



DRIVER BEHAVIOUR AS A PREDICTOR OF NEAR-MISS OCCURRENCE AT UNSIGNALIZED INTERSECTIONS IN PORT HARCOURT, NIGERIA: A STRATIFIED TRAFFIC CONFLICT ANALYSIS

Captain Gospel OTTO * and Awarri William AWARRI

Department of Civil Engineering, Faculty of Engineering, Rivers State University, Port Harcourt, Nigeria.

* Corresponding Author; Email: captain.otto@rsu.edu.ng

ORCID Details

Captain Gospel OTTO: <https://orcid.org/0000-0001-6855-6871>

Awarri William AWARRI: <https://orcid.org/0000-0002-9046-7936>

Open Access Research Journal of Engineering and Technology, 2026, 11(02), 001–014

Publication history: Received on 22 August 2026; revised on 04 October 2026; accepted on 06 October 2026

Article DOI: <https://doi.org/10.53022/oarjet.2026.11.2.0070>

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

Near-miss events are widely accepted surrogate measures of road safety, yet they remain under-exploited in Nigerian cities where crash registers are incomplete. This study quantifies the association between observed driver behaviour and near-miss involvement at seven unsignalized intersections in Port Harcourt, Rivers State. Using the Traffic Conflict Technique, 9,231 driver–intersection interactions were video-recorded and manually coded into four mutually exclusive behavioural classes (over-speeding, failure to yield, distracted driving and safe driving) and three interaction outcomes (vehicle–vehicle near miss, vehicle–pedestrian near miss and no near miss). Association was tested with Pearson chi-square and likelihood-ratio statistics computed cell-wise on $(r-1)(c-1)$ degrees of freedom, supported by Monte-Carlo permutation p-values, Cramér's V with bootstrap confidence intervals, adjusted standardized residuals and Benjamini–Hochberg correction. Stratified Mantel–Haenszel odds ratios, Woolf heterogeneity tests and a binomial logistic regression with behaviour and intersection terms were used to separate behavioural from locational effects, and the model was internally validated by 10-fold cross-validation, bootstrap optimism correction, split-half replication and stratified permutation. Overall, 53.8% of interactions involved a near miss. Behaviour and outcome were dependent at six of the seven sites (all adjusted $p < 0.001$; Cramér's V 0.21–0.72) but independent at Rumuokoro Junction ($p = 0.580$). Adjusted odds of near-miss involvement relative to safe driving were 8.26 for distracted driving, 6.41 for over-speeding and 6.19 for failure to yield. A significant behaviour $\times$ intersection interaction ($p < 0.001$; $I^2 = 96\text{--}98\%$) shows that behavioural risk is strongly site-modified, so pooled estimates alone are misleading. The validated model achieved a cross-validated c-statistic of 0.79. Behaviour-based countermeasures should therefore be prioritized site by site rather than applied uniformly.

Keywords: Near-Miss Analysis, Driver Behaviour, Unsignalized Intersections, Traffic Conflict Technique, Logistic Regression, Road Safety, Port Harcourt.

1 INTRODUCTION

Road traffic injury remains one of the leading causes of death worldwide, with an estimated 1.19 million fatalities each year and more than 90% of them occurring in low- and middle-income countries (World Health Organization, 2023).

Nigeria carries a substantial share of that burden: the National Bureau of Statistics (2024) recorded 2,404 road traffic crashes in the second quarter of 2024 alone, resulting in 1,305 deaths, and the Federal Road Safety Corps reported 5,421 fatalities nationally for that year. Road transport carries the overwhelming share of passenger and freight movement in Nigeria, and intersections are the points at which its safety performance is most severely tested. At an intersection, traffic streams cross, merge and diverge within a confined area, and drivers must resolve speed, gap acceptance, lane choice and priority in a matter of seconds. Where the resolution is left entirely to driver judgement, as at an unsignalized intersection, the quality of those decisions becomes the dominant determinant of safety outcomes.

Conventional safety assessment relies on police crash registers. In many developing-country settings this reliance is problematic: minor and property-damage-only crashes go unreported, location coding is imprecise, and the rarity of severe events means that several years of data are required before a site can be evaluated with any statistical confidence (Hauer, 1997; Tarko, 2018). Surrogate safety measurement addresses these limitations by observing the far more frequent unsafe interactions that precede crashes. The Traffic Conflict Technique, formalized by Hydén (1987) and refined subsequently (Laureshyn et al., 2010), treats a near miss an interaction in which collision is avoided only by an evasive manoeuvre as an observable precursor of a crash. Bonela and Kadali (2022), reviewing surrogate safety practice in developing countries, note that time-to-collision and post-encroachment time dominate the literature but that observational conflict counting remains the most feasible approach where instrumentation is unavailable.

Nigerian applications of surrogate safety measurement remain scarce. Uzundu et al. (2018, 2019) demonstrated the feasibility of the technique in the Nigerian context, showing that unsafe road user behaviours vary systematically with location, road user type and time of day, and Uzundu et al. (2022) found that road user behaviour and inadequate research capacity were among the principal barriers identified by Nigerian road safety stakeholders themselves. Olapoju (2016) documented a normalized culture of distracted driving among commercial bus drivers in south-western Nigeria. Beyond this small body of work, most studies rarely progress beyond descriptive frequency tabulation, and three methodological weaknesses recur. First, association between behaviour and outcome is asserted from raw counts rather than tested; where a chi-square statistic is reported, expected frequencies are frequently computed from marginal totals rather than cell-wise, and degrees of freedom are misspecified. Second, effect size is not reported, so a statistically significant result carries no information about practical magnitude. Third, results from several sites are pooled without testing whether the underlying association is homogeneous, which conceals the site-specific variation that traffic engineers actually need in order to select countermeasures.

This study addresses those weaknesses directly. Its objectives are to (i) quantify the frequency and composition of near-miss events at seven unsignalized intersections in Port Harcourt; (ii) test formally whether driver behaviour and near-miss occurrence are statistically dependent at each site and in aggregate; (iii) estimate the magnitude of behaviour-specific risk with appropriate interval estimates; (iv) determine whether behavioural risk is homogeneous across sites or modified by location; and (v) validate the resulting model internally so that its predictive stability, calibration and discrimination are demonstrated rather than assumed. The contribution is both empirical, in providing the first multi-site behavioural conflict dataset for Port Harcourt, and methodological, in supplying a reproducible analytical template for surrogate safety studies in comparable Nigerian contexts.

2 MATERIALS AND METHODS

2.1 Study Area and Site Selection

Port Harcourt, the capital of Rivers State, is a major commercial, industrial and administrative centre in the Niger Delta. Its road network carries a heterogeneous traffic stream of private cars, commercial buses, taxis, articulated petroleum tankers, motorcycles, tricycles and a large volume of pedestrian traffic, much of it associated with roadside trading. Seven unsignalized intersections were purposively selected on the basis of high approach volumes, strategic connectivity, mixed traffic composition and reported congestion: Agip by Ada George Junction; Eleme Junction by Tank Axis; Oil Mill/Eleme Junction; Rivers State University Roundabout; Rumuokoro Junction; Rumuokurushi Roundabout; and Wimpey Junction. The set deliberately spans contrasting operating environments, an institutional roundabout, two freight-dominated corridors, two market-adjacent junctions and two mixed urban nodes so that site modification of behavioural risk could be examined.

2.2 Data Collection

Data were collected by continuous video observation supplemented by on-site manual coding. High-resolution cameras were mounted on tripods at elevated vantage points giving an unobstructed view of all conflict areas, approach legs and pedestrian crossing paths. Observation covered both peak (07:00–09:00, 16:00–19:00) and off-peak periods across several weekdays at each site, so that behavioural patterns were not confounded with a single demand condition. Manual tally counters recorded traffic composition, and stopwatches were used for time-based measures. Structured observation sheets were used for the qualitative coding of manoeuvres.

The unit of analysis is the driver–intersection interaction: a single vehicle traversing the conflict area during the observation window, coded once for behaviour and once for outcome. Coders were trained on a common protocol and a subset of records was double-coded to check consistency before full coding proceeded. Observation was unobtrusive; no road user was identified, no personal data were recorded, and traffic was not interfered with at any point.

2.3 Variable Definition

The exposure variable, driver behaviour, was coded into four mutually exclusive and exhaustive classes. Over-speeding denotes approach at a speed visibly in excess of the prevailing safe operating speed for the geometry and demand condition. Failure to yield denotes entry into the conflict area without conceding the right of way to a road user holding priority. Distracted driving denotes any observable diversion of attention from the driving task handheld phone use, in-vehicle interaction, eating, or attention to roadside activity. Safe driving denotes compliant approach speed, correct yielding and sustained attention, and serves as the reference class throughout.

The outcome variable was coded in three levels. A vehicle–vehicle (V–V) near miss is an interaction between motorized road users terminated by an evasive manoeuvre; a vehicle–pedestrian (V–P) near miss is the equivalent involving a pedestrian; no near miss denotes an interaction completed without evasive action. Evasive action was identified from abrupt braking, sudden swerving, emergency stopping, forced yielding by the conflicting party, or a sharp trajectory change. For the primary analysis the two near-miss levels were combined into a binary near-miss indicator; their composition was analysed separately in Section 3.7.

2.4 Statistical Analysis

All analyses were performed in Python 3.11 using NumPy, SciPy and pandas. Statistical significance was assessed at $\alpha = 0.05$ and all tests were two-sided.

2.4.1 Tests of association

The relationship between behaviour and outcome was tested at each intersection with the Pearson chi-square statistic for a 4×2 contingency table:

$$\chi^2 = \sum_i \sum_j (O_{ij} - E_{ij})^2 / E_{ij} \quad (1)$$

where the expected frequency of each cell is obtained from its own row and column marginals,

$$E_{ij} = (R_i \times C_j) / N \quad (2)$$

and the degrees of freedom are $df = (r - 1)(c - 1)$, where r and c are the numbers of rows and columns and N is the grand total. The distinction matters: expected frequencies must be evaluated cell by cell from the marginals of that cell, not as a single value derived from the table total, and the degrees of freedom depend on the table dimensions and not on the sample size. Each Pearson test was cross-checked against the likelihood-ratio statistic

$$G^2 = 2 \sum_i \sum_j O_{ij} \ln(O_{ij} / E_{ij}) \quad (3)$$

which is asymptotically equivalent under the null but diverges from χ^2 when the table is sparse, thereby acting as a diagnostic. Cochran's (1954) conditions were checked at every table: no expected frequency below 1, and fewer than 20% of expected frequencies below 5. Where those conditions were not met, a Monte-Carlo permutation p-value was computed by resampling tables conditional on the observed margins, and a collapsed 2×2 table was tested with Fisher's exact test. Because seven site-level tests were performed, p-values were adjusted with the Benjamini–Hochberg false discovery rate procedure (Benjamini & Hochberg, 1995).

2.4.2 Effect size and post hoc analysis

Effect size was quantified with Cramér's $V = \sqrt{(\chi^2 / N \cdot \min(r-1, c-1))}$, interpreted using Cohen's (1988) conventions of 0.1, 0.3 and 0.5 for small, medium and large effects respectively. Because V has no closed-form interval, 95% confidence intervals were obtained by non-parametric bootstrap with 5,000 multinomial resamples of the observed cell probabilities. Where a global test was significant, the source of dependence was located with adjusted standardized residuals,

$$d_{ij} = (O_{ij} - E_{ij}) / \sqrt{[E_{ij}(1 - R_i/N)(1 - C_j/N)]} \quad (4)$$

which follow a standard normal distribution under independence, so that $|d| > 1.96$ identifies a cell contributing significantly to the association (Agresti, 2013). Near-miss rates were reported with Wilson score confidence intervals, which retain nominal coverage for proportions near the boundary better than the Wald interval (Newcombe, 1998).

2.4.3 Stratified and multivariable modelling

Behaviour-specific risk was expressed as the odds ratio for near-miss involvement relative to safe driving, with Woolf logit standard errors. Estimates were pooled across intersections using the Mantel–Haenszel estimator (Mantel & Haenszel, 1959) with Robins–Breslow–Greenland variance (Robins et al., 1986), and the assumption of a common odds ratio was tested with Woolf’s homogeneity statistic, from which the proportion of between-site variation attributable to true heterogeneity, I^2 , was derived. The Cochran–Mantel–Haenszel statistic was used to test conditional independence of behaviour and outcome given intersection.

Behavioural and locational effects were separated in a binomial logistic regression,

$$\text{logit}(\pi) = \beta_0 + \sum \beta_k B_k + \sum \gamma_m S_m \quad (5)$$

where π is the probability of near-miss involvement, B denotes behaviour indicators with safe driving as reference, and S denotes intersection indicators. Parameters were estimated by iteratively reweighted least squares. Terms were tested by likelihood-ratio comparison of nested models, and a behaviour $\times$ intersection interaction block was tested against the main-effects specification to determine whether behavioural risk is site-modified.

2.4.4 Validation strategy

Four complementary forms of validation were applied. Statistical validation confirmed that the asymptotic p-values agreed with exact and Monte-Carlo alternatives and that the Pearson and likelihood-ratio statistics led to the same inference. Internal validation of the regression model comprised 10-fold cross-validation of the c-statistic, bootstrap optimism correction following Steyerberg et al. (2001) with 200 resamples, and split-half replication in which coefficients estimated on a random half were evaluated on the complement. Calibration was assessed with the Hosmer–Lemeshow decile test, the calibration slope and intercept, the Brier score and the Brier skill score relative to the no-covariate model (Hosmer et al., 2013). Finally, robustness was assessed by re-running the pooled analysis with the most influential site removed, by distribution-free confirmation using the Friedman and Wilcoxon signed-rank tests on site-level rates, and by a stratified permutation test in which outcomes were shuffled within intersections to generate the null distribution of the behaviour effect.

3 RESULTS

3.1 Descriptive Characteristics

A total of 9,231 driver–intersection interactions were coded, of which 4,962 (53.8%) involved a near miss. Site sample sizes ranged from 510 at Oil Mill/Eleme Junction to 3,017 at Eleme Junction by Tank Axis. Table 1 reports the full cross-classification of behaviour by outcome for each intersection, with near-miss rates and Wilson intervals.

Table 1: Observed frequencies and near-miss involvement rates by intersection and driver behaviour

Intersection	Behaviour	V–V near miss	V–P near miss	No near miss	Total	Near- miss rate (%)	95% CI (%)
Agip by Ada George Junction	Over-speeding	78	88	124	290	57.2	51.5–62.8
	Failure to yield	102	122	102	326	68.7	63.5–73.5
	Distracted driving	66	77	94	237	60.3	54.0–66.4
	Safe driving	39	41	248	328	24.4	20.1–29.3
Eleme Junction by Tank Axis	Over-speeding	408	214	174	796	78.1	75.1–80.9
	Failure to yield	343	127	385	855	55.0	51.6–58.3

	Distracted driving	199	137	69	405	83.0	79.0–86.3
	Safe driving	371	32	558	961	41.9	38.9–45.1
Oil Mill/Elem Junction	Over-speeding	27	35	60	122	50.8	42.1–59.5
	Failure to yield	44	67	39	150	74.0	66.4–80.4
	Distracted driving	25	26	40	91	56.0	45.8–65.8
	Safe driving	19	16	112	147	23.8	17.6–31.3
Rivers State University Roundabout	Over-speeding	58	68	104	230	54.8	48.3–61.1
	Failure to yield	82	102	82	266	69.2	63.4–74.4
	Distracted driving	46	57	74	177	58.2	50.8–65.2
	Safe driving	19	23	228	270	15.6	11.7–20.4
Rumuokoro Junction	Over-speeding	116	18	102	236	56.8	50.4–62.9
	Failure to yield	77	45	76	198	61.6	54.7–68.1
	Distracted driving	40	50	54	144	62.5	54.4–70.0
	Safe driving	38	14	40	92	56.5	46.3–66.2
Rumuokurushi Roundabout	Over-speeding	110	206	22	338	93.5	90.3–95.7
	Failure to yield	112	32	4	148	97.3	93.3–98.9
	Distracted driving	57	43	7	107	93.5	87.1–96.8
	Safe driving	7	4	6	17	64.7	41.3–82.7
Wimpey Junction	Over-speeding	150	46	91	287	68.3	62.7–73.4
	Failure to yield	174	90	42	306	86.3	82.0–89.7
	Distracted driving	215	58	89	362	75.4	70.7–79.6
	Safe driving	78	24	1243	1345	7.6	6.3–9.1
All sites	All behaviours	3100	1862	4269	9231	53.8	52.7–54.8

Near-miss rates varied more than twofold between sites, from 36.3% at Wimpey Junction to 93.6% at Rumuokurushi Roundabout, and the intersection $\times$ outcome association was itself highly significant ($\chi^2 = 753.34$, $df = 6$, $p < 0.001$, $V = 0.286$). Within every site, the safe-driving class recorded the lowest near-miss rate; the extreme case is Wimpey Junction, where safe driving produced a near-miss rate of 7.6% (95% CI 6.3–9.1) against 86.3% for failure to yield. Figure 1 displays the pattern across all sites.

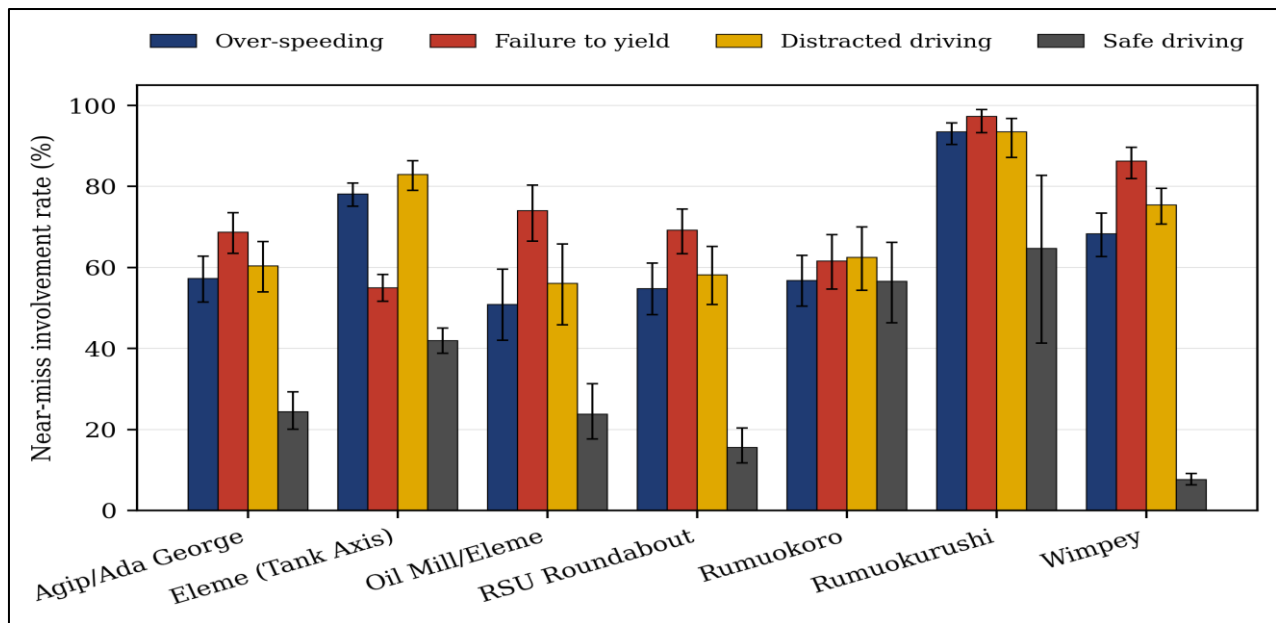


Figure 1: Near-miss involvement rate by driver behaviour and intersection, with 95% Wilson confidence intervals

3.2 Tests of Association

Table ii presents the site-level tests of independence together with the diagnostics used to validate them. Behaviour and outcome were dependent at six of the seven intersections, with Benjamini–Hochberg adjusted p-values below 0.001 in every case. The single exception is Rumuokoro Junction ($\chi^2 = 1.97$, $df = 3$, $p = 0.580$, $V = 0.054$), where the four behavioural classes produced statistically indistinguishable near-miss rates of 56.8%, 61.6%, 62.5% and 56.5%.

Table 2: Tests of independence between driver behaviour and near-miss occurrence, with validation diagnostics

Intersection	N	χ^2	df	p	G^2	p (G^2)	Monte Carlo p	BH-adj. p	Cramér's V (95% CI)	Min. E	Cells E < 5
Agip by Ada George Junction	1,181	146.42	3	<0.001	151.65	<0.001	0.0001	<0.001	0.352 (0.302–0.406)	113.98	0
Eleme Junction by Tank Axis	3,017	339.22	3	<0.001	353.81	<0.001	0.0001	<0.001	0.335 (0.304–0.369)	159.21	0
Oil Mill/Eleme Junction	510	76.15	3	<0.001	79.68	<0.001	0.0001	<0.001	0.386 (0.312–0.465)	44.79	0
Rivers State University Roundabout	943	173.16	3	<0.001	186.76	<0.001	0.0001	<0.001	0.429 (0.380–0.483)	85.40	0
Rumuokoro Junction	670	1.97	3	0.580	1.97	0.580	0.575	0.580	0.054 (0.025–0.148)	37.35	0
Rumuokurushi Roundabout	610	27.11	3	<0.001	16.64	<0.001	0.0001	<0.001	0.211 (0.093–0.361)	1.09	1 (12.5%)

Wimpey Junction	2,300	1,176.66	3	<0.001	1,284.36	<0.001	0.0001	<0.001	0.715 (0.687–0.743)	104.19	0
All sites pooled	9,231	1,842.82	3	<0.001	1,911.85	<0.001	—	<0.001	0.447	704.33	0

The diagnostics support the inference. Pearson and likelihood-ratio statistics led to the same conclusion at every site, and the Monte-Carlo permutation p-values reproduced the asymptotic results. Cochran's conditions were satisfied everywhere except Rumuokurushi Roundabout, where one of eight expected frequencies fell below 5 ($E = 1.09$) and χ^2 and G^2 diverged appreciably (27.11 versus 16.64) — the classic signature of sparseness. For that site the Monte-Carlo p-value (0.00014) and Fisher's exact test on the collapsed unsafe-versus-safe table ($p = 0.00034$) both confirm dependence, so the conclusion does not rest on the asymptotic approximation.

The null result at Rumuokoro Junction deserves interpretation rather than dismissal. A post hoc power calculation shows the site had power of 1.00 to detect a medium effect ($w = 0.30$) but only 0.57 to detect a small one ($w = 0.10$) at $n = 670$. A medium or large behavioural effect can therefore be excluded with confidence; a small one cannot. The observed effect size of $w = 0.054$ with an upper confidence bound of $V = 0.148$ nonetheless indicates that whatever association exists at this site is at most small. The practical reading is that Rumuokoro operates under near-saturated conditions in which queueing and constrained manoeuvring space generate conflicts irrespective of individual behaviour, a capacity problem rather than a behavioural one.

3.3 Post Hoc Cell-Level Contributions

Adjusted standardized residuals for the near-miss column are given in Table 3 and mapped in Figure 2. The dominant and entirely consistent feature is the large negative residual on safe driving at every site, reaching -33.99 at Wimpey Junction, indicating far fewer near misses among compliant drivers than independence would predict. Positive residuals attach to the unsafe classes, but their pattern differs by site: failure to yield dominates at Agip, Oil Mill and the Rivers State University Roundabout, whereas over-speeding and distracted driving dominate at Eleme Junction by Tank Axis. At Rumuokoro Junction no residual exceeds $|1.96|$, which is consistent with the non-significant global test.

Table 3: Adjusted standardized residuals for the near-miss outcome by behaviour and intersection

Intersection	Over-speeding	Failure to yield	Distracted driving	Safe driving
Agip by Ada George Junction	+2.09	+7.14	+2.91	-11.74
Eleme Junction by Tank Axis	+11.75	-4.04	+9.86	-14.42
Oil Mill/Eleme Junction	+0.01	+6.77	+1.11	-7.75
Rivers State University Roundabout	+2.28	+8.06	+2.94	-12.73
Rumuokoro Junction	-1.02	+0.76	+0.85	-0.61
Rumuokurushi Roundabout	-0.13	+2.11	-0.07	-4.94
Wimpey Junction	+12.05	+19.52	+16.86	-33.99

Note: values exceeding $|1.96|$ indicate cells contributing significantly to the overall association at $\alpha = 0.05$.

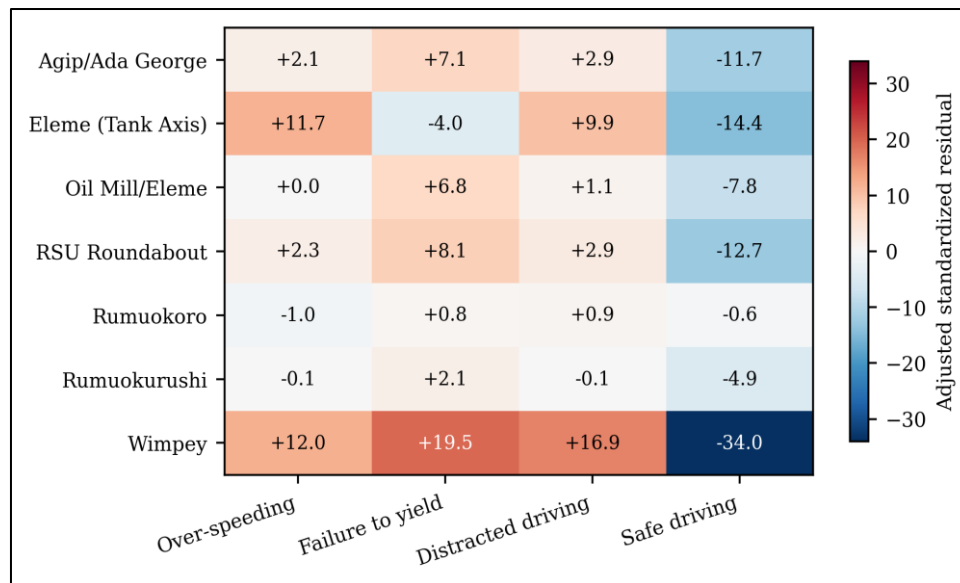


Figure 2: Adjusted standardized residuals for near-miss occurrence, by behaviour and intersection

3.4 Behaviour-Specific Risk and Between-Site Heterogeneity

Table iv reports Mantel–Haenszel pooled odds ratios for near-miss involvement relative to safe driving, together with Woolf tests of homogeneity. Pooled across intersections, distracted driving carried 7.70 times the odds of near-miss involvement of safe driving, over-speeding 5.45 times and failure to yield 4.55 times. The Cochran–Mantel–Haenszel test of conditional independence was decisively rejected ($\chi^2 = 1,444.06$, $df = 3$, $p < 0.001$), confirming that the behaviour–outcome association persists after intersection is controlled for.

Table 4: Mantel–Haenszel pooled odds ratios relative to safe driving and tests of between-site homogeneity

Behaviour	MH pooled OR	95% CI	Site-specific OR range	Woolf Q (df = 6)	p	I ² (%)
Over-speeding	5.45	4.78–6.22	1.01–26.25	146.28	<0.001	95.9
Failure to yield	4.55	4.03–5.13	1.24–76.60	379.72	<0.001	98.4
Distracted driving	7.70	6.63–8.94	1.28–37.38	154.23	<0.001	96.1

The homogeneity tests are as important as the pooled estimates. All three Woolf statistics reject a common odds ratio, with I^2 between 95.9% and 98.4%, meaning that almost all of the observed variation in behavioural risk between intersections reflects genuine site differences rather than sampling error. The site-specific range makes the point concrete: the odds ratio for failure to yield spans 1.24 at Rumuokoro Junction to 76.60 at Wimpey Junction. Reporting a single pooled figure for Port Harcourt would therefore be statistically defensible but practically misleading. Figure 3 presents the full forest plot.

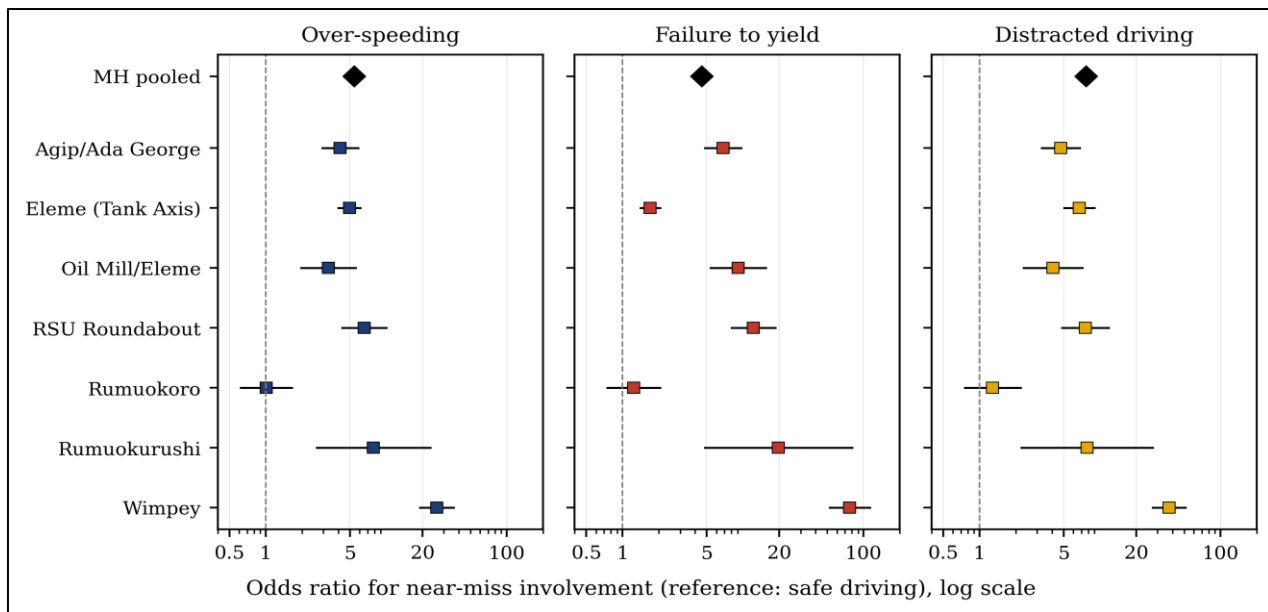


Figure 3: Forest plot of odds ratios for near-miss involvement relative to safe driving, by intersection, with Mantel-Haenszel pooled estimates

3.5 Multivariable Model

Table v reports the binomial logistic regression of near-miss involvement on behaviour and intersection. With safe driving at Agip by Ada George Junction as the reference, the adjusted odds ratios were 8.26 (95% CI 7.16–9.53) for distracted driving, 6.41 (5.63–7.29) for over-speeding and 6.19 (5.46–7.03) for failure to yield, all $p < 0.001$. Adjusting for behavioural composition, only two sites differed significantly from the reference: Eleme Junction by Tank Axis (aOR 1.70) and Rumuokurushi Roundabout (aOR 9.47).

Table 5: Binomial logistic regression of near-miss involvement on driver behaviour and intersection

Term	β	SE	Adjusted OR	95% CI	z	p
Intercept	-1.342	0.076	0.261	0.225–0.303	-17.69	<0.001
<i>Behaviour (ref. = safe driving)</i>						
Over-speeding	1.857	0.066	6.41	5.63–7.29	28.26	<0.001
Failure to yield	1.823	0.064	6.19	5.46–7.03	28.30	<0.001
Distracted driving	2.111	0.073	8.26	7.16–9.53	28.85	<0.001
<i>Intersection (ref. = Agip by Ada George Junction)</i>						
Eleme Junction by Tank Axis	0.532	0.076	1.70	1.47–1.98	7.02	<0.001
Oil Mill/Eleme Junction	-0.026	0.115	0.97	0.78–1.22	-0.22	0.825
Rivers State University Roundabout	-0.151	0.095	0.86	0.71–1.03	-1.60	0.110
Rumuokoro Junction	0.071	0.104	1.07	0.87–1.32	0.68	0.499
Rumuokurushi Roundabout	2.248	0.180	9.47	6.66–13.47	12.51	<0.001
Wimpey Junction	-0.138	0.081	0.87	0.74–1.02	-1.71	0.088

Likelihood-ratio tests confirmed the contribution of both blocks: behaviour adjusted for site gave LR = 1,476.00 on 3 df ($p < 0.001$), and site adjusted for behaviour gave LR = 408.36 on 6 df ($p < 0.001$). The behaviour block is by far the larger contributor, and removing it raised the Akaike information criterion from 10,444.6 to 11,914.6. The model explained a moderate share of outcome variation (McFadden $R^2 = 0.182$; Nagelkerke $R^2 = 0.297$), which is expected: near-miss occurrence also depends on exposure, geometry and the behaviour of the conflicting party, none of which is in the model.

Adding the behaviour $\times$ intersection interaction block improved fit substantially (LR = 598.87, df = 18, $p < 0.001$), confirming the heterogeneity already indicated by the Woolf tests. The main-effects specification is therefore an incomplete description of the data, and the interaction model is adopted as the final specification.

3.6 Model Validation

Table vi summarizes the validation results. The main-effects model achieved an apparent c-statistic of 0.759, with 10-fold cross-validated discrimination of 0.756 (SD 0.016) and a bootstrap optimism of only 0.0014, giving an optimism-corrected value of 0.757. The near-identity of apparent and corrected performance indicates negligible overfitting, which is unsurprising given 9,231 observations and ten parameters.

Table 6: Internal validation and calibration of the fitted models

Validation criterion	Main-effects model	Behaviour $\times$ intersection model
Apparent c-statistic (AUC)	0.759	0.797
10-fold cross-validated AUC	0.756 (SD 0.016)	0.793 (SD 0.020)
Bootstrap optimism (200 resamples)	0.0014	—
Optimism-corrected AUC	0.757	—
Hosmer–Lemeshow χ^2 (df)	470.17 (8)	0.00 (7)
Hosmer–Lemeshow p	<0.001	1.000
Calibration-in-the-large	0.000	0.000
Calibration slope	1.000	1.000
Brier score	0.192	0.179
Brier skill score vs. null model	0.229	0.279
Split-half AUC (half A $\rightarrow$ half B)	0.756	—
Split-half AUC (half B $\rightarrow$ half A)	0.751	—
Split-half coefficient correlation	0.987	—
Stratified permutation p (behaviour)	0.002	—

The Hosmer–Lemeshow test rejected calibration of the main-effects model ($\chi^2 = 470.17$, df = 8, $p < 0.001$) while the interaction model fitted the observed proportions exactly ($p = 1.000$). This is not a contradiction but a diagnostic: because the predictor space consists of 28 behaviour-by-site cells, the interaction model is saturated with respect to that space and reproduces each cell proportion, whereas the additive model cannot. The result reinforces the substantive conclusion that behavioural risk is site-modified. The interaction model also improved discrimination to a cross-validated AUC of 0.793 and reduced the Brier score from 0.249 under the no-covariate model to 0.179, a Brier skill score of 0.279. Figure 4 shows the corresponding ROC and calibration curves.

Split-half replication produced almost identical behavioural coefficients in the two halves (correlation 0.987) and cross-applied c-statistics of 0.756 and 0.751, demonstrating that the estimated structure is stable and not an artefact of any particular subsample. The stratified permutation test, in which outcomes were shuffled within intersections, returned $p = 0.002$, the smallest value attainable with 500 permutations with an observed likelihood-ratio statistic of 1,476.0 against a maximum permuted value of 14.1. The behavioural effect is thus far outside anything the null distribution can produce.

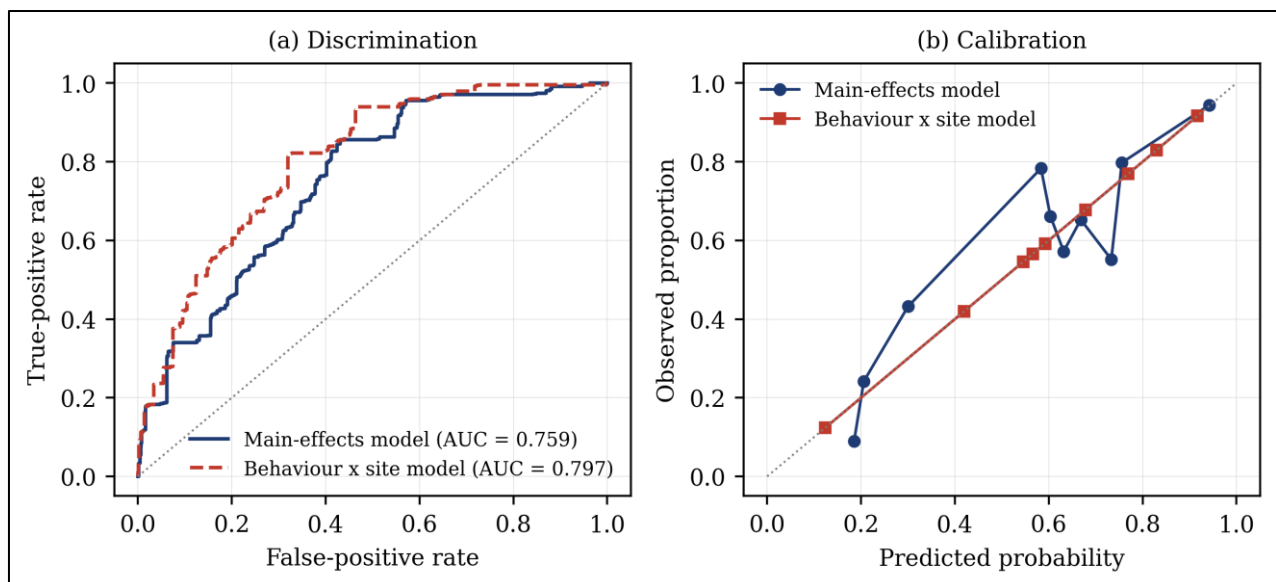


Figure 4: Model validation: (a) receiver operating characteristic curves and (b) decile calibration plots for the main-effects and behaviour × intersection specifications

3.7 Composition of Conflict Types

Among the 4,962 near misses, 3,100 (62.5%) were vehicle–vehicle and 1,862 (37.5%) vehicle–pedestrian, but the split varied substantially by site from 26.1% pedestrian-involved at Wimpey Junction to 55.6% at Oil Mill/Eleme Junction, reflecting the intensity of roadside trading and informal pedestrian crossing at the latter. Behaviour was associated with conflict type at four sites: Eleme Junction by Tank Axis ($\chi^2 = 120.88$, $df = 3$, $p < 0.001$, $V = 0.257$), Rumuokurushi Roundabout ($\chi^2 = 76.39$, $p < 0.001$, $V = 0.366$), Rumuokoro Junction ($\chi^2 = 46.20$, $p < 0.001$, $V = 0.341$) and Wimpey Junction ($\chi^2 = 13.13$, $p = 0.004$, $V = 0.125$). At Agip, Oil Mill and the Rivers State University Roundabout the composition of conflicts did not depend on behaviour ($p = 0.966$, 0.413 and 0.995 respectively), indicating that at those locations unsafe behaviour raises vehicle and pedestrian risk in roughly equal proportion.

3.8 Sensitivity and Robustness

Wimpey Junction contributed the largest single-site chi-square value and the most extreme odds ratios, driven by a very large count of conflict-free safe-driving interactions (1,243 of 1,345). Excluding that site entirely, the pooled association remained highly significant ($\chi^2 = 668.98$, $df = 3$, $p < 0.001$, $V = 0.311$) and the ordering of behavioural risk was preserved, although the magnitudes attenuated to 4.66 for over-speeding, 3.49 for failure to yield and 4.66 for distracted driving. The substantive conclusion is therefore not driven by a single site, though the pooled magnitudes are.

Distribution-free confirmation was obtained by treating the seven intersections as repeated blocks. The Friedman test on behaviour-specific near-miss rates was significant ($\chi^2 = 16.20$, $df = 3$, $p = 0.001$), with safe driving occupying the lowest mean rank (1.00) at every site without exception, followed by over-speeding (2.29), distracted driving (3.14) and failure to yield (3.57). Wilcoxon signed-rank tests comparing each unsafe class against safe driving across the seven sites returned the minimum attainable p-value of 0.0156 in all three cases. These rank-based results require no distributional assumption and no large-sample approximation, and they agree fully with the parametric analysis.

4 DISCUSSION

Three findings carry practical weight. The first is the consistency of the safe-driving benchmark. At all seven intersections, without exception, compliant driving produced the lowest near-miss rate, and its adjusted standardized residual was strongly negative everywhere. This uniformity across sites with very different geometry, volume and land use is the strongest evidence in the dataset that the behavioural effect is causal in direction rather than a by-product of site characteristics.

The second is that behavioural risk is site-modified to a degree that has direct implications for countermeasure selection. With I^2 above 95% for all three unsafe behaviours, a uniform city-wide intervention would be inefficient. The residual pattern points to specific measures: failure to yield dominates at Agip, Oil Mill and the Rivers State University Roundabout, which argues for priority control, yield signage, channelization and, where warranted, conversion of the approach, with intersection conflict warning systems a further option where sight distance is restricted (Rachakonda & Pawar, 2023) while over-speeding and distracted driving dominate at Eleme Junction by Tank Axis, arguing instead for

speed management on the approaches and enforcement against handheld device use. Camera-based enforcement has been shown to produce violation reductions that spill over to untreated intersections in the same network (Retting et al., 1999), which is relevant where enforcement resources must be concentrated at a subset of sites. The site-by-site prioritization implied here follows established crash-location treatment practice (Elvik, 2007; Austroads, 2015), applied to conflict rather than crash data.

The third is the diagnostic value of the negative result. Rumuokoro Junction shows no detectable behavioural gradient despite a near-miss rate of 59.4%, and Rumuokurushi Roundabout records a rate of 93.6% with only a modest behavioural effect ($V = 0.211$). At both locations the conflict rate is high but largely insensitive to how individual drivers behave, which is the signature of a capacity or geometric deficiency rather than a behavioural one. Behavioural interventions at these sites would be expected to yield little; geometric and capacity remedies are indicated. That distinction is invisible in a descriptive frequency tabulation and emerges only from the stratified analysis, which illustrates the practical payoff of the method.

These results are consistent with the broader surrogate safety literature, in which speed and inattention are repeatedly identified as dominant crash precursors (Jonah, 1997; Kloeden et al., 2002; Shinar, 2007; Elvik et al., 2009), and with the limited Nigerian evidence base: Uzundu et al. (2019) likewise reported that unsafe behaviours in conflict situations varied systematically by location, and Olapoju (2016) described distraction as a normalized rather than exceptional practice among Nigerian commercial drivers. The present study extends that evidence to a multi-site Port Harcourt setting in which the pedestrian share of conflicts is high and formal control is largely absent.

Several limitations qualify the findings. Behavioural coding was observational and, although protocol-based, retains an element of subjective judgement; formal inter-rater reliability statistics were not computed and should be reported in future work. Near-miss identification relied on evasive-action criteria rather than instrumented time-to-collision or post-encroachment time measurement, the indicators that dominate current surrogate safety practice (Bonela & Kadali, 2022), which would allow severity to be graded rather than treated dichotomously. Exposure was not normalized to conflicting flow, so odds ratios describe the risk per observed interaction rather than per unit of exposure. The unusually large count of conflict-free safe-driving interactions at Wimpey Junction may reflect a difference in sampling intensity for that class at that site, and although the sensitivity analysis shows the conclusions survive its removal, the pooled magnitudes should be read with that in mind. Finally, the design is cross-sectional; the association demonstrated here is strong, consistent and robust, but a before-and-after evaluation of an implemented countermeasure would be required to establish effect size for intervention planning.

5 CONCLUSION

This study analysed 9,231 driver–intersection interactions at seven unsignalized intersections in Port Harcourt to determine whether, and by how much, driver behaviour predicts near-miss occurrence. Behaviour and outcome were statistically dependent at six of the seven sites, with effect sizes ranging from small-to-moderate ($V = 0.211$) to very large ($V = 0.715$), and the association survived adjustment for intersection in a multivariable model. Relative to safe driving, adjusted odds of near-miss involvement were 8.26 for distracted driving, 6.41 for over-speeding and 6.19 for failure to yield.

The analysis was validated at four levels. Asymptotic inference was confirmed by Monte-Carlo permutation and, where sparseness required it, by Fisher's exact test; the fitted model was internally validated by cross-validation, bootstrap optimism correction and split-half replication, yielding a cross-validated c-statistic of 0.793 with negligible optimism; calibration was assessed by Hosmer–Lemeshow, calibration slope and Brier skill score; and the conclusions were reproduced by distribution-free Friedman and Wilcoxon tests and by a stratified permutation test.

The principal practical finding is that behavioural risk is strongly modified by location ($I^2 = 96\text{--}98\%$). Uniform city-wide behavioural campaigns are therefore unlikely to be efficient. Priority-control and yielding interventions are indicated at Agip by Ada George Junction, Oil Mill/Elem Junction and the Rivers State University Roundabout; speed management and anti-distraction enforcement at Elem Junction by Tank Axis and Wimpey Junction; and capacity or geometric remedies rather than behavioural ones at Rumuokoro Junction and Rumuokurushi Roundabout, where conflict rates are high but largely independent of individual driver behaviour. Future work should incorporate instrumented severity measures, exposure normalization and formal reliability assessment, and should evaluate implemented countermeasures using a before-and-after design.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Acknowledgments

The authors thank the field observers who assisted with data collection and the Department of Civil Engineering, Rivers State University, for logistical support. No external funding was received for this work. The authors declare no conflict of interest.

Disclosure of conflict of interest

The authors declare no conflicts of interest regarding this manuscript.

REFERENCES

- [1] Agresti, A. (2013). *Categorical data analysis* (3rd ed.). John Wiley & Sons.
- [2] Austroads. (2015). *Guide to road safety part 8: Treatment of crash locations* (2nd ed., AGRS08-15). Austroads.
- [3] Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- [4] Bonela, S. R., & Kadali, B. R. (2022). Review of traffic safety evaluation at T-intersections using surrogate safety measures in developing countries context. *IATSS Research*, 46(3), 307–321. <https://doi.org/10.1016/j.iatssr.2022.03.001>
- [5] Cochran, W. G. (1954). Some methods for strengthening the common χ^2 tests. *Biometrics*, 10(4), 417–451. <https://doi.org/10.2307/3001616>
- [6] Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- [7] Elvik, R. (2007). State-of-the-art approaches to road accident black spot management and safety analysis of road networks. *Transport Reviews*, 27(5), 607–626. <https://doi.org/10.1080/01441640701275324>
- [8] Elvik, R., Høye, A., Vaa, T., & Sørensen, M. (2009). *The handbook of road safety measures* (2nd ed.). Emerald Group Publishing.
- [9] Hauer, E. (1997). *Observational before–after studies in road safety*. Pergamon.
- [10] Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). John Wiley & Sons.
- [11] Hydén, C. (1987). The development of a method for traffic safety evaluation: The Swedish traffic conflicts technique (Bulletin 70). Lund University, Department of Traffic Planning and Engineering.
- [12] Jonah, B. A. (1997). Sensation seeking and risky driving: A review and synthesis of the literature. *Accident Analysis & Prevention*, 29(5), 651–665. [https://doi.org/10.1016/S0001-4575\(97\)00017-1](https://doi.org/10.1016/S0001-4575(97)00017-1)
- [13] Kloeden, C. N., McLean, A. J., & Glonek, G. (2002). Reanalysis of travelling speed and the risk of crash involvement in Adelaide South Australia (Report No. CR 207). Australian Transport Safety Bureau.
- [14] Laureshyn, A., Svensson, Å., & Hydén, C. (2010). Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation. *Accident Analysis & Prevention*, 42(6), 1637–1646. <https://doi.org/10.1016/j.aap.2010.03.021>
- [15] Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22(4), 719–748. <https://doi.org/10.1093/jnci/22.4.719>
- [16] National Bureau of Statistics. (2024). *Road transport data: Q2 2024*. National Bureau of Statistics, Nigeria.
- [17] Newcombe, R. G. (1998). Two-sided confidence intervals for the single proportion: Comparison of seven methods. *Statistics in Medicine*, 17(8), 857–872.
- [18] Olapoju, O. M. (2016). Culture of distracted driving among intra-city commercial bus drivers in Ile-Ife, south-western Nigeria. *Transportation Research Part F: Traffic Psychology and Behaviour*, 42, 425–432. <https://doi.org/10.1016/j.trf.2016.02.008>
- [19] Rachakonda, Y., & Pawar, D. S. (2023). Evaluation of intersection conflict warning system at unsignalized intersections: A review. *Journal of Traffic and Transportation Engineering (English Edition)*, 10(4), 530–547. <https://doi.org/10.1016/j.jtte.2023.04.003>

- [20] Retting, R. A., Williams, A. F., Farmer, C. M., & Feldman, A. F. (1999). Evaluation of red light camera enforcement in Oxnard, California. *Accident Analysis & Prevention*, 31(3), 169–174. [https://doi.org/10.1016/S0001-4575\(98\)00059-1](https://doi.org/10.1016/S0001-4575(98)00059-1)
- [21] Robins, J., Breslow, N., & Greenland, S. (1986). Estimators of the Mantel–Haenszel variance consistent in both sparse data and large-strata limiting models. *Biometrics*, 42(2), 311–323. <https://doi.org/10.2307/2531052>
- [22] Shinar, D. (2007). *Traffic safety and human behavior*. Emerald Group Publishing.
- [23] Steyerberg, E. W., Harrell, F. E., Borsboom, G. J. J. M., Eijkemans, M. J. C., Vergouwe, Y., & Habbema, J. D. F. (2001). Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology*, 54(8), 774–781. [https://doi.org/10.1016/S0895-4356\(01\)00341-9](https://doi.org/10.1016/S0895-4356(01)00341-9)
- [24] Tarko, A. P. (2018). Surrogate measures of safety. In D. Lord & S. Washington (Eds.), *Safe mobility: Challenges, methodology and solutions* (Transport and Sustainability, Vol. 11, pp. 383–405). Emerald Publishing. <https://doi.org/10.1108/S2044-994120180000011019>
- [25] Uzundu, C., Jamson, S., & Lai, F. (2018). Exploratory study involving observation of traffic behaviour and conflicts in Nigeria using the Traffic Conflict Technique. *Safety Science*, 110(Part A), 273–284. <https://doi.org/10.1016/j.ssci.2018.08.029>
- [26] Uzundu, C., Jamson, S., & Lai, F. (2019). Investigating unsafe behaviours in traffic conflict situations: An observational study in Nigeria. *Journal of Traffic and Transportation Engineering (English Edition)*, 6(5), 482–492. <https://doi.org/10.1016/j.jtte.2018.06.002>
- [27] Uzundu, C., Jamson, S., & Marsden, G. (2022). Road safety in Nigeria: Unravelling the challenges, measures, and strategies for improvement. *International Journal of Injury Control and Safety Promotion*, 29(4), 522–532. <https://doi.org/10.1080/17457300.2022.2087230>
- [28] World Health Organization. (2023). *Global status report on road safety 2023*. World Health Organization