

GIS-based spatial analysis of fatal road traffic incidents in Medellin, Colombia: Road segment burden, facility buffer zone proximity, and spatial autocorrelation

Noorah Osama Abid^{1*}, Abeer Salim Alwan Alnajjar², Shaymaa Khazaal Shnaishel³

^{1,2,3} College of Administration and Economics, University of Baghdad, Iraq

*Corresponding author E-mail: noura.os@coadec.uobaghdad.edu.iq

Received Apr. 3, 2026
Revised Oct. 2, 2026
Accepted Oct. 20, 2026
Online Oct. 30, 2026

Abstract

This study characterized the spatial distribution of fatal road traffic incidents in Medellín, Colombia, from January 2008 to August 2025. Of 702,540 records, 2,762 were fatal; 2,507 spatially eligible incidents were retained. Proximity to 2,065 education facilities and 78 strict hospital and emergency-care facilities was evaluated using 100-, 250-, and 500-m buffers. Each incident was assigned to the nearest segment of an all-road network (68,308 segments); segment burden and road-class distribution were assessed. Global and Local Moran's I were calculated using k-nearest-neighbor weights (k=8) on segment centroids, with sensitivity analyses for repeated coordinates and alternative networks. 19.55%, 68.77%, and 94.22% of incidents were within 100, 250, and 500 m of an education facility, versus 3.91%, 14.84%, and 36.14% for hospital and emergency-care facilities. Only 1,331 segments (1.95%) received an incident. Residential roads had the largest burden, but several major-road and link classes held larger shares of incidents than of network length. Global Moran's I was positive but weak (0.0245), whereas the unique-location hospital-proximity sensitivity analysis showed strong autocorrelation (0.7624–0.8875). Local Moran's I identified 140 High–High and 1,866 Low–High segments. High-burden rankings were stable across network definitions. The findings provide a screening and prioritization framework, not causal or exposure-adjusted risk estimates.

© The Author 2026.
Published by ARDA.

Keywords: GIS, Fatal road traffic incidents, Buffer zone analysis, Road segment analysis, Spatial autocorrelation, Facility proximity, OpenStreetMap, Colombia

1. Introduction

Road traffic injuries remain a major global public-health and transport challenge. The World Health Organization estimates that approximately 1.19 million people die annually as a result of road traffic crashes. Road traffic injuries are also the leading cause of death among people aged 5–29 years, while pedestrians, cyclists, and motorcyclists account for more than half of global road traffic deaths [1]. These patterns demonstrate that road safety cannot be addressed solely through changes in individual behavior. They require transport systems that recognize the inevitability of human error and reduce the probability that such errors will result in death or serious injury [2]. This principle is the basis of the Safe System approach and the international target for a substantial reduction in road traffic deaths and injuries by 2030 [1, 2].

This work is licensed under a [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>) that allows others to share and adapt the material for any purpose (even commercially), in any medium with an acknowledgement of the work's authorship and initial publication in this journal.



Fatal road traffic incidents within cities are seldom evenly distributed over the road network. The spatial pattern of them may reflect differences in road function, traffic movement, speed environment, intersection configuration, pedestrian activity, land use, and distribution of transport infrastructure [3]. Analyses based on large administrative areas only may hide important variations between individual roads and road segments. A geographic information system (GIS) is capable of tracking the geographic locations of incidents and connecting them to road geometries, facility locations, and adjacent spatial units [4]. Analysis of road segments is especially useful as it allows for identification of particular parts of a road network that accrue incident burden and may require detailed local investigation.

Therefore, the research on spatial networks has highlighted the need to represent crashes on road segments and not only as events happening in large administrative regions [5]. Spatial autocorrelation is another way to check whether similar incident burden is near each other spatially. Global Moran's I describes the overall direction and strength of spatial association for the entire study network, while Local Moran's I , a local indicator of spatial association, identifies specific locations that contribute to local clusters or spatial contrasts [6, 7]. This distinction matters because weak spatial dependence across the network can coexist with limited local clusters of high burden.

Therefore, the combination of observed road segment counts and global and local spatial statistics can be more informative for spatial screening than descriptive mapping or a single global statistic alone [6, 7]. The spatial relationship between fatal incidents and nearby education facilities also has planning potential. The layer of analytical education facilities consists of schools, kindergartens, colleges, and universities. The travel activities around them may vary depending on the type of facility. Therefore, proximity to the combined education-facility layer was assumed to indicate general spatial co-location. School-specific interpretation: The fact that an incident occurred within a school buffer does not mean that pupils were involved or that the school was involved in the incident.

Detailed evaluation requires information on school entrances, pedestrian routes, crossing facilities, operating speeds, traffic volumes, and the micro-level street environment [8]. Similarly, straight-line distance to a hospital or emergency-care facility can describe geographic separation but cannot measure ambulance response, route accessibility, congestion, treatment capability, or the facility that actually received an injured road user. Network distance and travel-time models require additional transport and emergency-service data and address a different question from initial geographic-proximity analysis [9, 10].

Medellín is the capital of the Department of Antioquia and is located in the Aburrá Valley in the central Andean region of Colombia [11]. Its urban setting provides a suitable context for examining fatal road traffic incidents at a detailed geographic scale. These considerations are particularly relevant because the municipal government has formally adopted the Vision Zero approach to road safety. This policy treats fatal and serious road traffic injuries as preventable outcomes and places responsibility on the broader transport system to accommodate human vulnerability and error [12].

Official road-traffic incident data are published by Medellín's mobility authority through Colombia's open-data portal [13]. The database used in this study is from 1 January 2008 to 31 August 2025 [14, 15]. Each source record is a reported road traffic incident, not an individual fatality. The outcome studied in this paper is therefore called a fatal road traffic incident and is an incident-level outcome and not a person-level count of deaths. When these records are combined with mapped education facilities, strict hospital and emergency-care facilities, and a prepared road network, several related aspects of the spatial distribution can be examined within one analytical framework.

The underlying incident database has been described before in a dedicated data descriptor, reporting on the structure, the spatial completeness, descriptive summaries, and an example of a grouped logistic regression of fatal outcomes [14], while the curated dataset is available in the public domain through Mendeley Data [15]. Previous applications that have addressed other analytical questions have used the same underlying Medellín

incident records. Sánchez Corredor et al. quantified the social burden and spatiotemporal concentration of fatal incidents using Years of Potential Life Lost and spatial-clustering methods, with spatial robustness analyzes including Moran's I, Getis-Ord Gi, and kernel density estimation [16]. Afterward, Arango Uribe et al. explored the spatiotemporal patterns, pedestrian vulnerability, and spatial inequalities from the governance perspective by using regression and spatial analysis, including Moran's I and Getis-Ord Gi [17]. Most recently, López-López et al. used the incident registry to develop and evaluate standard ML models for rare event fatality prediction under time-aware validation and calibration [18].

As opposed to those earlier applications, this study incorporates fatal-incident locations with three external GIS layers: mapped education facilities, a rigorously defined set of hospital and emergency-care facilities, and the OpenStreetMap road network. The analysis quantified facility proximity at 100-, 250-, and 500-m thresholds, assigned fatal incidents to their nearest road segments, and measured road segment burden and road class distribution. It also assessed the sensitivity of hospital proximity spatial autocorrelation to repeated coordinate locations, tested the robustness of high-burden road segment rankings under alternative road network definitions, and examined Global Moran's I and Local Moran's I of road-segment fatal-incident burden.

Accordingly, this study aimed to characterize the spatial distribution of fatal road traffic incidents in Medellín from January 2008 to August 2025. The specific objectives were to: quantify proximity to mapped education facilities and strict hospital and emergency-care facilities using cumulative 100-, 250-, and 500-m thresholds; assign each spatially eligible fatal incident to its nearest mapped road segment; describe road segment fatal-incident burden and its distribution across OpenStreetMap (OSM) road classes; compare raw and road-length-normalized road class measures; evaluate global and local spatial autocorrelation in road segment burden; assess the sensitivity of spatial autocorrelation in strict hospital and emergency-care proximity to repeated coordinate locations; and assess the robustness of high-burden road segment rankings to alternative road network definitions. The analysis was designed as a spatial screening and prioritization framework rather than as a causal or traffic-exposure-adjusted assessment of road-safety risk.

2. Materials and methods

2.1. Study design and setting

This study used a retrospective geospatial observational design. Administrative road traffic incident records were integrated with spatial data on education facilities, hospitals and emergency-care facilities, and the road network. No exposure level, road condition, or facility location was assigned or manipulated by the study. The analysis used incidents and infrastructure that had already been recorded or mapped.

The study area was Medellín, Antioquia, Colombia. The road traffic incident database covered the period from 1 January 2008 to 31 August 2025 [14,15]. The 2014 municipal boundary supplied by the Alcaldía de Medellín (Medellín Mayor's Office) was retained as the primary municipal reference boundary for mapping and the boundary audit [19]. The Medellín municipality polygon (municipality code 05001) from the 2017 Municipios de Antioquia [Municipalities of Antioquia] shapefile issued by the Departamento Administrativo Nacional de Estadística (DANE; National Administrative Department of Statistics) was retained only as a secondary boundary-reference layer and was not used to exclude incidents or calculate distances [20].

GIS was used to represent the study entities as spatial objects. A spatial object combines an attribute record with a geometry. Point, line, and polygon geometries were handled within the simple features framework. Simple features provide a standardized representation of vector spatial data [21]. Fatal road traffic incidents and education facilities were represented as points. Strict hospital and emergency-care facilities were represented by their original point, polygon, or multipolygon geometries. Roads were represented as line or multiline geometries, with each retained geometry treated as one analytical road segment [21].

The spatial units were kept distinct throughout the analysis. The first unit was the fatal-incident point, which was used for education-facility-proximity analysis, hospital and emergency-care proximity analysis, and

nearest-road assignment. The second unit was the road segment, which was used to calculate the number of fatal incidents assigned to each segment and to examine global and local spatial autocorrelation in road segment fatal-incident burden.

Throughout the analysis, $F = \{p_i\}_{i=1}^{n_F}$ denotes the final fatal-incident point set, where $n_F = 2507$. Unless otherwise stated, $d(a, b)$ denotes the shortest planar distance between projected spatial objects a and b , expressed in meters. The education-facility, strict hospital and emergency-care, and road segment sets are defined in Sections 2.5.1–2.5.3.

2.2. Road-traffic incident data

The source database consisted of 702540 administrative road traffic incident records reported in Medellín. The records originated from the Secretaría de Movilidad de Medellín (Medellín Mobility Secretariat). The structure and georeferenced content of the database are documented in a published data descriptor and an associated public repository. The repository describes more than 700000 incident reports and identifies 2762 fatal-incident records [14,15].

Each row represented one reported road traffic incident. The source database was therefore incident-based and was not a person-level mortality register. This distinction is important because the present study counted fatal incident records. It did not estimate the number of individual decedents when more than one death may have occurred in a single incident [14,15].

The source data contained 13 principal fields describing the record identifier, calendar year, incident date, incident time, incident class, incident severity, interpolated address, latitude, longitude, planning-comuna code and name, and planning-neighborhood code and name. A comuna is an urban administrative district that is used in Medellín's municipal planning system. The original Spanish field names and their English labels are reported in Table 1 [14–15]. In the analytical code, field names were standardized to prevent problems in handling the data caused by accents, spaces, capitalization, or punctuation. This standardization only altered the coding labels, not the recorded values.

The definition of the analytical outcome was especially important for two source fields. CLASE_INCIDENTE recorded the type or mechanism of the incident (e.g., collision, pedestrian strike, occupant fall, rollover, fire, or another incident class), and GRAVEDAD_INCIDENTE recorded the severity of the incident. We only employed GRAVEDAD_INCIDENTE to define fatal incidents, not CLASE_INCIDENTE to infer if a death occurred. Therefore, a collision, pedestrian strike, or rollover was considered fatal only if GRAVEDAD_INCIDENTE was coded as MUERTO (“dead” in Spanish) [14,15].

The spatial coordinates of the published repository are expressed in decimal-degree longitude and latitude relative to WGS84 [15]. WGS84 stands for World Geodetic System 1984, and is a geographic coordinate reference system for expressing locations on the Earth in terms of longitude and latitude. Administrative-area labels were not used as proxies for coordinates. The absence of a planning-comuna or planning-neighborhood tag did not necessarily exclude an incident if a valid longitude-latitude pair was available. The main location variables in the GIS analyzes were geographic coordinates. The structure and missingness of the source variables are summarized in Table 1.

Table 1. Structure, definitions, and missingness of variables in the original road traffic incident dataset.

Original variable	English label	Statistical type	Measurement level	Unit / Format	What it actually represents	Missing n, (%)
LLAVE	Unique record identifier	Qualitative identifier	Nominal	None	Unique administrative key identifying the record	0(0)

Original variable	English label	Statistical type	Measurement level	Unit / Format	What it actually represents	Missing n, (%)
ANIO	Calendar year	Quantitative discrete / temporal	Ordered temporal	Year	Calendar year in which the incident was recorded	0(0)
FECHA_INCIDENTE	Incident date	Temporal variable	Date / interval-time	Calendar Date	Calendar date on which the incident occurred	0(0)
HORA_INCIDENTE	Incident time	Temporal variable	Time-of-day	HH: MM	Clock time at which the incident occurred	0(0)
CLASE_INCIDENTE	Incident class	Qualitative	Nominal	None	Type or mechanism of the road traffic incident	0(0)
GRAVEDAD_INCIDENTE	Incident severity	Qualitative	Ordinal	None	Recorded severity category of the incident	0(0)
DIRECCION_INTERPOLADA	Interpolated address	Qualitative / text	Nominal free text	None	Interpolated street-address description	11,100 (1.580)
Latitud	Latitude	Quantitative continuous	Continuous	Decimal Degrees	Geographic latitude of the incident location	44,785 (6.375)
Longitud	Longitude	Quantitative continuous	Continuous	Decimal Degrees	Geographic longitude of the incident location	44,785 (6.375)
Codigo_Comuna_Planeacion	Planning comuna code	Qualitative code	Nominal	Administrative code	Code identifying the planning comuna	28,154 (4.007)
Comuna_Planeacion	Planning comuna name	Qualitative	Nominal	None	Name of the planning comuna	28,154 (4.007)
Codigo_Barrio_Planeacion	Planning neighborhood code	Qualitative code	Nominal	Administrative code	Code identifying the planning neighborhood	46,732 (6.652)
Barrio_Planeacion	Planning neighborhood name	Qualitative	Nominal	None	Name of the planning neighborhood	46,732 (6.652)

2.3. Fatal-incident definition and spatial eligibility

Severity text was cleaned before the fatal-incident filter was applied. Leading and trailing spaces were removed, and the text was converted to uppercase.

For incident i , fatal status was defined as

$$\delta_i = 1\{\text{GRAVEDAD_INCIDENTE}_i = \text{"MUERTO"}\},$$

where $1\{\cdot\}$ is the indicator function. The indicator equals 1 when the stated condition is true and 0 otherwise.

Equivalently,

$$\delta_i = \begin{cases} 1, & \text{if incident } i \text{ was recorded as MUERTO,} \\ 0, & \text{otherwise.} \end{cases}$$

The severity filter identified 2762 fatal incident records among the 702540 source records. The same fatal-incident count is documented in the public repository [15].

Coordinate values were cleaned after fatal-record selection. Decimal commas were replaced with decimal points when required, and the cleaned longitude and latitude strings were converted to numeric values.

The spatial eligibility was assessed in four steps. First, the records required numeric longitude and latitude values, so 119 records with missing or nonnumeric coordinates were removed. Second, coordinates were required to be between longitude -75.700000 and -75.450000 and latitude 6.100000 and 6.450000 , and 84 records were excluded at this stage. Third, the remaining 2,559 points were compared to the entire municipal boundary. Of these, 2,497 points were located within or on the boundary, and 62 points were located outside the boundary. Finally, the 62 outside-boundary points were checked against the official road-hierarchy layer; 10 were retained, and 52 were excluded, resulting in the final spatial sample of 2507 fatal incidents [22]. The absence of planning-comuna or planning-neighborhood labels did not exclude records that met the criteria based on these coordinates.

After coordinate cleaning and spatial screening, the final spatial sample contained. $n_F = 2507$

fatal incident records. These 2507 records formed the set. F and were used in all subsequent proximity and nearest-road assignment procedures.

The spatial completeness of the fatal subset was calculated as

$$C_F = 100 \times \frac{n_F}{n_{F,0}},$$

where $n_{F,0} = 2762$ is the number of fatal incident records before spatial screening and $n_F = 2507$ is the final spatial sample. The quantity C_F stands for a coordinate-completeness measure and is not a mortality rate.

The full source database was used to define and audit the source frame, whereas the subsequent GIS analyses used the final 2507 spatially eligible fatal incidents.

2.4. Geometry creation, coordinate reference system, and analytical extent

The cleaned geographic coordinates were converted to point geometries. The longitude-latitude pair in the record for fatal incident i was written as

$$(\lambda_i, \phi_i),$$

where λ_i denotes longitude and ϕ_i denotes latitude. This coordinate pair was converted to the point geometry. p_i .

The source coordinates were given in angular units referenced to WGS84 [15]. Geographic longitude and latitude are convenient for recording global positions, but the current study required distances and buffer radii in meters.

The fatal-incident points and the external infrastructure layers were then transformed to MAGNA-SIRGAS / Colombia Bogota zone (EPSG:3116). The EPSG codes are standardized identifiers in the EPSG Geodetic Parameter Dataset for coordinate reference systems (CRS) and other geodetic definitions. The projected CRS used in this study is EPSG:3116. MAGNA, Marco Geocéntrico Nacional de Referencia, and SIRGAS, Sistema de Referencia Geocéntrico para las Américas; MAGNA-SIRGAS is the Colombian geodetic reference framework [23].

After projection, the fatal incident i was represented as

$$x_i = (E_i, N_i),$$

where E_i is the easting coordinate and N_i is the northing coordinate. Both coordinates were expressed in meters. For two point locations a and b , with projected coordinates $x_a = (E_a, N_a)$ and $x_b = (E_b, N_b)$, the planar Euclidean distance was calculated as

$$d(a, b) = \sqrt{(E_a - E_b)^2 + (N_a - N_b)^2}.$$

For a fatal-incident point p_i and a road segment r_ℓ , $d(p_i, r_\ell)$ denotes the shortest planar distance from p_i to any position on the line geometry of r_ℓ . It does not denote the distance from the incident point to the midpoint or centroid of the road segment.

All distance-based and buffer-based operations were conducted after transformation to EPSG:3116. The projected workflow used planar geometry. Simple-features operations were used for coordinate transformation and spatial-vector processing [21, 24].

A fatal-incident-centered analytical extent was defined from the final fatal-incident points.

Let

$$E_{\min} = \min_{1 \leq i \leq n_F} E_i, E_{\max} = \max_{1 \leq i \leq n_F} E_i,$$

and

$$N_{\min} = \min_{1 \leq i \leq n_F} N_i, N_{\max} = \max_{1 \leq i \leq n_F} N_i$$

denote the minimum and maximum projected coordinates of the final fatal-incident point set $\mathcal{F}$. The bounding rectangle of the fatal-incident points was defined as

$$\mathcal{E}_F = [E_{\min}, E_{\max}] \times [N_{\min}, N_{\max}].$$

For local spatial processing, this rectangle was expanded by 3000 m in each direction:

$$\mathcal{E}_{F,3000} = [E_{\min} - 3000, E_{\max} + 3000] \times [N_{\min} - 3000, N_{\max} + 3000].$$

The expanded rectangle established the local analytical extent for preparing and inspecting the spatial datasets. The 3000-m expansion was merely a local processing margin, and not a reported proximity threshold or the legal or administrative boundary of Medellín.

The national road data were processed using a larger processing window. The bounding extent was derived by merging local reference layers and expanded by 5000 m before spatially filtering the national road layer. This extra margin reduced the chance of edge truncation prior to the nearest-road assignment. The health-facility layer was also retained within 5000 m of the local analytical extent. This margin ensured that potentially relevant facilities, just outside the core analytical extent, were available during nearest-facility calculations. It was a data extraction safeguard, not one of the reported 100-, 250-, or 500-m buffer thresholds. In Figure 1, we present the final points of fatal road traffic incidents over the total municipal boundary of Medellín.

The municipal boundary was obtained from the official Límite Municipio de Medellín shapefile, a 2014 layer of administrative boundaries published by the Alcaldía de Medellín [19]. The municipal boundaries and road hierarchy were audited as described in Section 2.3. The municipal polygon was kept in the geographic context of Figure 1, but was not used as an automatic exclusion boundary. Unless otherwise specified, all distances in this work are the shortest planar Euclidean distances between the relevant projected geometries.

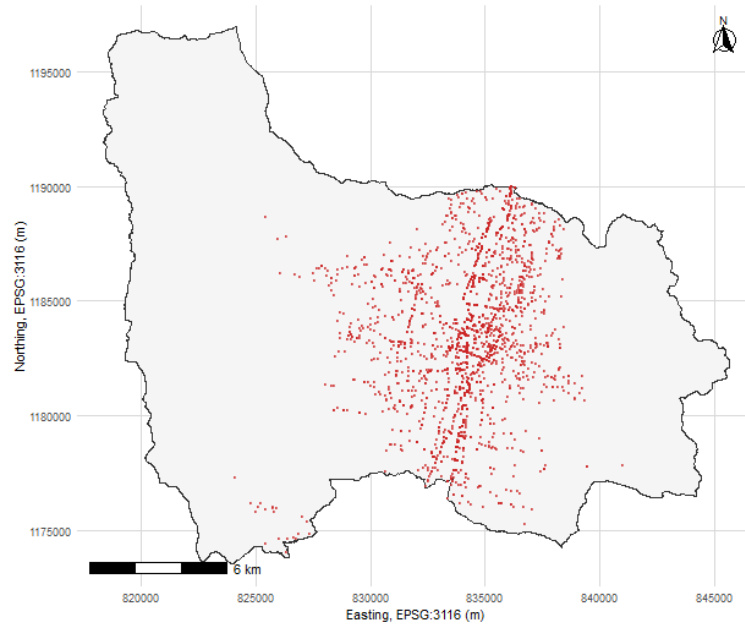


Figure 1. Study area and spatial distribution of the 2507 georeferenced fatal road traffic incident points overlaid on the full municipal boundary of Medellín, Colombia. The boundary was obtained from the Limite Municipio de Medellín 2014 shapefile. Ten incident points located outside the municipal-boundary polygon were retained after the road-hierarchy audit described in Section 2.3. All mapped data were projected to MAGNA-SIRGAS / Colombia Bogotá zone (EPSG:3116).

2.5. External spatial datasets and analytical geometries

Three external spatial datasets were prepared for the analysis: education facilities, strict hospital and emergency-care facilities, and roads. Each dataset was cleaned, spatially restricted to the relevant local processing extent, transformed to EPSG:3116, and converted to the geometry type required for the corresponding analysis.

2.5.1. Education facilities

The education-facility layer was obtained from the Humanitarian OpenStreetMap Team (HOT) through the Humanitarian Data Exchange (HDX), using the Colombia Education Facilities (OpenStreetMap Export) dataset. This HOT/HDX dataset is a dedicated spatial extract of education-related OpenStreetMap (OSM) features [25].

The national source dataset contained 45667 education-facility features before local spatial preparation. The HOT/HDX layer included OpenStreetMap features identified as kindergartens, schools, colleges, or universities through the amenity or building attributes. Empty or unusable geometries were removed, and the remaining geometries were checked for validity.

The geometries of the retained education facilities were represented as point locations for the point-based proximity analysis. The generated point layer was spatially limited to the processing extent of Medellín and converted to EPSG:3116. The final analytic education-facility layer was 2065 point locations. These points were used only as mapped facility locations for calculating nearest-education-facility-distance and buffer zone. Because the dataset was not standardized for information on enrollment, student population, capacity of the facility, time of operation, or pedestrian activity, these characteristics were not included in the proximity measures.

$$E = \{e_j\}_{j=1}^{n_E}, n_E = 2065.$$

2.5.2. Strict hospital and emergency-care facilities

The health-facility layer was obtained from HOT using the HDX, Health Facilities of Colombia (OpenStreetMap Export) dataset. The HDX dataset uses OSM as the source and has health related features under

the health, healthcare, and amenity attributes [26]. The raw national health-facility layer had 9546 features. The source consisted of heterogeneous health-related facilities with different functions. A broad health-facility category was therefore not used as a proxy for hospitals or emergency care services.

A stricter subset of hospitals and emergency care was created using the available health-related attributes and descriptive facility names. Facility names retained matched terms indicating hospital or emergency-care functions, including hospital, clínica (clinic), unidad hospitalaria (hospital unit), unidad intermedia (intermediate-care unit), urgencias (emergency services), and policlínica (polyclinic). The rule did not cover health-related features that did not qualify as a hospital or emergency care facility; these included many pharmacies, dental facilities, routine outpatient services, and other non-emergency health facilities.

This dataset is the strict hospital and emergency-care layer. The term “strict” refers to the narrower feature-selection rule. This does not mean that all retained facilities had the same clinical capacity, trauma capability, emergency departments, operating hours, or service levels.

Points, polygons and multipolygons were the retained health-facility geometries. The original geometries were retained for the proximity calculation. Distances from fatal-incident points to point facilities were thus calculated to the mapped point locations, while distances to polygon or multipolygon facilities were calculated as the shortest planar distance to the corresponding facility geometry. When a fatal-incident point intersected a retained polygonal facility geometry, a distance of zero was assigned. Representative internal points were generated for cartographic display only and are not used in calculating the reported nearest-facility distances. All geometries were converted to MAGNA-SIRGAS / Colombia Bogotá zone (EPSG:3116) prior to distance calculation. The final strict hospital and emergency care layer had 78 facilities: 20 point features and 58 polygon or multipolygon features.

The facility set was defined as:

$$\mathcal{H} = \{g_h\}_{h=1}^{n_H}, n_H = 78.$$

where g_h is an original mapped geometry of strict hospital or emergency-care facility h , represented as a point, polygon, or multipolygon.

The spatial selection was based on the pre-defined analytical extent around the locations of the fatal incident and not only on the Medellín municipal boundary. Hence, at the last analytical layer, no clipping was applied to the municipal boundary. Of the 78 retained facilities, 57 were within or intersected the municipal boundary, and 21 were outside it. Facilities outside the administrative boundary were retained because they may still represent the nearest mapped hospital or emergency-care facility to a fatal-incident point.

2.5.3. Road network

The road layer was obtained from HOT through HDX, using the Colombia Roads (OpenStreetMap Export) dataset. The HOT Export Tool converts selected OSM features into standard GIS formats, including GeoPackage [27].

There were 1184448 features in the national source road dataset. A spatial read filter was applied before doing the full local processing to limit the amount of national road data brought in for the Medellín analysis. Road geometries were kept, both lines and multiline roads. The selected road features were projected to EPSG:3116 and clipped to the wider local processing extent as described in Section 2.4. Invalid or unusable geometries were dealt with during spatial preparation prior to road-length measurement and nearest-road assignment.

The final all-road analytical network consisted of road segments:

$$n_R = 68308$$

In this study, a road segment denotes one retained line or multiline geometry in the prepared road network.

The road segment set was defined as

$$\mathcal{R} = \{r_\ell\}_{\ell=1}^{n_R}.$$

The OpenStreetMap highway attribute was retained as the descriptive road class field [27]. It was subsequently used for road-class summaries and for the alternative road network definitions evaluated in the sensitivity analyses.

The term all-road analytical network is used throughout the primary analysis to distinguish the complete retained road layer from the OpenStreetMap road class named primary. The primary analysis was based on all 68308 retained road segments rather than only on roads classified as primary.

2.6. Education-facility proximity and buffer zone analysis

Education-facility proximity was measured as the shortest projected distance between each fatal-incident point and the final education-facility-point layer described in Section 2.5.1.

For a fatal incident i , the nearest-education-facility distance was defined as

$$d_i^{(E)} = \min_{1 \leq j \leq n_E} d(p_i, e_j),$$

where p_i stands for the location of fatal incident i , e_j denotes an education facility j , $n_E = 2065$ is the number of education facilities in the final analytical layer, and $d(p_i, e_j)$ is the planar Euclidean distance between the two projected point geometries. The superscript E identifies quantities associated with the education-facility layer. The value $d_i^{(E)}$ therefore represents the distance, in meters, from a fatal incident i to the nearest mapped education facility.

A buffer zone was defined as the set of locations lying within a specified distance of a spatial feature. For all education facilities combined, the education-facility buffer zone of radius r was defined as

$$\mathcal{Z}_E(r) = \bigcup_{j=1}^{n_E} \{x: d(x, e_j) \leq r\}.$$

The union operator indicates that the individual buffers surrounding all education facilities were treated as one combined buffer zone. Overlapping education-facility buffers were therefore not counted as separate areas.

Three buffer radii were evaluated:

$$r \in \{100, 250, 500\} \text{ m.}$$

For fatal incident i , membership in the education-facility buffer zone of radius r was defined as

$$I_i^{(E)}(r) = 1 \left\{ d_i^{(E)} \leq r \right\},$$

where $1\{\cdot\}$ is the indicator function. Thus,

$$I_i^{(E)}(r) = 1$$

when a fatal incident i was located within distance r of at least one education facility, and

$$I_i^{(E)}(r) = 0.$$

Because $d_i^{(E)}$ is the minimum distance from fatal incident i to any education facility, the condition

$$d_i^{(E)} \leq r$$

is equivalent to

$$p_i \in \mathcal{Z}_E(r).$$

The buffer classifications were verified against the corresponding nearest-education-facility distances. The number of fatal incidents located within a radius r of at least one education facility was calculated as

$$N_E(r) = \sum_{i=1}^{n_F} I_i^{(E)}(r),$$

where $n_F = 2507$ is the number of fatal incidents in the final spatial sample.

The corresponding percentage was calculated as

$$P_E(r) = 100 \times \frac{N_E(r)}{n_F}.$$

The three buffer thresholds were cumulative rather than mutually exclusive. A fatal incident located within 100 m of an education facility was therefore also included in the 250-m and 500-m categories.

Education-facility proximity was measured using straight-line planar distance in EPSG:3116. The analysis did not estimate walking distance, road network distance, travel time, pedestrian-route accessibility, or barriers between a fatal-incident point and an education facility. The resultant measures were suggestive of spatial co-location with mapped education facilities and were not construed as evidence that an education facility caused, contributed to, or was otherwise directly related to a specific fatal incident.

2.7. Strict hospital and emergency-care proximity

Nearest-facility distance was calculated from every fatal-incident point to the final strict hospital and emergency-care geometry layer described in Section 2.5.2. All 78 retained facilities were included in the calculation, irrespective of their location relative to the municipal boundary.

For fatal incident i , the nearest-facility distance was defined as

$$d_i^{(H)} = \min_{1 \leq h \leq n_H} d(p_i, g_h),$$

where p_i denotes the fatal-incident point i , g_h denotes the original mapped geometry of the facility h , and $d(p_i, g_h)$ is the planar Euclidean distance from the incident point to that geometry in EPSG:3116. For polygon and multipolygon facilities, this was the shortest distance to the mapped facility geometry and was zero when the incident point intersected that geometry.

The superscript H identifies quantities associated with the strict hospital and emergency-care layer. The value $d_i^{(H)}$ therefore represents the distance, in meters, between a fatal incident i and the nearest mapped strict hospital or emergency-care facility.

For a specified radius r , the combined hospital and emergency-care buffer zone was defined as

$$\mathcal{Z}_H(r) = \bigcup_{h=1}^{n_H} \{x: d(x, g_h) \leq r\}.$$

Three cumulative distance thresholds were evaluated:

$$r \in \{100, 250, 500\} \text{ m}.$$

For fatal incident i , membership in the hospital and emergency-care buffer zone was defined as

$$I_i^{(H)}(r) = 1 \{d_i^{(H)} \leq r\}.$$

Thus,

$$I_i^{(H)}(r) = 1$$

when a fatal incident i was located within distance r of at least one strict hospital or emergency-care facility, and

$$I_i^{(H)}(r) = 0.$$

The number of fatal incidents located within the threshold r was calculated as

$$N_H(r) = \sum_{i=1}^{n_F} I_i^{(H)}(r),$$

and the corresponding percentage was calculated as

$$P_H(r) = 100 \times \frac{N_H(r)}{n_F}.$$

The three distance thresholds were cumulative rather than mutually exclusive. A fatal incident located within 100 m of a strict hospital or emergency-care facility was therefore also included in the 250-m and 500-m categories.

Hospital and emergency-care proximity represented straight-line spatial co-location. It did not represent road network travel distance, ambulance-routing distance, travel time, emergency-response time, or realized access to medical services. The nearest mapped facility was not assumed to be the facility that received or treated the people involved in a particular fatal incident. No clinical, ambulance, referral, or patient-transport records were linked to the road-traffic incident data.

2.8. All-road network preparation for the primary analysis

The all-road network described in Section 2.5.3 was retained as the primary analytical road layer and included all locally prepared road classes, including carriageways, links, service roads, paths, cycleways, and other mapped line features. Retaining the all-road network avoided restricting nearest-road assignment to a selected group of road classes. Excluding mapped classes before assignment could have forced a fatal-incident point to be assigned to a more distant eligible road segment even when a closer mapped segment was present.

The primary analysis therefore used the all-road analytical network. $n_R = 68308$ road segments. OpenStreetMap road class values were retained as descriptive attributes for subsequent road class summaries and sensitivity analyses. A road class label did not determine whether a segment was eligible to receive a fatal-incident assignment in the primary analysis.

The final road geometries were transformed to EPSG:3116 and retained as line or multiline geometries. Road segment geometry was used directly for nearest-road assignment and road-length calculation. Segment centroids were created only for the spatial-neighborhood structure used in the global and local spatial-autocorrelation analyses described in Sections 2.11 and 2.12.

For a road segment r_ℓ , let $L_\ell^{(m)}$ denote its mapped length in meters. The corresponding length in kilometers was calculated as

$$L_\ell^{(km)} = \frac{L_\ell^{(m)}}{1000}.$$

Road lengths were calculated from the projected EPSG:3116 geometries using the simple-features implementation in R [21,24].

2.9. Nearest-road assignment and road segment fatal-incident burden

Each fatal-incident point was assigned to the nearest road segment in the all-road analytical network. Distance was measured from the incident point to the complete line geometry of each candidate road segment rather than to its midpoint or centroid.

For fatal incident i , the nearest-road assignment function was defined as

$$a_i = \operatorname{argmin}_{\ell \in \{1, \dots, n_R\}} d(p_i, r_\ell)$$

where a_i identifies the index of the road segment with the smallest planar distance to the fatal incident point p_i .

Nearest-road assignment was used as a geometric linkage procedure. It did not reconstruct vehicle trajectory, route choice, direction of travel, legal road access, police-confirmed carriageway membership, or network travel distance.

After assignment, the fatal-incident burden of road segment r_ℓ was defined as

$$y_\ell = \sum_{i=1}^{n_F} 1\{a_i = \ell\},$$

where $1\{\cdot\}$ is the indicator function and y_ℓ is the number of fatal incident records assigned to road segment r_ℓ .

The road segment fatal-incident burden was the outcome used in the subsequent road class summaries and spatial-autocorrelation analyses. It was a count of fatal incident records and not a person-level death count. Because the source database was incident-based, the assigned value did not estimate the number of individual decedents associated with each incident.

Each fatal incident was assigned to one road segment. The assignment therefore satisfied the conservation identity:

$$\sum_{\ell=1}^{n_R} y_\ell = n_F = 2507.$$

The left-hand side represents the total number of assigned fatal incidents after aggregation across all road segments, whereas the right-hand side represents the total number of fatal incidents in the final spatial sample.

For descriptive comparison among individual road segments, a road-length-normalized measure was calculated as

$$\rho_\ell = \frac{y_\ell}{L_\ell^{(km)}}, \quad L_\ell^{(km)} > 0,$$

where $L_\ell^{(km)}$ is the mapped length of road segment r_ℓ , expressed in kilometers. The quantity ρ_ℓ represents assigned fatal incidents per kilometer of mapped segment.

This road-length-normalized measure was descriptive and was not interpreted as an exposure-adjusted road-safety risk estimate. Traffic volume, vehicle kilometers traveled, pedestrian activity, cyclist activity, person-time at risk, operating speed, and other exposure denominators were not available for the road segments.

2.10. Road class distribution measures

After nearest-road assignment, the distribution of assigned fatal incidents was summarized across the 30 OpenStreetMap road classes represented in the all-road analytical network. The retained OpenStreetMap highway value was used as a descriptive road class attribute after assignment. Road class did not determine which road segment was selected as the nearest segment and did not exclude any retained segment from receiving a fatal-incident assignment in the primary analysis.

Let $\mathcal{C} = \{c_1, \dots, c_{30}\}$, $c_\ell \in \mathcal{C}$ denote the set of retained road classes and let c_ℓ denote the road class assigned to segment r_ℓ . For road class c , the corresponding set of road segments was defined as

$$\mathcal{R}_c = \{r_\ell \in \mathcal{R} : c_\ell = c\}.$$

The total number of mapped road segments belonging to class c was $n_c = |\mathcal{R}_c|$.

The raw fatal-incident burden of road class c was calculated as

$$Y_c = \sum_{r_\ell \in \mathcal{R}_c} y_\ell,$$

where Y_c is the total number of fatal incident records assigned to all mapped segments belonging to road class c .

The class-specific counts satisfied the conservation identity:

$$\sum_{c \in \mathcal{C}} Y_c = \sum_{\ell=1}^{n_R} y_\ell = n_F = 2507$$

Thus, aggregation by road class preserved the total number of fatal incidents assigned in the nearest-road analysis.

The total mapped length of road class c was calculated as

$$L_c = \sum_{r_\ell \in \mathcal{R}_c} L_\ell^{(km)},$$

where L_c is expressed in kilometers.

The road-length-normalized fatal-incident measure for road class c was calculated as

$$\rho_c = \frac{Y_c}{L_c}, L_c > 0.$$

The quantity ρ_c represents the number of assigned fatal incidents per kilometer of mapped road in class c . It was calculated from the aggregate fatal-incident count and aggregate mapped length of the class. It was not calculated as the arithmetic mean of the segment-specific values. ρ_ℓ .

The value Y_c measures the total number of assigned fatal incidents in road class c . To compare this burden with the class's representation in the mapped road network, two class-specific percentage shares were calculated.

The share of all assigned fatal incidents associated with road class c was

$$P_c^{(\text{fatal})} = 100 \times \frac{Y_c}{\sum_{u \in \mathcal{C}} Y_u} = 100 \times \frac{Y_c}{n_F}.$$

The corresponding share of total mapped road network length was

$$P_c^{(\text{length})} = 100 \times \frac{L_c}{\sum_{u \in \mathcal{C}} L_u}.$$

A fatal-share-to-length-share ratio was then calculated as

$$Q_c = \frac{P_c^{(\text{fatal})}}{P_c^{(\text{length})}}.$$

A value of $Q_c > 1$ indicated that road class c accounted for a larger share of assigned fatal incidents than its share of the total mapped road network length. A value of $Q_c < 1$ indicated that its share of assigned fatal incidents was smaller than its share of network length, whereas $Q_c = 1$ represented equality between the two shares.

The relationship between $P_c^{(\text{fatal})}$ and $P_c^{(\text{length})}$ was also displayed graphically. Each road class was represented by one point, with mapped-network-length share on the horizontal axis and assigned-fatal-incident share on the vertical axis. A dashed 1:1 reference line represented equality between the two shares. Road classes above the reference line had $Q_c > 1$, whereas classes below the line had $Q_c < 1$.

For concise presentation in the main manuscript, road classes accounting individually for at least 1% of all assigned fatal incidents were reported separately. The remaining classes were combined into a single summary category.

Road length was used in the present analysis as an available spatial denominator because class-specific traffic-volume and mobility-exposure measures were unavailable. However, road length does not account for differences in traffic intensity, road function, road capacity, speed environment, or the number of road users exposed to risk. Traffic- or mobility-based denominators are preferable for estimating exposure-adjusted road-safety risk when suitable data are available [28, 29]. Accordingly, Y_c , ρ_c , $P_c^{(\text{fatal})}$, $P_c^{(\text{length})}$, and Q_c were interpreted as descriptive road class distribution measures and not as exposure-adjusted risk estimates.

2.11. Global spatial autocorrelation of road segment fatal-incident burden

Global Moran's I was used to examine whether road segments with similar fatal-incident burden values tended to occur near one another within the primary all-road analytical network. Moran's I is a global measure of spatial autocorrelation whose value depends on both the observed spatial variable and the definition of the spatial weights structure [6,30]. Different spatial-weight specifications can influence the estimated spatial association and its statistical interpretation [30].

The spatial variable analyzed for road segments r_ℓ was

$$y_\ell, \ell = 1, \dots, n_R,$$

where y_ℓ is the number of fatal incident records assigned to that segment.

The mean road segment burden was

$$\bar{y} = \frac{1}{n_R} \sum_{\ell=1}^{n_R} y_\ell.$$

A single representative coordinate was required to construct the spatial-neighborhood structure among road segments. Segment centroids were calculated from the projected EPSG:3116 road geometries [21,24]. The centroid coordinates were checked for exact duplication before k-nearest-neighbor construction; no duplicate-coordinate rows were identified.

Let

$$z_\ell = (E_\ell^{\text{cent}}, N_\ell^{\text{cent}})$$

denote the projected centroid coordinate of road segment r_ℓ .

For each road segment, the eight nearest road-segment centroids were identified ($k = 8$). This centroid-based k-nearest-neighbor (k-NN) structure represented geographical proximity between road segments. It did not represent direct road network connectivity, shared intersections, traffic-flow direction, legal movement between segments, or topological adjacency. Let $b_{\ell m}$ denote the initial binary neighbor weight between road segments r_ℓ and r_m :

$$b_{\ell m} = \begin{cases} 1, & \text{if the centroid of } r_m \text{ was among the eight nearest centroids to } r_\ell, \\ 0, & \text{otherwise.} \end{cases}$$

Self-neighbor relations were excluded:

$$b_{\ell\ell} = 0.$$

The resulting neighbor structure was converted to a spatial-weights object. Binary weights were row-standardized using the W coding style [30]. The row-standardized weight was

$$w_{\ell m} = \frac{b_{\ell m}}{\sum_{q=1}^{n_R} b_{\ell q}},$$

when the corresponding row contained at least one neighbor. Consequently,

$$\sum_{m=1}^{n_R} w_{\ell m} = 1.$$

Let $W = [w_{\ell m}]$ denote the resulting row-standardized spatial-weights matrix.

Global Moran's I was defined as

$$I = \frac{n_R}{S_0} \frac{\sum_{\ell=1}^{n_R} \sum_{m=1}^{n_R} w_{\ell m} (y_\ell - \bar{y})(y_m - \bar{y})}{\sum_{\ell=1}^{n_R} (y_\ell - \bar{y})^2},$$

Where:

$$S_0 = \sum_{\ell=1}^{n_R} \sum_{m=1}^{n_R} w_{\ell m}.$$

The term $w_{\ell m}$ is the spatial weight linking road segments r_ℓ and r_m . A positive value of Moran's I indicates that road segments with similar fatal-incident burden values tend to occur near one another under the specified spatial-weights structure. A negative value indicates a tendency for neighboring road segments to have dissimilar burden values, whereas a value near zero indicates limited global spatial association under the specified null model [6,30].

Global Moran's I was calculated using the original row-standardized $k = 8$ spatial weights.

Statistical significance for the primary road segment analysis was evaluated using a Monte Carlo permutation test with 9999 permutations. A one-sided test with fixed random seed 20260629 was used to check for positive spatial autocorrelation. Due to a large number of road segments, strong statistical evidence could be observed even when the magnitude of spatial dependence was small. Therefore, the magnitude of Moran's I was interpreted separately from the statistical significance.

2.12. Local Moran analysis

Global Moran's I provides one summary statistic for the complete road network. It does not identify the individual road segments that contribute to local spatial clustering or local spatial contrast. Local Moran's I, a local indicator of spatial association, was therefore used to examine segment-specific spatial patterns [7].

For road segment r_ℓ , Local Moran's I was defined as

$$I_\ell = \frac{y_\ell - \bar{y}}{m_2} \sum_{m=1}^{n_R} w_{\ell m} (y_m - \bar{y}),$$

Where:

$$m_2 = \frac{1}{n_R} \sum_{\ell=1}^{n_R} (y_\ell - \bar{y})^2.$$

The first term describes the deviation of the focal road segment's fatal-incident burden from the network mean. The weighted term describes the burden pattern among the segment's spatial neighbors.

The primary Local Moran analysis used the same centroid-based $k = 8$ nearest-neighbor structure and row-standardized W weights as the primary Global Moran analysis. Local Moran statistics and analytical local p-values, denoted p_l , were calculated using this spatial-weights structure [7,30].

For cluster classification, the road segment fatal-incident burden was first standardized across the complete all-road network. For road segment r_l , the standardized burden was defined as

$$y_l^* = \frac{y_l - \bar{y}}{s_y}$$

where s_y denotes the standard deviation of road segment fatal-incident burden across the complete analytical network. Positive values of y_l^* therefore indicated above-average burden, whereas negative values indicated below-average burden.

The spatial lag of the standardized burden for road segment r_l was calculated as

$$(Wy^*)_l = \sum_{m=1}^{n_R} w_{lm} y_m^*$$

using the same row-standardized W spatial weights employed in the Local Moran calculation. Because the lag was calculated from the centered and standardized burden values, zero represented the reference value for cluster-quadrant classification.

nominal local significance threshold of $p_l \leq 0.05$ was applied. Significant road segments were classified according to the signs of the standardized focal burden y_l^* and its spatial lag $(Wy^*)_l$. A High-High pattern was

defined by $y_l^* > 0$ and $(Wy^*)_l > 0$, a Low–Low pattern was defined by $y_l^* < 0$ and $(Wy^*)_l < 0$, a High–Low pattern was defined by $y_l^* > 0$ and $(Wy^*)_l < 0$, a Low–High pattern was defined by $y_l^* < 0$ and $(Wy^*)_l > 0$.

Road segments that did not meet the nominal local significance threshold were labeled Not significant. No multiplicity correction was applied to the local p values. The Local Moran classifications were therefore treated as exploratory spatial evidence. They were used to identify and map local spatial patterns requiring interpretation, but they did not replace the observed road segment fatal-incident burden y_l and were not interpreted as causal effects.

2.13. Sensitivity and robustness analyses

Two separate sensitivity procedures were performed. The first examined point-based spatial autocorrelation in strict hospital and emergency-care proximity after exact duplicate coordinates were collapsed into unique spatial locations. The second examined whether the ranking of high-burden road segments depended on the road classes retained as eligible candidates during nearest-road assignment.

2.13.1. Unique-location sensitivity analysis for strict hospital and emergency-care proximity

Several fatal-incident records shared exactly the same recorded coordinates. Repeated coordinates can create zero-distance ties and duplicate spatial positions during point-based nearest-neighbor construction. Fatal-incident records were therefore grouped by exact coordinate pair for this sensitivity analysis.

Let

$$\mathcal{U} = \{u_q\}_{q=1}^{n_U}$$

denote the set of unique fatal-incident locations. The unique-location layer contained $n_U = 1568$ spatial sites.

For each unique location u_q , the number of fatal-incident records sharing that coordinate was

$$m_q = \sum_{i=1}^{n_F} 1\{p_i = u_q\}.$$

The distance to the nearest strict hospital or emergency-care facility was aggregated at each unique location using the arithmetic mean:

$$\tilde{d}_q^{(H)} = \frac{1}{m_q} \sum_{i:p_i=u_q} d_i^{(H)}.$$

Because records grouped at the same unique site had identical coordinates, they also had the same geometric distance to the nearest facility. The averaging step therefore preserved the site-specific nearest-facility distance while assigning one analytical value to each unique spatial location.

The spatial variable analyzed in this sensitivity procedure was

$$\tilde{d}_q^{(H)},$$

rather than road segment fatal-incident burden. Four k-NN specifications were examined:

$$k \in \{8, 12, 16, 24\}.$$

For each value of k , the nearest-neighbor relation was symmetrized before the spatial weights were created. The resulting weights were row-standardized using the W coding style [30].

Global Moran's I was evaluated for the unique-location nearest-facility distances using a Monte Carlo permutation procedure. The analysis used 9999 random permutations and a fixed random seed of 20260629 to support computational reproducibility. The permutation alternative was specified as greater, corresponding to a one-sided assessment of positive spatial autocorrelation.

This point-based sensitivity analysis was analytically distinct from the primary road segment Moran analysis. It used a different spatial unit, a different variable, and multiple nearest-neighbor specifications to address the sensitivity of hospital-proximity clustering to exact duplicate coordinates and the choice of k .

2.13.2. Road network-definition sensitivity analysis

The all-road network containing 68308 segments was retained as the primary analytical network. Two restricted motorized road networks were used only for sensitivity analysis.

The motorized core network retained the following OpenStreetMap highway classes:

motorway; trunk; primary; secondary; tertiary; residential; living_street; unclassified; road;
motorway_link; trunk_link; primary_link; secondary_link; and tertiary_link.

The motorized broad network contained all road classes retained in the motorized core network and additionally included the service class. These restricted definitions did not replace the primary all-road analysis. Nearest-road assignment was repeated independently within each network scenario so that every fatal-incident point was assigned to the nearest segment eligible under that scenario.

Let $s \in \{A, B, C\}$ index the road network scenario, where A scenario represented the primary all-road network, scenario B represented the motorized broad network, and scenario C represented the motorized core network.

For each scenario s , let

$$\mathcal{R}^{(s)} \subseteq \mathcal{R}, s \in \{A, B, C\}$$

denote the set of road segments eligible for nearest-road assignment under that scenario. The scenario-specific assignment function was defined as

$$a_i^{(s)} \in \arg \min_{\ell: r_\ell \in \mathcal{R}^{(s)}} d(p_i, r_\ell).$$

The scenario-specific fatal-incident burden of eligible road segment r_ℓ was

$$y_\ell^{(s)} = \sum_{i=1}^{n_F} 1\{a_i^{(s)} = \ell\}, \quad r_\ell \in \mathcal{R}^{(s)}.$$

The conservation identity was checked for every scenario:

$$\sum_{r_\ell \in \mathcal{R}^{(s)}} y_\ell^{(s)} = n_F = 2507.$$

This check ensured that differences among the road network definitions represented reassignment among eligible road segments rather than the loss of fatal-incident records.

Ranking robustness was evaluated using the road segments with the highest scenario-specific fatal-incident burdens. Let $\mathcal{T}_K^{(s)}$ denote the set of the K highest-burden road segments under scenario s . For two scenarios s and the percentage overlap was calculated as

$$O_K^{(s,t)} = 100 \times \frac{|\mathcal{T}_K^{(s)} \cap \mathcal{T}_K^{(t)}|}{K},$$

where $|\cdot|$ denotes the number of elements in a set. Pairwise comparisons were conducted between the primary all-road scenario and each restricted scenario:

$$(s, t) \in \{(A, B), (A, C)\}.$$

Ranking overlap was evaluated at $K \in \{10, 20\}$. The top 10 comparison assessed the stability of the most highly burdened road segment group, whereas the top 20 comparison assessed stability across a broader group of high-burden road segments.

2.14. Descriptive summaries

The nearest-education-facility distances d_i^E and the distances to the nearest strict hospital or emergency-care facility d_i^H were summarized using the minimum, median, arithmetic mean, and maximum. The median was reported because spatial-distance distributions may be asymmetric. Education-facility and hospital buffer zone results were reported as counts and percentages using the fixed denominator $n_F = 2507$. Road segment and road class summaries were calculated using the measures defined in Sections 2.9 and 2.10. Because traffic volume, vehicle kilometers traveled, pedestrian activity, cyclist activity, and person-time at risk were unavailable, ρ , ρ_c , and Q_c were interpreted as descriptive spatial-distribution measures rather than exposure-adjusted estimates of road-safety risk.

2.15. Software and reproducibility

Data preparation, spatial processing, statistical analysis, and figure production were conducted in R 4.5.1. The principal R packages used were sf 1.1.1, dplyr 1.1.4, spdep 1.4.2, ggplot2 4.0.0, openxlsx 4.2.8.1, ggspatial 1.1.10, scales 1.4.0, grid 4.5.1, and ragg 1.5.0. All distance and road-length calculations were performed using geometries projected to MAGNA-SIRGAS / Colombia Bogotá zone (EPSG:3116). Reproducibility checks verified the retained numbers of fatal incidents, facilities, and road segments and confirmed conservation of all 2507 fatal-incident assignments during the primary and sensitivity road network analyses.

The primary road segment Global Moran's I test and the unique-location Global Moran's I sensitivity analyses used Monte Carlo permutation inference with 9999 permutations and a fixed random seed of 20260629. The complete analytical sequence, from fatal-incident selection and spatial screening through facility-proximity, nearest-road, spatial-autocorrelation, and sensitivity analyses, is summarized in Figure 2.

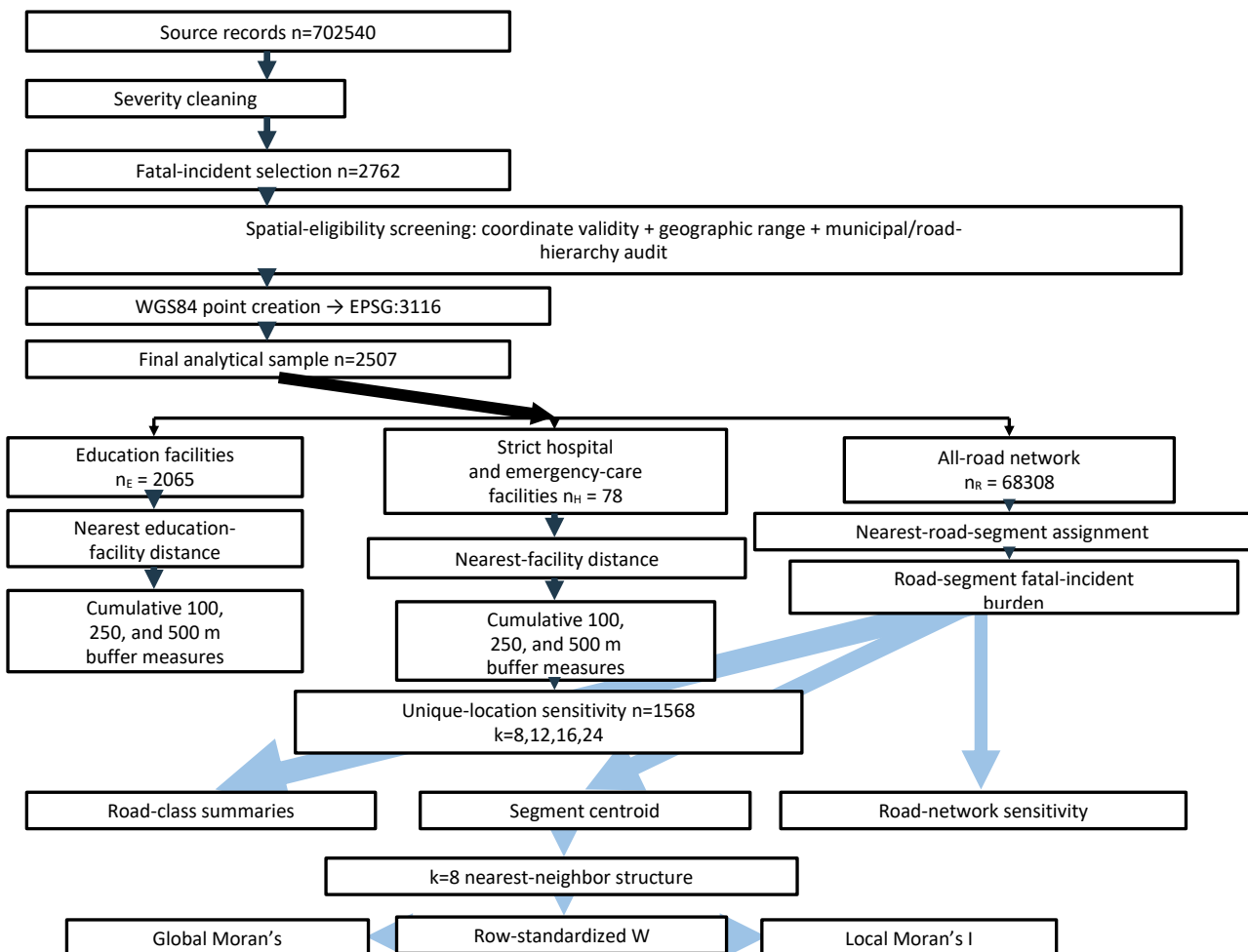


Figure 2. GIS analytical workflow

3. Results

3.1. Analytical sample and spatial coverage

The original Medellín road traffic incident database contained 702540 records covering the period from 1 January 2008 to 31 August 2025. Filtering the cleaned GRAVEDAD_INCIDENTE field identified 2762 fatal-incident records. After spatial-eligibility screening, 2507 fatal-incident points (90.77%) were retained for spatial analysis. These records occurred at 1568 unique coordinate locations; 400 were repeated-coordinate locations containing 1339 records, with a maximum of 24 records at one coordinate. After the coordinate checks, 2559 points remained, of which 2497 were inside or on the municipal boundary, and 62 were outside it. The road-hierarchy audit retained 10 of these 62 points and excluded 52. Figure 1 displays the final analytical sample, including the 10 retained outside-boundary points.

3.2. Proximity of fatal incidents to education facilities

The final analytical education facility layer contained 2065 point locations. The distance from each fatal-incident point to the nearest education facility ranged from 13.48 m to 1281.98 m, with a median of 190.57 m and a mean of 217.61 m. Of the 2507 analyzed fatal incidents, 490 (19.55%) were located within 100 m of the nearest education facility, 1724 (68.77%) within 250 m, and 2362 (94.22%) within 500 m. These results describe spatial proximity to mapped education facilities and do not represent education facility-related traffic exposure. Figure 3 shows the spatial pattern, cumulative proximity percentages, and distribution of nearest-education facility distances.

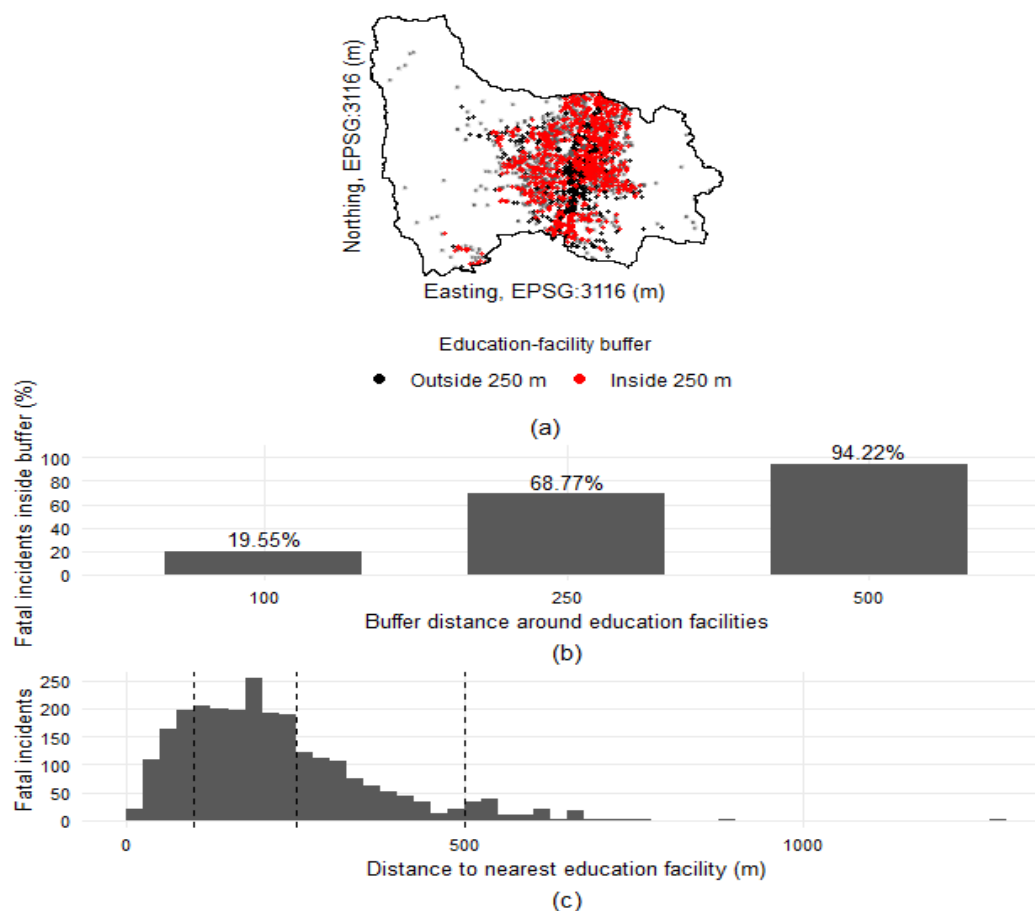


Figure 3. Proximity for fatal road traffic incidents to education facilities in Medellín. (a) Fatal incidents within and beyond 250 m based on the nearest education facility; gray areas show 250-m buffers. (b) Cumulative percentages within 100, 250, and 500 m. (c) Distribution of distances to the nearest education facility; dashed lines mark the three thresholds. Distances were calculated in EPSG:3116 using all 2507 fatal incidents and 2065 education facilities. Map layers in panel (a) had been clipped to the municipal boundary for display only.

3.3. Proximity of fatal incidents to strict hospital and emergency-care facilities

The strict hospital and emergency-care analytical layer contained 78 facilities: 20 point facilities and 58 polygon or multipolygon facilities. Of these, 57 were located inside or intersected the Medellín municipal boundary and 21 were located outside it; all 78 were retained in the nearest-facility analysis. Fig. 4 uses representative internal points for display purposes only.

The distance to the nearest facility ranged from 0 m to 6293.62 m, with a median of 660.36 m and a mean of 753.07 m. 98 fatal incidents (3.91%) were located within 100 m of the nearest strict hospital or emergency-care facility, 372 (14.84%) within 250 m and 906 (36.14%) within 500 m. These ratios were less than the respective education facility-proximity measures. Hospital proximity is a measure of the location of a facility relative to other included facilities and not a direct measure of emergency response time, treatment availability, or survival after a crash. The four mutually exclusive distance classes and the locations of the 78 retained facilities are shown in Figure 4.

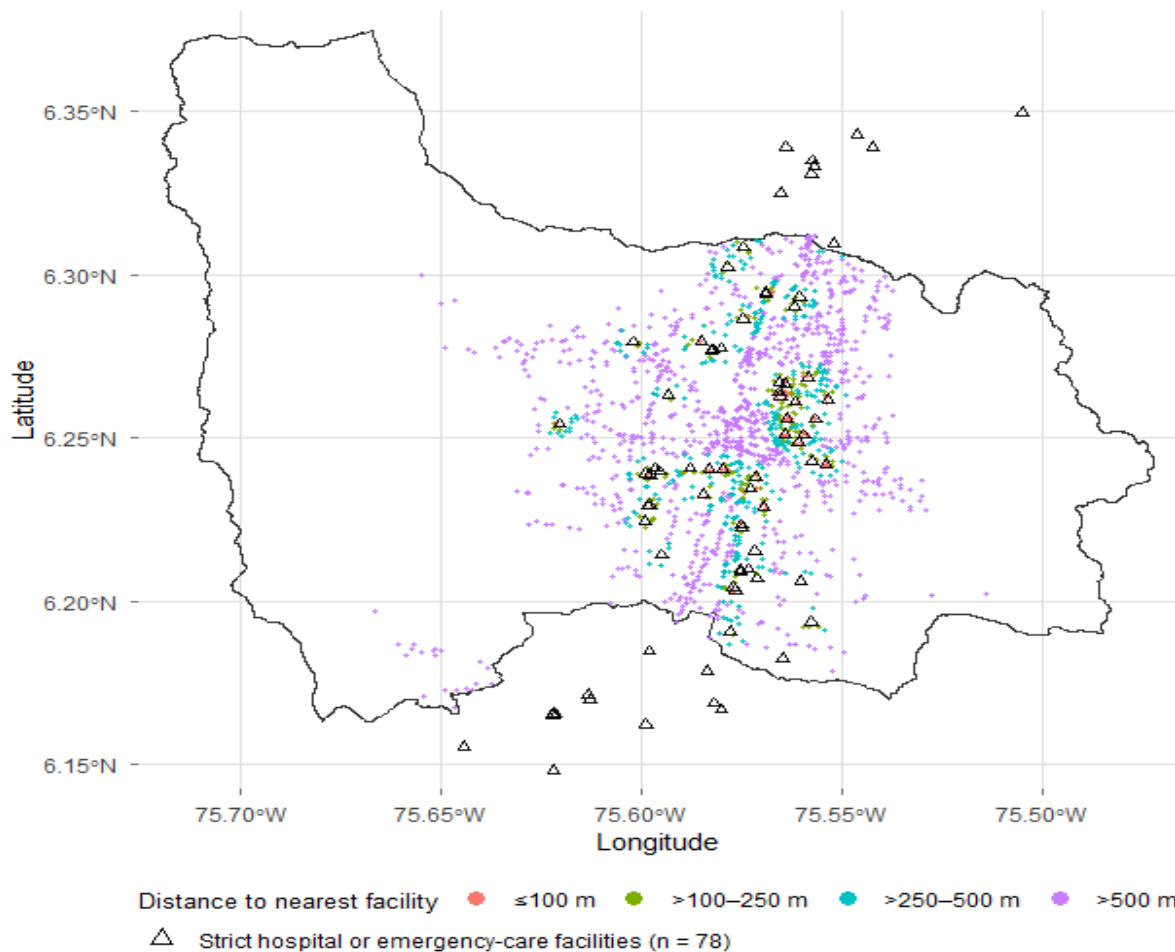


Figure 4. Spatial distribution of fatal road traffic incidents according to distance from the nearest strict hospital or emergency-care facility in Medellín, 2008–2025. Fatal incidents (n = 2507) were classified into four mutually exclusive distance classes: ≤ 100 m, $>100\text{--}250$ m, $>250\text{--}500$ m, and >500 m. Hollow triangles show the 78 facilities included in the proximity analysis; 57 were located inside or intersected the municipal boundary and 21 were located outside it. Euclidean distances were calculated in MAGNA-SIRGAS / Colombia Bogotá zone (EPSG:3116), whereas geographic coordinates are shown for map display.

3.4. Nearest-road assignment, road segment burden, and road class distribution

The primary road-based analysis used the complete all-road network of 68308 road segments with a combined mapped length of 9354.690 km, and all 2507 geocoded fatal incidents were assigned to their nearest road segments. The median assignment distance was 0.983 m, and all incidents were within 100 m of the nearest

mapped segment. At least one fatal incident was assigned to 1331 segments (1.95% of the network). The network-wide burden was 0.268 assigned fatal incidents per kilometer, with a maximum segment burden of 28. Figure 5 shows the spatial distribution of assigned fatal-incident burden across the road network.

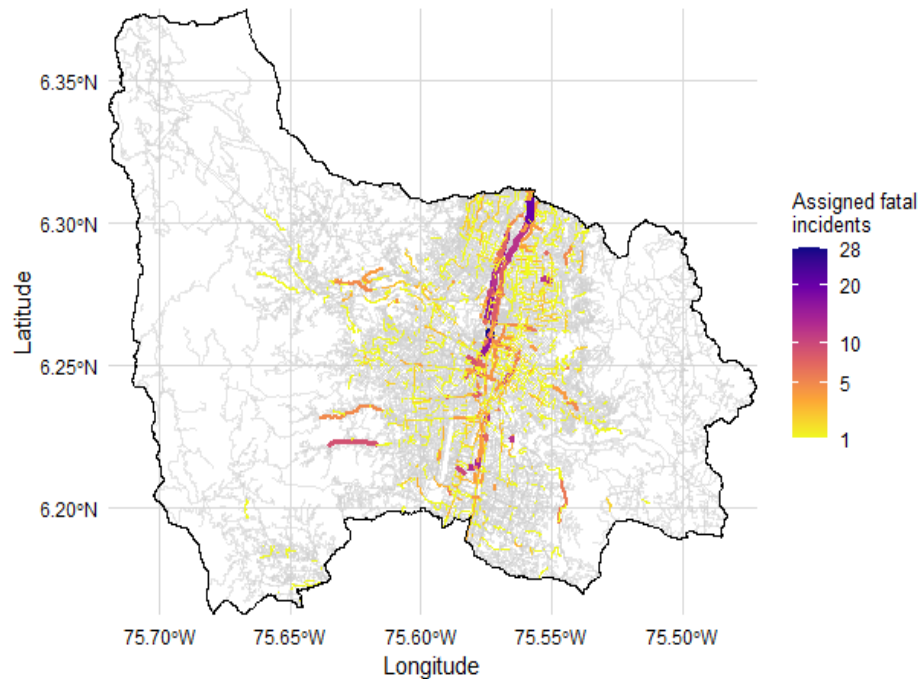


Figure 5. Spatial distribution of fatal road traffic incident burden across road segments in Medellín, 2008–2025. Each fatal incident was assigned to its nearest road segment, and line color represents the number of assigned fatal incidents.

The road network contained 30 OpenStreetMap road classes, of which 17 received at least one fatal-incident assignment. Residential roads had the largest raw burden, with 550 assigned incidents (21.94%), followed by tertiary, secondary, primary, and trunk roads. Together, these five classes comprised 85.08% of all assigned fatal incidents. Table 2 summarizes the main road class distribution measures and the count of High–High and Low–High Local Moran classifications in the displayed road classes.

Table 2. Distribution of assigned fatal incidents and Local Moran cluster classifications across selected OSM road classes in the all-road network.

	OSM road class	Assigned fatal incidents, n (%)	Road network length, km (%)	Assigned fatal incidents per km	Fatal-share-to-length-share ratio	Local Moran clusters, HH/LH (n)
1	residential	550 (21.94%)	2736.734 (29.26%)	0.201	0.75	38/316
2	tertiary	522 (20.82%)	650.220 (6.95%)	0.803	3.00	22/115
3	secondary	483 (19.27%)	384.633 (4.11%)	1.256	4.69	19/99
4	primary	368 (14.68%)	231.134 (2.47%)	1.592	5.94	21/91
5	trunk	210 (8.38%)	236.170 (2.52%)	0.889	3.32	19/47

	OSM road class	Assigned fatal incidents, n (%)	Road network length, km (%)	Assigned fatal incidents per km	Fatal-share-to-length-share ratio	Local Moran clusters, HH/LH (n)
6	footway	140 (5.58%)	1029.038 (11.00%)	0.136	0.51	9/587
7	service	56 (2.23%)	1478.594 (15.81%)	0.038	0.14	6/334
8	primary_link	55 (2.19%)	31.080 (0.33%)	1.770	6.60	2/34
9	trunk_link	44 (1.76%)	34.953 (0.37%)	1.259	4.70	2/55
10	cycleway	26 (1.04%)	121.051 (1.29%)	0.215	0.80	0/65
	All remaining 20classes	53 (2.11%)	2421.082 (25.88%)	0.022	0.08	2/123
	Total network	2507 (100%)	9354.690 (100.00%)	0.268	1.00	140/1866

Abbreviations: OSM, OpenStreetMap; HH, High–High; LH, Low–High. These Local Moran classifications are examined further in Section 3.5.

The rankings obtained with road-length normalized measures and raw fatal incident counts were different. The largest absolute burden was on residential roads. Primary_link, primary, and several other major or link classes had higher burdens after normalization by mapped road length. Figure 6 compares the proportion of assigned fatal incidents for each road class to its share of the total mapped network length. Above the dashed equality line, the classes had a larger share of fatal incidents than the share of network length, while below the line, the classes had a smaller share of fatal incidents. Some major-road and link classes were above the equality line, while residential roads were below the raw fatal-incident count. This distinction differentiates absolute burden from road-length-normalized concentration.

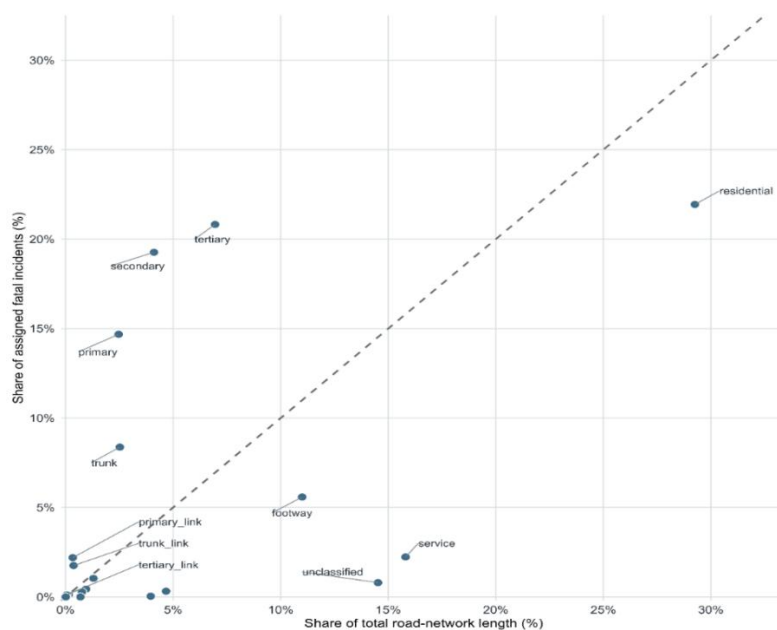


Figure 6. Fatal-incident share versus road network length share by OSM road class

Each point represents one of the 30 OSM road classes in the all-road primary analysis. The horizontal axis shows the class-specific share of total road network length, and the vertical axis shows the corresponding share of all assigned fatal incidents. The dashed 1:1 line denotes equality between the two shares. Classes above the line accounted for a greater share of assigned fatal incidents than of network length, whereas classes below the line accounted for a smaller share. Selected classes are labeled for readability. These comparisons are descriptive and road-length-normalized; they are not exposure-adjusted risk estimates.

3.5. Spatial autocorrelation and sensitivity analyses

Global spatial autocorrelation of road segment fatal-incident burden was positive but weak. Using the $k = 8$ row-standardized spatial-weights structure, Global Moran's I was 0.0245 (permutation $p = 0.0001$), indicating statistically significant but limited overall spatial dependence.

The unique-location sensitivity analysis of distance to the nearest strict hospital or emergency-care facility showed strong positive spatial autocorrelation across all four neighborhood definitions (Moran's $I = 0.7624$ – 0.8875 ; permutation $p = 0.0001$ for all specifications), with Moran's I decreasing as k increased. These values are not directly comparable with the weaker road segment result because the analyses used different spatial units and different variables.

The alternative road network definitions showed high ranking stability: both the motorized broad and motorized core networks retained 8 of the 10 highest-burden segments (80%) and 19 of the 20 highest-burden segments (95%) relative to the all-road network.

Local Moran's I identified 140 High–High road segments (0.20% of the network) and 1866 Low–High segments (2.73%). The remaining 66302 segments (97.06%) were not significant at the nominal $p \leq 0.05$ threshold, and no Low–Low or High–Low segments met this threshold. Together, the 140 High–High segments received 327 assigned fatal incidents. Figure 7 shows the spatial distribution of these Local Moran classifications.

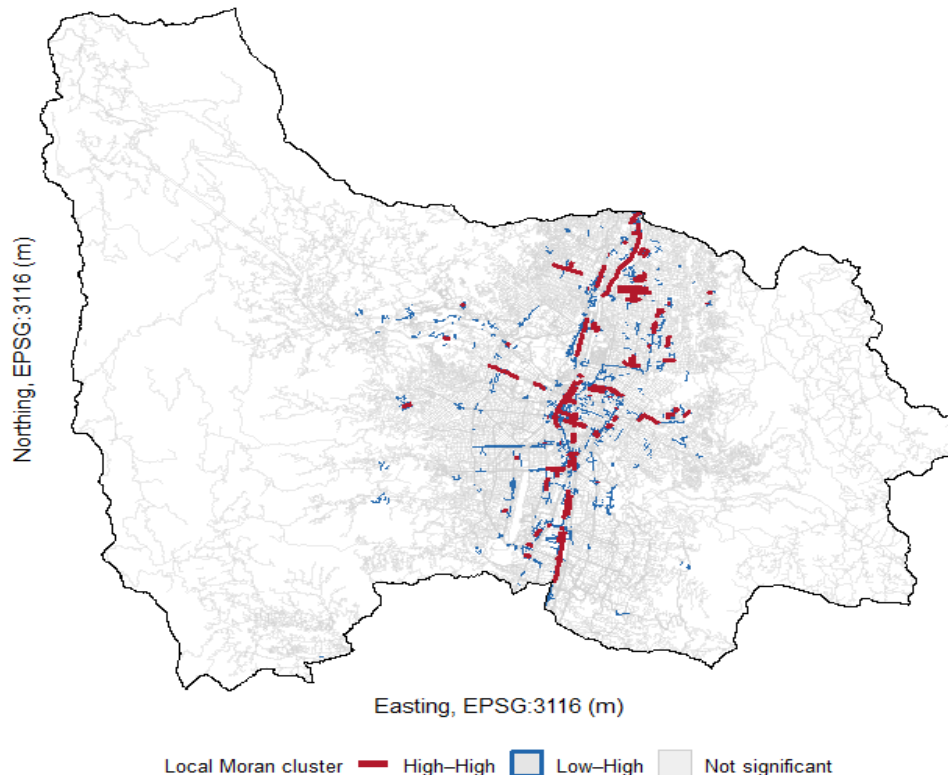


Figure 7. Local Moran's I cluster classification of road segment fatal-incident burden in Medellín using an eight-nearest-neighbor spatial-weights structure

High–High segments indicate localized concentrations of elevated fatal-incident burden, whereas Low–High segments indicate lower-burden road segments surrounded by higher-burden neighbors. Non-significant road segments are shown as the background road network. The principal quantitative findings from Sections 3.1–3.5 are consolidated in Table 3.

Table 3. Summary of principal quantitative findings

Analysis	Measure	Result
Analytical sample	Fatal incidents retained for spatial analysis	2507 of 2762 (90.77%)
Coordinate audit	Unique coordinate locations	1568
Coordinate audit	Fatal records at repeated-coordinate locations	1339 (53.41%)
Education facility proximity	Median nearest distance	190.57 m
	Within 100 m	490 (19.55%)
	Within 250 m	1724 (68.77%)
	Within 500 m	2362 (94.22%)
Strict hospital and emergency-care proximity	Median nearest distance	660.36 m
	Within 100 m	98 (3.91%)
	Within 250 m	372 (14.84%)
	Within 500 m	906 (36.14%)
Road network	Road segments, n	68308
	Total mapped length	9354.690 km
Road-segment burden	Segments with ≥ 1 assigned fatal incidents	1331 (1.95%)
	Segments with no assigned fatal incident	66977 (98.05%)
	Assigned fatal incidents per km	0.268
	Maximum segment burden	28 assigned fatal incidents
Primary road segment spatial autocorrelation	Global Moran's I, k=8	I=0.0245, permutation p=0.0001
Local Moran's I	High–High	140 (0.20%)
	Low–High	1866 (2.73%)
	Not significant	66302 (97.06%)
High–High burden	Assigned fatal incidents	327
Road network sensitivity	Top 10 ranking overlap	80% for both restricted networks
	Top 20 ranking overlap	95% for both restricted networks
Unique-location hospital-proximity sensitivity	Global Moran's I, k=8	I=0.8875, permutation p=0.0001
	Global Moran's I, k=12	I=0.8457, permutation p=0.0001
	Global Moran's I, k=16	I=0.8125, permutation p=0.0001
	Global Moran's I, k=24	I=0.7624, permutation p=0.0001

4. Discussion

For this study, we integrated the locations of fatal road traffic incidents with education-facility, hospital and emergency-care, and road-network data to analyze spatial patterns in Medellin. Fatal-incident burden was

concentrated on a small part of the mapped network. Only 1331 of 68308 road segments (1.95%) got at least one assignment. The main quantitative results of the study are shown in Table 3.

Facility proximity patterns were strikingly different: 94.22% of fatal incidents occurred within 500 m of the nearest mapped education facility, compared to 36.14% for the nearest strict hospital or emergency-care facility. These measures are based on geographic co-location only and do not imply that education facilities played a role in fatal incidents or that distance to hospital had an impact on survival. Road class comparisons distinguished absolute burden from road-length normalized concentration. The biggest raw burden fell on residential roads, and several major-road and link classes had larger shares of assigned incidents than of mapped network length. These differences are descriptive, not exposure-adjusted risk estimates, because exposure data were unavailable. Corresponding road-class measures are given in Table 2 and Figure 6.

The global Moran's I was significant, but weak, at 0.0245 for the overall spatial autocorrelation, and the local Moran's I identified 140 High-High and 1866 Low-High segments. This combination is consistent with localized concentrations in a network dominated by zero-burden segments. Road segment analysis can thus reveal local concentrations that might be obscured by aggregation to larger administrative areas [5]. Figure 7 shows the Local Moran classifications.

The sensitivity analyzes of the road network also revealed stable high-burden rankings, with the restricted motorized networks sharing 80% of the top 10 segments and 95% of the top 20 segments in common with the all-road network. This means that the leading locations were not influenced solely by the inclusion of non-core road classes.

Residential roads had the greatest raw burden but a fatal-share-to-length-share ratio of <1 (0.75), indicating a smaller share of assigned incidents than their share of mapped network length (Table 2). Some major-road and link classes showed the opposite pattern with fatal-share-to-length-share ratios greater than 1, as shown in Table 2 and Figure 6. The traffic volume, speed environment, intersection configuration, and the interaction with vulnerable road users may be different on these roads. Such factors align with the Safe System approach's focus on speed management and forgiving infrastructure for human error [1,2]. These mechanisms were not directly measured. Thus, the observed road class differences are descriptive normalized patterns of road length and not demonstrated causes or exposure-adjusted risk estimates. The results of road segments are also sensitive to the spatial scale, network representation, and the procedure to associate incident points with road lines. Street segment analysis has been shown to provide useful local information, but can be sensitive to network segmentation and spatial support [5].

Figure 3 shows the results for the education-facility proximity, which must be interpreted in a combined layer of schools, kindergartens, colleges, and universities. Proximity, therefore, does not mean student engagement or a causal role of an education facility. Additional information on entrances, pedestrian routes, traffic conditions, and the local street environment [8] would be required for school-specific assessment. Figure 4 presents initial metrics of geographic proximity for hospital and emergency-care access, but the straight-line distance does not account for route configuration, barriers, congestion, or facility capability. Where the research question concerns realized accessibility or emergency transport, network distance or modeled travel time is more appropriate [9, 10].

The strong spatial autocorrelation in nearest-facility distance reported in Table 3 is consistent with geographic continuity in facility proximity, and does not demonstrate that hospital distance caused the spatial distribution or fatal outcome of incidents.

These results complement previous analyzes on the same incident database in Medellin on spatiotemporal concentration, pedestrian vulnerability, spatial inequalities, and fatality prediction [16–18]. This is complemented by the present study with a road-segment and external-infrastructure perspective, integrating

mapped facility proximity, road-class distribution and road-segment spatial autocorrelation. The results provide a spatial screening framework for the identified locations for further investigation. Segments with a combination of high raw burdens, High-High classifications, and stable rankings across alternative road network definitions are the best candidates for detailed assessment, however these indicators do not specify what the intervention is.

Site-specific assessment requires information about traffic and pedestrian exposure, operating speeds, road-user characteristics, and roadway design. This accords with the Safe System approach and encourages multidisciplinary investigation rather than reliance on road-user education or enforcement alone [1,2]. Future school-specific assessment could isolate mapped schools from the combined education-facility layer, and examine their coincidence with high-burden road segments using more detailed pedestrian, traffic and street environment data.

The hospital results could identify areas of more geographic separation from mapped facilities, but cannot evaluate emergency-response performance. Such assessment requires transport, ambulance, facility-capability, and clinical data as effective post-crash care depends on coordinated prehospital and hospital systems rather than physical distance alone [31].

The findings also complement Medellin's Vision Zero policy by providing an empirical basis for geographically targeted investigation of a relatively small group of consistently high-burden road segments [12]. Methodological strengths included consistently using the final 2507 fatal incidents, distance and length calculations in EPSG:3116, using original hospital geometries to measure proximity, verifying assignment conservation across road-network scenarios, and collapsing exact duplicate coordinates to unique locations for the unique-location sensitivity analysis.

Interpretation is affected by several limitations. The source data did not reflect individual deaths but fatal incidents. A total of 255 fatal records were removed during spatial screening. The infrastructure layers may not be representative of the entire 2008-2025 incident period.

Road and facility layers derived from OpenStreetMap are crowdsourced and may be incomplete [25-27]. Nearest-road assignment was geometric (not police confirmed), so misassignment may occur when multiple mapped road geometries are close to each other.

Geographic separation was represented by straight-line facility distances rather than route accessibility or emergency response. The only available normalization denominator was road length. The lack of traffic and road-user exposure data therefore limits interpretation of the road class measures as exposure-adjusted risks. The eight-nearest-neighbor structure based on centroids reflected geographic proximity rather than road connectivity or traffic flow, and the unadjusted Local Moran's I threshold of $p \leq 0.05$ indicates that the local classifications remain exploratory.

5. Conclusion and future trends

Fatal road traffic incidents in Medellin were concentrated in a small proportion of the mapped road network, with only 1331 of 68308 segments (1.95%) receiving at least one assignment. Residential roads carried the largest raw burden, and several major-road and link classes carried larger shares of assigned incidents than of mapped network length.

Patterns of proximity to facilities also differed, with 94.22% of fatal incidents within 500 m of the nearest mapped education facility compared with 36.14% for strict hospital or emergency-care facilities. These results reflect spatial co-location only and do not imply causal relationships with incident occurrence or post-crash outcomes. The global spatial autocorrelation of burden across road segments was weak (Moran's $I=0.0245$), but Local Moran's I identified 140 High-High and 1866 Low-High segments. The high-burden rankings also held up to alternative definitions of the road network, with 80% overlap between the top 10 segments and 95%

overlap between the top 20 segments. Taken together, these findings warrant geographically targeted investigation of consistently high-burden segments and local High-High concentrations.

The results are a framework for spatial screening and prioritization, not causal or exposure-adjusted estimates of road-safety risk. Limitations include incident-level records, spatial exclusions, temporal differences between incident and infrastructure data, limitations of OpenStreetMap data, geometric road assignment, straight-line facility distances, missing exposure denominators, and exploratory Local Moran classifications. Site-level assessment with additional traffic, road-user, roadway, temporal, and emergency-response data is needed to evaluate identified locations and support interventions in line with Safe System and Vision Zero principles.

Future trends will be urban networks as high-frequency circuit boards, to maximize safety and efficiency. Metamaterial Urbanism will involve cities designing roads that passively separate transit flows and erase noise pollution using microstrip coupler principles. Spatial S-Parameters will be used to determine traffic hotspots for predictive crash modeling, as physical impedance mismatches and power reflections occur at multi-port crossings. Digital Twin Diplexing will offer a real-time safety solution by using 6G spatial buffers to dynamically isolate autonomous transit channels from pedestrians to eliminate the physical cross-talk. Fractal Network Urbanism will use scale-invariant geometries to construct resilient emergency routing and drainage infrastructure that can manage structural stress seamlessly from micro-streets to macro-metropolitan scales [32, 33]. Optimization by differential fuzzy equations is also possible [34].

Declaration of competing interests

The authors declare that they have no known financial or non-financial competing interests in any material discussed in this paper.

Funding information

No funding was received from any financial organization to conduct this research.

Author contribution

The contribution to the paper is as follows: Noorah Osama Abid: study conception, data sourcing, acquisition and processing, data analysis, and drafting the manuscript; Abeer Salim Alwan Alnajjar and Shaymaa Khazaal Shnaishel: literature review, study design, formal analysis, visualization, and validation. Both authors approved the final version of the manuscript.

References

- [1] World Health Organization, *Global Status Report on Road Safety 2023*. Geneva: World Health Organization, 2023. ISBN: 978-92-4-008651-7.
- [2] International Transport Forum, *The Safe System Approach in Action*. Paris: OECD Publishing, 2022. doi: 10.1787/ad5d82f0-en.
- [3] R. Ewing and E. Dumbaugh, "The built environment and traffic safety: A review of empirical evidence," *Journal of Planning Literature*, vol. 23, no. 4, pp. 347–367, 2009. doi: 10.1177/0885412209335553.
- [4] S. Erdogan, I. Yilmaz, T. Baybura, and M. Gullu, "Geographical information systems aided traffic accident analysis system case study: City of Afyonkarahisar," *Accident Analysis & Prevention*, vol. 40, no. 1, pp. 174–181, 2008. doi: 10.1016/j.aap.2007.05.004.
- [5] A. Gilardi, J. Mateu, R. Borgoni, and R. Lovelace, "Multivariate hierarchical analysis of car crashes data considering a spatial network lattice," *Journal of the Royal Statistical Society: Series A*, vol. 185, no. 3, pp. 1150–1177, 2022. doi: 10.1111/rssa.12823.
- [6] P. A. P. Moran, "Notes on continuous stochastic phenomena," *Biometrika*, vol. 37, no. 1–2, pp. 17–23, 1950. doi: 10.1093/biomet/37.1-2.17.

- [7] L. Anselin, "Local indicators of spatial association—LISA," *Geographical Analysis*, vol. 27, no. 2, pp. 93–115, 1995. doi: 10.1111/j.1538-4632.1995.tb00338.x.
- [8] K. Zhang, R. Tamakloe, M. Cao, and I. Kim, "Exploring fatal/severe pedestrian injury crash frequency at school zone crash hotspots: Using interpretable machine learning to assess the micro-level street environment," *Journal of Transport Geography*, vol. 121, p. 104034, 2024. doi: 10.1016/j.jtrangeo.2024.104034.
- [9] R. C. Nesbitt et al., "Methods to measure potential spatial access to delivery care in low- and middle-income countries: A case study in rural Ghana," *International Journal of Health Geographics*, vol. 13, p. 25, 2014. doi: 10.1186/1476-072X-13-25.
- [10] A. Mizen, R. Fry, D. Grinnell, and S. E. Rodgers, "Quantifying the error associated with alternative GIS-based techniques to measure access to health care services," *AIMS Public Health*, vol. 2, no. 4, pp. 746–761, 2015. doi: 10.3934/publichealth.2015.4.746.
- [11] Alcaldía de Medellín, *Información General de Ciudad: Ubicación, división geográfica y principales características de la ciudad*. Medellín: Alcaldía de Medellín, 2023.
- [12] Alcaldía de Medellín, Decreto 0261 de 2019: Por el cual se adopta el enfoque denominado “Visión Cero” en materia de seguridad vial para el municipio de Medellín y se dictan otras disposiciones, *Gaceta Oficial* No. 4598, Medellín, 2019.
- [13] Alcaldía de Medellín, "Incidentes viales registrados en Medellín" [Dataset], Datos Abiertos Colombia.
- [14] M. L. Arango Uribe, E. Quiceno Rúa, and C. D. Correa Álvarez, "Georeferenced dataset on road traffic incidents and fatalities in Medellín, Colombia (2008–2025)," *Data*, vol. 11, no. 4, p. 67, 2026. doi: 10.3390/data11040067.
- [15] C. Correa, "Dataset on Road Traffic Fatalities in Medellín, Colombia (2008–2025)" [Dataset], Version 1, Mendeley Data, 2026. doi: 10.17632/r6g5dfnpgh.1.
- [16] J. Sánchez Corredor, M. L. Arango Uribe, and C. D. Correa Álvarez, "Quantifying the social burden and spatiotemporal concentration of fatal road traffic incidents in Medellín (2008–2025)," *Sustainability*, vol. 18, no. 5, p. 2628, 2026. doi: 10.3390/su18052628.
- [17] M. L. Arango Uribe, J. Sánchez Corredor, and C. D. Correa Álvarez, "Governance of road-safety inequality: Spatiotemporal patterns and pedestrian vulnerability in Medellín, Colombia," *Urban Science*, vol. 10, no. 6, p. 329, 2026. doi: 10.3390/urbansci10060329.
- [18] E. M. López-López, O. E. Bru-Cordero, and C. D. Correa-Álvarez, "Rare-event road-traffic fatality prediction: A reproducible machine-learning benchmark with time-aware validation, calibration, and interpretability," *Future Transportation*, vol. 6, no. 4, p. 146, 2026. doi: 10.3390/futuretransp6040146.
- [19] Alcaldía de Medellín, Departamento Administrativo de Planeación, *Limite Municipio—Capa Limite del Municipio de Medellín 2014*. Medellín: Alcaldía de Medellín, 2014.
- [20] Departamento Administrativo Nacional de Estadística (DANE), *Marco Geoestadístico Nacional (MGN), versión 2017—Municipio (MGN_MPIO_POLITICO)*. Bogotá: DANE, 2017.
- [21] E. Pebesma, "Simple features for R: Standardized support for spatial vector data," *The R Journal*, vol. 10, no. 1, pp. 439–446, 2018. doi: 10.32614/RJ-2018-009.
- [22] Alcaldía de Medellín, Departamento Administrativo de Planeación, "Jerarquía Vial," *GeoMedellín*.
- [23] EPSG Geodetic Parameter Dataset, "MAGNA-SIRGAS / Colombia Bogota zone—EPSG:3116," 2023.
- [24] E. Pebesma and R. Bivand, *Spatial Data Science: With Applications in R*. Boca Raton: Chapman & Hall/CRC, 2023. doi: 10.1201/9780429459016.
- [25] Humanitarian OpenStreetMap Team (HOT), "Colombia Education Facilities (OpenStreetMap Export)" [Dataset], Humanitarian Data Exchange, 2026. [Online]. Available: OpenStreetMap contributors. Data snapshot: 19 June 2026.
- [26] Humanitarian OpenStreetMap Team (HOT), "Colombia Health Facilities (OpenStreetMap Export)" [Dataset], Humanitarian Data Exchange, 2026. [Online]. Available: OpenStreetMap contributors. Data snapshot: 19 June 2026.

-
- [27] Humanitarian OpenStreetMap Team (HOT), "Colombia Roads (OpenStreetMap Export)" [Dataset], Humanitarian Data Exchange, 2026. [Online]. Available: OpenStreetMap contributors. Data snapshot: 19 June 2026.
 - [28] G. A. Golembiewski and B. Chandler, Roadway Departure Safety: A Manual for Local Rural Road Owners. Washington, DC: Federal Highway Administration, 2011, Report No. FHWA-SA-11-09.
 - [29] Federal Highway Administration, Safety Performance Measures Fact Sheet. Washington, DC: U.S. Department of Transportation, Federal Highway Administration, 2016, Report No. FHWA-SA-16-044.
 - [30] R. S. Bivand and D. W. S. Wong, "Comparing implementations of global and local indicators of spatial association," *TEST*, vol. 27, pp. 716–748, 2018. doi: 10.1007/s11749-018-0599-x.
 - [31] World Health Organization, Post-crash Response: Supporting Those Affected by Road Traffic Crashes. Geneva: World Health Organization, 2016, WHO/NMH/NVI/16.9.
 - [32] Y. S. Mezaal, H. T. Eyyuboğlu, and J. K. Ali, "A new design of dual band microstrip bandpass filter based on Peano fractal geometry: Design and simulation results," in *2013 13th Mediterranean Microwave Symposium (MMS)*, Saida, Lebanon, 2013, pp. 1–4, doi: 10.1109/MMS.2013.6663140
 - [33] Y. S. Mezaal, "State of Art on Microstrip Resonators, Filters, Diplexers and Triplexers," *Journal of Mechanics of Continua and Mathematical Sciences*, vol. 19, no. 2, pp. 1–24, Feb. 2024, doi: [10.26782/jmcms.2024.02.00001](https://doi.org/10.26782/jmcms.2024.02.00001).
 - [34] S. K. Khaleel, I. Rikan Hassan, and M. Yousif Hmood, "Optimization of Fuzzy Differential Equation-Based FARDL Models for Engineering Applications: Comparative Analysis of Linear and Quadratic Estimators via Parallel Monte Carlo Simulations," *Journal of Mechanics of Continua and Mathematical Sciences*, vol. 21, no. 6, Jun. 2026, doi: 10.26782/jmcms.2026.06.00011.