Addressing Data Gaps in Pedestrian Safety at Bus Stops: A Review of Datasets and Case Study of Minnesota

Allison Rewalt*
Department of Civil and Environmental Engineering
University of Tennessee, Knoxville, TN, 37916
Email: arewalt@vols.utk.edu
ORCID: https://orcid.org/0009-0004-5645-4623

Amelia Thomson
Department of Civil and Environmental Engineering
University of Tennessee, Knoxville, TN, 37916
Email: athomso6@vols.utk.edu
ORCID: https://orcid.org/0009-0008-9433-9393

Candace Brakewood
Department of Civil and Environmental Engineering
University of Tennessee, Knoxville, TN, 37916
Email: cbrakewo@utk.edu
ORCID: https://orcid.org/0000-0003-2769-7808

Word Count: 6476 words + 4 tables (250 words per table) = 7476 words

Submitted [November 19, 2025]
*Corresponding Author

ABSTRACT

Transit passengers are a vulnerable group of pedestrians, exposed to vehicular traffic as they walk to/from or wait at bus stops. However, research on pedestrian safety at bus stops in the United States remains limited, raising concerns about whether crash data adequately captures details pertaining to transit bus stop-related pedestrian crashes. This study reviewed publicly available datasets to assess their utility in analyzing pedestrian safety at bus stops. The review identified sixteen datasets containing transit bus stop-related criteria beyond the proximity of the crash to a bus stop. From this review, Minnesota's crash database was selected for a case study due to its unique variable identifying 38 crashes involving pedestrians "*going to or from public transit*". Most crashes in the Minnesota dataset resulted in non-incapacitating injuries (39%), while 29% involved serious injuries and 3% were fatal. Crashes frequently occurred on higher-speed roads (61%) and at intersections (55%), with "dart/dash" being the most commonly documented pedestrian behavior (18%). Then, crash narratives from the Minnesota dataset were analyzed to manually categorize crashes into six groups: accessing/egressing from transit, bus as contact vehicle, LRT-related, pedestrian waiting at bus stop, bus driver witness, and no mention of transit. These categories demonstrate how unstructured text data provides context not usually captured by structured crash data fields. Overall, this study highlights the limited availability of detailed transit bus stop-related information in crash data. Improving the collection of transit bus stop-related crash reporting can enhance future research and support strategies to improve pedestrian safety at bus stops.

Keywords: Crash Data, Bus Stops, Pedestrian Safety

INTRODUCTION

Pedestrian safety remains a critical concern in the United States, with pedestrian fatalities trending upwards in recent years [1]. One particularly vulnerable group is transit passengers. Many transit users must walk to and from bus stops and often wait at bus stops on the roadside [2-4], exposing them to vehicular traffic during key stages of their transit journeys. As such, improving pedestrian safety for transit passengers is an integral part of promoting safer public transportation systems and reducing overall roadway fatalities in the United States.

Crash data has been instrumental in addressing safety issues and assisting decision-makers implement effective countermeasures [5]. However, research utilizing crash data to analyze the safety of transit passengers during the pedestrian stages of their trips remains limited in the United States, raising concerns about the quality of crash data in this context and whether it adequately captures details about pedestrians walking to or from, or waiting at, bus stops. This paper addresses these gaps by asking: What types of data are currently available in the United States to analyze pedestrian safety at bus stops? By reviewing data already used to study pedestrian safety at bus stops and identifying existing datasets that include transit bus stop-related information, this study aims to enhance the tools available for understanding pedestrian safety at bus stops.

This paper is organized as follows. First, a brief literature review examined research on the relationship between pedestrian safety and bus stops in the United States, with a specific focus on the data sources utilized in these studies. Next, a review of data sources is presented to summarize datasets containing transit bus stop-related details that can be leveraged to understand pedestrian safety at bus stops. Last, a Minnesota database identified during the review of datasets was selected for a case study analyzing crash narratives to better understand the nature of these crashes and how they are reported.

LITERATURE REVIEW

A review of research from the last 20 years on the relationship between pedestrian safety and bus stops in the United States was conducted, primarily focused on the data sources used to analyze this connection [3, 4, 6-16]. Pedestrian crash data has been the most commonly utilized measure of pedestrian safety at bus stops, with 10 prior studies relying on this source [3, 4, 6, 8, 9, 11-13, 15, 16]. In many cases, pedestrian crash data was supplemented with additional datasets such as ridership levels, street network data, roadway inventory information, and Annual Average Daily Traffic (AADT). However, these studies predominantly used a common methodology that measured the distance between a crash and the nearest bus stop as an indicator of relevance. This approach introduces limitations, as it assumes a direct relationship between crash location and bus stop proximity. For instance, crashes occurring near bus stops could be coincidental rather than directly related to activity at the stop, which may introduce bias into the results.

A few studies used alternative data sources to analyze the relationship between pedestrian crashes and bus stops. These included direct observations of pedestrian and motorist behavior through video footage [14] or field visits [7, 9, 10, 15], and unstructured text data such as police officer-reported crash narratives [9, 15]. These methods offer the advantage of more accurate identification of transit bus stop-related crashes but are often more time-consuming and less readily available than pedestrian crash data. Despite the potential, these data sources are underutilized in pedestrian safety research at bus stops.

Only one prior study explicitly identified transit bus stop-related pedestrian crashes using criteria that went beyond simply proximity to a bus stop. This nationwide analysis of fatal crashes in the Fatality Analysis Reporting System (FARS) compared transit bus stop-related pedestrian crashes to other nearby pedestrian crashes, isolating risks specific to pedestrians at transit bus stops [17]. This is the most relevant reference to the current study, as it also utilizes crash data with parameters for bus stops rather than relying solely on proximity.

In summary, most prior research in the United States has not utilized datasets with specific criteria categorizing a crash as directly related to a bus stop. This approach introduces some uncertainty about whether the findings reflect the direct impact of bus stops on pedestrian safety or are influenced by

unrelated crashes occurring nearby. The remainder of this study is therefore guided by the question: What transit bus stop-related safety data sources are available to support future research on pedestrian safety at bus stops?

Given that most prior literature has not used data explicitly indicating that a bus stop contributed to a pedestrian crash, this suggests a lack of transit bus stop-related information in reported data. To address this, this study first conducted a review of databases to identify those with fields specifying whether a crash occurred at or was related to a bus stop. Based on the review, Minnesota's database was selected for a case study analyzing "*going to or from public transit*" pedestrian crashes. This approach helps address data gaps and improves understanding of safety risks associated with accessing, egressing, and waiting at bus stops.

REVIEW OF DATASETS

The following section includes a review of transit bus stop-related safety datasets. First, crash typing frameworks were analyzed to determine whether they had dedicated categories for crashes occurring at or related to bus stops. Next, publicly available national, state, and local crash databases were evaluated for incorporation of these frameworks or other transit bus stop-related information beyond standard crash typing frameworks. For each state, crash databases, data dictionaries, and dashboards were assessed for relevant data fields, and sample sizes were recorded when available. Additional safety datasets not primarily focused on crash data were also reviewed for their relevance to pedestrian safety at bus stops.

Transit Bus Stop-Related Crash Typing Frameworks

Crash data plays an integral role in traffic safety, offering insights into crash trends and informing infrastructure and policy decisions. Reporting typically follows standardized formats to ensure consistency across agencies. Two commonly used frameworks containing transit bus stop-related criteria are the Pedestrian and Bicyclist Crash Analysis Tool (PBCAT) and the Model Minimum Uniform Crash Criteria (MMUCC).

Pedestrian and Bicyclist Crash Analysis Tool (PBCAT)

Introduced in 1999, PBCAT supports agencies in categorizing vulnerable road user crashes. A "*bus-related*" pedestrian crash group (#340), defined as crashes where pedestrians are struck while crossing to or waiting at a bus stop, was introduced in 2006 (PBCAT 2.0) [18]. Although the term "*bus stop*" is not explicitly used in the crash type name, this category is exclusive to crashes occurring at or related to bus stops. This ensures that the crash type avoids general bus-involved pedestrian crashes, which could misrepresent safety conditions at bus stops.

This group is further divided into "*commercial bus-related*" (#341) and "*school bus-related*" (#342) categories [18]. Distinguishing between commercial and school bus stops allows countermeasures tailored to the unique characteristics of each bus stop. While PBCAT does not specify which types of bus services are classified as commercial bus-related, it is presumed that transit services are included within this scope. This may also include services like charter buses, shuttle services, or tourist buses, making it challenging to draw independent conclusions about transit bus stop-related crashes without the ability to isolate these crashes. Additionally, the definition of commercial bus-related crashes requires the presence of a bus at the time of the crash, thereby excluding incidents that may involve bus stops but do not depend on the bus being present.

The 2021 update (PBCAT 3.0) addressed these limitations by replacing "*commercial*" with "*transit*" to focus exclusively on transit bus stops [19]. It introduced a hierarchical structure for transit-related crashes, categorizing crashes into scenarios like "*obstructed view*", "*bus pullover*", and "*other*". This update allowed for an even more direct approach to countermeasures, such as improving visibility at

bus stops (obstructed view crash type) or installing pedestrian detection systems on buses (bus pullover crash type).

Model Minimum Uniform Crash Criteria (MMUCC)

In 2017, MMUCC (5th edition) adopted the crash factor "*related to a bus stop*", which includes various activities in which a bus stop contributed to a crash (e.g. vehicles maneuvering around a bus waiting at the bus stop, pedestrians walking to or from the bus stop, and pedestrians crossing lanes of traffic to access the bus stop) [20]. The latest version (6th edition, 2024) maintained this definition [21]. While the MMUCC framework is used by many state DOTs for crash reporting compared to PBCAT, the "*related to a bus stop*" data element does not differentiate between school and transit bus stop-related crashes, which is a significant limitation of this framework.

National, State and Local Databases

Many national, state, and local agencies incorporate these frameworks into their crash data, while some agencies report transit bus stop-related information outside the scope of these frameworks. The following sections discuss sixteen datasets identified during this review. Additional datasets may exist but were either not publicly available or not identified during the online search of publicly available crash datasets, data dictionaries, and crash dashboards, which was conducted in 2024.

National Databases

The review identified three national databases reporting transit bus stop-related information: the Fatality Analysis Reporting System (FARS), the Crash Report Sampling System (CRSS), and the National Transit Database (NTD). Managed by federal agencies, these databases differ in scope, sample size, and focus but collectively contribute to understanding pedestrian safety at bus stops (see Table 1).

FARS, maintained by NHTSA since 1975, documents police-reported fatal motor vehicle crashes in the United States [22]. To be reported in FARS, a crash must involve a motor vehicle on a trafficway and result in at least one fatality within 30 days. Data is collected from multiple state sources, including police crash reports, crash report supplements, vehicle registration files, driver records, roadway classification files, death certificates, toxicology reports, and emergency medical service reports, and standardized into over 170 elements.

In 2014, FARS adopted a modified version of the PBCAT 2.0 framework for reporting transit bus stop-related crashes, redefining "*commercial bus-related*" crashes as "*transit bus stop-related*" [23]. This definition applies to all pedestrians involved in crashes related to transit bus stops, regardless of the intent to use the bus. Since its adoption in 2014, FARS has reported 216 pedestrians involved in fatal transit bus stop-related crashes. In 2020, FARS implemented the "*related to a bus stop*" crash factor from the MMUCC framework, reporting 163 pedestrians involved in fatal crashes under this criterion (2020-2023) [24].

CRSS, also maintained by NHTSA, is a nationally representative sample of motor vehicle crashes in the United States [8]. The database was introduced in 2016 as an expansion of its predecessor, the National Automotive Sampling System General Estimates System (NASS GES). CRSS samples police-reported motor vehicle crashes that occur in 60 predefined locations in the United States, which reflect the geography, population, travel behavior, and crash characteristics of the United States. The survey is designed to sample crashes of greater relative importance (e.g. fatal crashes and pedestrian-involved crashes). CRSS uses trained coders to translate crash reports into 120 data elements and assigns case weights that are representative of the annual estimate of crashes. Like FARS, CRSS incorporates transit bus stop-related data based on PBCAT 2.0 and MMUCC definitions. Since 2016, CRSS has documented 95 pedestrian-involved transit bus stop-related crashes based on PBCAT definitions through 2023.

Additionally, since 2020, CRSS has recorded 208 crashes categorized as related to a bus stop under MMUCC definitions through 2023.

Unlike FARS and CRSS, NTD focuses on Safety and Security incidents in public transportation systems rather than traditional crash data [25]. The National Transit Database (NTD) Safety and Security database, managed by the Federal Transit Administration (FTA), collects agency-reported data on fatalities, injuries, property damage, and other safety events that occur within transit systems, such as those involving transit vehicles, transit facilities, or within transit right-of-way.

Since 2014, the NTD Safety and Security database has documented 5,270 pedestrian-involved crashes that met the reporting threshold for a major safety event through 2025 [25]. While it lacks detailed infrastructure, crash, and person-level data fields found in FARS or CRSS, NTD does specify whether the pedestrian was in a crosswalk at the time of the incident and whether the pedestrian was a transit passenger. NTD also includes a narrative for each event, offering additional insights that are not available in FARS or CRSS.

While few national crash databases report transit bus stop-related information, FARS, CRSS, and NTD are valuable resources to analyze pedestrian safety at bus stops due to their detailed information, larger sample sizes (compared to individual state and local crash databases), and ability to capture nationwide trends, as summarized in Table 1. However, the small sample of transit bus stop-related crashes in FARS and CRSS limits its utility for detailed analysis. While NTD offers a larger sample of pedestrian-involved crashes, it lacks specific fields for transit bus stop-related crashes.

Table 1 Summary of national transit bus stop-related safety databases

Database	Agency	Years	Description of Data	Transit Stop-Related Definition	Sample Size
Fatality Analysis Reporting System (FARS)	NHTSA	2014-2023	Population of fatal police-reported crashes	· **Transit bus stop-related** (2014): adapted version of PBCAT 2.0: transit bus must be present at the bus stop at the time of the crash · **Related to a bus stop** (2020): definition consistent with MMUCC	n = 216 (pedestrians involved)
Crash Report Sampling System (CRSS)	NHTSA	2016-2023	Nationally representative sample of police-reported crashes	· **Transit bus stop-related** (2014): adapted version of PBCAT 2.0: transit bus must be present at the bus stop at the time of the crash · **Related to a bus stop** (2020): definition consistent with MMUCC	n = 95 (pedestrians involved)
National Transit Database (NTD) – Safety and Security	FTA	2014-2024	Collection of annual financial, asset, and operating data from public transit agencies nationwide	· Meets reporting threshold (per Safety and Security manual) OR · Occurs at a transit revenue facility, maintenance facility, or rail yard · Happens on transit right-of-way infrastructure · Occurs during a transit-related maintenance activity	n = 5097 (pedestrians involved)

Note:
2023 is the most recent year of data available for FARS and CRSS; 2024 is the most recent year for NTD

State Databases

The review of state databases identified six publicly available online crash databases that use the PBCAT or MMUCC framework to report transit bus stop-related information. North Carolina and Florida use the PBCAT 2.0 framework ("*commercial bus-related*"), while Illinois, Louisiana, Nebraska, and Utah

use the MMUCC framework ("*related to a bus stop*") [26-31]. While several states use these frameworks, sample sizes were small or the crash data was private and only available upon request.

At least three state agencies report transit bus stop-related safety information using classifications beyond the PBCAT and MMUCC frameworks. Idaho crash data, managed by the Idaho Transportation Department (ITD), includes a field for damaged infrastructure, with "*bus shelter*" as an option [32]. Georgia crash data, managed by the Georgia Department of Transportation (GDOT), reports crashes involving pedestrians "*entering/exiting a bus*" and specifies the transit agency involved, when applicable [33]. Managed by the Minnesota Department of Transportation (MnDOT), Minnesota crash data reports pedestrians who were "*going to or from public transit*" prior to the crash [34].

Although several state datasets were identified (see Table 2), it is presumed that many more states use the "*related to a bus stop*" data element under the MMUCC framework, as MMUCC is adopted by many states. However, this was not always evident because many states do not publish their crash data or data dictionaries online, and this review relied only on publicly available online information. Additionally, some states vary in their implementation, with some states using earlier versions of MMUCC. Similarly, while PBCAT 2.0 is commonly implemented, no states were identified as openly using PBCAT 3.0 for detailed transit stop-related crash reporting.

Table 2 Summary of statewide transit bus stop-related safety databases from publicly available information found online

Database	Agency	Years	Description of Data	Transit Bus Stop-Related Definition	Sample Size
Florida Traffic Safety Dashboard (Signal4Analytics)	UF Geoplan Center	2014-2025	Statewide crash dashboard for police-reported crashes	**Commercial bus-related**: definition consistent with PBCAT 2.0	Unavailable (information not public)
Georgia Crash Dashboard	GDOT	2013-2023	Statewide crash dashboard for police-reported crashes	· **Entering/exiting bus:** no details provided; data dictionary not public Transit Agency Involvement	n = 681 (pedestrians involved)
Idaho Crash Dashboard	ITD	2008-2025	Statewide crash dashboard for police-reported crashes	· **Item damaged = bus shelter:** indicates whether property was damaged as a result of the crash	Unavailable (information not public)
Illinois Crash Data	HSIS	2019-2022	Multistate database that contains crash, roadway, and traffic characteristics for a select group of states	· **Related to a bus stop:** definition consistent with MMUCC	Available upon request
Louisiana Crash Data	CARTS	Unknown	Raw statewide police-reported crash data	· **Related to a bus stop:** definition consistent with MMUCC · **Going to or from public transit:** prior action of the non-motorist	Unavailable (information not public)
Minnesota Crash Data	MnDOT	2016-2023	Raw statewide police-reported crash data; publicly available upon request	· **Going to or from public transit:** prior action of the non-motorist	n = 39 (crashes)
Nebraska Crash Data	NDOT	Unknown	Raw statewide police-reported crash data	· **Related to a bus stop:** definition consistent with MMUCC · **Going to or from public transit:** prior action of the non-motorist	Unavailable (information not public)
North Carolina Non-Motorist Crash Map	NCDOT	2007-2022	Map of police-reported pedestrian and bicyclist crashes in North Carolina	· **Commercial bus-related**: definition consistent with PBCAT 2.0	n = 126 (crashes)
Utah Pedestrian Crash Dashboard	UHSO	2010-2025	Statewide crash dashboard for police-reported crashes	· **Related to a bus stop:** definition consistent with MMUCC	Unavailable (information not public)

Local Databases

The review identified at least four local databases reporting transit bus stop-related details outside the PBCAT and MMUCC frameworks. These databases are not strictly crash databases but instead broadly address transit bus stop-related safety concerns, including crashes and other incidents. The Highway Safety Information System (HSIS) provides detailed bus route and stop amenity information for Charlotte, North Carolina, which is supplemented by crash data for the city [31]. Cleveland, Ohio piloted the Enhanced Transit Safety Retrofit Package (E-TRP) to assess pedestrian detection warning systems on buses [35]. The Los Angeles Police Department's Traffic Collision dataset specifies incident locations, including "*bus stop*" as an option [36]. The Washington Metropolitan Area Transit Authority (WMATA) implemented a "close call" reporting system for transit employees to confidentially report near-misses [37]. This allows WMATA to proactively address safety concerns, with the agency reporting 300 preventable actions since the adoption of the program.

While these databases provide insights on a local scale, none use PBCAT or MMUCC (see Table 3). However, many of these relevant databases can be linked with local crash data to help address pedestrian safety at bus stops, particularly in Charlotte, Cleveland, and Washington, D.C.

Table 3 Summary of local transit bus stop-related safety databases

Database	Agency	Years	Description of Data	Transit Bus Stop-Related Definition	Sample Size
Charlotte, NC	HSIS	2018-2022	Multistate database that contains crash, roadway, and traffic characteristics for a select group of States	· Detailed information regarding bus route and bus stop amenities	Available upon request
Enhanced Transit Safety Retrofit Package – Cleveland, OH	FTA	2018	Data from 24 transit buses 3 locations over a 6 month test period	· Pedestrian detection warning systems at crosswalks (Pedestrian in Crossing Warning – PCW)	n = 15,609 (PCW alerts)
City of Los Angeles, CA Traffic Collision Data	LAPD	2010-2025	Traffic collision incidents in the City of Los Angeles; transcribed from original traffic reports	· Crashes occurring at a bus stop: reported in the "premise codes" element: the type of structure or location where the incident took place	n = 7 (crashes)
Confidential Close Call Reporting System	WMATA	2013-2025	Voluntary reporting of "Close Calls" or near miss safety and security events by WMATA employees	· Any event observed by a rail transit or bus system employee that was potentially dangerous but did not result in any serious harm	n = 300 (close call reports)

Selection of Data

The review of databases for transit bus stop-related information revealed that data to understand the relationship between pedestrian safety and bus stops is limited in crash databases and other relevant datasets. National databases like FARS and CRSS provide valuable insights through their broad coverage, standardized reporting formats, and ability to identify nationwide trends, but they are relatively limited in sample size. Several state and local agencies maintain public crash databases with transit bus stop-related information, supporting a more localized approach. However, these generally have very small sample sizes or include information outside the scope of PBCAT and MMUCC. Notably, only one prior study has used a dataset from this review to analyze pedestrian safety at bus stops, focusing on a national scale using FARS [17]. To build on this, a case study analyzed Minnesota crashes labeled "*going to or from public transit*" to better understand crashes involving pedestrians walking to, from, or waiting at bus stops.

CASE STUDY: "GOING TO OR FROM PUBLIC TRANSIT" PEDESTRIAN CRASHES IN MINNESOTA

MnDOT introduced "*going to or from public transit*" into their crash data in 2016, which was selected for this local analysis due to the agency's willingness to share the data and the availability of crash narratives. This information is reported at the person level in the dataset, under the NMACTION variable, which categorizes the pedestrian's actions or circumstances immediately preceding the crash. While it is not part of the PBCAT or MMUCC frameworks, this unique crash type has several advantages: it captures only transit passengers, is not restricted to the immediate vicinity of the bus stop like FARS and CRSS and reflects the full trip to or from a bus stop or transit station. Additionally, it records the full range of injury severities, which can support analysis of crash outcomes.

First, descriptive statistics were generated to examine the distribution of key variables in this unique dataset. Next, because the sample was relatively small, it was feasible to review each crash narrative and categorize crashes based on the transit bus stop-related terminology used by the reporting officers, as well as identify crash details not captured in structured fields. Thus, the crash narrative analysis sought to assess whether any patterns emerged in the narratives that may suggest subtypes of transit bus stop-related pedestrian crashes, such as walking to or from the stop, crossing the street, or waiting at the stop.

Descriptive Analysis

From 2016 to 2023, there were a total of 38 pedestrian crashes reported as "*going to or from public transit*" in the raw Minnesota crash data. A descriptive analysis was conducted to examine the distribution of key variables, including injury severity, vehicle speed, relation to intersection, and contributing circumstances of both the motorists and the pedestrians. Table 4 presents the results, followed by a discussion of key findings.

Using the KABCO scale, where K represents a fatality, A an incapacitating injury, B a non-incapacitating injury, C a possible injury, and O no apparent injury, most crashes were non-incapacitating (39%). Notably, 29% of crashes were incapacitating, and one crash (3%) was fatal (see Table 4). The range of severity outcomes offers insight into the potential factors that influence injury severity in crashes involving pedestrians going to or from transit, supporting further analysis of the risks experienced by transit passengers.

The majority of "*going to or from public transit*" crashes occurred on higher-speed roads, with 61% taking place on roads with speed limits over 30 mph (see Table 4), consistent with findings from prior studies [3, 8, 16]. Most crashes (55%) occurred at intersections, particularly at four-way intersections (50%). While a previous study using FARS data found that fatal transit bus stop-related pedestrian crashes more commonly occurred at midblock locations, the Minnesota dataset shows more crashes happening at intersections [17]. Unlike FARS, which only includes fatal crashes, the Minnesota data includes all levels of injury severity. Additionally, the crashes in the Minnesota data encompass the entire walking trip to or from transit, rather than just crashes near or at the stops themselves, as is typically the case in FARS. This difference suggests that intersection-related crashes may occur more frequently but tend to result in less severe injuries, but further analysis is needed to validate this.

Forty-three percent of cases reported no contributing circumstance for the motorist (see Table 4). Among those that did, no single factor was stood out. Contributing circumstances varied and included failure to yield, disregarding traffic signals, failure to maintain lane, carelessness, and speeding. For pedestrians, 41% of crashes listed no clear contributing factor. However, the most common cited behavior was darting or dashing into traffic (18%), which may reflect the urgency of attempting to catch a bus (see Table 4). This risky behavior aligns with findings from prior studies, where many pedestrians were struck

while crossing the roadway [9, 14, 15, 17]. Other circumstances for pedestrians included failure to yield, disregarding signals, entering or exiting parked vehicles, and poor visibility conditions. For both motorists and pedestrians, some crashes reported more than one contributing circumstance.

Table 4. Descriptive analysis of "*going to or from public transit*" pedestrian crashes in MN

Variable		Frequency	Percentage
Crash severity	K- Fatality	1	3%
	A- Incapacitating	11	29%
	B- Non-Incapacitating	15	39%
	C- Possible Injury	10	26%
	O- No apparent injury	1	3%
Speed Limit	0-19 mph	2	5%
	20-29 mph	5	13%
	30-39 mph	22	58%
	40-49 mph	1	3%
	Not Reported	8	21%
Location - Relative to Intersection	Entrance/Exit Ramp	1	3%
	Four-Way Intersection	19	50%
	Intersection Related	4	11%
	Not at Intersection, Interchange or Junction	10	26%
	T Intersection	2	5%
	Other	1	3%
	Unknown	1	3%
Motorist Contributing Circumstances	Driver Speeding	1	2%
	Failed to Keep in Proper Lane	1	2%
	Failure to Yield Right-of-Way	1	2%
	Improper Turn/Merge	2	4%
	No Clear Contributing Action/No Improper Action	20	43%
	Operated Motor Vehicle in Careless, Negligent, or Erratic Manner	2	4%
	Ran Red Light	2	4%
	Other	1	2%
	Not Reported	11	24%
	Unknown	3	7%
Pedestrian Contributing Circumstances	Dart/Dash	8	18%
	Entering/Exiting Parked/Standing Vehicle	1	2%
	Failure to Obey Traffic Signs, Signals, or Officer	2	5%
	Failure to Yield Right-of-Way	5	11%
	In Roadway Improperly (Standing, Lying, Working, Playing)	1	2%
	No Clear Contributing Action/No Improper Action	18	41%
	Not Visible (Dark Clothing, No Lighting, etc.)	1	2%
	Other	2	5%
	Unknown	8	18%

Notes:
Percentages may not sum to 100% due to rounding.
Motorist and Pedestrian Contributing Circumstances variables allow multiple inputs per case.

Crash Narrative Analysis

The analysis of crash narratives identified six types of "*going to or from public transit*" pedestrian crashes (see Figure 1). Crash narratives were reviewed and classified into six categories reflecting the type of interaction with transit. The initial categorization was conducted by one analyst and then reviewed by two others for validation. Upon reviewing the narratives, it was determined that one of the crashes

involved a limousine or taxi rather than a transit bus and was unrelated to public transit. As a result, this crash was excluded from the discussion and removed from Figure 1. The following sections provide descriptive examples for each category, demonstrating the range of crashes associated with the pedestrian stages of a transit trip. When available, street-level imagery from KartaView and georeferenced crash maps from OpenStreetMap are included to visually represent the crash locations [38-40].

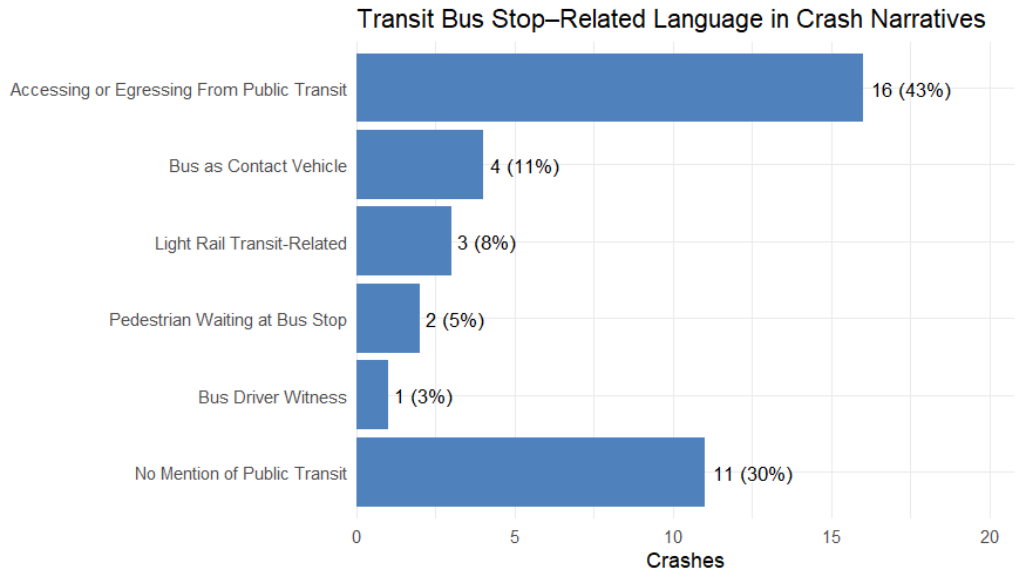

Figure 1 Transit bus stop-related language used in the crash narratives

Accessing or Egressing From Public Transit

Crashes were most commonly categorized as "*accessing or egressing from public transit*" due to a pedestrian who was hit while accessing or egressing from a bus stop, with 43% of crash narratives using this terminology (see Figure 1). This was the most common category, suggesting that most crashes occurred on the pedestrian's journey to or from a bus stop or station, rather than at the stop itself.

Bus as Contact Vehicle

There were 4 (11%) crashes where a bus was the contact vehicle. In one instance (see Figure 2), a transit bus traveling on a 30 mph roadway stopped at an intersection to pick up passengers. As it departed, a pedestrian attempted to flag it down, fell under the passenger side, and sustained minor injuries. Such scenarios often involve pedestrians engaging in risky behaviors while trying to board departing buses,

occasionally leading to crashes. This example highlights how urgency when approaching bus stops can negatively impact safety, a nuance not easily captured in structured crash data.

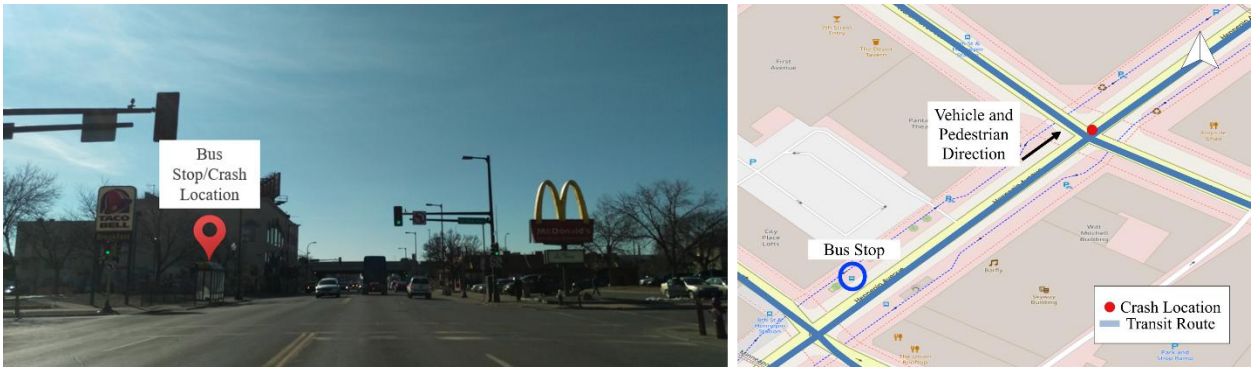

Figure 2. Example of pedestrian crash from the "bus as contact vehicle" category (Map created using QGIS 3.40 [39] with basemap data © OpenStreetMap (ODbL 1.0) [40] and street-level imagery obtained from © KartaView (CC BY-SA 4.0) [38])

Light Rail Transit-Related

Light Rail Transit (LRT)-related crashes were categorized separately from transit bus-related crashes due to differing infrastructure characteristics. Although less common (8%), this distinction allows for a more detailed examination of each type. For example, one crash involved an SUV traveling at 30 mph through a green light at the intersection, striking a child who ran into the road (see Figure 3). The narrative specified that the driver believed the pedestrians were attempting to catch the nearby LRT.

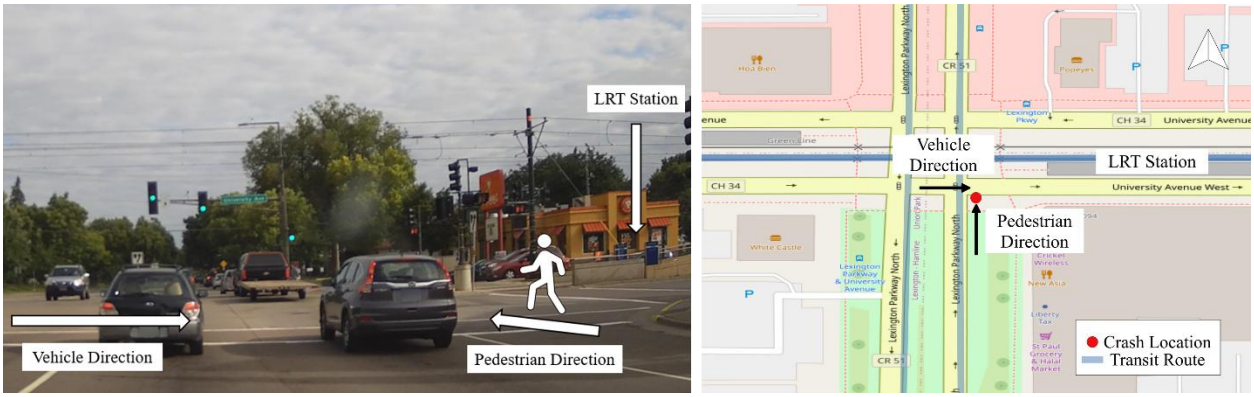

Figure 3 Example of pedestrian crash from the "LRT-related" category (Map created using QGIS 3.40 [39] with basemap data © OpenStreetMap (ODbL 1.0) [40] and street-level imagery obtained from © KartaView (CC BY-SA 4.0) [38])

Pedestrian Waiting at Bus Stop

Crash narratives with keywords indicating that a pedestrian was waiting at a bus stop were categorized separately from those suggesting the pedestrian was walking to or from public transit. Specifically, all crashes in this category had specified that a pedestrian was waiting at a bus shelter in the narrative. Although less common in this sample (5%), a prior study using FARS data also identified this as an important distinction [17].

Bus Driver Witness

There was one crash (3%) where a bus driver served as a witness, likely leading to its classification as "*going to or from public transit*". However, the case materials did not confirm whether the pedestrian was a transit passenger or if the bus was actively picking up or dropping off passengers. This highlights ambiguity in this crash type, as officer interpretations vary. Such nuances are often overlooked in structured data fields, which demonstrates the importance of analyzing crash narratives for more context.

No Mention of Public Transit

The crashes categorized as "*no mention of public transit*" were those where the crash narrative lacked transit bus stop-related details. This was the second most common category, comprising 11 crashes (30%). All crashes in the dataset were classified as "*going to or from public transit*" based solely on an officer checking the corresponding box on the crash report form, regardless of whether transit bus stop-related information appeared in the narrative. For example, one crash involved a vehicle striking a pedestrian in a crosswalk at an intersection (see Figure 4). The pedestrian initially reported no injuries but later contacted police about a minor injury. While the narrative lacked transit bus stop-related details, it is possible the pedestrian mentioned going to or from a bus stop or train station, and the officer omitted this detail, relying only on the checkbox. This lack of detail challenges narrative-based analyses, as it may exclude crashes without direct references to transit.

Figure 4. Example of pedestrian crash from the "no mention of public transit" category (Map created using QGIS 3.40 [39] with basemap data © OpenStreetMap (ODbL 1.0) [40] and street-level imagery obtained from © KartaView (CC BY-SA 4.0) [38])

CONCLUSIONS

This study sought to evaluate the availability and application of data in the U.S. for analyzing pedestrian safety at bus stops. First, a literature review was conducted to examine how pedestrian safety has been measured in relation to bus stops, with a focus on the data sources utilized. Next, publicly available online national, state, and local databases were reviewed to determine whether they included specific criteria for transit bus stop-related crashes. Last, a crash dataset from Minnesota with specific criteria for bus stops was selected for a case study, which features a unique crash type for pedestrians categorized as "*going to or from public transit*". Both a descriptive analysis and an analysis of crash narratives offered insights into the characteristics of these crashes and revealed details often overlooked by structured data fields. These findings provide a foundation for improving how transit bus stop-related pedestrian crashes are identified in future research.

Key Findings

The review of literature and transit bus stop-related safety databases revealed limitations in the data previously used and are currently available to analyze pedestrian safety at bus stops. Numerous prior studies correlated the crash proximity to a bus stop with direct involvement [3, 4, 6, 8, 9, 11-13, 15, 16],

which may impact the accuracy of the results, since the extent of the pedestrian involvement in such crashes may not be accurately captured simply by the distance between the location of a crash and the nearest bus stop. This method has been widely used due to the availability of pedestrian crash data and bus stop locations, but it overlooks whether the crash was related to a bus stop or simply occurred nearby. The review of datasets further showed that those explicitly identifying crashes that occur at or are related to a bus stop are scarce, limited by small sample sizes and public accessibility. While commonly used crash typing frameworks (e.g., PBCAT and MMUCC) specify whether a crash was related to a bus stop, many datasets have not adopted the most recent versions of these frameworks. As a result, this type of crash may go unreported or miscategorized.

Even databases using the PBCAT and MMUCC frameworks may underreport transit bus stop-related pedestrian crashes. For example, major nationwide databases like FARS and CRSS have reported only 216 and 95 transit bus stop-related pedestrian crashes, respectively (2016-2023). These databases rely on an adapted PBCAT 2.0 definition, which requires a transit bus to be present at the stop during the crash and limits the scope to the immediate area around bus stops, excluding other potentially related crashes. Although PBCAT 3.0 addresses some of these limitations, FARS and CRSS have not yet adopted this updated version.

To further investigate the available data, a case study was conducted using Minnesota's crash data, which included 38 "*going to or from public transit*" pedestrian crashes reported from 2016 to 2023. This crash type offered a unique opportunity to focus on crashes specifically involving transit passengers, rather than just pedestrian crashes in close proximity to bus stops. With the full range of injury severity available, the data allowed for an initial understanding of the conditions influencing severity. The majority of crashes resulted in non-incapacitating injuries (39%), but nearly a third involved serious injuries or a fatality. Crashes most often occurred on higher-speed roads (over 30 mph) and at intersections, building on existing research that identified midblock locations as higher-risk locations for fatal transit bus stop-related pedestrian crashes [17]. Darting or dashing into the roadway was the most frequently documented pedestrian behavior (18%). This finding suggests that measures such as reduced speed limits and improved crossing infrastructure could address this behavior.

The crash narrative analysis identified six types of "*going to or from public transit*" pedestrian crashes, with the most common being accessing or egressing from transit (43%). These findings suggest that improvements to pedestrian safety for transit users should extend beyond the immediate area of bus stops, particularly at intersections on higher-speed roads. Additionally, this case study illustrates the value of crash narratives for identifying and categorizing transit bus stop-related pedestrian crashes, with 70% of narratives containing some form of transit bus stop-related language, demonstrating that future research could leverage unstructured text data to address data gaps in this area.

Limitations and Future Research

The Minneosta crash data used in this study was limited because of the relatively small sample size. Additionally, its statewide focus may not generalize to nationwide trends. Future research should use national or multistate crash databases to validate these findings. While transit bus stop-related pedestrian crashes have been analyzed nationally, prior research has focused exclusively on fatalities [17]. Future research should address this gap by using national data to examine the full range of transit bus stop-related pedestrian crash injury severity.

The analysis of crash narratives from the Minnesota case study also highlighted the potential to improve the identification of transit bus stop-related crashes using unstructured text data. These narratives provided valuable details often overlooked in structured data fields. Future research should consider leveraging crash narratives to improve the reporting of transit bus stop-related pedestrian crashes. Additionally, state agencies should consider modernizing their crash data reporting procedure by adopting PBCAT v3.0, which provides detailed and standardized criteria for documenting transit bus stop-related pedestrian crashes. This adoption would improve the consistency of data collection and support more effective state- and national-level analyses aimed at enhancing pedestrian safety at bus stops.

In summary, the absence of bus stop-related information in crash data does not imply a lack of pedestrian safety concerns. Research indicates that perceived safety influences walking behavior [41], suggesting underreporting may occur as pedestrians avoid unsafe bus stops or forgo transit use altogether. Ferenchak and Marshall [42] explored safety concerns arising from trip suppression using data on suppressed child pedestrian and bicyclist trips, derived from a survey of parental perceptions and trip frequencies. A similar framework could also be applied to foster safer access to bus stops. While crash data is essential for understanding pedestrian safety at bus stops, incorporating additional data sources can offer a more complete picture of the factors shaping pedestrian safety at bus stops. Addressing these data gaps by improving how transit bus stop-related crashes are reported is a key step toward advancing research and developing countermeasures to protect pedestrians when accessing, egressing, or waiting at bus stops.

ACKNOWLEDGMENTS

Thank you to the Minnesota Department of Transportation for providing the case study crash data. During the preparation of this work, the authors used ChatGPT in order to edit and proofread the manuscript.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: A. Rewalt, C. Brakewood; data collection: A. Rewalt and A. Thomson; analysis and interpretation of results: A. Rewalt, C. Brakewood, A. Thomson; draft manuscript preparation: A. Rewalt and C. Brakewood; All authors reviewed the results and approved the final version of the manuscript.

DECLARATION OF CONFLICTING INTERESTS

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

This research was supported by the Center for Pedestrian and Bicyclist Safety (Award No. 23UTK01) and the Dwight D. Eisenhower Transportation Fellowship Program (Award No. 693JJ32545013).

REFERENCES

1. NHTSA, *Traffic Safety Facts: 2022 Data - Pedestrians* 2024, U.S. Department of Transportation. .
2. Goughnour, E., et al., *Improving Safety for Pedestrians and Bicyclists Accessing Transit*. 2022, United States. Federal Highway Administration. Office of Safety.
3. Pulugurtha, S.S. and V.K. Vanapalli, *Hazardous bus stops identification: An illustration using GIS.* Journal of Public Transportation, 2008. **11**(2): p. 65-83.
4. Ulak, M.B., et al., *A stop safety index to address pedestrian safety around bus stops.* Safety science, 2021. **133**: p. 105017.
5. Kirley, B., et al., *Countermeasures that work: A highway safety countermeasure guide for state highway safety offices, 2023*. 2023, United States. Department of Transportation. National Highway Traffic Safety ….
6. Anis, M., S.R. Geedipally, and D. Lord, *Pedestrian crash causation analysis near bus stops: Insights from random parameters NB-Lindley models.* arXiv preprint arXiv:2410.22253, 2024.
7. Craig, C.M., et al., *Pedestrian safety and driver yielding near public transit stops.* Transportation research record, 2019. **2673**(1): p. 514-523.
8. Hess, P.M., A.V. Moudon, and J.M. Matlick, *Pedestrian safety and transit corridors.* Journal of Public Transportation, 2004. **7**(2): p. 73-93.
9. Pessaro, B., et al., *Impact of transit stop location on pedestrian safety.* 2017.
10. Pilla, M., et al., *Neighborhood Walking Environments Around Bus Stops: A Community-Based Participatory Approach for Streetscape Data Collection.* 2024.
11. Pulugurtha, S.S. and E.N. Penkey, *Assessing use of pedestrian crash data to identify unsafe transit service segments for safety improvements.* Transportation research record, 2010. **2198**(1): p. 93-102.
12. Salum, J.H., et al., *Toolkit for the assessment of safety at bus stops.* Transportation research record, 2024. **2678**(2): p. 279-292.
13. Singh, M., et al., *Effect of transit-oriented design on pedestrian and cyclist safety using bivariate spatial models.* Journal of safety research, 2022. **83**: p. 152-162.
14. Singleton, P.A., et al., *Pedestrian Safety and Traffic Operations Near Transit Stops*. 2024, Mountain-Plains Consortium.
15. Voorhes Transportation Center, *Pedestrian safety at bus stops*. 2012, Rutgers University.
16. Yu, C.-Y., *The impact of built environments on pedestrian safety around bus stops.* Journal of Urban Affairs, 2024. **46**(2): p. 267-280.
17. Rewalt, A., C. Brakewood, and C. Cherry, *An Analysis of Pedestrian Safety at Bus Stops Using FARS Data.* Journal of Safety Research, 2025. **95**: p. 147-159.
18. Harkey, D.L., et al., *PBCAT-Pedestrian and Bicycle Crash Analysis Tool Version 2.0*. 2006, United States. Federal Highway Administration. Office of Research ….
19. Hamilton, I., *PBCAT-Pedestrian and Bicycle Crash Analysis Tool Version 3.0 [techbrief]*. 2021, United States. Federal Highway Administration.
20. NHTSA, *Model Minimum Uniform Crash Criteria–5th Edition.* 2017.
21. NHTSA, *MMUCC guideline: Model minimum uniform crash criteria, 6th edition*. 2025.
22. NHTSA, *Fatality Analysis Reporting System (FARS)*. 2023, U.S. Department of Transportation.
23. NHTSA, *2014 FARS/CRSS Pedestrian Bicyclist Crash Typing Manual* 2016.
24. NHTSA, *2020 FARS/CRSS Pedestrian Bicyclist Crash Typing Manual: A Guide for Coders Using the FARS/CRSS Ped/Bike Typing Tool*. 2022.
25. Federal Transit Administration, *National Transit Database*. 2025.
26. North Carolina Department of Transportation. *NCDOT Non-Motorist Crash Map*. 2025; Available from:

https://ncdot.maps.arcgis.com/apps/mapviewer/index.html?webmap=b4fcdc266d054a1ca075b60715f88aef.
27. Nebraska Department of Transportation, *Investigator's Motor Vehicle Crash Report Manual*. 2023.
28. Louisiana Highway Safety Commission, *Louisiana Crash Data Element Manual*. 2023.
29. GeoPlan Center (University of Florida), *Signal Four Analytics Data Dictionary*. 2021.
30. Christofferson, B., Bitner, B., *Utah Crash Report Instruction Manual*. 2021, The Utah Highway Safety Office- Traffic Records Team.
31. Federal Highway Administration, *Highway Safety Information System (HSIS)*. 2025.
32. Olsen, G. *Idaho Transportation Department Crash Data Dictionary*. n.d.; Available from: https://support.numetric.com/en/articles/8715919-itd-data-dictionary.
33. Georgia Department of Transportation. *Georgia Crash Data*. 2025; Available from: https://gdot.aashtowaresafety.net/crash-data#/.
34. Minnesota Department of Transportation, *Minnesota Crash Data*. 2025.
35. U.S. Department of Transportation, *Enhanced Transit Safety Retrofit Package (E-TRP)*. 2025: Transportation.gov.
36. Los Angeles Police Department, *Traffic Collision Data from 2010 to Present*. 2025.
37. Bureau of Transportation Statistics, *Confidential Close Call Transit Data*. 2025.
38. *KartaView*. 2025; Available from: https://kartaview.org/landing.
39. QGIS, *QGIS Geographic Information System (Version 3.34)*. 2025, Open Source Geospatial Foundation.
40. OpenStreetMap, *OpenStreetMap*. 2025, OpenStreetMap Foundation.
41. Kweon, B.-S., et al., *The Effects of Pedestrian Environments on Walking Behaviors and Perception of Pedestrian Safety.* Sustainability, 2021. **13**(16): p. 8728.
42. Ferenchak, N.N. and W.E. Marshall, *Suppressed child pedestrian and bicycle trips as an indicator of safety: Adopting a proactive safety approach.* Transportation Research Part A: Policy and Practice, 2019. **124**: p. 128-144.