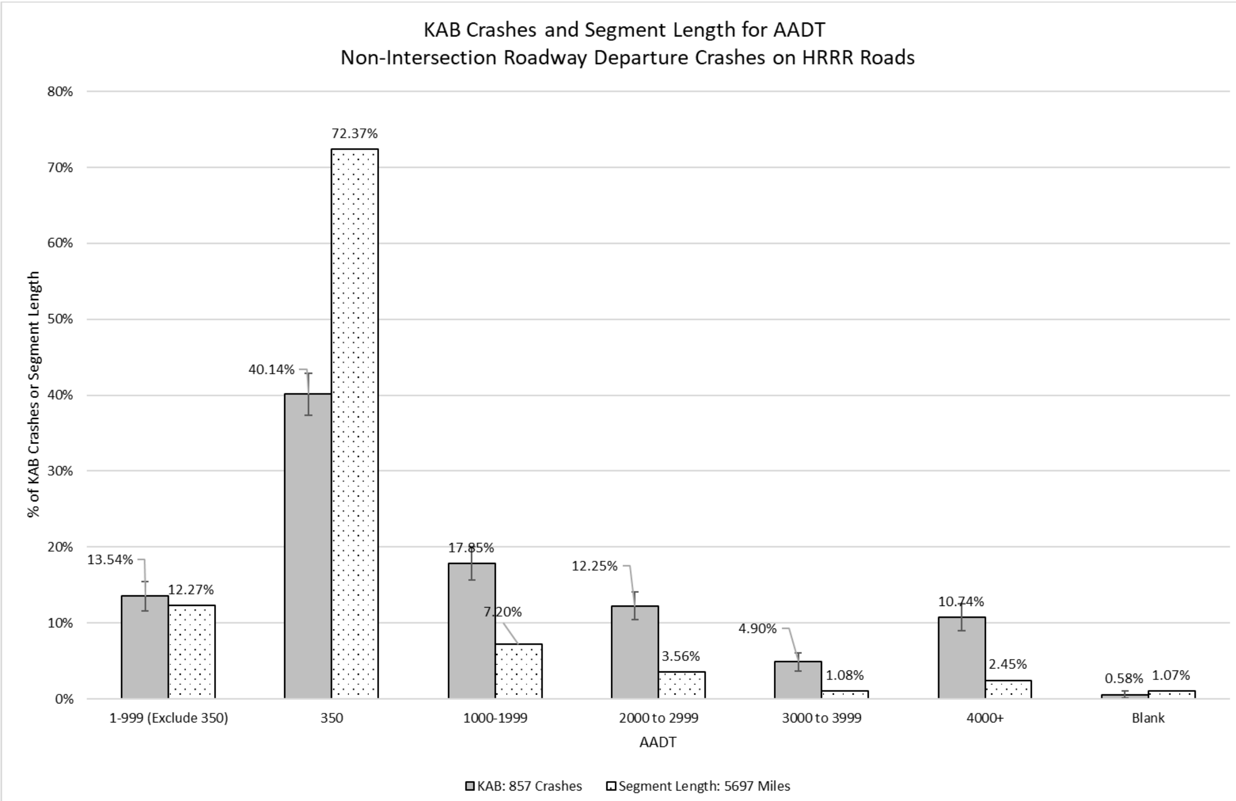

Figure B2. Risk factor analysis plot for AADT on HRRR roads.

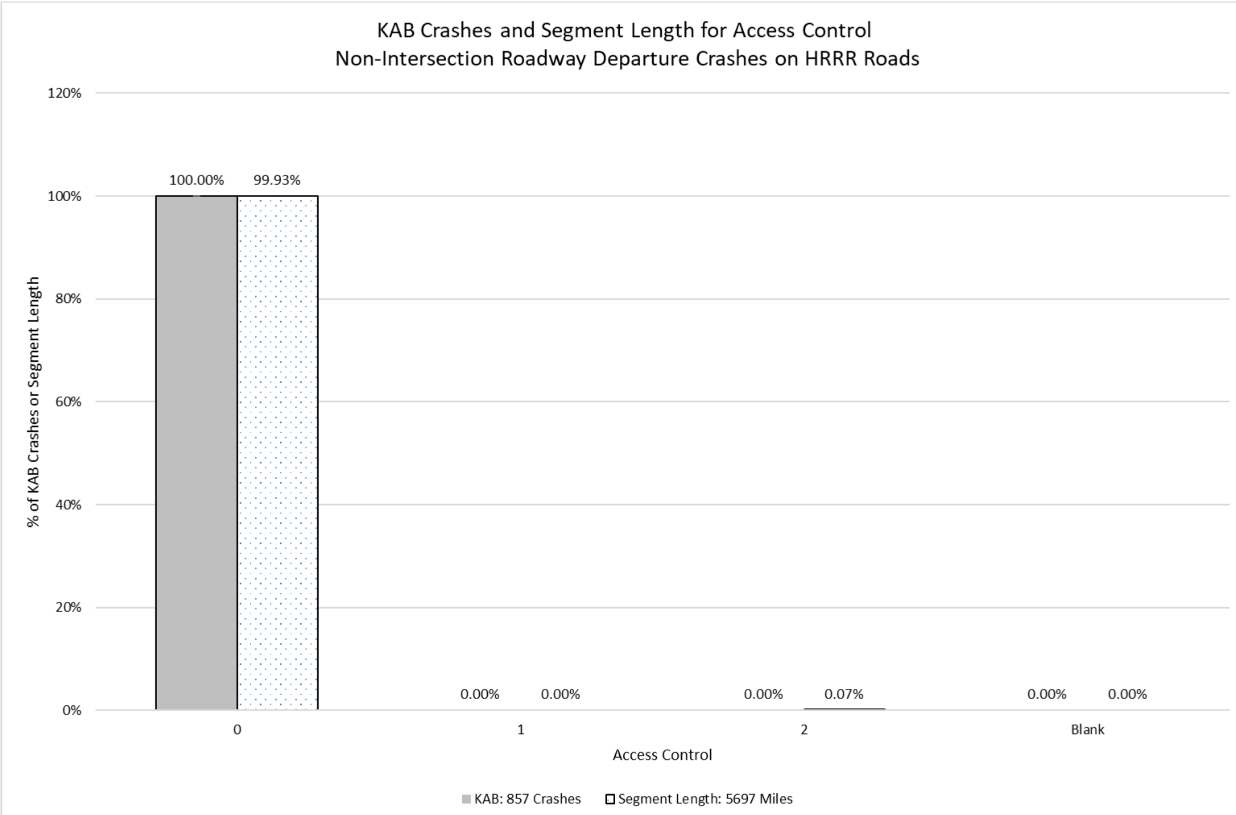

Figure B3. Risk factor analysis plot for access control on HRRR roads.

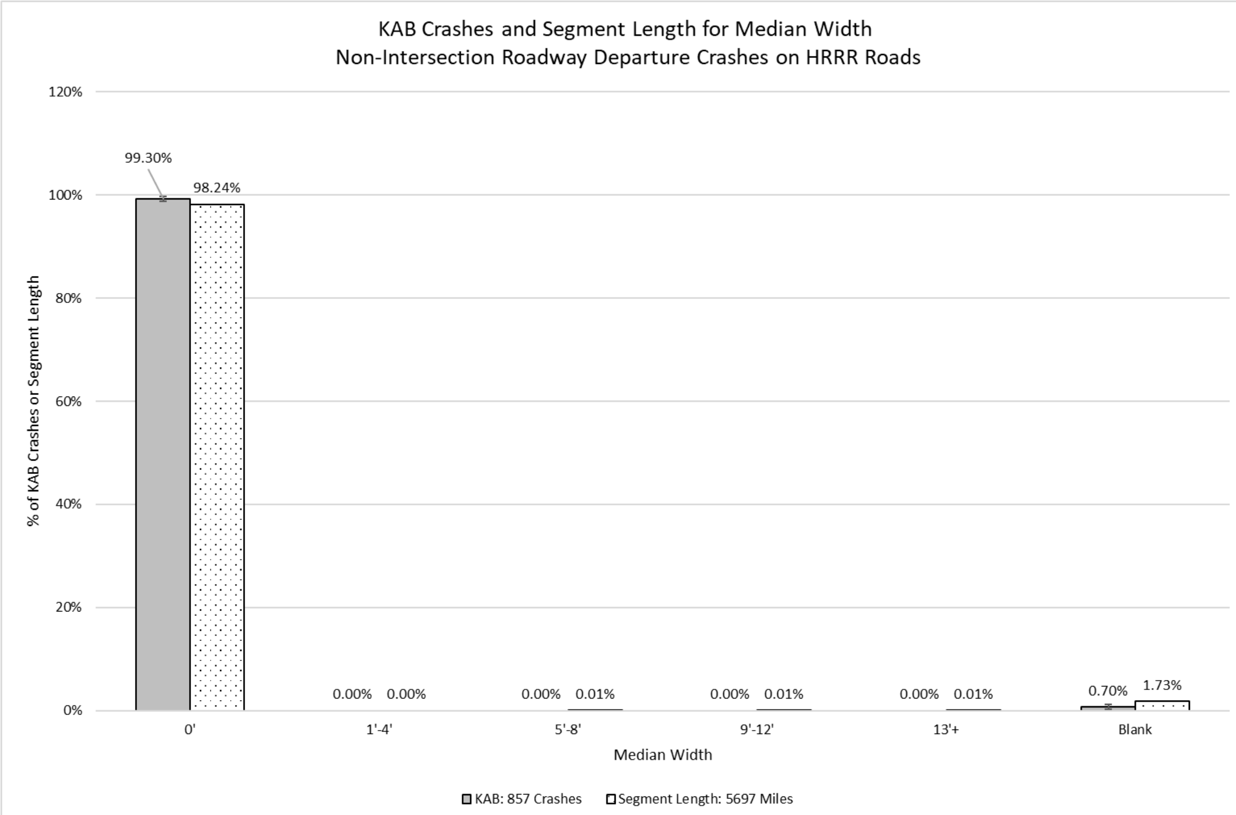

Figure B4. Risk factor analysis plot for median width on HRRR roads.

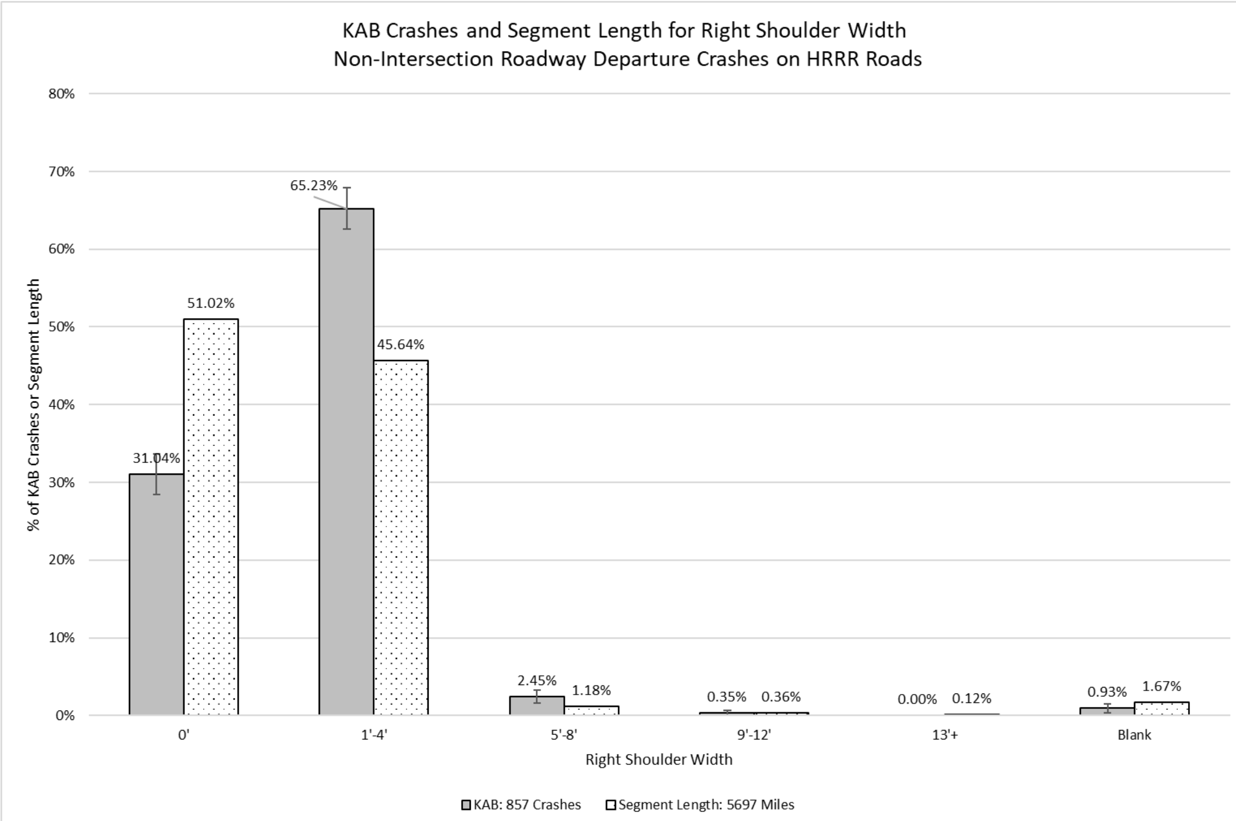

Figure B5. Risk factor analysis plot for right shoulder width on HRRR roads.

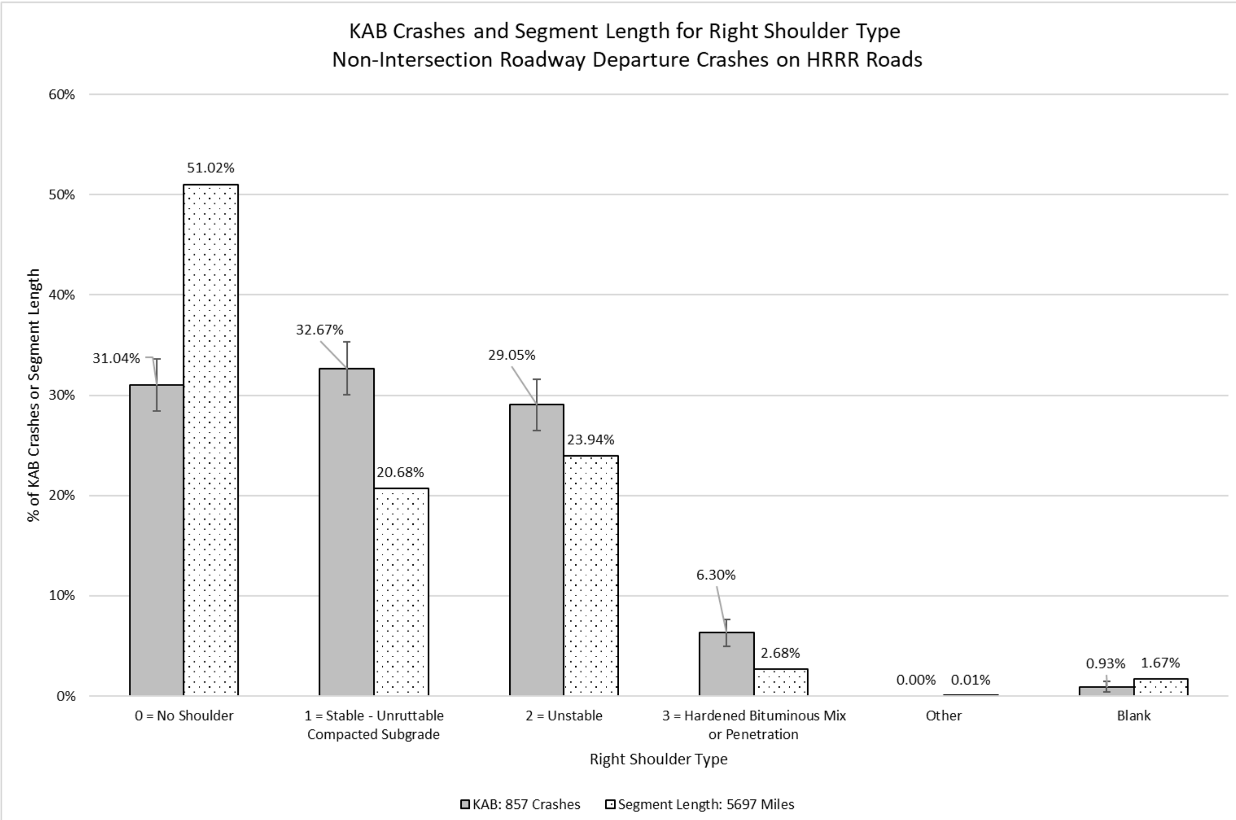

Figure B6. Risk factor analysis plot for right shoulder type on HRRR roads.

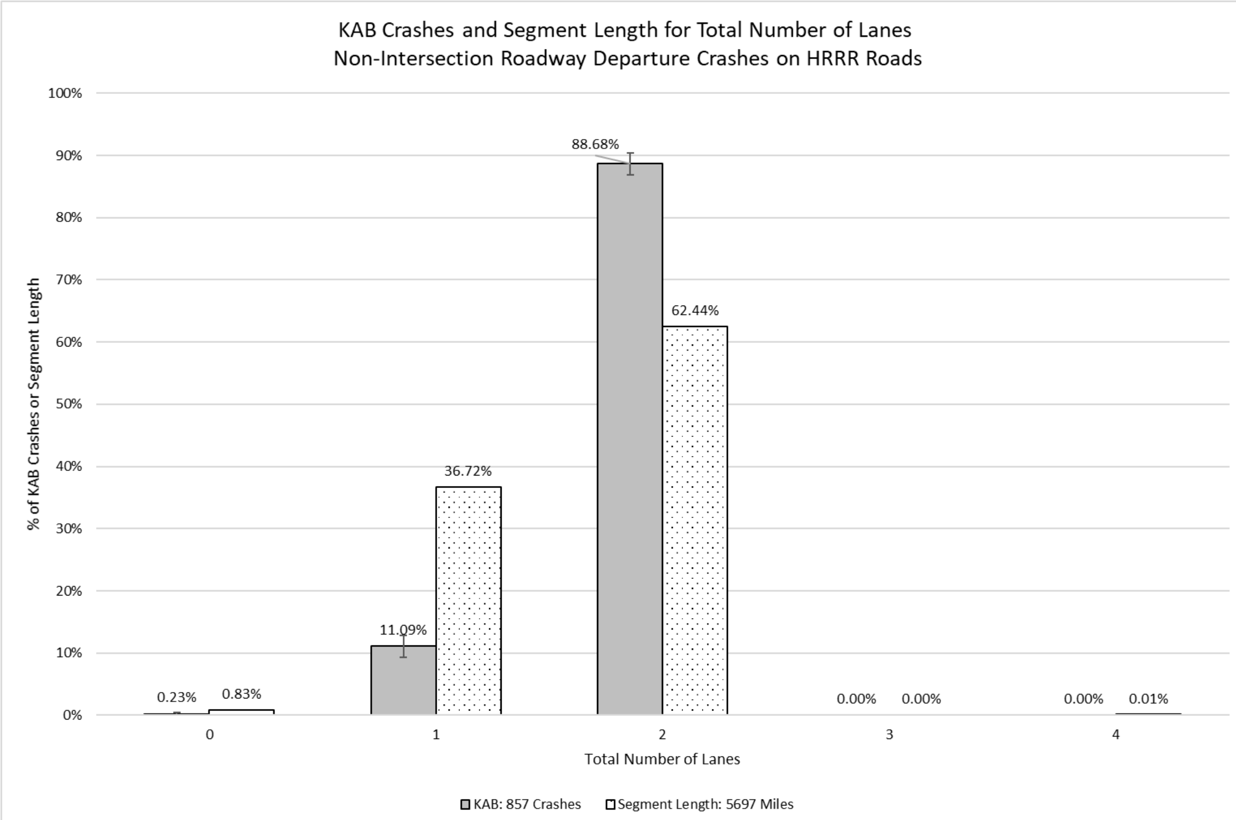

Figure B7. Risk factor analysis plot for total number of lanes on HRRR roads.

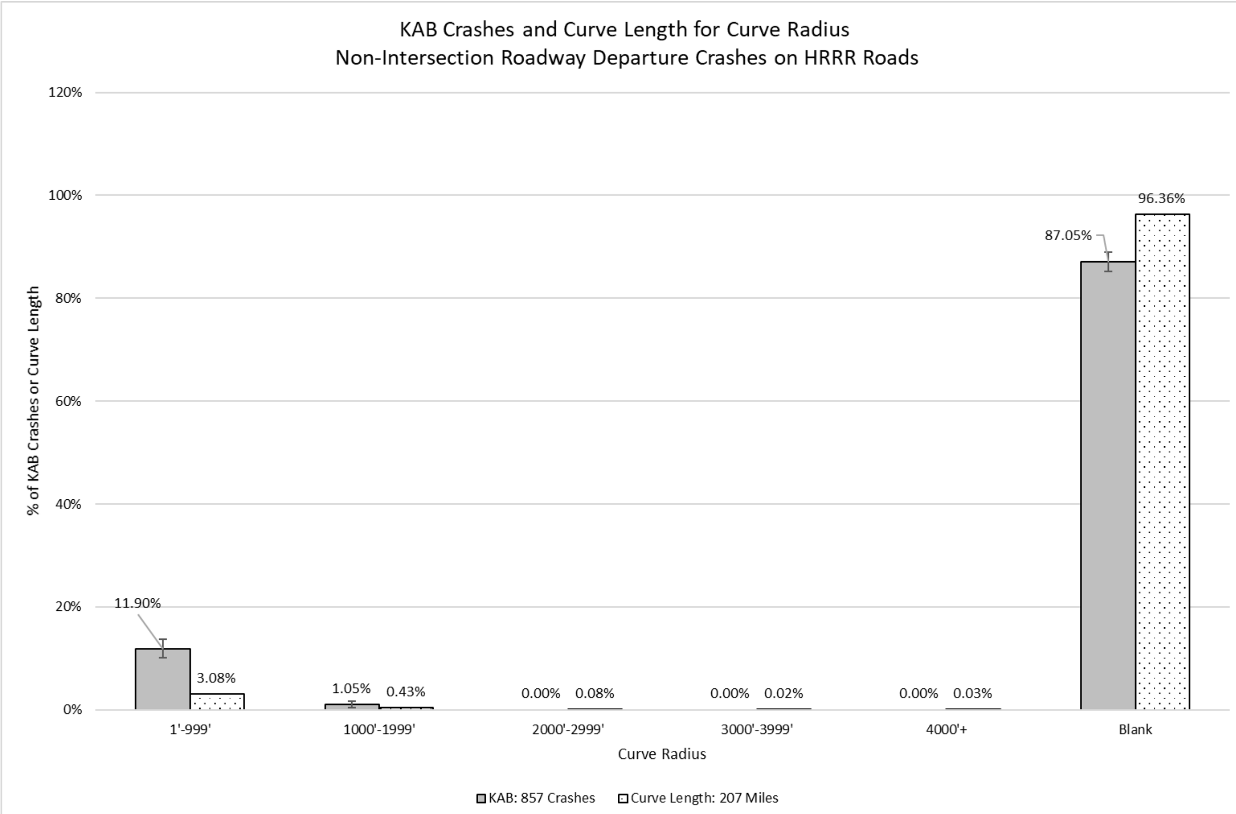

Figure B8. Risk factor analysis plot for curve radius on HRRR roads.

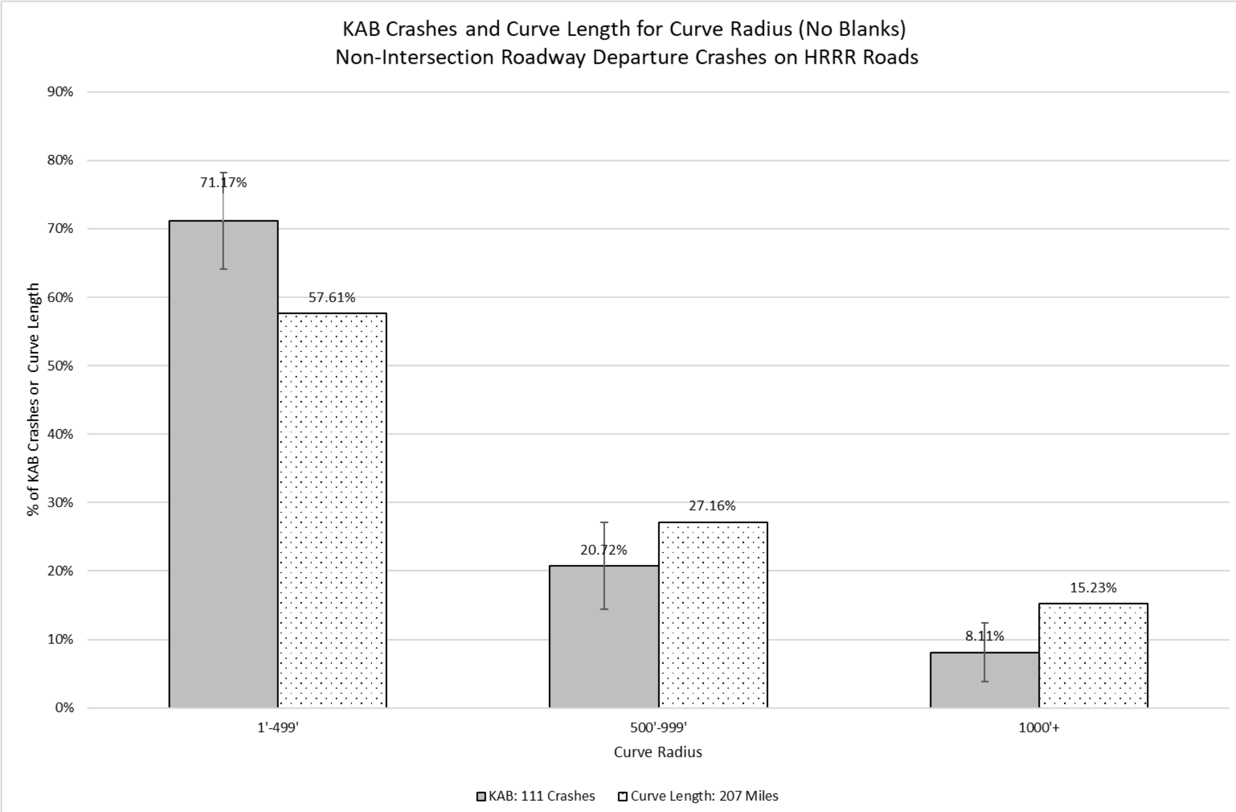

Figure B9. Risk factor analysis plot for curve radius (no blanks) on HRRR roads.

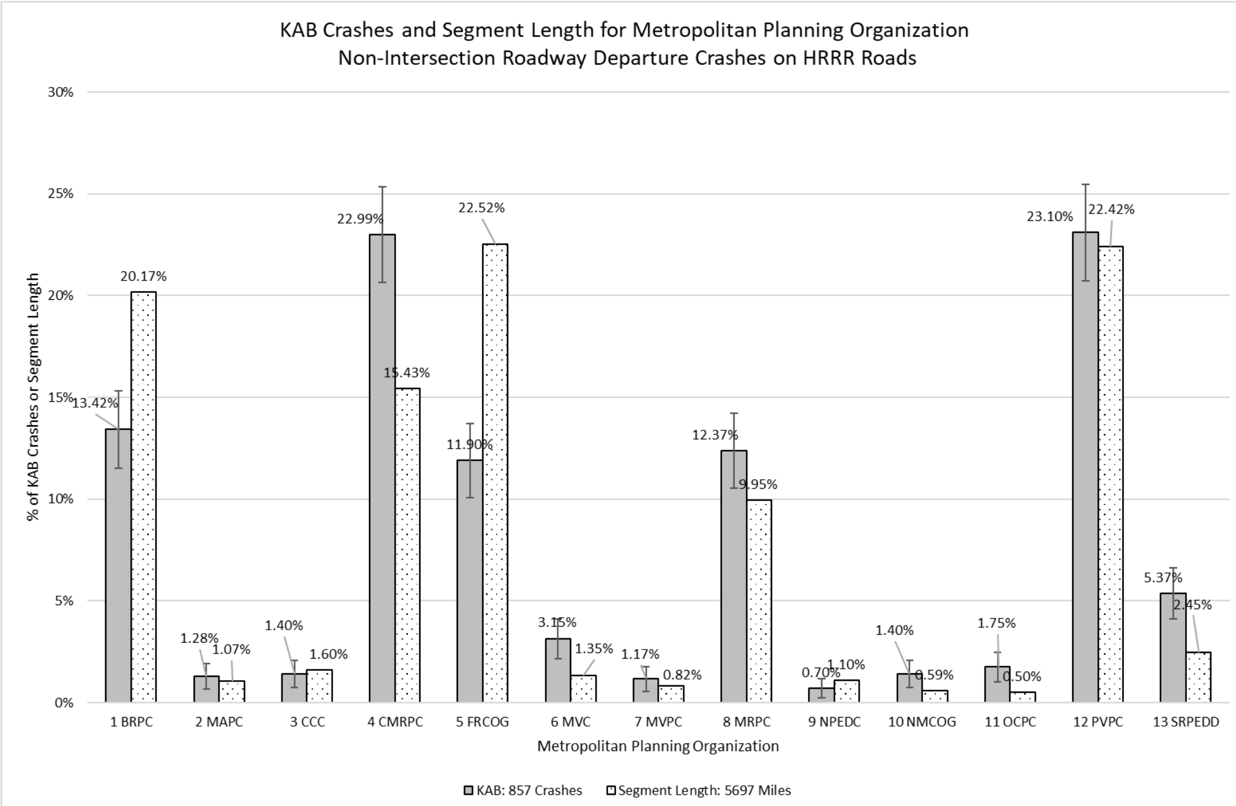

Figure B10. Risk factor analysis plot for metropolitan planning organization on HRRR roads.

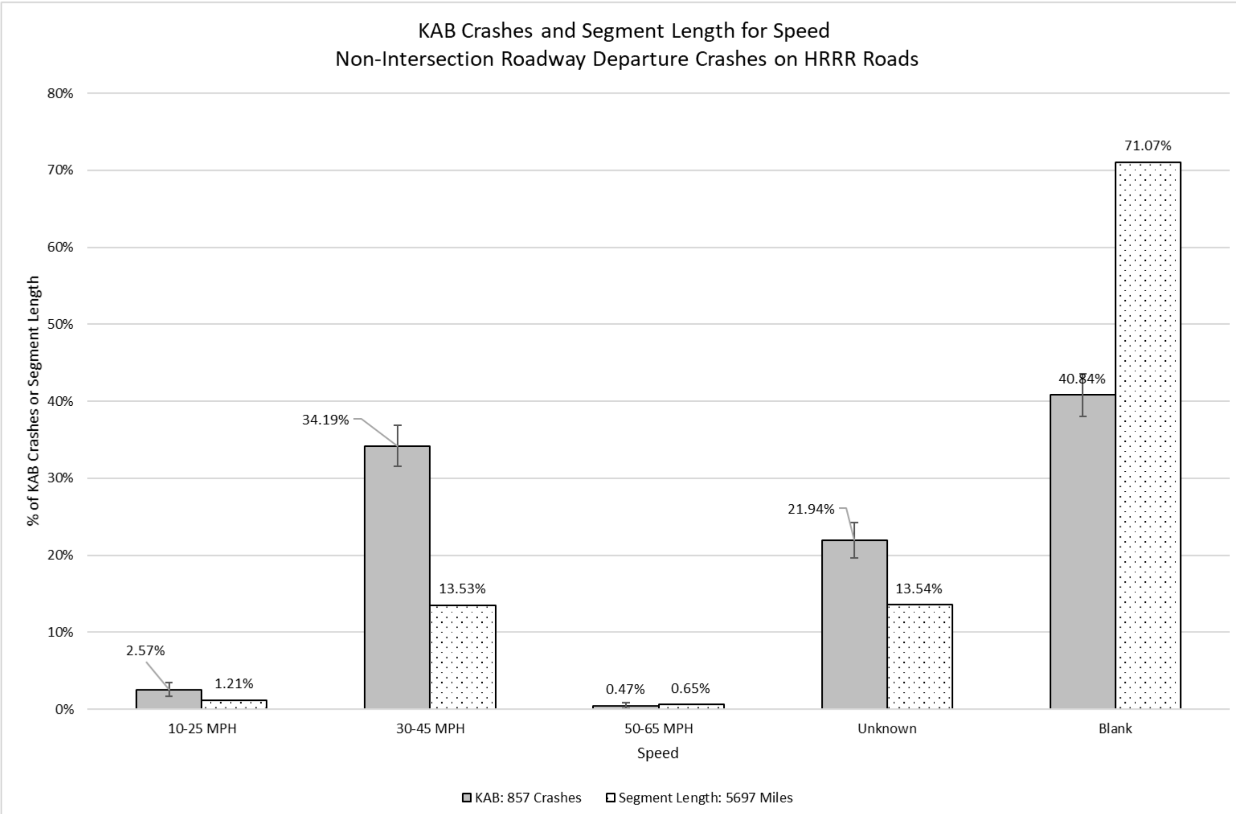

Figure B11. Risk factor analysis plot for posted speed limit on HRRR roads.

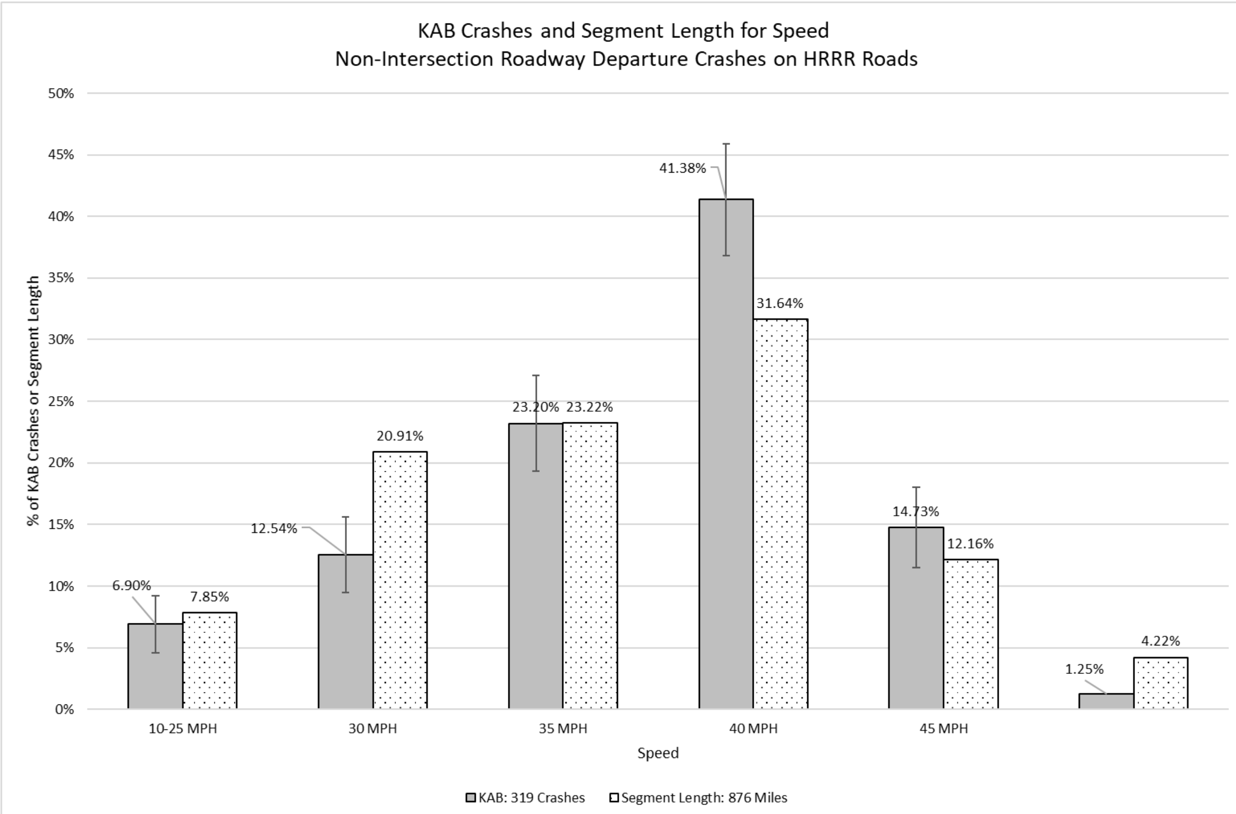

Figure B12. Risk factor analysis plot for posted speed limit (with no blanks or unknowns) on HRRR roads.

Appendix C – Non-Intersection RwD Crashes in Urban Areas: Risk Factor Plots

The following plots were used to identify risk factors for non-intersection RwD crashes on urban principal arterial and urban minor arterial roads.

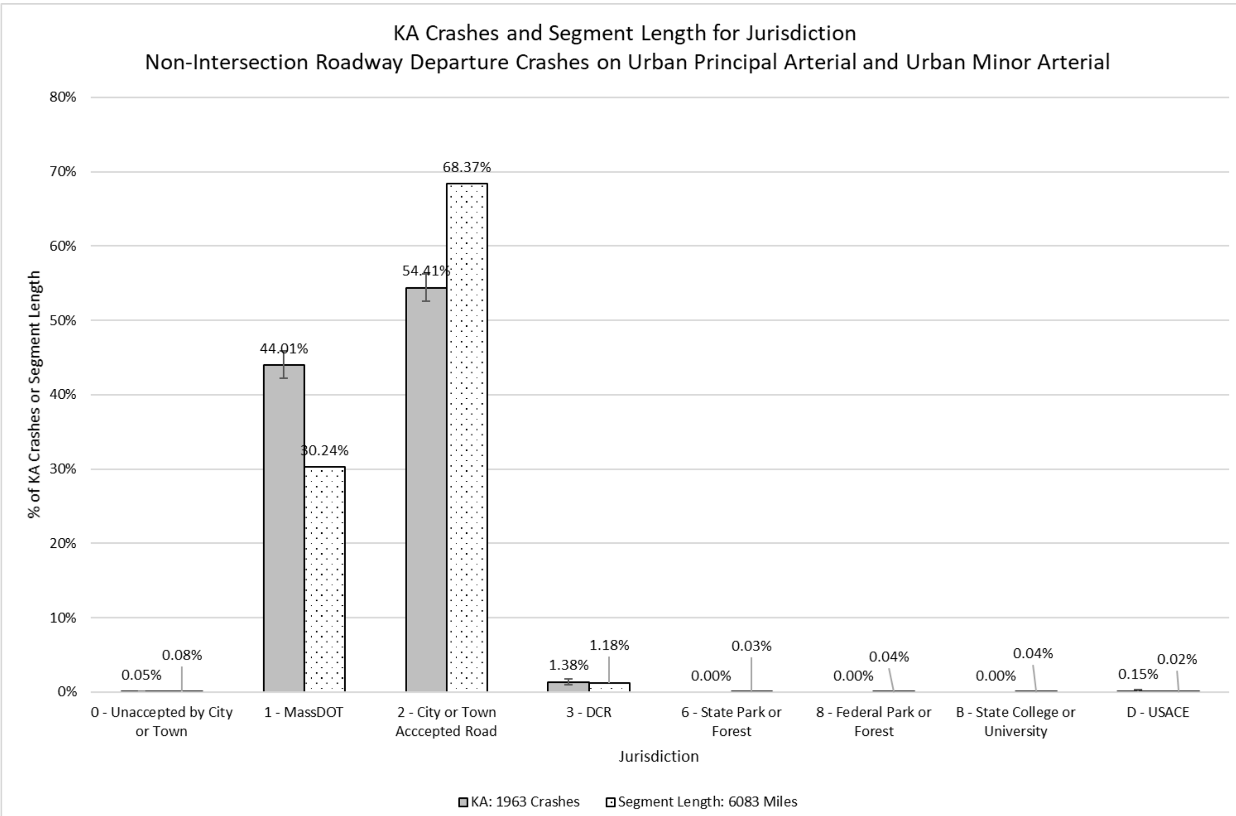

Figure C1. Risk factor analysis plot for Jurisdiction on urban principal arterial and minor arterial roads.

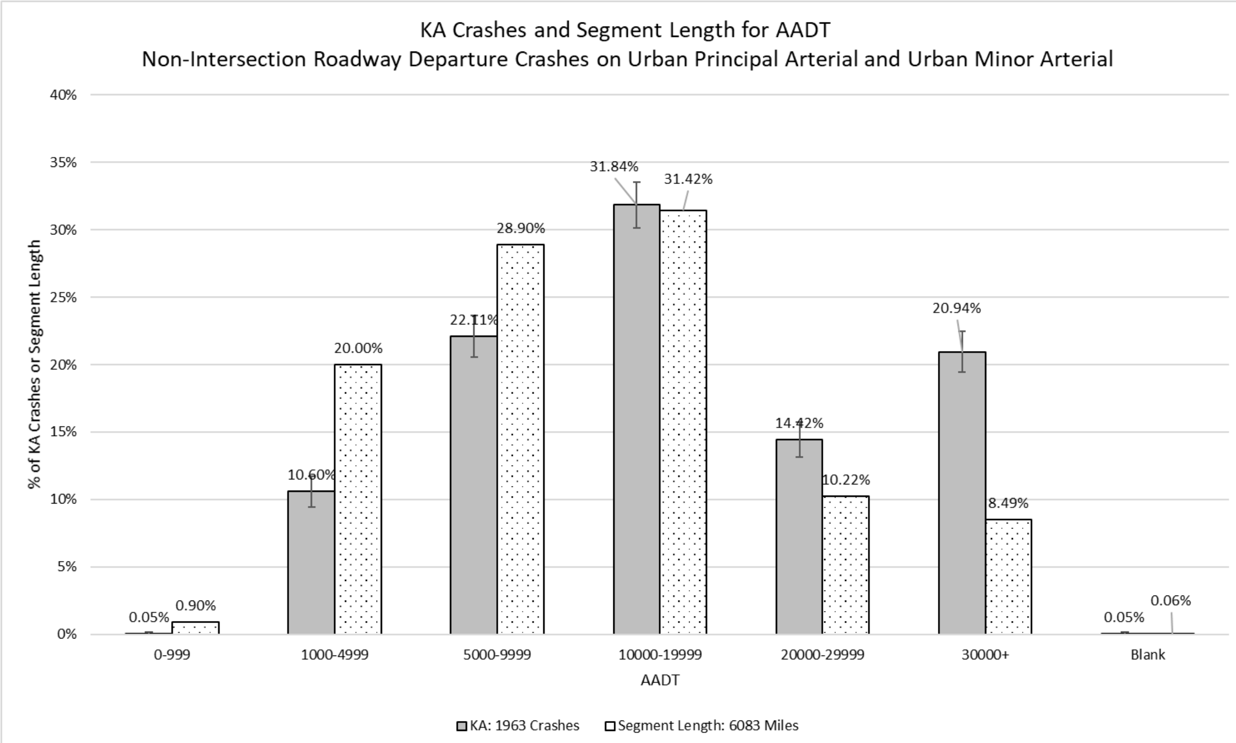

Figure C2. Risk factor analysis plot for AADT on urban principal arterial and minor arterial roads.

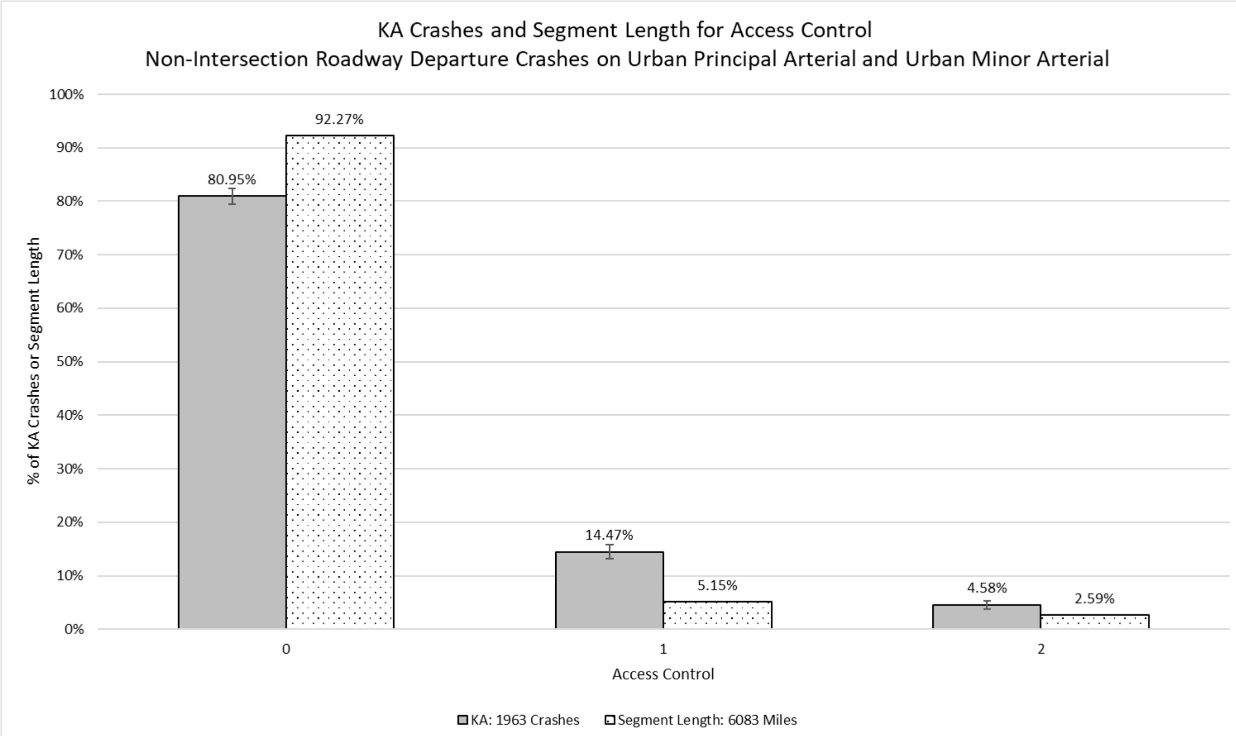

Figure C3. Risk factor analysis plot for access control on urban principal arterial and minor arterial roads.

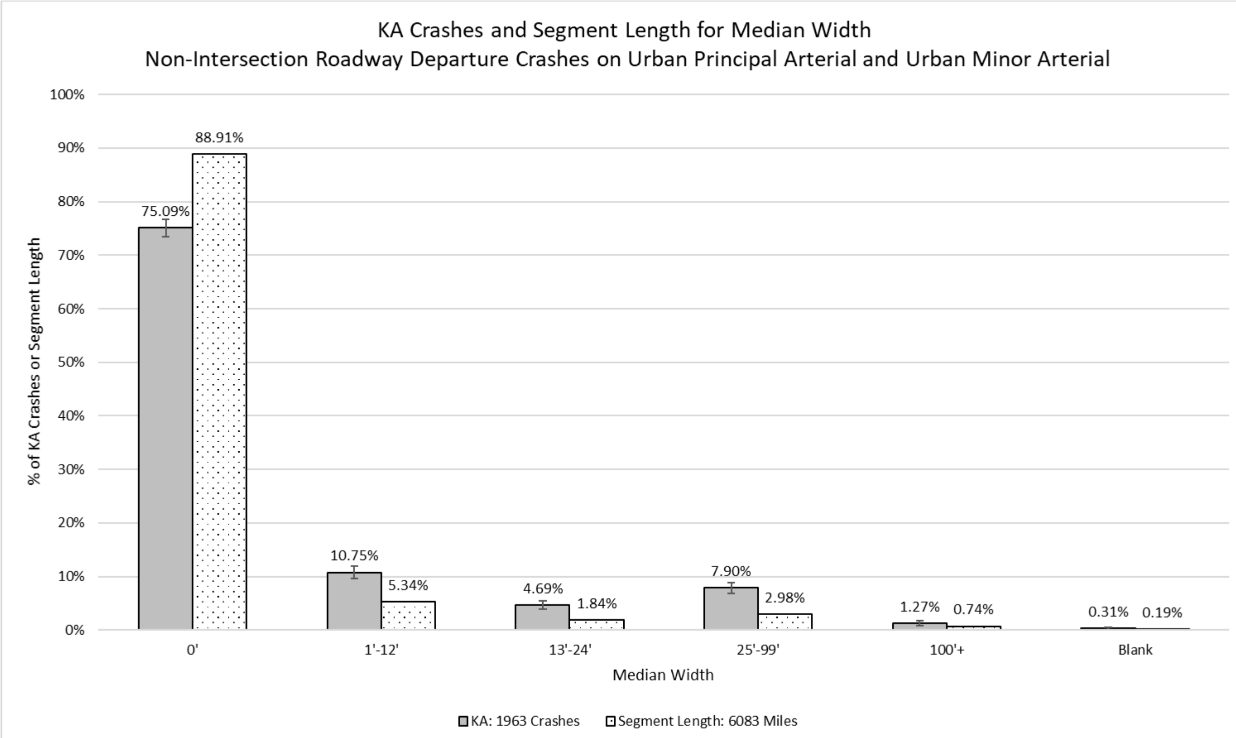

Figure C4. Risk factor analysis plot for median width on urban principal arterial and minor arterial roads.

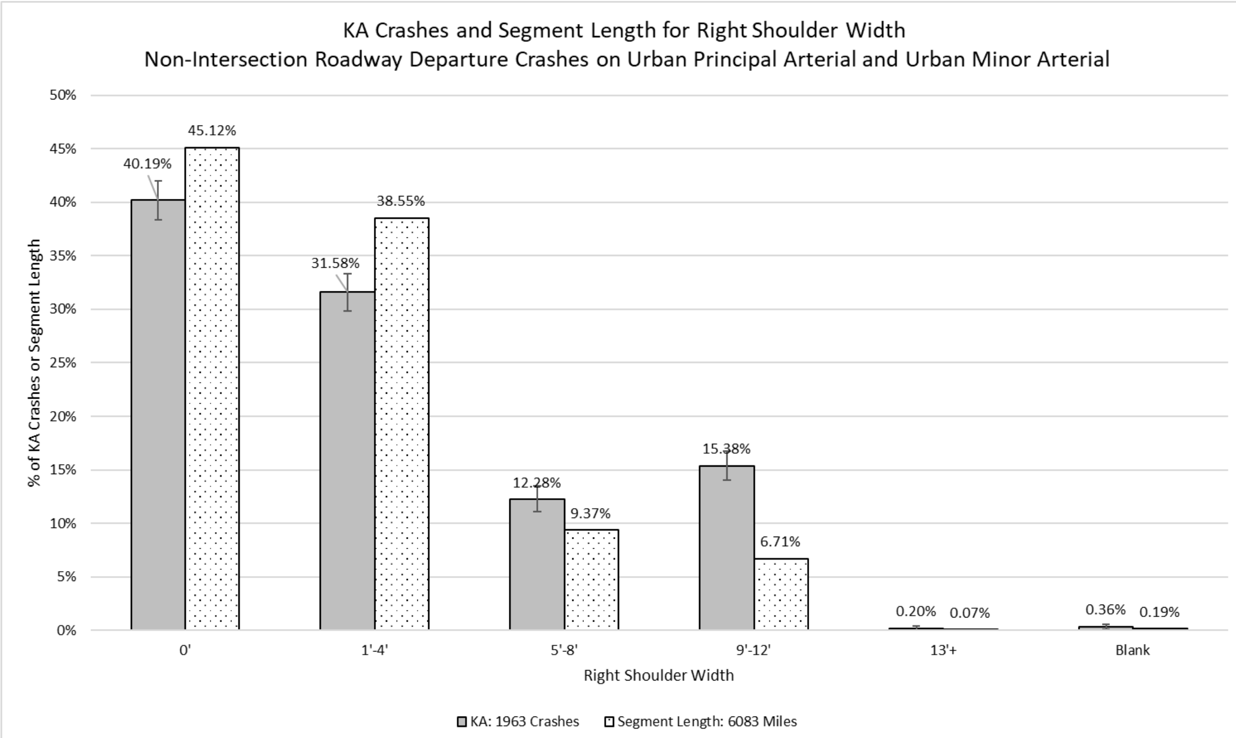

Figure C5. Risk factor analysis plot for right shoulder width on urban principal arterial and minor arterial roads.

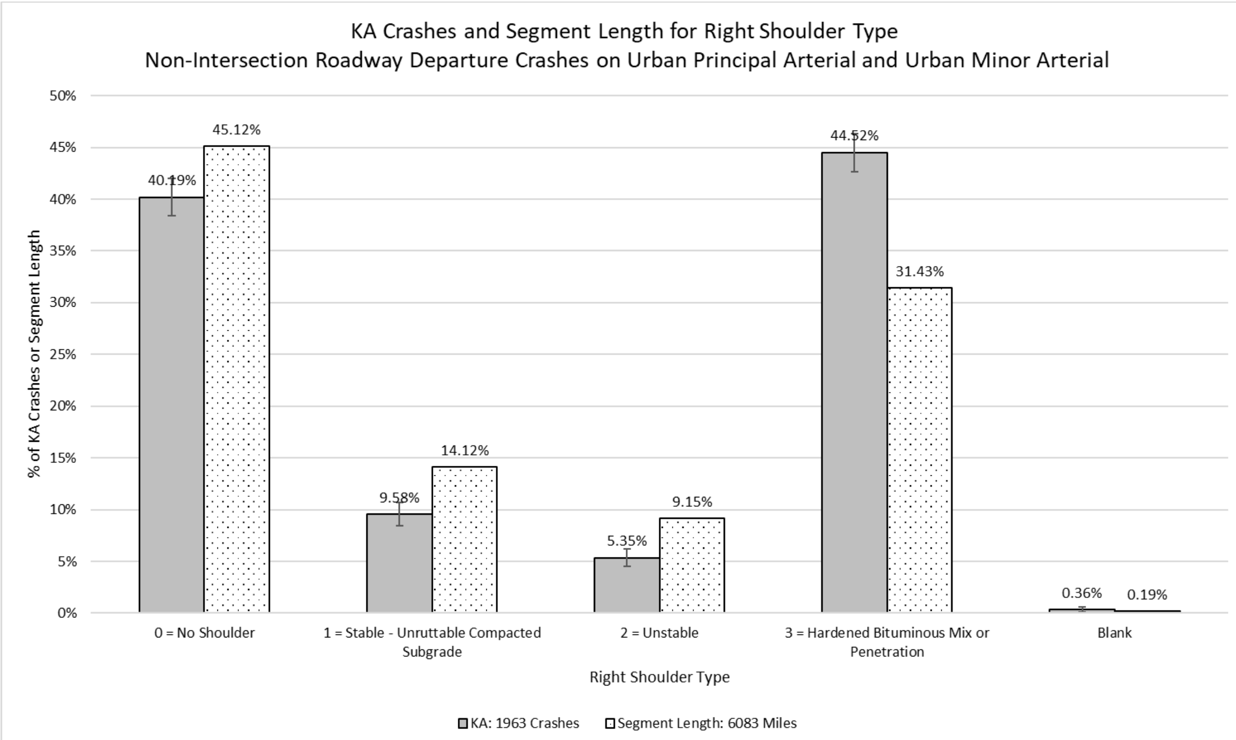

Figure C6. Risk factor analysis plot for right shoulder type on urban principal arterial and minor arterial roads.

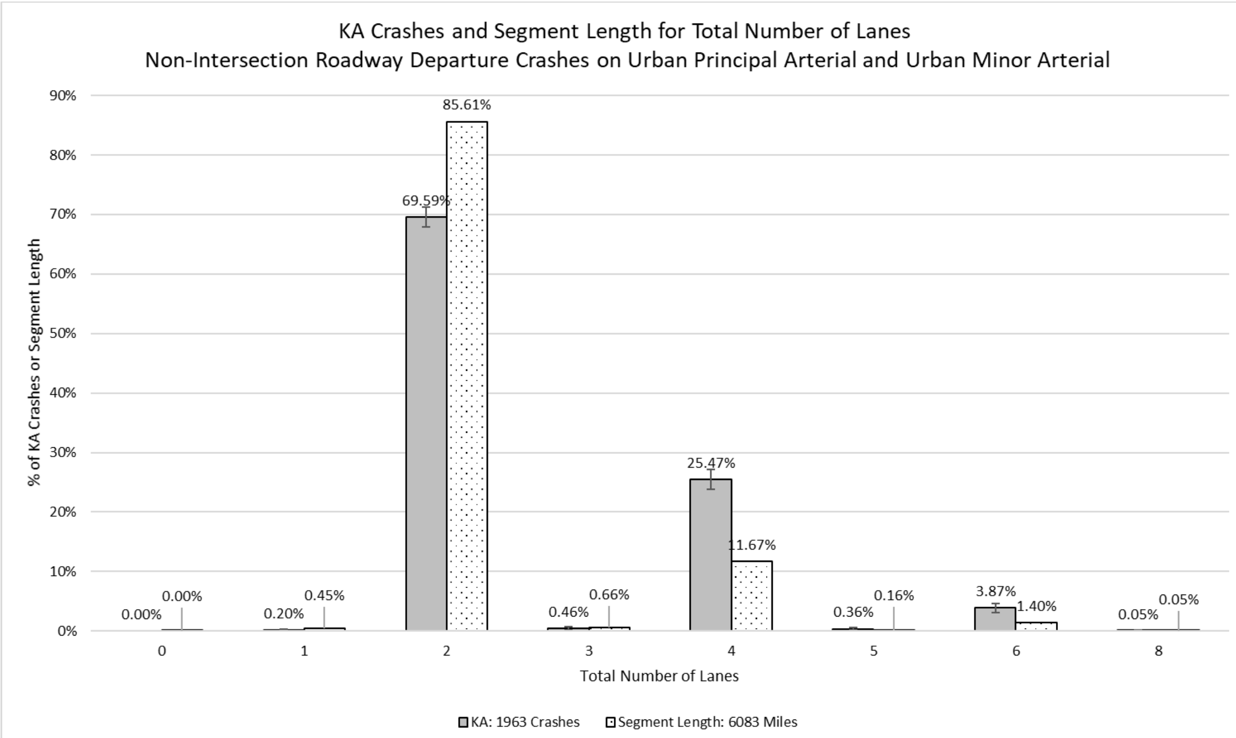

Figure C7. Risk factor analysis plot for total number of lanes on urban principal arterial and minor arterial roads.

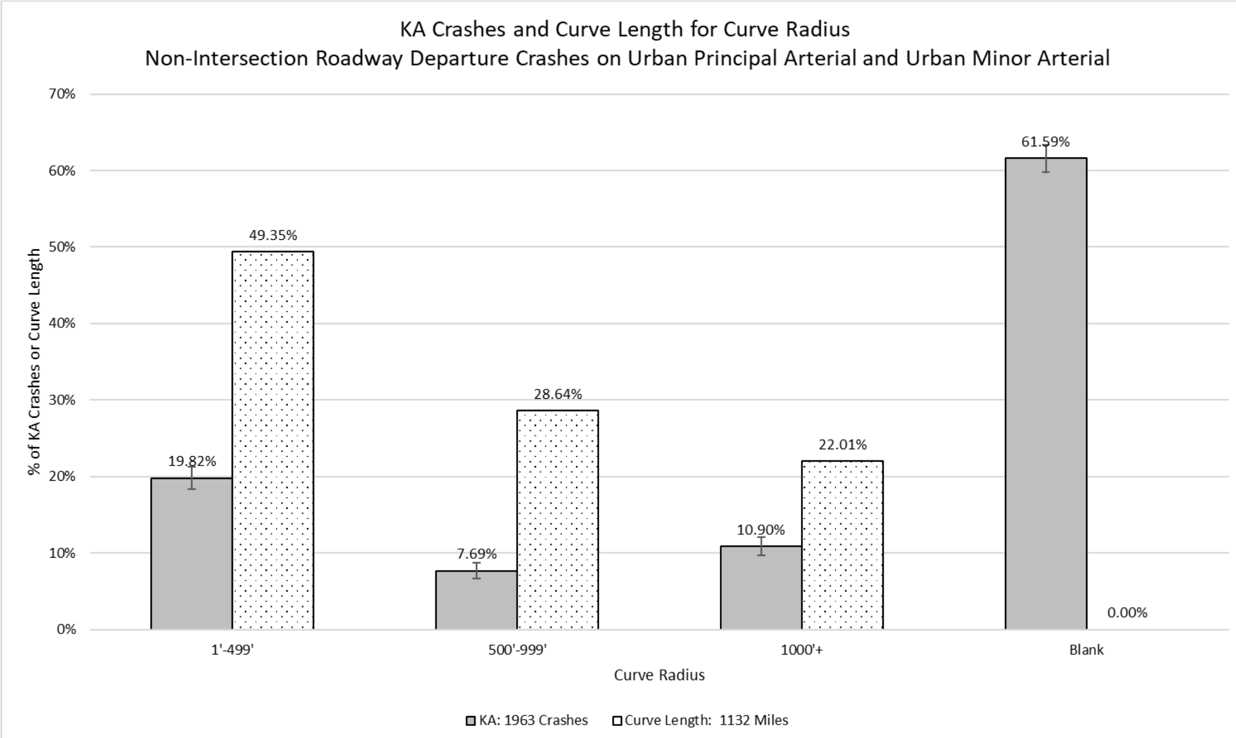

Figure C8. Risk factor analysis plot for curve radius on urban principal arterial and minor arterial roads.

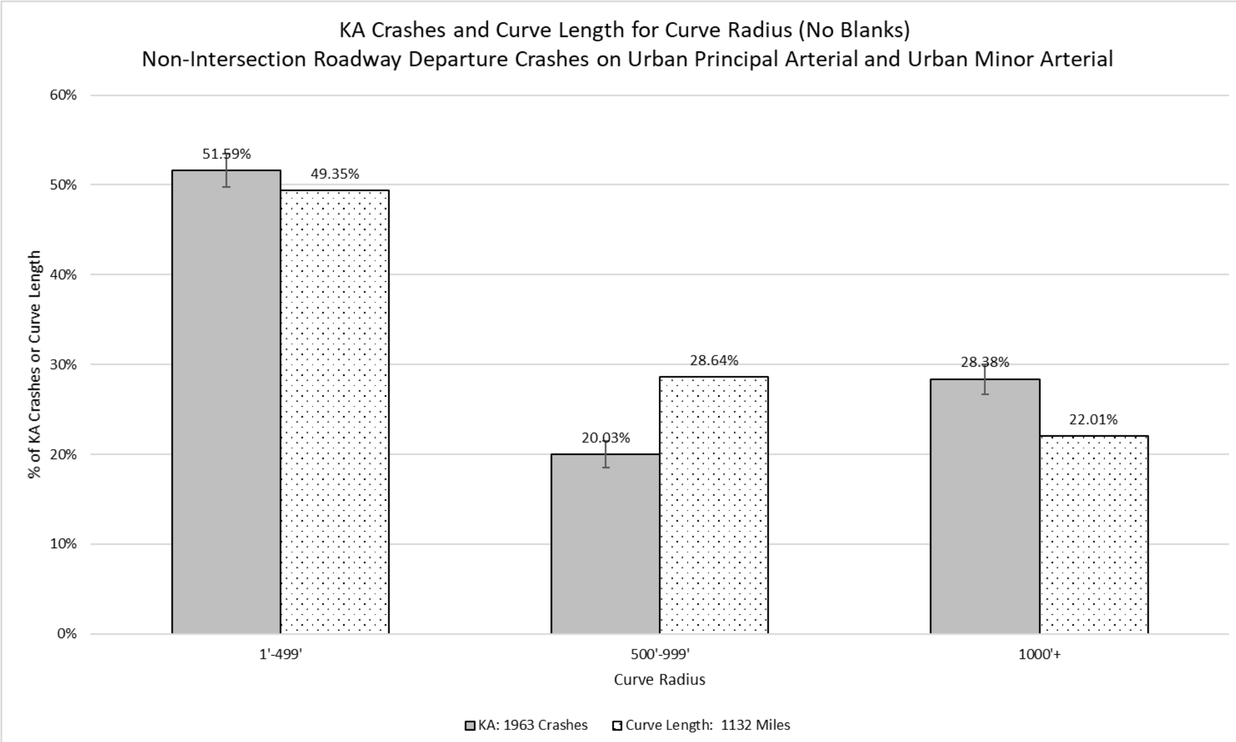

Figure C9. Risk factor analysis plot for curve radius (no blanks) on urban principal arterial and minor arterial roads.

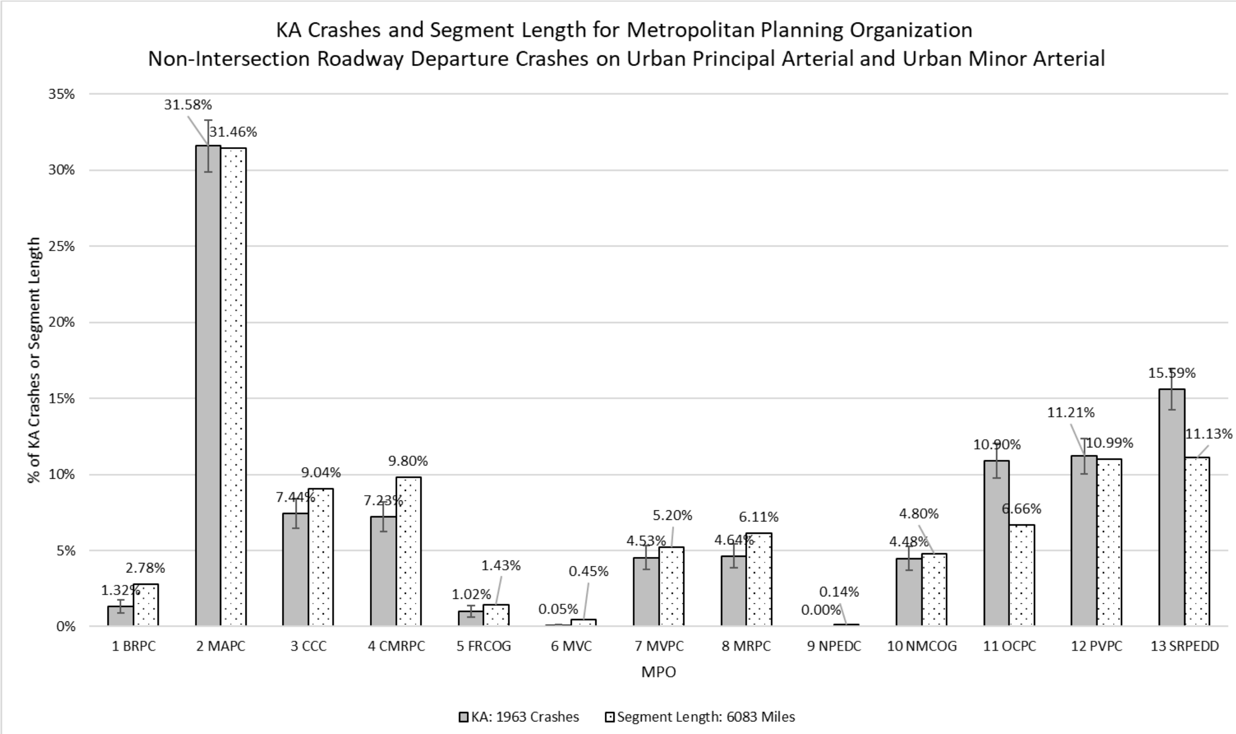

Figure C10. Risk factor analysis plot for metropolitan planning organization on urban principal arterial and minor arterial roads.

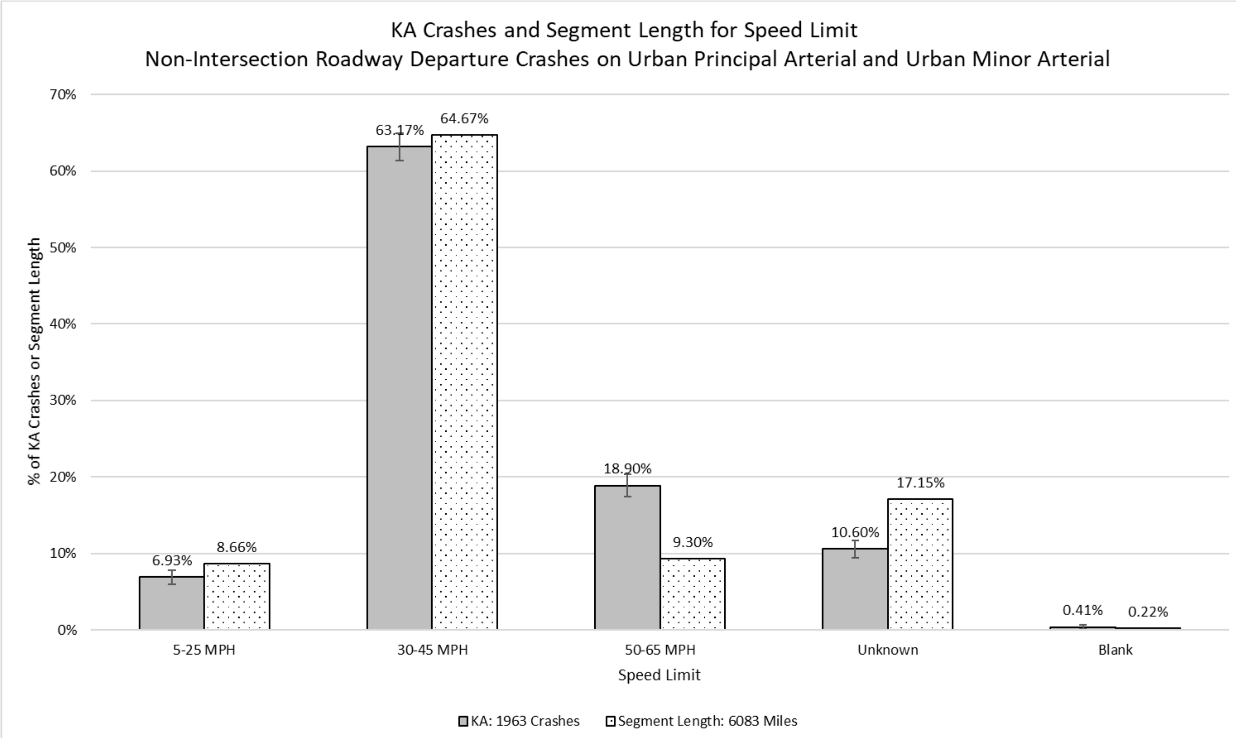

Figure C11. Risk factor analysis plot for posted speed limit on urban principal arterial and minor arterial roads.