REFINEMENTS IN BENCHMARKING FOR VULNERABLE ROAD USER COLLISIONS BY COMPARING NATURALISTIC DRIVING AND POLICE-REPORT DATA

Eamon T. Campolettano
John M. Scanlon
Kristofer D. Kusano
Waymo, LLC
United States

Paper Number ESV26-294

ABSTRACT

It has been posited that naturalistic driving studies (NDS) offer an improvement of capturing lower severity events that are often missing from police-reported datasets and aid in filling in the overall collision severity distribution. As such, this study compared VRU collision benchmarks for New York City generated from NDS data and police-reported data to determine whether the NDS data offered a sufficient supplemental signal for benchmarking applications.

The NDS dataset consisted of collision data collected from vehicles outfitted with forward-facing dash cameras (Nexar) and driving over 100 million miles across New York City, and the police-reported dataset consisted of publicly-available collision and mileage data from the New York State Departments of Transportation and Motor Vehicles. Collision rates were generated for each dataset, with unique rates calculated for each VRU agent type (pedestrian, cyclist, motorcyclist) and each New York City county, as well as an aggregate city-level rate.

After correcting for baseline exposure differences in driving, the city-level collision rates for the NDS data were 0.27 per million miles (95% CI: 0.08-0.53) for pedestrians, 0.44 (95% CI: 0.23-0.77) for cyclists, and 0.36 (95% CI: 0.14-0.63) for motorcyclists. By comparison, the police-reported rates were 0.90 per million miles (95% CI: 0.88-0.91) for pedestrians, 0.45 (95% CI: 0.43-0.46) for cyclists, and 0.17 (95% CI: 0.16-0.18) for motorcyclists. While the hypothesis was that NDS data collision rates should be higher than police-reported crash rates, the results from this study indicated an inconclusive result for cyclists and motorcyclists and a lower rate in the NDS data for pedestrians. Continued work in characterizing these data and identifying potential geographic differences remains a viable research direction for better informing the total VRU crash burden in dense urban areas for ADS applications.

Keywords: ADS benchmarking, crash severity, severity mapping

INTRODUCTION

Overall traffic fatalities in the United States fell for the second consecutive year in 2023, though the overall fatality burden and fatality rate (per 100 million vehicle miles traveled) are approximately equal to figures from nearly 20 years ago [1]. Pedestrian fatalities saw a small decrease (1.3%) from 7,593 fatalities in 2022 to 7,314 in 2023 [2]. Cyclists (4.4%) and motorcyclists (1.3%) saw small increases in fatalities in 2023 [3-4]. These vulnerable road users (VRUs) thus comprise over 36% of all traffic fatalities in the United States. Beyond fatalities, traffic collisions involving VRUs can strain the resources of the healthcare system, with approximately 156,000 pedestrians, 118,000 cyclists, and 230,000 motorcyclists seeking medical treatment at a hospital emergency department following collisions with motor vehicles in 2023 [5]. Better understanding the aspects of VRU collisions is a key toward achieving meaningful reductions in traffic injuries and fatalities in the United States and working toward the goals of Vision Zero.

Vision Zero is an approach toward the elimination of serious and fatal injuries within the road transport system [6-9]. As adopted by the United States, as the Safe System Approach, there are five key objectives aimed toward the elimination of crashes that result in fatalities and serious injuries: Safer People, Safer Vehicles, Safer Speeds, Safer Roads, and Post-Crash Care [10]. An underlying, implicit aspect of Vision Zero is that the biomechanical tolerance of the human body to injury represents the limiting factor in a Safe System [6]. Beyond these objectives, the Safe System Approach has several principles. From the lens of VRU safety, relevant principles include the idea that death and serious injuries are unacceptable, that humans are vulnerable, and that humans make mistakes. Whether that be a distracted driver or a pedestrian crossing at an intersection without the right of way, that mistake should not mean someone has to be seriously or fatally injured. In other words, the design of the transportation system and/or vehicle safety features can be designed to limit the potential for these types of injuries in the event of a crash. Leveraging data from the current state of traffic collisions as they relate to VRUs can help provide insights into how modifications to the traffic safety network may ease the harm and burden associated with these types of collisions.

Police report data and national collision databases are the primary data sources generally used to provide meaningful information on these kinds of events, such as safety trends and collision rates. These collision rates often serve as benchmarks that can be used to assess the effectiveness of safety systems. For higher severity crashes, those involving serious injury or fatality, most events are reported to police and included in the databases; however, there is fundamental missingness associated with lower severity events (non-injury or minor injury), particularly for VRUs. There exists a well-documented underreporting of VRU collisions in police-reported data, where crashes were reportable (e.g., an injury occurred) but there was not an associated crash report; estimates from around the world have observed reporting rates ranging from 44-75% among pedestrians, 7-46% among cyclists, and less than 20% among motorcyclists [11-15]. In other words, while police-report data may produce reasonable collision rates for VRUs for higher severity collisions, it is expected that these same data would under-estimate the overall collision incidence for VRUs when considering the entire severity spectrum. Naturalistic driving data offers one potential avenue to fill in these lower severity events, as contact and collision events across the severity spectrum may be captured, with the caveat that representativeness of the data sample must also be assessed. Naturalistic driving data also offer the opportunity for additional qualitative and quantitative insights that are necessarily absent from the limited encoded information available from police report data. Additionally, collision scenarios may only be present in the police report data due to the observed outcome, while events with equivalent risk may be absent (e.g., two pedestrians struck by a vehicle traveling at 8 mph, one falls to the ground and sustains an injury while the other is able to maintain their balance and avoid a ground-induced injury). In this example scenario, the underlying crash configuration is the same, but the true incidence would not be captured because of the lack of an injury outcome for the second pedestrian. The capacity for a more complete perspective of lower severity events from naturalistic driving data sources is potentially key in more refined benchmarking comparisons.

Previous research has characterized VRU collisions from naturalistic driving data and sought to estimate overall collision rates from available mileage data, though to date a comparison to police-reported data has not been completed. Based on the information presented above, the operating hypothesis is that naturalistic driving data would be associated with higher recall (correct retrieval of true collision events) compared to police-reported data. This would predominantly be expected at the lower severity levels, as the more modest mileage associated with naturalistic driving data may not adequately capture higher severity events that are more rare in occurrence.

To date, the largest aggregation of VRU collision events from naturalistic driving data primarily involved collisions occurring in New York City (NYC) [16]. With the primary goal of creating a denser collision benchmark inclusive of lower severity events involving VRUs, this study supplemented that dataset with additional events, pairing naturalistic driving data from VRU collisions with publicly-available police report data from NYC. To that end, the principal objective was to leverage a city-specific collision severity distribution to generate an aggregate crash rate benchmark for pedestrians, cyclists, and motorcyclists in NYC. With these defined for the naturalistic driving data, the primary question of interest was to consider how the VRU crash benchmark compared to one defined using NYC police report data (i.e., are the NDS data credible for a benchmarking application?). With naturalistic driving data potentially providing increased capacity to better capture lower severity, injury-relevant collisions, that is, those crashes associated with meaningful risk of injury even in absence of an observed injury outcome, the comparisons developed in this study aim to improve the credibility of VRU collision benchmarks, an essential part of assessing ADS effectiveness relative to the current status quo [17]. A secondary aim of this study was to evaluate whether the continuous signal available through the reconstructed naturalistic driving data events could be used in the development of collision threshold benchmarks, as a supplement to collision outcome benchmarks available through police-reported data.

METHODOLOGY

This study leveraged data from several sources in order to address the outlined research questions. Naturalistic driving data were used principally to map out the observed severity distributions for collisions involving pedestrians, cyclists, and motorcyclists. Driving mileage data provided as part of the naturalistic driving data supplemented this analysis to allow for computation of collision rates. Police report data were also analyzed to generate corresponding VRU collision benchmarks for comparison to the naturalistic driving data. These police report data benchmarks were reported for several geographic regions to serve as a basis for comparison across geographic regions.

Naturalistic Driving Data

Naturalistic driving data for this study came from vehicles equipped with Nexar dash cameras operating in the five counties of New York City (Bronx, Kings [Brooklyn], New York [Manhattan], Queens, and Richmond [Staten Island]). This data sample is a multi-year sample spanning from July 2020 through October 2024. These data represented a census of all collisions involving VRUs in these counties identified by Nexar. The Nexar network, which contributes approximately 200 million miles of new driving monthly in the U.S., offers an opportunity to capture a wide range of driving events and environments, with drivers coming from rideshare drivers, corporate/municipal fleets, and the general public.

Collision Data Collision data from Nexar consisted of anonymized video of the event from a forward-facing camera and sensor data (Global Positioning Systems (GPS) and accelerometers) from the filming vehicle involved in the collision. Only data from collisions involving VRUs were included in this study. The expanded dataset presented herein represents an expansion of previously-described research, where collisions were identified using a combination of video and kinematic triggers, as well as individual reporting [16]. The available video, GPS, and accelerometer data were used as part of the reconstruction process to calculate vehicle speed at the time of impact, with video and GPS data being used to estimate VRU speed and impact time/location. The reader is directed to the original publication for more information on the reconstruction protocol [16]. All collision data were additionally manually reviewed by the authors to visually verify collisions and evaluate collision features as part of input into a collision speed assessment for severity estimation. Given the Nexar data is exclusively collected from a forward-facing camera, this dataset of VRU collisions did not include any events in which the rear of the vehicle was struck. While this would tend to result in some data missingness, these rear contact events are reported to occur relatively infrequently in the police-reported data, representing fewer than 2% of fatalities [1-4].

Severity Estimation For each event in the dataset, the reconstruction provided time-synced vehicle and VRU speeds at the time of impact. Review of the collision video provided supplemental signal on the impact location and configuration. While vehicle speed is a primary factor in collision energy for impacts involving VRUs, injury risk models for VRUs have increasingly incorporated the contribution of the VRU's speed and the collision configuration into the risk assessment to better capture the associated severity for a given event [18-21]. While this study did not carry forward the analysis through to injury risk assessment as has been done previously, the same principles were used to quantify collision speed to define the severity distribution.

Specifically, collision speed was defined for each event as the maximum of vehicle speed, VRU speed, and closing speed between the two agents. In accordance with previous research, side contacts are often associated with incomplete engagement between the VRU and the vehicle, where energy transfer can be more limited than in a frontal contact. Neale et al. (2021) performed photogrammetric analysis of video recordings of pedestrian sideswipe and minor overlap impact events, observing that level of engagement with the vehicle was found to vary depending on the nature of the collision and was lower than would be expected in a frontal collision [22-23]. For this study, a scaling factor of 1.5 was applied for pedestrian collisions involving side contact to account for this decrease in engagement (i.e., a side collision at 15 mph would be modeled as if it were a 10 mph frontal collision). Based on this phenomenon, Schubert et al. applied a similar methodology in their development of injury risk models for cyclists and motorcyclists, finding that non-frontal contacts were associated with a 30% reduction in engagement relative to a frontal contact (i.e., a side collision at 15 mph would be considered as equivalent to a 10.5 mph frontal collision) [20-21]. It is noteworthy that these two independent methodologies and data sources for capturing this effect of reduced engagement for non-frontal contacts arrived at such consistent solutions.

Data Filtering As part of this video and data review, three criteria were used to flag specific events for filtering as part of the benchmark comparisons. Firstly, events in which there was no direct contact between the VRU and the filming vehicle were flagged for exclusion. As part of the automatic process for identifying events, those in which there is a vehicle response (e.g., braking and/or steering) and a proximity buffer between the VRU and the vehicle may be initially categorized as a collision event. This lends itself to a conservative assessment of collision risk but with the potential for lower levels of precision. Secondly, events in which there was limited direct contact between the VRU or their bicycle/motorcycle and the vehicle were flagged. These events were associated with limited engagement, with very low to negligible impulse transfer associated with the collision event [24]. Examples include a pedestrian placing their hands on the hood of a braking vehicle as part of their avoidance maneuver or a cyclist slowly rolling into a stopped vehicle to bring their bicycle to rest rather than using the brakes. Events of this nature are not expected to be associated with injury risk and would certainly not be expected to appear in police report databases. Exclusion of events such as these results in a higher confidence naturalistic dataset consisting of those events with injury relevance. Lastly, events were binned based on the reconstructed collision speed for the event, with 1-mph increments used to define different thresholds from 5 mph to 30 mph to serve as a supplemental collision rate signal. By way of example, previous research has shown that 10 mph approximately represents a limit of VRU post-collision stability, in that nearly all VRUs are knocked down following a collision at 10 mph or greater [24-25]. This is noteworthy in that additional research has shown that contact with the ground was the identified source of injury for more than half of pedestrian events at impact speeds below 30 kph (18.6 mph) [26]. As such, it is not unreasonable to expect that events at speeds of 10 mph and greater would be more likely to show up in police report data than events at lower speeds.

Driving Mileage Mileage data was provided by Nexar in two formats: in aggregate summary and as a sample of overall driving from trip data in 1-minute increments. The aggregate summary consisted of monthly breakdowns of the overall mileage, as well as the total mileage within the provided sample for the given month. This was provided at the borough level for this analysis. The target sampling percentage was approximately 20% each month and was sampled with uniform probability to provide an estimate for the overall driving distribution. The trip data samples consisted of latitude, longitude, and speed data for the vehicle's trip in 1-second increments for a total of 1-minute per trip. These were used to (1) identify on what types of roads the sampled driving mileage occurred on and (2) where in the city the driving occurred. These two factors were used as proportional factors in scaling the overall aggregate driving mileage. Driving mileage data was supplied from November 2018 through the end of June 2022.

Police Report Data

Police report data for this study consisted of all police-reported collisions in New York State. These data are publicly available from the New York State government (data.ny.gov). The data provided include the most recent complete year of collisions, as well as the two years prior to that. Thus, for this study, data from 2022, 2023, and 2024 were available.

Collision Data The police report data were filtered to only include those collisions involving VRUs and occurring in one of the five boroughs of NYC. The collision data consist of four tables providing information at the overall crash level, vehicle level, person level, and then a final table including information related to traffic violations associated with the crash (not used for this analysis) [27]. New York's crash data reports severity across two variables (MaxInjurySeverity and CrashSeverity). These two variables were consolidated into a single assessment of severity leveraging the 'KABCO' classification system. Encoded latitude, longitude, and road name information were used to assign each crash to the appropriate road type based on the Federal Highway Administration's (FHWA) functional classification system and classification designation within the FHWA's Highway Performance Monitoring System (HPMS) data [28]. For this analysis, road type was bifurcated as surface street (local, collectors, and non-freeway arterials) or freeway (interstates and "principal arterials: other freeways and expressway"). Further, only crashes involving in-transport (not parked) passenger vehicles were included given that this is the population of vehicles for which Nexar collision and driving mileage data were provided. This classification reflects vehicles with a gross vehicle weight rating of under 10,000 pounds [29]. Given this study's focus on VRU collisions, motorcycles were also retained in the dataset. This was achieved by using the *VehicleType, VehicleBodyType*, and *PreCrashAction* ('PARKED') fields to identify relevant vehicles. VRU collisions were identified using the *CrashType* field for pedestrian and cyclist crashes and leveraging the vehicle typings for cases involving motorcyclists.

Driving Mileage Driving mileage consisted of a combination of data from the Federal Highway Administration and the New York State Department of Transportation [30]. Driving mileage was assigned road typing in the same manner as the crash data described above. The FHWA reports driving mileage attributable to different vehicle classes (e.g., motorcycles, passenger vehicles, heavy vehicles) at the state level for different road types. This was used to scale the overall driving mileage appropriately to only consider that of passenger vehicles.

Rates Calculations

Given that the collision and mileage data from Nexar have differing time windows, collision rates using the Nexar data are only computed using the overlapping time period where both sources of data are present: the two-year period from July 2020 through June 2022. While not a perfect comparison, the benchmark data from police-reported data for NYC presented herein is from the calendar year 2022. Further, given that the dataset consisted of so few freeway collision events and the relative rarity of these events in general, for this analysis, only a comparison for collisions occurring on surface streets was conducted.

All computed rates were crashed vehicle rates (vehicle-level crash rates), which is defined as the number of crashed vehicles per vehicle mile traveled. In aiming to be compliant with the RAVE Checklist, particularly item 1C, a vehicle-level crash rate is best for direct comparison of systems [17]. It can be thought of more simply as how many times did a vehicle have a collision involving a VRU rather than how many total VRUs were involved in collisions with a vehicle, adjusted for driving mileage. By not indexing on the total number of involved VRUs in a given collision, the resulting rate represents the absolute estimate for how frequently these types of collisions occur, allowing for a more apples-to-apples comparison between different sources. Approximately 4% of events in this dataset involved contact with more than one VRU in a single collision (e.g., striking two riders on a single bicycle). This is not to suggest that the potential increased harm associated with collisions involving multiple VRUs should be discounted; rather, there are other mechanisms and metrics where absolute numbers of involved persons/injuries/fatalities are appropriate.

In light of the hypothesis stated at the outset of the paper (i.e., that expected higher levels of recall for lower severity events in the naturalistic driving data should result in higher collision rates than police-reported data), the observed rate of these collisions in the Nexar dataset were compared against two human driving benchmarks for VRU collisions: (1) all police-reported collisions and (2) all suspected serious injury or worse collisions, identified by any involved person reported as having a serious (A) of fatal (K) injury on the KABCO classification system as noted on the police report. It should be noted that, relying on the police-reported outcome level is not without limitations for safety impact analyses, largely attributable to the underreporting phenomenon highlighted in the introduction [17]. That being said, for this study leveraging naturalistic driving data, a comparison to this benchmark level offers some insight into the expected recall (proportion of collisions in the dataset out of total observed collisions) of the naturalistic driving data. At very low severity levels, particularly for VRU collisions, there is some unquantifiable degree of expected data missingness given the more limited kinematic signals present in these events.

All rates presented in this study are on a per million mile basis. Confidence intervals (95%) for these rates were developed using Byar's method, which represents an exact approximation to the Poisson distribution and retains high levels of accuracy for both small and large counts, to the count data and normalizing by the relevant mileage to compute the rate [31].

RESULTS

Collision Data

Naturalistic Driving Data Overall, this naturalistic driving dataset consisted of 674 total events identified as collisions involving VRUs. After initial video review, 592 contact events involving VRUs in NYC remained in the dataset (Table 1). Approximately 14% of verified contacts were deemed to not have injury potential (e.g., limited engagement scenarios like a pedestrian placing their hands on the hood of a braking vehicle as part of their avoidance maneuver or a cyclist slowly rolling into a stopped vehicle to bring their bicycle to rest rather than using the brakes) based on video review, as outlined in the methods.

Table 1. Event count filtering for naturalistic driving data events by agent type

	Pedestrian	Cyclist	Motorcyclist
Nexar-identified	246	302	126
Video-verified Contacts	194	279	119
Potential for injury	154	250	105

Given these distributions of event counts, it is not surprising that the median pedestrian collision speed (5.1 mph [8.2 kph]) was lower than that of either cyclists (7.2 mph [11.6 kph]) or motorcyclists (8.3 mph [13.4 kph]). This trend was also observed for higher severity events at the 75th and 95th percentiles (Figure 1). Collision speeds varied from 0 to 21.7 mph [34.9 kph] for collisions involving pedestrians, 0 to 38.6 mph [62.1 kph] for motorcyclists, and 0 to 24.2 mph [38.9 kph] for cyclists.

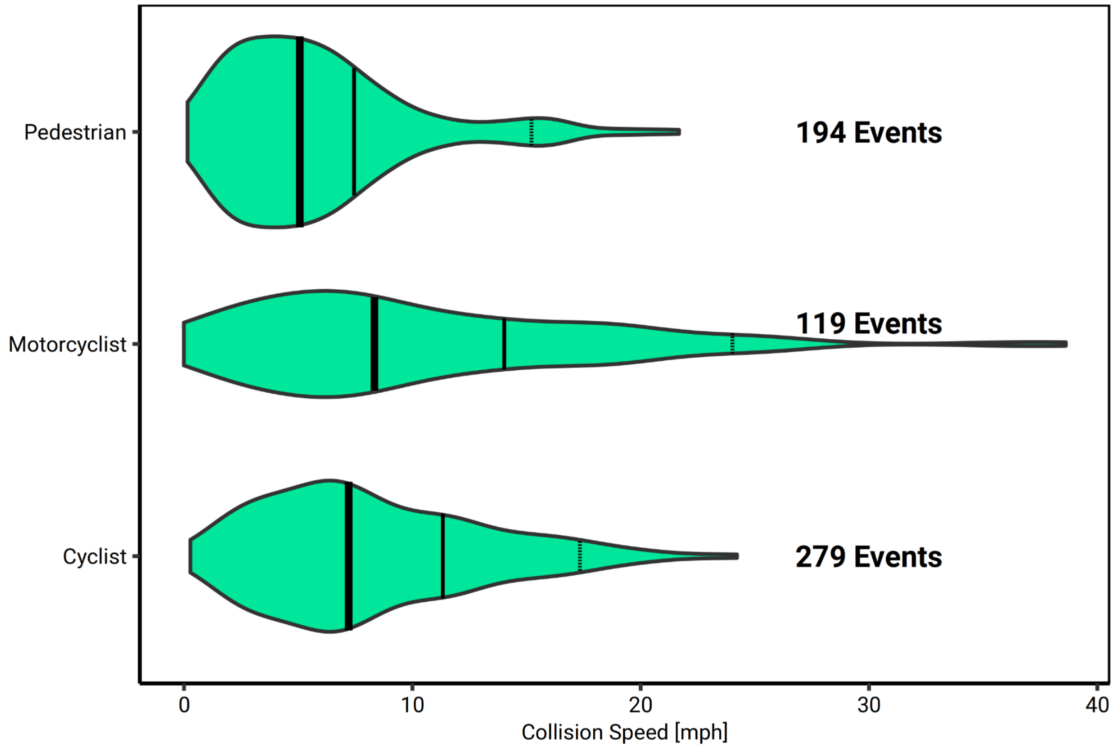

Figure 1. Distribution of collision speed by VRU agent type for all video-verified contact events in the dataset. Thick, solid line represents the median, solid line is the 75th percentile, and dashed line is the 95th percentile collision speed value.

Police Report Data Across the five boroughs of New York City, passenger vehicles were reported to drive nearly 1.4 billion miles in 2022. Surface street collision rates at two severity levels (police-reported and suspected serious injury or worse) are presented in Table 2. Collisions with pedestrians were most frequently observed among all VRU agent types, occurring more than 5x as often as collisions involving motorcyclists. Aggregation by borough for both mileage and collisions is presented in the Appendix (Table A2).

Table 2. Selected surface street benchmarks for human drivers in New York City for 2022

Total mileage: 9494.58374 Mmi	**Pedestrians**	**Cyclists**	**Motorcyclists**
Police-reported Collisions			
Crashes	8501.58	4239.95	1610.95
IPMM [95% CI]	0.895 [0.877 - 0.915]	0.447 [0.433 - 0.460]	0.170 [0.161 - 0.178]]
'KA' Collisions (Suspected Serious Injury+)			
Crashes	835.45	296.15	200.89
IPMM [95% CI]	0.088 [0.082 - 0.094]	0.031 [0.028 - 0.035]]	0.021 [0.018 - 0.024]

Driving Mileage

For this study, the two years of driving mileage from Nexar in the five boroughs of New York City consisted of 139 million miles of driving. The most driving was observed in Queens County (75.7 million miles), with Richmond County the least amount of driving, only contributing 2.4 million miles (Table 3). From the sampled mileage data, the proportion of driving on surface streets was observed to be 66.7% of all driving across the five boroughs, ranging from 56% in Queens County to 82% in Kings County. Given this, the aggregate surface street driving mileage may be estimated to be 92.7 million miles. In comparison, the reported FHWA driving mileage for New York City skews toward a much lower proportion of surface street driving mileage.

Table 3. Summary of driving mileage data for both the naturalistic driving data and police-reported collisions

	Nexar Data (July 2020 through June 2022)			New York City Driving Mileage (2022)		
County	**Total Mileage (Mmi)**	**Surface Street Proportion**	**Surface Street Mileage (Mmi)**	**Total Mileage (Mmi)**	**Surface Street Proportion**	**Surface Street Mileage (Mmi)**
Bronx	12.1	72.0%	8.7	2923.3	44.3%	1294.9
Kings	31.9	83.6%	26.6	4092.8	63.4%	2596.9
New York	17.0	80.7%	13.7	2579.9	53.3%	1375.1
Queens	75.7	55.9%	42.3	6944.9	46.2%	3209.5
Richmond	2.4	79.0%	1.9	1857.4	54.8%	1018.2
New York City	139.0	66.9%	92.9	18,398.3	51.6%	9494.6

The approximate driving distributions also vary between two datasets. While Queens represents the county with the largest surface street driving mileage in both, driving there comprises approximately 46% of overall NYC driving in the Nexar data and only 34% in the FHWA driving data. Based on the differences observed in Table 3, a simple scaling factor was applied to the Nexar driving mileage so that the appropriate proportions as measured by FHWA were maintained for more consistent aggregate comparisons between the two datasets.

Collision Rates

As outlined in the methods, the naturalistic driving event data comes from a longer timeframe than the mileage data, so it is necessary to align the two for rates-based comparisons to the police report data. For the two year period (July 2020 through June 2022) with mileage data, there were 104 total contact events, with 96 (92%) being considered as having injury potential. Beyond that, defining the collision rate based on a collision speed threshold was expectedly shown to vary considerably at different speeds (Figure 2). For example, using a 10 mph collision speed threshold would be associated with a reduction in collision rates compared to a 5 mph threshold of 49% for cyclists, 39% for motorcyclists, and 65% for pedestrians. It should be noted that while the confidence intervals appear larger, in absolute magnitude, at lower collision speed thresholds where there is more data, the size of the confidence interval relative to the estimated collision rates does increase at higher collision speed thresholds as expected.

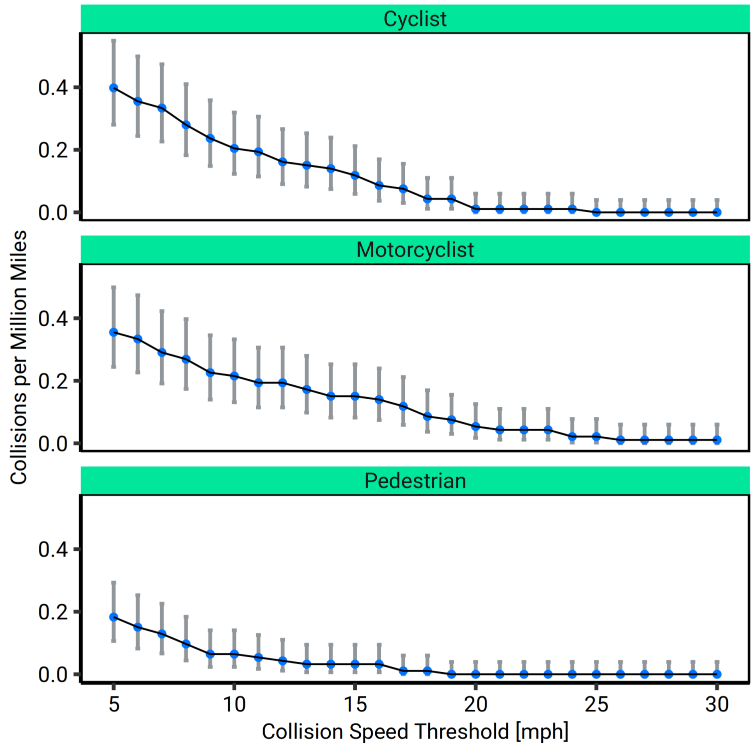

Figure 2. Variation in collision rate by VRU agent type as a function of increasing collision speed threshold.

In order to make comparisons between the naturalistic driving data and the police-reported data, consideration of the driving exposure differences outlined in Table 3 is necessary. As shown in Table 4, the driving data for Kings and New York Counties were approximately proportional between the two datasets, with Queens County being overrepresented in Nexar, and the Bronx and Richmond Counties being considerably underrepresented. Aggregate estimates of collision rates for NYC as a whole would be scaled based on these factors rather than simply taking the overall rate.

Table 4. Re-weighting of surface street driving mileage by county for naturalistic driving data to align with FHWA mileage.

County	Nexar Mileage Proportion	FHWA Driving Mileage Proportion	Scale Factor
Bronx	9.4%	13.6%	1.45
Kings	28.7%	27.4%	0.95
New York	14.8%	14.5%	0.98
Queens	45.6%	33.8%	0.74
Richmond	2.0%	10.7%	5.35

As driving mileage varies across the different boroughs of New York City, so too do collision rates (Figure 3). Tabulated data for this comparison is available in the Appendix (Tables A1 and A2). While the absolute rates from the naturalistic driving data do not align with the police report data, the borough-level comparison is directionally consistent, with Kings and New York counties having the highest collision rates, followed by the Bronx, and then Queens. Within New York County, the overall collision rate from the naturalistic driving data for cyclists and motorcyclists expectedly exceeds the police-reported rates for those VRU types. In general, a higher rate of motorcyclist events was observed in the naturalistic driving data than in the police-reported data, while the opposite trend was observed for pedestrian events.

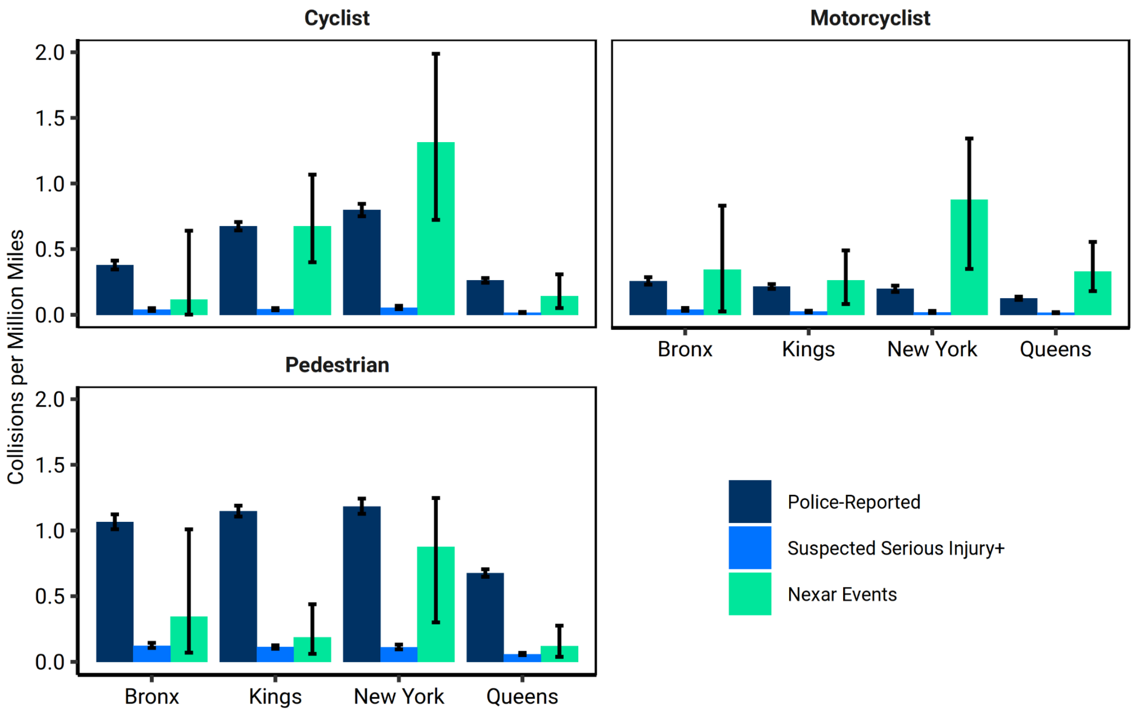

Figure 3. Summary of collision rates at various severity levels across New York City boroughs, broken down by VRU type. Note that Richmond County (Staten Island) is not included in the above plot due to a lack of collision data from the naturalistic driving dataset.

Even with the variability in collision rates and exposure by county, it is also relevant to investigate how the aggregate city-level collision rate compares between the naturalistic driving data and the police-reported data. All collision rates from the Nexar data were adjusted using the scale factors presented in Table 4. The visually-verified collision rate from the naturalistic driving data is approximately 2x as large as the police-reported rate for motorcyclists, approximately equal for cyclists, and approximately 3x lower for pedestrians when aggregated across all the counties that make up New York City (Figure 4). Given that it would not be expected that all of the collisions identified in the naturalistic driving dataset would be police-reported, a demonstrative example using a 10-mph collision speed-based threshold for rates estimation is also presented in Figure 4. This necessarily reduces the number of collision events included in the comparison, resulting in lower rates across the board. While the 10 mph+ event criteria results in comparable crash rates between the naturalistic driving data from this study and the total police-reported rate for motorcyclists, the same comparison is associated with approximately 2.5x as many collisions for cyclists and nearly 15x as many for pedestrians (Figure 4).

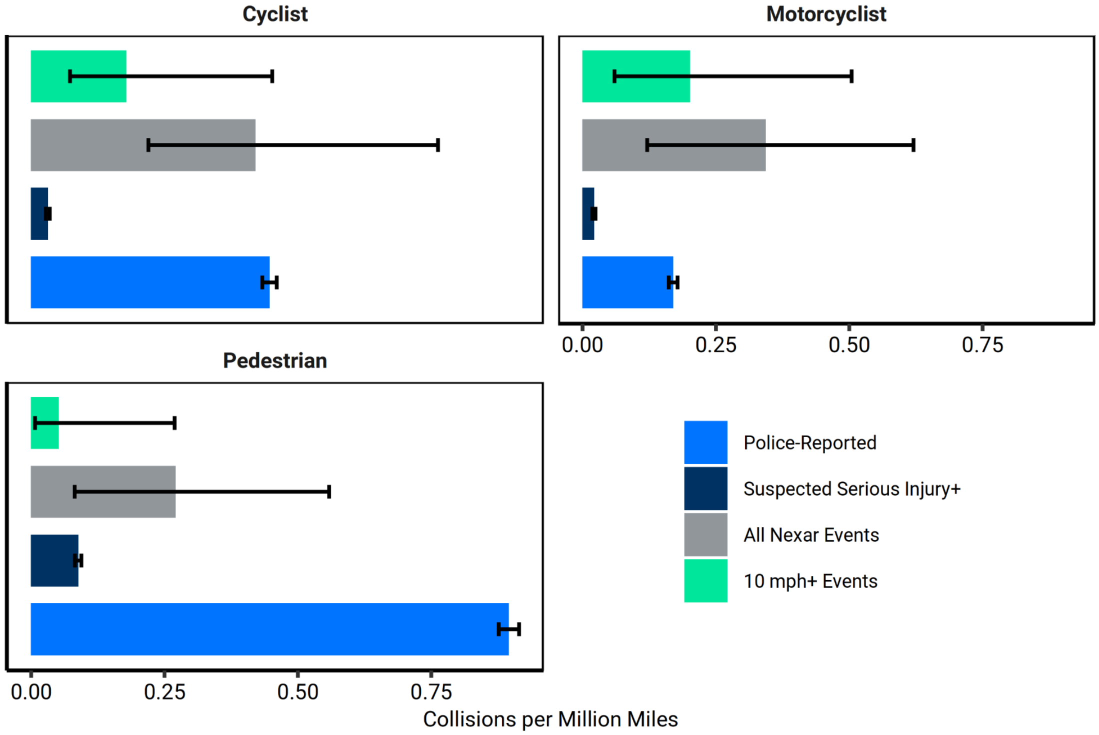

Figure 4. Summary of aggregate collisions per million miles at various severity levels, grouped by VRU agent type. Shades of blue correspond to the police report data, while the green and gray bars come from the naturalistic driving data.

DISCUSSION

This study presented a comparison of VRU collision rates in New York City between naturalistic driving data and police-reported data, leveraging one of the largest collections of naturalistic VRU collision events observed in the literature. After correcting for baseline exposure differences in driving, observed differences in collision rates between these datasets would seem to indicate that, despite such a high concentration of VRU collision events in this naturalistic driving dataset, the cyclist and motorcyclist crash rate comparison was inconclusive, and the pedestrian collision rate exhibited the opposite trend (i.e., police-reported crash rate was higher).

While the aggregate city-level collision rates suggest that the naturalistic driving dataset may be limited for overall benchmark comparisons across all VRU types, there may be utility for motorcyclists (Figure 4). Beyond that, the county-level analysis that comparisons involving cyclists may also be possible for Kings and New York Counties. It does appear as though benchmarking applications for pedestrians in New York City using this dataset are not viable at this time given the large difference in rates observed here. These observed differences can likely be attributed to a few different factors: (1) differences in driving exposure and behavior between the Nexar dash camera users and the overall driving population of New York City; (2) underreporting and/or reporting biases in the police-reported data; and (3) possible recall issues related to the identification of collision events in the naturalistic driving dataset, as well as other potential factors. Specific to the first factor listed, potential explanations for these differences include a higher share of ride-hailing or fleet service vehicles than the general driving population, which could be associated with both a differing level of safe driving, as well as differing driving patterns in regards to time of day, day of week, and areas of the city in which driving occurs.

Higher event recall is generally hypothesized to be present in naturalistic driving datasets, compared to police-reported data, though this study largely only observed that effect for motorcyclist events across the board and cyclist events in New York County. This denser signal for lower severity events can provide supplemental information that police-reported data may be missing. Using thresholds based on reconstructed collision speeds

(Figures 2 and 4) allows for a quantifiable severity component to be attached to these naturalistic driving events for which there is no objective outcome data available. Given previous research showing that experiencing a collision exceeding 10 mph (16 kph) approximately represents the limits of human postural stability, a demonstrative comparison using a 10 mph based threshold for collision rates was carried out (Figure 4) [24]. This was not to suggest that all events above this threshold would result in injury, nor that all contacts below 10 mph (16 kph) would be non-injurious. In reality, some of these lower speed contacts may result in injury, particularly if there is a loss of balance or fall to the ground as a result of the impact or in trying to avoid direct contact with the oncoming vehicle. The vast swath of literature related to the biomechanics of injury associated with falls from standing height serve as an additional support for this [eg., Verma et al., 2016 [32]]. The use of other collision speed thresholds for rates estimation may be appropriate depending on the severity level of interest and the underlying data.

Despite these limitations in the aggregate rates comparisons, this is not to suggest that these naturalistic collision data are without utility in ADS and traffic safety applications. This sample of VRU collision events from dash cameras represents a rich set of qualitative and quantitative data that has previously not been possible with more systematic naturalistic driving studies in the past [16]. The collision severity distributions provide meaningful insights into the relative frequency of higher severity events as a function of all contact events that may be useful in re-sampling or re-weighting other datasets. Further, the granularity that is possible from video-based analysis provides for additional insights that codified police-reported datasets may not offer. Specifically, the encoded information in NYC police report data involving VRUs allows for some scene information related to how the contact came to pass (i.e., what was the VRU doing, how were they traveling in the roadway network); however, additional insights such as whether the vehicle/VRU saw the other actor prior to the event or whether evasive action was taken is not included. Previous research has shown that over 90% of VRUs attempt an avoidance maneuver when they see the vehicle prior to contact; to say nothing of the potential for vehicles to perform pre-impact swerving or braking [24]. Lastly, because naturalistic driving data provide information on observed contacts and collisions across the severity spectrum, insights on the most commonly observed conflict types or crash configurations are also attainable. From a traffic safety standpoint, identifying and indexing on the most commonly observed injury-causing configurations is key toward safety interventions geared at reducing the injury burden; from an ADS evaluation standpoint, there is an additional component necessary of being able to quantify the aggregate effect on the overall collision space. Changes to the crash distribution are important to quantify to ensure that any increasingly observed collision configurations are not associated with high severity injury potential.

While the overall collision rates were different between the police-reported data and the naturalistic driving data, the two data sources were directionally consistent at the county level, with New York County associated with the highest rates among the New York City counties (Figure 3). Further, it is reasonable to expect that collision risk and severity vary within a given borough as well, as has been previously shown in other cities [33]. In fact, a geographic-based visualization of the naturalistic driving collision and mileage data illustrate this effect (Figure 5). While there is a larger concentration of driving mileage in Queens, those same areas are not associated with the highest concentrations of collisions. The denser areas of Manhattan (New York County) are expectedly associated with a higher overall occurrence of collisions involving VRUs. According to data from the 2024 national American Community Survey, New York City has the largest proportion of active commuting, with approximately 60% of workers biking, walking, or taking public transport (which necessarily involves arriving at these transit stops) to work [34]. The average commuter in NYC thus is exposed to potential conflict points in the road traffic network as a VRU on a daily basis, often multiple times per day.

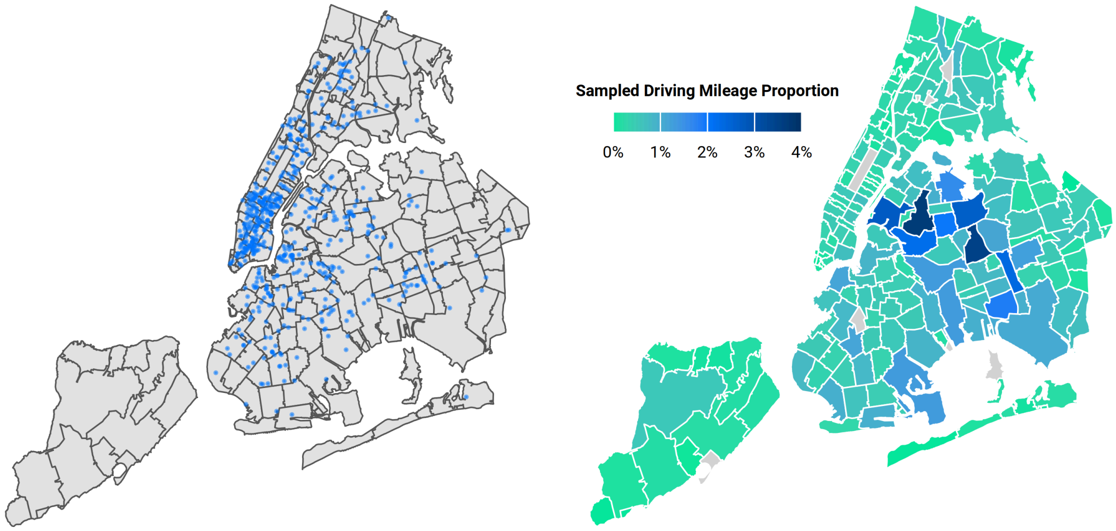

Figure 5. Left: Geographic visualization of collision locations (blue dots) from naturalistic driving dataset, with zip code borders shown. Larger concentrations of events are identified by clusters. Right: Geographic visualization of driving mileage by zip code (darker blues represent higher concentrations of mileage, while green is indicative of lower mileage. Grayed out zones are park areas where there is expectedly no mileage.

While computing collision rates on a per vehicle mile traveled basis is quite common and particularly useful for retrospective vehicle safety analyses, when it comes to VRUs, they perhaps present an incomplete picture. Consider two hypothetical scenarios. First, a vehicle traveling on a multi-lane arterial roadway with no bike lanes or pedestrian crosswalks at intersections. Second, consider a vehicle traveling on a city street with separated bike lanes and pedestrian crosswalks at each intersection. Independent of any collision data, it is clear that these two roadways are very different from the perspective of a VRU. While unlikely, it is possible that these two roads could have similar VRU collision rates from the vehicle perspective; however, the exposure to VRUs clearly would be expected to vary considerably, with far more opportunity for VRUs and vehicles to come into conflict in the second scenario. By way of a real-world example, the suspected serious injury or worse collision rate for passenger vehicles with pedestrians in Los Angeles in 2022 was 0.0298 per million vehicle miles traveled (approximately ⅓ as much as the rate in New York City), and the fatal injury rate for pedestrians was approximately 25% lower in LA than NYC for 2022 from the vehicle perspective [35]. While not a perfect estimate for exposure, the fatality rate on a per 100,000 population basis was 1.32 in NYC, compared to 4.00 in Los Angeles [2]. This example illustrates the importance of including VRU exposure into safety evaluations when dealing with VRU rates; depending on which rate calculation carried out, in isolation, one would come to very different conclusions regarding pedestrian safety between these two cities. Efforts to estimate the exposure of VRUs to traffic conflicts/collisions/injuries are an important part of the overall safety picture, and low VRU collision rates from the vehicle perspective are not necessarily an indication of risk from the VRU perspective. Examples of ways to incorporate this include modeling of pedestrian walk miles or cyclist miles traveled as a measure of their exposure [36], as well as using smaller geographic areas, such as zip codes or S2 cell geometries to better capture variation in exposure to VRUs (e.g., business district vs. industrial zone differences that may be masked when looking at a larger region) [33].

Taken together, the variability in collision severity, collision location, and driving mileage presented in this study exemplify that driving data are not a monolith. Collisions and driving mileage can vary considerably across different cities, but also within cities, as shown. Accordingly, there are differing levels of collision and injury risk associated with driving in various parts of cities that should be considered in the evaluation of relative comparisons between cities or in benchmarking applications of ADS performance.

LIMITATIONS

There are several limitations of this study to note. Firstly, the collision identification schema and collision reconstruction protocol are associated with some degree of error. While collisions are identified using a combination of kinematic and visual data with high degrees of specificity, as well as the capacity for end users to manually trigger events, the potential for a missed collision does remain non-zero. Similarly, the reconstruction data is dependent on the mapping between the video and location data, with estimates indicating accuracy to within 0.5 m for dynamic objects, like other vehicles or VRU agents. Further, all the events in this study are from VRU collision events in New York City; these results may not be generalizable to other dense urban locations or to non-dense urban locations. As evidenced by the differences in driving mileage distribution between the naturalistic driving data and the FHWA data, there may be population-level differences between Nexar dash camera users and the overall driving population in New York City that are at play. Again, as noted above, the differences between the naturalistic driving data and police-reported data benchmarks is an indication that these may not be appropriate for full-scale benchmarking. The mismatch in collision data and driving mileage within the naturalistic driving data also presented a limitation that restricted the number of events to be considered for comparison to the police-reported data and resulting in wider confidence intervals than would be expected with a larger dataset. Similarly, the police-reported data are not without limitations. Specifically, the limited number of encoded variables for each crash can be limiting in making comparisons or gleaning additional insights from given events. Lastly, the mileage data used in this study, from both state data and the naturalistic driving data, represent estimates based on data samples and the extrapolation of these data to aggregate estimates may be associated with variability and bias that is not adequately captured with existing methodologies.

CONCLUSIONS

This study performed the first comprehensive comparison of naturalistic driving VRU collision data to police-reported collision benchmarks. While the hypothesis was that naturalistic driving data collision rates should be higher than police-reported crash rates, the comparison between these two datasets for collisions in NYC indicated an inconclusive result for cyclists and motorcyclists and a lower rate in the naturalistic driving data for pedestrians. The observed collision rate differences may be attributable to differences in driving exposure (time of day, day of week, area of city) and behavior (safe driving) between the Nexar dash camera users and the overall driving population of New York City, possible recall issues related to the identification of collision events in the naturalistic driving dataset, and any potential reporting biases present in the police-reported data; though, it should be noted that fully accounting for these differences may not be viable. Despite this, the utility of these novel collision data continue to be in the additional insights regarding some of the qualitative components of the collision events that police reports do not encode. Continued work in characterizing these data and identifying potential geographic differences remains a viable research direction for better informing the total VRU crash burden in dense urban areas for ADS applications, as well as broader traffic safety goals of reducing the rising contribution of VRUs traffic injuries and fatalities.

REFERENCES

[1] National Center for Statistics and Analysis. (2025, April). Overview of motor vehicle traffic crashes in 2023 (Traffic Safety Facts Research Note. Report No. DOT HS 813 705). National Highway Traffic Safety Administration.
[2] National Center for Statistics and Analysis. (2025, June). Pedestrians: 2023 data (Traffic Safety Facts. Report No. DOT HS 813 727). National Highway Traffic Safety Administration.
[3] National Center for Statistics and Analysis. (2025, July). Motorcycles: 2023 data (Traffic Safety Facts. Report No. DOT HS 813 732). National Highway Traffic Safety Administration.
[4] National Center for Statistics and Analysis. (2025, July). Bicyclists and other cyclists: 2023 data (Traffic Safety Facts. Report No. DOT HS 813 739). National Highway Traffic Safety Administration.
[5] Centers for Disease Control and Prevention. Web-based Injury Statistics Query and Reporting System (WISQARS). Atlanta, GA: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. Available at cdc.gov/injury/wisqars. Accessed on 2025/12/01.
[6] Larsson P, Dekker SW, Tingvall C. The need for a systems theory approach to road safety. Safety science, 2010, 48(9), 1167-1174. https://doi.org/10.1016/j.ssci.2009.10.006
[7] Lie A, Tingvall C. Governmental status report, Sweden (No. 2001-06-0121). SAE Technical Paper. 2001
[8] Tingvall C. (1997). The zero vision: A road transport system free from serious health losses. In Transportation, traffic safety and health: the new mobility (pp. 37-57). Berlin, Heidelberg: Springer Berlin Heidelberg.
[9] Tingvall C. (2022). Vision zero: how it all started. In The Vision Zero Handbook: Theory, Technology and Management for a Zero Casualty Policy (pp. 1-22). Cham: Springer International Publishing.

[10] United States Department of Transportation. (2022). National roadway safety strategy. https://www.transportation.gov/NRSS/SafeSystem
[11] Watson A, Watson B, Vallmuur K. Estimating under-reporting of road crash injuries to police using multiple linked data collections. Accident Analysis & Prevention, 2015, 83:18-25. https://doi.org/10.1016/j.aap.2015.06.011
[12] Gildea K, Simms C. Characteristics of cyclist collisions in Ireland: Analysis of a self-reported survey. Accident Analysis & Prevention, 2021, 151. https://doi.org/10.1016/j.aap.2020.105948
[13] Shinar D, Valero-Mora P, et al. Under-reporting bicycle accidents to police in the COST TU1101 international survey: Cross-country comparisons and associated factors. Accident Analysis & Prevention, 110, 177-186, 2018. https://doi.org/10.1016/j.aap.2017.09.018
[14] Janstrup KH, Kaplan S, Hels T, Lauritsen J, Prato CG. Understanding traffic crash under-reporting: linking police and medical records to individual and crash characteristics. Traffic injury prevention, 17(6), 580-584, 2016. https://doi.org/10.1080/15389588.2015.1128533
[15] Doggett S, Ragland DR, Felschundneff G. Evaluating research on data linkage to assess underreporting of pedestrian and bicyclist injury in police crash data. UC Berkeley: Safe Transportation Research & Education Center. Safe Transportation Research and Education Center, University of California, Berkeley, 2018. https://escholarship.org/uc/item/0jq5h6f5
[16] Campolettano, E. T., Scanlon, J. M., Kadar, I., Lavy, L. Y., Moura, D. C., & Kusano, K. D. (2024). Baseline vulnerable road user injury risk in multiple US dense urban driving environments. Traffic Injury Prevention, 25(sup1), S94-S104. https://doi.org/10.1080/15389588.2024.2364050
[17] Scanlon, J. M., Teoh, E. R., Kidd, D. G., Kusano, K. D., Bärgman, J., Chi-Johnston, G., Di Lillo, L., Favaro, F., Flannagan, C., Liers, H., Lin, B., Lindman, M., McLaughlin, S., Perez, M., & Victor, T. (2024). RAVE Checklist: Recommendations for Overcoming Challenges in Retrospective Safety Studies of Automated Driving Systems. Traffic Injury Prevention. 1–14. https://doi.org/10.1080/15389588.2024.2435620
[18] Lubbe N, Wu Y, Jeppsson H. Safe speeds: fatality and injury risks of pedestrians, cyclists, motorcyclists, and car drivers impacting the front of another passenger car as a function of closed speed and age. Traffic Safety Research, 2022, 2. https://doi.org/10.55329/vfma7555
[19] Schubert A, Babisch S, et al. Passenger and heavy vehicle collisions with pedestrians: assessment of injury mechanisms and risk. Accident Analysis & Prevention, 190, 107139, 2023. https://doi.org/10.1016/j.aap.2023.107139
[20] Schubert, A., Campolettano, E. T., Scanlon, J. M., McMurry, T. L., & Unger, T. (2024). Bridging the gap: Mechanistic-based cyclist injury risk curves using two decades of crash data. Traffic Injury Prevention, 25(sup1), S105-S115. https://doi.org/10.1080/15389588.2024.2400276
[21] Schubert, A., Campolettano, E. T., Scanlon, J. M., McMurry, T. L., & Unger, T. (2025). A mechanistic approach to modeling omnidirectional motorcyclist injury risk. Traffic Injury Prevention, 1-11. https://doi.org/10.1080/15389588.2025.2570829
[22] Neale, W. T., Danaher, D., Donaldson, A., & Smith, T. (2021). Pedestrian Impact Analysis of Side-Swipe and Minor Overlap Conditions. *SAE Technical Paper*, 01-0881. https://doi.org/10.4271/2021-01-0881
[23] Toor, A., & Araszewski, M. (2003). Theoretical vs. empirical solutions for vehicle/pedestrian collisions. *SAE transactions*, 853-865. https://doi.org/10.4271/2003-01-0883
[24] Campolettano, E. T., Scanlon, J. M., & Kusano, K. D. (2024). Characterising vulnerable road user evasive manoeuvring in real-world crashes: injury risk implications. In Proceedings of the 2024 IRCOBI Conference. Stockholm, Sweden. https://www.ircobi.org/wordpress/downloads/irc24/pdf-files/24117.pdf
[25] Campolettano, E. T., Scanlon, J. M., & Victor, T. (2023). Representative pedestrian collision injury risk distributions for a dense-urban us odd using naturalistic dash camera data. injury, 5, 10. https://index.mirasmart.com/27esv/PDFfiles/27ESV-000075.pdf
[26] Guillaume, A., Hermitte, T., Hervé, V., & Fricheteau, R. (2015). Car or ground: which causes more pedestrian injuries. In *24th International Technical Conference Enhanced Safety Vehicle*. https://www-nrd.nhtsa.dot.gov/departments/esv/24th/files/24ESV-000084.PDF
[27] New York State Department of Motor Vehicles. 2025. Motor Vehicle Crashes - Case Information: Three Year Window.
https://data.ny.gov/Transportation/Motor-Vehicle-Crashes-Case-Information-Three-Year-/e8ky-4vqe/about_data
[28] Federal Highway Administration. 2023a. Highway functional classification concepts, criteria and procedures. 2023 ed.
[29] "Vehicle Identification Number (VIN) Requirements." Code of Federal Regulations, title 49, part 565 (2025). https://www.ecfr.gov/current/title-49/subtitle-B/chapter-V/part-565.

[30] New York State Department of Transportation. 2023. Vehicle Miles of Travel 2022. https://www.dot.ny.gov/divisions/engineering/technical-services/highway-data-services
[31] Breslow NE, Day NE, Heseltine E. 1980. Statistical methods in cancer research. Lyon (France): International Agency for Research on Cancer (IARC Scientific Publications No. 32).
[32] Verma SK, Willetts JL, Corns HL, Marucci-Wellman HR, Lombardi DA, Courtney TK. 2016. Falls and fall-related injuries among community-dwelling adults in the United States. PLoS one, 11(3):e0150939. https://doi.org/10.1371/journal.pone.0150939
[33] Chen, Y. H., Scanlon, J. M., Kusano, K. D., McMurry, T. L., & Victor, T. (2024). Dynamic benchmarks: spatial and temporal alignment for ADS performance evaluation. arXiv preprint arXiv:2410.08903. https://doi.org/10.48550/arXiv.2410.08903
[34] U.S. Census Bureau, “American Community Survey 1-Year Estimates, 2024. https://www.census.gov/programs-surveys/acs.html
[35] Kusano, K.D., Scanlon, J. M., Chen, Y., McMurry, T. L., Gode, T., Victor, T. (2025). Comparison of Waymo Rider-Only Crash Rates by Crash Type to Human Benchmarks at 56.7 Million Miles. Traffic Injury Prevention, 26(sup1), S8–S20. https://doi.org/10.1080/15389588.2025.2499887
[36] Vellimana, M., & Kockelman, K. (2023). Darkness and Death in the US: Walking Distances Across the Nation by Time of Day and Time of Year. *Transport Findings*. https://doi.org/10.32866/001c.89418

APPENDIX

Table A1. Collision rates by county from Nexar data, grouped by VRU agent type.

PEDESTRIANS				
	Nexar July 2020 - June 2022 (Video-verified)			
	Mmiles	Count	IPMM	Adjusted Mmiles
Bronx County	8.69	3	0.35	12.60
Kings County	26.64	5	0.19	25.31
New York County	13.69	9	0.66	13.42
Queens County	42.34	5	0.12	31.33
Richmond County	1.87	0	0	10.00
CYCLISTS				
	Nexar July 2020 - June 2022 (Video-verified)			
	Mmiles	Count	IPMM	Adjusted Mmiles
Bronx County	8.69	1	0.12	12.60
Kings County	26.64	18	0.68	25.31
New York County	13.69	17	1.24	13.42
Queens County	42.34	6	0.14	31.33
Richmond County	1.87	0	0	10.00
MOTORCYCLISTS				
	Nexar July 2020 - June 2022 (Video-verified)			
	Mmiles	Count	IPMM	Adjusted Mmiles
Bronx County	8.69	2	0.23	12.60
Kings County	26.64	6	0.23	25.31
New York County	13.69	10	0.73	13.42
Queens County	42.34	14	0.33	31.33
Richmond County	1.87	0	0	10.00

Table A2. Collision rates by county from police-reported data, grouped by VRU agent type.

PEDESTRIANS						
	NYC Police Reports 2022 (Police-reported)			NYC Police Reports 2022 (Suspected Serious Injury+)		
	Mmiles	Count	IPMM	Mmiles	Count	IPMM
Bronx County	1294.87	1379.3	1.065	1294.87	160	0.124
Kings County	2596.89	2979.7	1.147	2596.89	290.8	0.112
New York County	1375.13	1627.6	1.184	1375.13	153.8	0.112
Queens County	3209.53	2169.1	0.676	3209.53	188.1	0.059
Richmond County	1018.17	345.9	0.340	1018.17	52.7	0.042

CYCLISTS						
	NYC Police Reports 2022 (Police-reported)			NYC Police Reports 2022 (Suspected Serious Injury+)		
	Mmiles	Count	IPMM	Mmiles	Count	IPMM
Bronx County	1294.87	490.3	0.379	1294.87	49.7	0.038
Kings County	2596.89	1751.9	0.675	2596.89	111.3	0.043
New York County	1375.13	1095.8	0.797	1375.13	75.6	0.055
Queens County	3209.53	839.5	0.262	3209.53	53.6	0.017
Richmond County	1018.17	62.4	0.061	1018.17	5.9	0.006

MOTORCYCLISTS						
	NYC Police Reports 2022 (Police-reported)			NYC Police Reports 2022 (Suspected Serious Injury+)		
	Mmiles	Count	IPMM	Mmiles	Count	IPMM
Bronx County	1294.87	333.2	0.257	1294.87	51.6	0.040
Kings County	2596.89	559.4	0.215	2596.89	62.5	0.024
New York County	1375.13	270.7	0.197	1375.13	28	0.020
Queens County	3209.53	403	0.126	3209.53	48.9	0.015
Richmond County	1018.17	44.7	0.044	1018.17	9.9	0.010