

UNIVERSIDAD POLITÉCNICA DE MADRID
Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos



**ANALYSIS OF THE ROAD SAFETY OF
OLDER PEDESTRIANS IN THE URBAN
ENVIRONMENT: APPLICATION TO THE
SPANISH CASE STUDY**

TESIS DOCTORAL

Presentada para optar al título de Doctor por:

Daniel Gálvez Pérez

Ingeniero de Caminos, Canales y Puertos

Madrid, 2025



UNIVERSIDAD POLITÉCNICA DE MADRID
Escuela Técnica Superior de Ingenieros de Caminos,
Canales y Puertos

Doctorado en Sistemas de Ingeniería Civil

**ANALYSIS OF THE ROAD SAFETY OF
OLDER PEDESTRIANS IN THE URBAN
ENVIRONMENT: APPLICATION TO THE
SPANISH CASE STUDY**

TESIS DOCTORAL

Presentada para optar al título de Doctor por:

Daniel Gálvez Pérez

Ingeniero de Caminos, Canales y Puertos

Bajo la dirección de:
Dra. Begoña Guirao Abad

Madrid, 2025

Título: ANALYSIS OF THE ROAD SAFETY OF OLDER PEDESTRIANS IN THE URBAN ENVIRONMENT: APPLICATION TO THE SPANISH CASE STUDY

Autor: Daniel Gálvez Pérez

Programa de Doctorado: Sistemas de Ingeniería Civil

Dirección de tesis: Dra. Begoña Guirao Abad, Catedrática de Universidad, Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos

Revisores externos:

Tribunal de tesis:

Fecha de defensa:

Esta tesis ha sido parcialmente financiada por el PROGRAMA PROPIO DE I+D+I 2020 de la Universidad Politécnica de Madrid.

Agradecimientos

Esta tesis doctoral no es solo obra de su autor; han sido muchas las personas que, durante los últimos años, me han ayudado a seguir adelante.

Quiero agradecer especialmente el papel de mi directora de tesis, Begoña Guirao, quien desde el principio confió en mí para llevar a cabo este trabajo. He sentido su continuo ánimo y apoyo en todas las etapas de la investigación. Gracias también por haber contribuido a mi desarrollo como docente y por incluirme en todos los proyectos que se te presentan. He aprendido mucho sobre cómo desarrollar una investigación, pero aún más sobre cómo trabajar en equipo y organizar mi tiempo.

Gracias a todas y cada una de las personas que, en algún momento, me han aconsejado y animado a continuar. En especial, gracias a mis amigos del Laboratorio de Carreteras, tanto a quienes ya son doctores como a quienes aún no lo son. Vuestra compañía y ayuda durante todo este tiempo me han permitido crecer como investigador y, en definitiva, como persona.

Gracias a los innumerables profesores que he conocido y que, en algún momento, me hicieron pensar críticamente sobre cualquier tema, fuera o no fuera esa su intención. Todas esas pequeñas reflexiones han contribuido a que sea quien soy hoy.

Gracias, cómo no, a mi familia y a mi pareja. Lleváis mucho tiempo escuchándome hablar sobre temas que ni yo mismo acabo de comprender. Gracias por haberme escuchado o, al menos, por haberlo intentado. No siempre es fácil. Vuestro incansable apoyo me ha ayudado, una vez más, a alcanzar la meta que me había propuesto.

Gracias a todos.

Abstract

With global demographic shifts leading to increased proportions of older adults (people aged above 65), urban areas are becoming hotspots for traffic crashes involving this vulnerable group. Despite their cautious behavior and the recognized health benefits of walking, older pedestrians face disproportionate traffic crash risks. In Spain, older adults represented 70% of pedestrian fatalities on urban roads in 2019. These risks and traffic crash outcomes are exacerbated by physical fragility and slow recovery processes if a vehicle collision, unfortunately, occurs. This doctoral thesis investigates the critical issue of road safety for older pedestrians in urban environments, focusing on Spain and its capital city, Madrid, as a case studies.

This research responds to many significant gaps in previous literature on the impact of population ageing on road safety. On the one hand, older drivers are remarkably more extensively studied than older pedestrians. On the other hand, most studies on older pedestrians rely on surveys and observations, with little analysis using traffic crash data. Moreover, research on older pedestrian crashes often focuses solely on this group, without comparing them to other pedestrian crashes, making it unclear whether variable effects on older pedestrian crash frequency apply to all pedestrian crashes. In consequence, the main objective of this thesis is to determine which factors influence the frequency of traffic crashes involving older pedestrians in urban settings and propose countermeasures to mitigate these risks.

Using a holistic methodological framework, the research employes traffic crash data from the Spanish Directorate-General for Traffic (DGT) and supplements it with socio-economic, land use, and infrastructure variables. The thesis develops a systematic approach for data acquisition, processing, enabling detailed statistical analysis. These processes were designed for a variety of spatial units of analysis, including macro- and micro-level units.

Comparative analyses between the two demographic groups demonstrate that a set of specific factors for older pedestrian crashes differ from those affecting non-older pedestrians, necessitating targeted interventions. Findings reveal that older pedestrian crashes are closely linked to specific built environment features, such as crosswalk design, traffic lights, sidewalk width, and elements located on the sidewalks, such as trees and bus stops. A key finding of this doctoral thesis implies that no infrastructure features influence both the frequency of older and non-older pedestrian traffic crashes having opposite effects simultaneously in any of the developed studies. In other words,

a modification of the infrastructure to seek an improvement in the road safety of older pedestrians is not expected to deteriorate the road safety of the rest of the pedestrians.

Regarding the impact of COVID-19 on older pedestrian urban crashes, the frequency was altered especially during the lockdown period, but the injury severity levels remained constant. In addition, crash related factors do not show remarkable changes when comparing crashes before and after the pandemic, suggesting that interventions proposed with pre-pandemic data to improve traffic safety are still suitable.

The thesis culminates in practical recommendations for policymakers, including infrastructure improvements, speed management in urban areas, and the integration of age-friendly design principles. By addressing these factors, the research contributes to safer urban mobility and promotes the concept of healthy ageing. This work represents a significant effort to improve the road safety of older pedestrians using traffic crash data, as it not only fills a critical research gap but also provides actionable insights for urban planning and road safety strategies, fostering safer urban environments for older pedestrians in Spain.

Resumen

Ante los cambios demográficos globales que conducen a un aumento en la proporción y número de personas mayores de 65 años, las áreas urbanas se están convirtiendo en puntos críticos de siniestralidad vial que afectan a este colectivo vulnerable. A pesar de una mayor prudencia en su comportamiento, los peatones mayores son más vulnerables a sufrir atropellos. En España, el 70% de los peatones fallecidos en vías urbanas tenía más de 65 años en 2019. Estos riesgos y las consecuencias de los accidentes de tráfico se agravan por su fragilidad física y unos procesos de recuperación más lentos en comparación con el resto de la población. Esta tesis doctoral aborda la problemática de la seguridad vial de los peatones mayores en entornos urbanos, con especial atención a España y a su capital, Madrid, como casos de estudio.

Esta investigación aborda una brecha en el conocimiento sobre la seguridad vial de peatones mayores. Los conductores mayores han sido más estudiados, y la mayoría de los trabajos sobre peatones mayores se basan en encuestas y observaciones, con pocos análisis de siniestralidad vial. Además, estos estudios suelen centrarse solo en este grupo, impidiendo determinar si los efectos de diversas variables sobre la frecuencia de accidentes son aplicables a todos los peatones. El objetivo de esta tesis es identificar los factores que influyen en la frecuencia de los atropellos sufridos por peatones mayores en entornos urbanos y proponer medidas para reducirlos.

Utilizando un marco metodológico holístico, esta investigación emplea datos de accidentes de tráfico suministrados por la Dirección General de Tráfico (DGT) y los complementa con variables socioeconómicas, de uso del suelo e infraestructura. La tesis desarrolla un enfoque sistemático para la adquisición y procesamiento de datos, que permite realizar análisis estadísticos detallados. Estos procesos se diseñaron para abarcar diversas unidades espaciales de análisis, incluyendo unidades a nivel macro y micro.

Los análisis comparativos entre ambos grupos demográficos demuestran que un conjunto de factores que influyen en los accidentes de peatones mayores difiere de los que afectan a los peatones no mayores, por lo que se requieren intervenciones específicas. Los atropellos a peatones mayores están estrechamente relacionados con características específicas del entorno construido, como el diseño de los pasos de peatones, los semáforos, el ancho de las aceras y los elementos ubicados en estas, como árboles y paradas de autobús. El principal hallazgo de esta tesis es que ninguna característica de la infraestructura viaria influye de manera opuesta en la frecuencia de

atropellos a peatones mayores y al resto de peatones. Es decir, mejorar la seguridad vial de los peatones mayores no afectaría negativamente al resto.

En cuanto al impacto de la COVID-19 en los atropellos urbanos de peatones mayores, la frecuencia de los atropellos se vio alterada especialmente durante el período de confinamiento, pero los niveles de gravedad de las lesiones se mantuvieron constantes. Además, los factores asociados a los siniestros no muestran cambios notables al comparar los atropellos antes y después de la pandemia, lo que sugiere que las intervenciones propuestas con datos previos a la pandemia para mejorar la seguridad vial siguen siendo adecuadas.

La tesis culmina con recomendaciones prácticas, que incluyen mejoras en la infraestructura, gestión de la velocidad en áreas urbanas y la integración de principios de diseño amigable con las personas mayores. Al abordar estos factores, la investigación contribuye a una movilidad urbana más segura y promueve el concepto de envejecimiento saludable. Este trabajo representa un esfuerzo significativo por mejorar la seguridad vial de los peatones mayores en ciudades mediante el uso de datos de siniestralidad, fomentando entornos urbanos más seguros para los peatones mayores en España.

Table of Contents

Agradecimientos	iii
Abstract.....	v
Resumen	vii
List of Figures.....	xiii
List of Tables	xv
List of Abbreviations.....	xvi
1 INTRODUCTION.....	1
1.1 Context.....	1
1.2 Objectives	4
1.3 Investigation phases.....	5
1.4 Key concepts and terminology.....	8
1.4.1 Older adults	8
1.4.2 Traffic crashes	8
1.5 Structure of the dissertation.....	9
2 THEORETICAL FRAMEWORK.....	11
2.1 A general understanding of road safety	11
2.1.1 Traffic crash definition.....	11
2.1.2 Types of traffic crashes.....	12
2.1.3 Traffic crash concurrent factors	13
2.1.4 Traffic crash phases	14
2.1.5 Traffic crash frequency and severity.....	15
2.1.6 Traffic crash underreporting.....	16
2.2 Road safety situation and evolution	17
2.2.1 Road safety situation and evolution worldwide	17
2.2.2 Road safety situation and evolution in Europe	19
2.2.3 Road safety situation and evolution in Spain	20
2.3 Road safety of older pedestrians in urban roads.....	23

2.3.1	Older pedestrian situation in Europe	23
2.3.2	Older pedestrian situation in Spain.....	27
2.3.3	Mega-trends: threads to older pedestrian urban road safety	33
2.4	Road safety strategies and their implications for older pedestrians.....	39
2.4.1	Sustainable Development Goals	39
2.4.2	Second Decade of Action for Road Safety 2021-2030.....	41
2.4.3	Vision Zero in road safety	41
2.4.4	Road Safety Strategy 2030 in Spain	42
2.5	Road safety of older pedestrians as a precursor of active and sustainable mobility	44
2.5.1	Healthy and active ageing	44
2.5.2	Age-friendly cities	46
2.5.3	Walking benefits: more than physical	48
2.5.4	Age-related functional limitations and illnesses of pedestrians	49
2.5.5	Studies on the behavior of older pedestrians	51
2.6	Methods and approaches in studies with traffic crash data	58
2.6.1	Outcome variable: crash frequency and crash injury severity.....	58
2.6.2	Spatial approaches in traffic crash studies	58
2.6.3	Traffic crash data modeling techniques.....	60
2.7	Studies on older pedestrians using traffic crash data.....	62
2.7.1	Traffic crash profile or exploratory studies.....	62
2.7.2	Traffic crash injury severity studies	65
2.7.3	Traffic crash frequency studies.....	66
2.7.4	Main countermeasure proposals to improve older pedestrian road safety	69
2.7.5	Scientific gap in the literature and needs.....	70
3	MATERIALS AND METHODS.....	73
3.1	Spanish traffic crash database	74
3.1.1	Definitions applicable to the Spanish traffic crashes	75
3.1.2	Communication of traffic crashes	77
3.1.3	Data collection systems: ARENA and ARENA2	78
3.1.4	Database structure	78
3.1.5	Relational databases	80
3.1.6	Database queries.....	81
3.1.7	Database deficiencies.....	82

3.2	Database elaboration for the research.....	84
3.2.1	Variable definition	84
3.2.2	Spatial unit of analysis	85
3.2.3	Data acquisition	87
3.2.4	Data processing.....	89
3.3	Statistical modeling	106
3.3.1	Traffic crash frequency modeling	106
3.3.2	Traffic crash injury severity modeling	112
3.3.3	Bayesian inference	114
3.3.4	Model assessment.....	116
3.3.5	Two independent models: older and non-older pedestrian traffic crashes.....	117
3.4	Application of the methodology to the publications	118
3.4.1	Chronology of publications	118
3.4.2	Article A Macro-level analysis using District as the spatial unit of study	122
3.4.3	Article B: Macro-level analysis using Municipalities as the spatial unit of study ..	123
3.4.4	Article C: Micro-level analysis using Street segments as the spatial unit of study	124
3.4.5	Article D: Impact of COVID-19 pandemic on older pedestrian road safety	127
4	PUBLICATIONS	131
4.1	Summary of publications	131
4.2	Article A: The Influence of Built Environment Factors on Elderly Pedestrian Road Safety in Cities: The Experience of Madrid	133
4.3	Article B: Analysis of the elderly pedestrian traffic accidents in urban scenarios: the case of the Spanish municipalities.....	155
4.4	Article C: Age-Friendly Urban Design for Older Pedestrian Road Safety: A Street Segment Level Analysis in Madrid.....	179
4.5	Article D: The effect of COVID-19 on older pedestrian road safety: A holistic analysis through pandemic phases	205
4.6	Additional publications.....	225
5	RESULTS	227
5.1	Results of Article A.....	227
5.2	Results of Article B	230
5.3	Results of Article C.....	234
5.4	Results of Article D.....	236

6	GENERAL DISCUSSION	239
6.1	Integrated solution and Proposed countermeasures.....	239
6.2	Conclusions.....	245
6.3	Further research.....	250
	References	253
	Annex 1 Code for extracting traffic crashes from the database.....	265

List of Figures

Figure	Title	Page
Figure 1.1.	Workflow of the doctoral thesis development across conceptualization, methodology, and reporting phases.	7
Figure 2.1.	Hydén's Pyramid.	15
Figure 2.2.	Evolution of the number of road traffic fatalities in the World, 2000-2021.	17
Figure 2.3.	Share of population, road fatalities, motor vehicles, and paved inter-urban roads by country income level, 2021.	18
Figure 2.4.	Rates of road fatalities per population by region, 2013 and 2016.	19
Figure 2.5.	Evolution of road fatalities in Spain, 1960-2023.	21
Figure 2.6.	Evolution of the mortality rate in road traffic crashes in Spain, 1993-2023.	22
Figure 2.7.	Triple scope of the analysis: older adults, pedestrians, and urban roads.	23
Figure 2.8.	Pedestrians and other road user fatalities in traffic crashes in Europe considering the age of the victim, 2022.	24
Figure 2.9.	Pedestrians and other road user fatalities in traffic crashes in Europe considering the road type, 2022.	24
Figure 2.10.	Evolution of older adult road fatalities and share in the total number of fatalities in Europe, 2012-2022.	25
Figure 2.11.	Distribution of older and non-older adult fatalities considering the transportation mode in Europe, 2022.	26
Figure 2.12.	Evolution of the distribution of pedestrians and other road user fatalities considering the age of the victim in Europe, 2012-2022.	26
Figure 2.13.	Number of fatalities, seriously and slightly injured pedestrians, older adults, and on urban roads in traffic crashes in Spain, 2019.	28
Figure 2.14.	Distribution of vulnerable road user fatalities in urban and other roads considering the transport mode of the victim in Spain, 2021.	28
Figure 2.15.	Evolution of pedestrian fatalities and seriously injured pedestrian in urban and other roads in Spain, 2012-2021.	29
Figure 2.16.	Pedestrian fatalities considering the age of the pedestrian and the road type in Spain, 2021.	31
Figure 2.17.	Number of pedestrian fatalities in urban roads in Spain, 2012, 2019 and 2021.	31
Figure 2.18.	Pedestrian fatality rate per million population in urban roads in Spain, 2018.	32
Figure 2.19.	Share of population aged 65 years and over by region, 1990-2050.	34
Figure 2.20.	Urban and rural population of the world, 1950-2050.	36
Figure 2.21.	The 17 Sustainable Development Goals.	39
Figure 2.22.	Functional capacity over the life course.	45
Figure 2.23.	Map of cities that were members of the Global Network for Age-friendly Cities and Communities in 2018.	47
Figure 2.24.	Responses of older adults to the survey regarding pedestrian habits and behavior.	56

Figure 2.25. Responses of older adults to the survey regarding their main road safety concerns as pedestrians.	56
Figure 2.26. Responses of older adults to the survey regarding proposed countermeasures to improve pedestrian road safety.	57
Figure 2.27. Classification of Machine learning algorithms.	61
Figure 3.1. Scheme of ARENA and ARENA 2 database formats.	79
Figure 3.2. Scheme of the ARENA Database.	80
Figure 3.3. Scheme of the ARENA2 Database.	81
Figure 3.4. Common SQL join types applicable to the road safety official database.	82
Figure 3.5. Examples of macro-level spatial units of analysis.	86
Figure 3.6. Process to filter vehicle-pedestrian collisions from the traffic crash database.	91
Figure 3.7. Process to filter single vehicle single pedestrian collisions from the traffic crash database.	92
Figure 3.8. Process to filter singles vehicle single pedestrian traffic crashes involving older and non-older adults from the traffic crash database.	93
Figure 3.9. Developed process for geocoding of traffic crashes at street segments.	97
Figure 3.10 Workflow with areal units of study and areal tabular input data.	100
Figure 3.11 Workflow with areal units of study and areal input data.	101
Figure 3.12 Workflow with areal units of study and linear input data.	102
Figure 3.13 Workflow with areal units of study and point input data.	102
Figure 3.14 Workflow with linear units of study and areal input data.	103
Figure 3.15 Workflow with linear units of study and line input data.	104
Figure 3.16 Workflow with linear units of study and point input data.	104
Figure 3.17 Example of data processing for linear units of study and point input data.	104
Figure 3.18. Representation of the prior, likelihood and posterior distributions used in Bayesian statistics.	115
Figure 3.19 Conceptual configuration of statistical modeling in the doctoral thesis.	117
Figure 3.20. Chronology of methodological decisions across scientific publications.	121
Figure 3.21. Process to assign each point to a street segment.	125
Figure 3.22. Infrastructure variables considered in the micro-level analysis.	125
Figure 3.23. Example of street segment with contour complexity B, and pedestrians' field of view.	127
Figure 3.24. Scheme of the buffer around each traffic crash.	129

List of Tables

Table	Title	Page
Table 2.1.	Road fatalities by country in EU27 and EFTA, 2021-2022.	19
Table 2.2.	Evolution of the number of road fatalities by age categories in EU27, 2012-2022.	20
Table 2.3.	Pedestrian fatalities in Spain and Europe, 2012-2022.....	23
Table 2.4.	Fatally, seriously and slightly injured pedestrians in urban and other road in Spain, 2021.	29
Table 2.5.	Traffic crashes involving pedestrians and pedestrian victims in urban roads in Spain, 2009-2018.	30
Table 2.6.	Fatality rate of injured road users aged 65 and over, and comparison with the rest of users.	30
Table 2.7.	Older pedestrian fatalities in urban roads in Spain, 2010-2019.....	32
Table 2.8.	Pedestrians injured considering the crash location in urban roads in Spain, 2016-2018.	33
Table 2.9.	Life expectancy at birth and at age 65 by sex and region, 2015-2019.	35
Table 2.10.	Number and proportion of population aged 65 and over, and old-age dependency ratio, 2019 and 2050.	36
Table 2.11.	Urban population, 1990, 2018, 2030, 2050.	37
Table 2.12.	Top 30 countries with the largest declines in rural population projected between 2018 and 2050.	37
Table 2.13.	Objectives of the Road safety strategy 2030 in Spain.	43
Table 2.14.	Reviewed literature on older pedestrian traffic crash profile or exploratory studies.	64
Table 2.15.	Reviewed literature on older pedestrian traffic crash injury severity.	65
Table 2.16.	Reviewed literature on older pedestrian traffic crash frequency.....	68
Table 3.1.	Main data sources used in the development of this doctoral thesis.	88
Table 3.2.	Dataset terminology in the traffic crash database extraction.	90
Table 3.3.	Examples of the parts of a street name considered in the geoprocessing of addresses.....	95
Table 3.4.	Data processing in terms of the spatial unit of study and the nature of the data.	99
Table 3.5.	Main features of the developed publications.	118
Table 3.6.	Combinations of possible types of intersections at street edges and definition of level of complexity of the road segment contour.	127
Table 6.1.	Comparison of results of the published studies on the frequency of older pedestrian traffic crashes.	240
Table 6.2.	Objectives of the thesis and their relationship with the published articles.....	245

List of Abbreviations

AADT	Average annual daily traffic
AIC	Akaike information criterion
ANN	Artificial neural networks
CC	Count component
DGT	Directorate-General for Traffic
DIC	Deviance information criterion
EFTA	European Free Trade Association
GDP	Gross Domestic Product
GIS	Geographical Information System
GLM	Generalized linear model
HHI	Herfindahl-Hirschman index
HNB	Hurdle negative binomial
IGN	National Geographic Institute
INE	National Statistics Institute
INLA	Integrated Nested Laplace Approximation
KNN	K-nearest neighbours
MCMC	Markov chain Monte Carlo
ML	Machine learning
NB	Negative binomial
NOP	Non-older pedestrian
OP	Older pedestrian
OSM	OpenStreetMap
PDF	Probability density function
PMF	Probability mass function
POI	Point of interest
RF	Random forest
SDG	Sustainable Development Goal
SQL	Structured Query Language
SUV	Sport utility vehicle
SVM	Support vector machine
WHO	World Health Organization
ZC	Zero component

1

Introduction

In this first chapter of the PhD dissertation, the Introduction of the developed work is presented. This introduction includes the **Context** inside of which this work is framed (Section 1.1), the main and specific **Objectives** (Section 1.2), the developed **Investigation phases** (Section 1.3), a clarification on Key **concepts and terminology** (Section 1.4), and, finally, the **Structure of the dissertation** is presented (Section 1.5).

1.1 Context

Road safety is a general health concern in our society. Each year, more than 1 million people die because of road traffic crashes, and roughly 50 million people are injured worldwide. In 2021, there were 1.19 estimated million road traffic deaths (World Health Organization, 2023a). Although this represents a reduction of approximately 5% compared to the 1.25 million road fatalities recorded in 2010, it still equates to an average of 3,300 lives lost daily, roughly 140 fatalities per hour, or one death every 27 seconds. It is the leading cause of death for individuals aged 5 to 29, and the 12th cause of deaths for the general population (World Health Organization, 2022). Of these road deaths, 21% are suffered by pedestrians. Apart from being a health issue, road fatalities also represent an economic impact on national budgets, as the estimated social cost of road traffic crashes is remarkable, totaling between 1% and 3% of the Gross Domestic Product (GDP). Road traffic deaths are unevenly distributed around the globe, and low- and middle-income countries constitute 92% of road traffic deaths, while accounting for 84% of the population and 72% of the vehicles worldwide.

Europe recorded the lowest rate of road fatalities globally in 2016, with 9.3 road deaths per 100,000 population, which is nearly half the global average of 18.2 road deaths per 100,000 population (World Health Organization, 2018b). The estimated cost of road crashes in Europe is 280€ billion, which is roughly 2% of EU's GDP (European Commission, 2021). In 2022, 20,634 road deaths were registered in Europe, representing a general long-term decrease since 2012 of 22%, and a general short-term decrease since 2019 of 9% (European Commission, 2024a). Regarding other injuries, in 2019

there were about 140,000 seriously injured victims, and a total of 1.2 million injured victims in 0.94 million traffic crashes in Europe.

Pedestrians are the second most present individuals among European road fatalities, only behind car occupants. In consequence, pedestrians are the vulnerable road user group with more road fatalities in Europe, constituting 18% of road fatalities. In general terms, most road deaths occur in rural roads (53%) and motorways (9%), while urban roads contain the rest (38%) of these fatalities. However, pedestrian fatalities occur in urban roads in 70% of the cases (European Commission, 2024a). Regarding the age of the pedestrians, it is notable that 47% of the pedestrians dead in traffic crashes were aged 65 and over in Europe in 2022. This proportion exceeds the average in certain countries. For instance, in Spain, 57% of the 343 pedestrian road fatalities in 2022 involved older adults.

In Spain, the scenario is similar to Europe, and each year in Spain roughly 1,500 people lose their lives in the roads. In 2023, the road fatality rate per million population in Spain was approximately 38, slightly above the European average of 46 fatalities per million population. The number of road fatalities in Spain has declined significantly over the decades, from a peak of 9,344 fatalities in 1989 to 1,680 fatalities in 2013. Since 2013, this figure has remained relatively stable on an annual basis. With respect to injury severity, the proportion of fatalities and seriously injured victims in traffic crashes has declined over the years. In 1965, fatalities accounted for approximately 5% of traffic crash victims, while 26% were seriously injured. By 2023, these figures had decreased to 1% and 7%, respectively. However, the total number of victims, particularly those with minor injuries, has shown an upward trend during this period.

Among vulnerable road users, pedestrians accounted for 39% of road fatalities in Spain in 2021 considering all roads, and 55% in urban roads. Although the mortality rate (number of fatalities by 100 traffic crash victims) is remarkably higher on non-urban roads (15.8) than in urban roads (1.8), most fatalities (61%), seriously injured (89%) and slightly injured (95%) pedestrians are recorded in urban roads, according to 2021 data. Older adults are disproportionately overrepresented among pedestrian traffic crashes, especially in urban roads. This demographic accounted for 70% of pedestrian fatalities and for 40% of seriously injured pedestrians in urban roads in 2019, even though older adults constituted 19% of the population.

Older adults are likely to have more physical and mental impairments than the rest of the population because of their age. This fact, along with their intrinsic fragility, might lead them to being more likely to be killed or seriously injured if a pedestrian-vehicle collision occurs. Moreover, recovery processes after a traffic crash are longer for older people (O'Hern et al., 2015). As pedestrians, older people show more cautious

behaviors than their younger counterparts because of their self-awareness of their limitation, which is known as self-regulation. On the other hand, physical activity, and especially walking, provides them with a wide variety of benefits that go beyond physical benefits, as they are also favored in mental and social terms, which strongly is related to their well-being.

Promoting walking among older adults is in line with the so-called process of healthy ageing promoted by WHO (World Health Organization, 2021a), which consists of “developing and maintaining the functional ability that enables well-being in older age”. Walking is a key contributor to physical activity among older adults, and it can promote the functional ability of older people so they can live their lives in a healthier way. However, as stated above, they are more vulnerable to a negative walking externality: suffering a vehicle collision.

There is a wide body of research addressing older adult traffic crashes as drivers (Tourner et al., 2016), even though this group of drivers do not pose a remarkable level of traffic crashes (Langford et al., 2006, 2008). On the other hand, older pedestrians have been more studied through observational and survey studies than through research based on traffic crash data using statistical modeling techniques. In consequence, there is a need to detect which factors influence the occurrence of older pedestrian traffic crashes in urban environments to propose countermeasures based on traffic crash data to reduce the number of these events.

Administrations should focus not only on training programs for older adults, but also on infrastructure measures to create and transform our streets towards the idea of age-friendly cities (World Health Organization, 2007). In addition, population ageing and urbanization (United Nations, 2019, 2020b), two demographic mega-trends that are especially found in developed countries, may pose a threat in the future, as both the number and proportion of older pedestrians in urban environments are expected to increase. This shift could potentially result in a higher incidence of pedestrian fatalities and serious injuries.

In this scenario of population ageing and urbanization, Spain and its capital city, Madrid, are both selected as case studies to analyze older pedestrian traffic crashes in urban environment. Spain requires attention on this demographic, as it is the fifth country in the world with the higher projected share of older adults by 2050 (United Nations, 2023). Moreover, in terms of expected growth from 2019 to 2050, Spain is the only European country among the top 10 countries with the largest projected increase in the proportion of older adults, and the third country in the world with the highest projected old-age dependency ratio by 2050 (United Nations, 2020b).

1.2 Objectives

Within this context, the **main objective** of this doctoral thesis is to analyze and assess the specific features influencing the occurrence of older pedestrian traffic crashes in urban environments. In other words, the focus is set on which factors influence the number of older pedestrian traffic crashes, in comparison with factors affecting traffic crashes suffered by the rest of pedestrians. The specific objectives of this doctoral thesis are as follows: In addition to this main objective, other complementary objectives defined for this doctoral thesis are as follows:

- Objective 1.** Identify suitable and complete data sources to gain insight into the urban built environment, conditioning road safety.
- Objective 2.** Develop a systematic and free procedure to geolocate traffic crashes considering the postal address where they were registered.
- Objective 3.** Develop dedicated tools to systematically process data on socioeconomics, land use, and infrastructure with different spatial units of analysis.
- Objective 4.** Design ad-hoc indicators of the urban built environment influencing frequency of traffic crashes, including street infrastructure and traffic variables, both at macro and micro level spatial analysis.
- Objective 5.** Develop statistical models suitable to predict the frequency and injury severity of older pedestrian traffic crashes.
- Objective 6.** Propose infrastructure countermeasures to improve the road safety scenario of older pedestrians based on the results of the statistical modeling of traffic crashes involving older adults.
- Objective 7.** Assess the effect of the COVID-19 pandemic on the frequency, injury severity and crash-related factors of older pedestrian traffic crashes compared to the pre-pandemic situation in urban scenarios.

1.3 *Investigation phases*

The development of this doctoral thesis is organized into three main phases: conceptualization, methodology, and reporting. These phases reflect the iterative nature of the research process, whereby progression from one phase to the next often involves revisiting and refining earlier stages. The workflow of the investigation is depicted in Figure 1.1, and the phases are defined as follows:

A. Conceptualization phase. The first phase was the conceptualization, which laid the foundation for defining the purposes of this doctoral thesis.

A.1. Observation. The investigation began by observing a series of social phenomena, particularly in developed countries. The ageing of the population, driven by increasing life expectancy and a decrease in the total fertility rate, has resulted in a higher proportion of older adults (people aged 65 and above) in public spaces. Older adults are particularly vulnerable road users within the already vulnerable group of pedestrians, being more likely to be killed or seriously injured in a traffic crash. In fact, the ageing population has already influenced road safety statistics concerning pedestrians. Additionally, active mobility, especially walking, offers significant physical, mental, and social benefits for older adults. These factors highlight the need to analyze the characteristics of traffic crashes involving older pedestrians in urban environments to promote healthy ageing within age-friendly cities.

A.2. Research aims and questions. The objective of this doctoral thesis was to analyze traffic crashes involving older pedestrians in urban environments. Three research questions were formulated. Questions 1 and 2 focused on identifying the factors influencing the frequency of traffic crashes involving older pedestrians, respectively. Question 3 aimed to determine whether these factors were the same before and after COVID-19. This approach not only identifies the factors affecting older pedestrians but also reveals whether these patterns are common to the rest of the population and whether some factors have opposite effects on the two groups.

A.3. Background research. A review of the existing literature on traffic crashes suffered by older pedestrians was carried out.

B. Methodological phase. The second phase is the methodological phase, which encompassed all the necessary operations to gain insight into older pedestrian crashes, from data acquisition and processing to data analysis.

B.1. Data acquisition. The first stage involved acquiring data, which included traffic crash data as the dependent variable and socioeconomic, land use, and infrastructure data as the independent variables.

- B.2. Data processing.** In the subsequent stage, the data underwent processing, which included data cleaning and wrangling. A significant task during this phase was the accurate geolocation of traffic crashes.
- B.3. Data analysis.** With the processed data, further analysis was conducted. This phase primarily included exploratory analysis, and more critically, the development of a statistical model to assess which factors influenced the frequency or injury severity of traffic crashes.
- C. Reporting phase.** The final phase involved extracting the results and conclusions derived from the methodological phase, and these results were then prepared for dissemination in academic formats, such as indexed journal articles.
- C.1. Results and conclusions.** This section presents the interpretation of the statistical models developed to identify the factors that influence pedestrian traffic crashes. Specifically, the analysis distinguishes between factors affecting older pedestrians and those that impact non-older pedestrians. The results highlight both common and distinct risk factors contributing to traffic crashes involving older pedestrians compared to the general population, providing a comprehensive understanding of the dynamics influencing these crashes.
- C.2. Diffusion.** The conclusions were disseminated in the form of scientific articles published in indexed journals.
- C.3. Further research** was proposed to address the limitations identified during the development of this doctoral thesis.

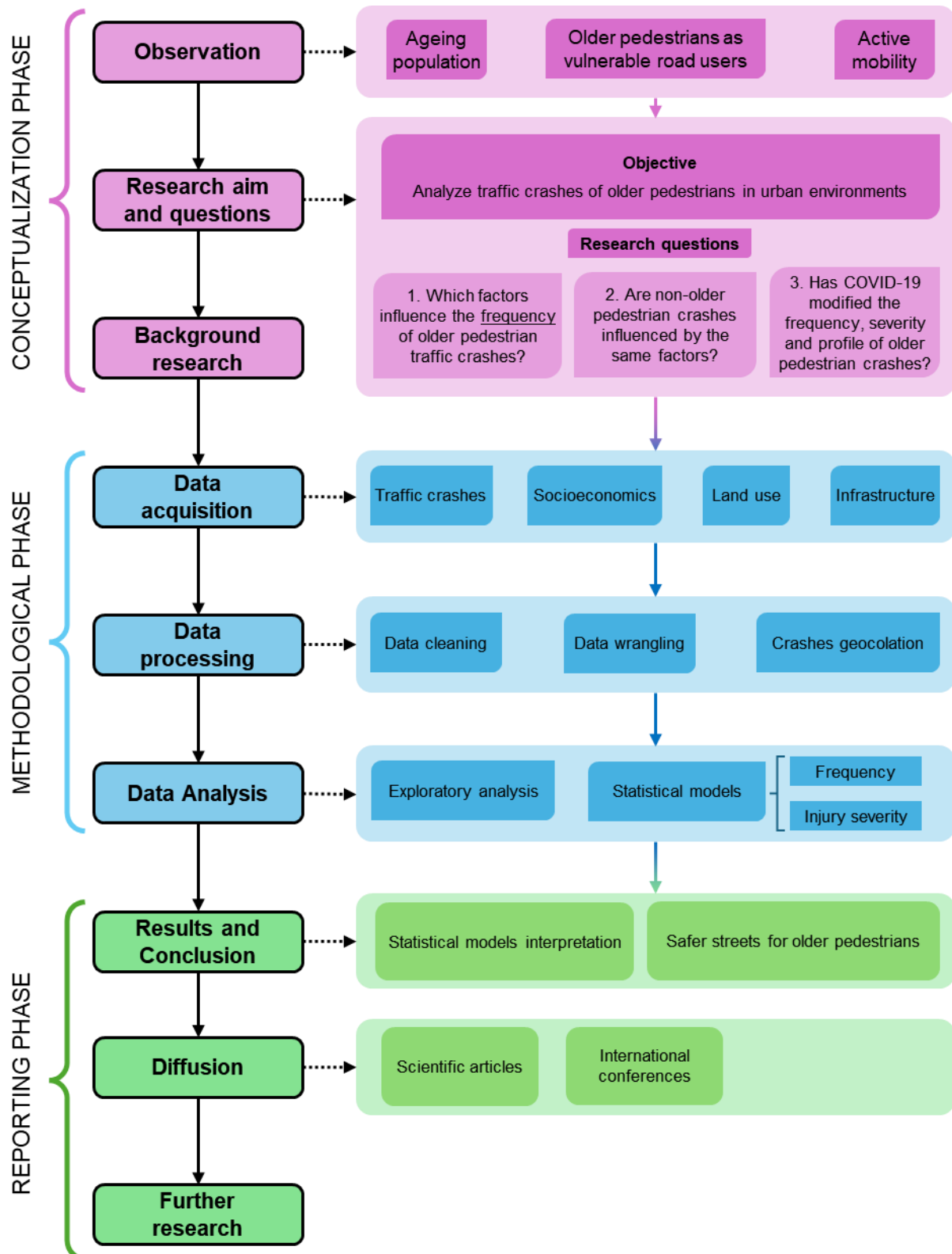


Figure 1.1. Workflow of the doctoral thesis development across conceptualization, methodology, and reporting phases.

1.4 *Key concepts and terminology*

1.4.1 Older adults

In academic research, the terminology used to categorize age groups is crucial to preventing the perpetuation of stereotypes or biases. In studies on older pedestrians, specifically those over 65, the term “older adults” is preferable to labels such as “elderly”, “aged”, or even “senior”. This distinction is important because “elderly” carries connotations of frailty and dependency, which do not accurately reflect the diverse abilities within this age group. In clinical settings, for instance, “elderly” has been associated with assumptions about physical and mental decline, which may influence treatment decisions based on age rather than individual health status (Avers et al., 2011).

In this PhD dissertation, the term “older adults” is used to highlight the diversity and autonomy of people aged 65 and over, who range from highly active to more physically dependent. By avoiding terms that carry negative connotations or are rejected by the population they describe, this research seeks to prevent the perpetuation of ageist assumptions and provide a more accurate representation of the population studied.

1.4.2 Traffic crashes

Traffic crashes or collisions are often termed “traffic accidents” or simply “accidents”. As Davis & Pless (2001) argued, the term “accident” suggests an unpredictable event or an “act of God”, fostering the misconception that such occurrences are inevitable. However, most injuries and their causes are both predictable and preventable. By avoiding the term “accident”, this PhD dissertation seeks to highlight the preventable nature of traffic crashes, in alignment with public health perspectives that advocate for precise terminology. Consequently, this thesis will consistently use the terms “traffic crash”, “crash” or “collision”, following the call to eliminate terminology that diminishes the preventability of such events.

1.5 *Structure of the dissertation*

This doctoral dissertation is structured as follows. **Section 1** Introduction provides the introduction, including the context, objectives, and investigation phases. **Section 2** Theoretical framework details the theoretical framework regarding older adult pedestrians, including their limitations, behavior, and walking benefits, the road safety situation of older pedestrians, the most usual methodological approaches in road safety studies, and the research gaps and questions found in the literature. **Section 3** Materials and Methods shows the materials and methods used during the development of this thesis, including the DGT's traffic crash database, database elaboration, and statistical models. **Section 3** concludes with the justification of the chronology of the published works, showing the reason leading to one work from the previous one. **Section 4** Publications includes the scientific contributions. **Section 5** Results outlines the results of the doctoral thesis. Finally, **Section 6** General Discussion presents the integrated solution, the main conclusions of the developed works and the avenues for future research.

2

Theoretical framework

The theoretical framework of this doctoral thesis covers wide topics to more specific ones. First, general definitions and ideas on road safety are exposed in the **Section 2.1 A general understanding of road safety**. Second, the most relevant figures on road safety worldwide, in Europe and Spain are presented in **Section 2.2 Road safety situation and evolution**. After this broad theoretical framework on road safety, the problem of older pedestrian road safety in urban environments is presented through the Main figures of Road safety of older pedestrians in urban roads (**Section 2.3**). The main Road safety strategies and their implications for older pedestrians are presented (**Section 2.4**). The importance of walking for older adults, along with their limitations and differentiate behaviors when walking, are shown in (**Section 2.5**). Third, the main approaches used in scientific road safety studies are briefly explained in **Section 2.6 Methods and approaches in studies with traffic crash data**. Finally, the main Studies on older pedestrians using traffic crash data are presented and summarized (**Section 2.7**).

2.1 A general understanding of road safety

2.1.1 Traffic crash definition

An injury **traffic crash**, road crash, or simply a crash, is defined as an event “**involving at least one road vehicle** in motion on a public road or private road to which the public has right of access, resulting in at least one injured or killed person” (UNECE et al., 2019). The definitions often include distinctions based on the severity of the crash (e.g., fatal, injury, property damage only) and the parties involved (e.g., motor vehicles, pedestrians, cyclists). In the mark of this doctoral thesis, it is important to note that **a pedestrian fall is not recorded as a road crash**, even if it occurs in public spaces due to the road infrastructure, for example, because of a cracked sidewalk.

Once traffic crashes have been defined, these crashes can be categorized considering different criteria, but the most common categorizations depend on the results, the

number of vehicles, the location, and the crash mechanism. A scheme of this classification is presented below.

2.1.2 Types of traffic crashes

Depending on the results. Traffic crashes can be categorized considering the highest injury severity level of the individuals involved in the crash.

- Fatal traffic crashes. Traffic crashes resulting in at least one fatal victim.
- Traffic crashes with injured victims. Traffic crashes resulting in at least one injured victim and not including any fatal victim. In terms of the injury severity level of the victims, these can be further categorized into:
 - Traffic crashes with seriously injured victims. Traffic crashes resulting in at least one seriously injured victim, and no fatal victims.
 - Traffic crashes with seriously injured victims. Traffic crashes resulting in at least one seriously injured victim, and no injuries of a higher level.
- Property damage only traffic crashes. Traffic crashes resulting in no injured victims.

Depending on the location. Traffic crashes can be categorized in terms of the location where they were registered.

- Urban crashes. Crashes registered on urban roads, inside of built-up areas.
- Non-urban or inter-urban crashes. Crashes registered on non-urban roads, such as rural roads or motorways, outside of built-up areas.

Depending on the crash mechanism. Traffic crashes can be categorized in terms of how they occur.

- Collision. A crash where two or more vehicles collide.
- Vehicle-pedestrian collision. A crash where a vehicle strikes a pedestrian.
- Run-off-road crash. A crash where a vehicle leaves the roadway.
- Rollover. A crash where a vehicle tips over onto its side or roof.

- Vehicle-fixed object collision. A crash where a vehicle hits a stationary object, such as a tree or barrier.

Depending on the number of vehicles. Traffic crashes can be categorized in terms of the number of vehicles involved.

- Single vehicle crashes. Crashes involving only one vehicle.
- Multiple crashes. Crashes involving more than one road user apart from a vehicle, either pedestrians or vehicles.

2.1.3 Traffic crash concurrent factors

Apart from traffic crash classification, it is important to underline that the generation of road crashes is usually due to concurrent factors contributing simultaneously. In this sense, researchers have studied, grouped and quantified the different factors in traffic crashes. Concurrent factors are all the possible characteristics that influence the generation of traffic crashes. Three groups of concurrent factors are commonly identified in literature (Treat et al., 1979): human, roadway and environmental, and vehicle factors.

- **Human factors.** The human factors are the most frequent factor in traffic crashes, and they are present in more than 90% of crashes. This factor includes human factor direct causes (recognition errors, decision errors, performance errors, critical non-performances and intentionality), and human condition or state (alcohol, drugs, fatigue, driver inexperience, in-hurry, emotional upset, vehicle unfamiliarity, pressure from other drivers, road area unfamiliarity, and reduced vision).
- **Roadway/environmental factors.** The roadway factors are the second most frequent in traffic crashes, being present in about 30% of crashes. These factors deal with the infrastructure, including view obstruction, slick roads, special hazards, design problems, control hindrances, inadequate signals, avoidance obstructions, ambient vision limitations, maintenance problems, and camouflage effect.
- **Vehicle factors.** These factors are the least common ones in traffic crashes, and they are present in about 10% of crashes. These factors include brake systems, tires and wheels, communication systems, body and doors, power train,

suspension problems, and driver seating and controls. Notably, this factor is influenced by the development of vehicle technology.

The crash results in terms of the final location of vehicles and victims involved together with the vehicles damage and the injury severity of victims, can be explained through the study of the crash phases.

2.1.4 Traffic crash phases

Although brief, a traffic crash is a dynamic process that unfolds over time and space, marked by specific points or zones and moments where events take place. The intersection of a particular moment and point is called a position, which defines a phase of the crash.

Perception phase: The perception phase includes the moment since the driver or pedestrian realizes the hazard to the moment prior to the driver/pedestrian making an action to avoid the traffic crash or minimize its consequences. This phase includes two positions or points:

- Real Perception Position (RPP). Actual point or position in which the driver perceives the hazard for the first time.
- Possible Perception Position (PPP). Point or position in which a driver without “special characteristics” should perceive the hazard for the first time.

Decision phase: The decision phase is the phase in which the driver or pedestrian reacts to the abnormal circumstance by executing the evasive maneuver that will be carried out to avoid the crash. The evasive maneuvers are classified as follows:

- Simple passive (e.g., flashing the lights).
- Simple active (e.g., slowing down, stopping the vehicle, accelerating, turning).
- Complex (i.e., combinations of the above).

Once real perception position is achieved, the decision phase is defined by the decision point (DP) and the critical point (CP).

- The decision point (DP) is the moment when the involved party begins to perform the evasive maneuver.
- The key point (KP) is the moment when the traffic crash is no longer avoidable.

Conflict phase: The conflict phase is the final period of the crash' progression, and it occurs between the critical point (CP) and the final position (FP). The following can be distinguished:

- Conflict zone: This is the area where the highest likelihood of the crash occurring exists.
- Point of impact: This belongs to the conflict zone and is the moment when the crash occurs.
- Final position: This is the stationary position that vehicles, people, and objects assume after the incident has occurred.

Before reaching the final position, there is still a possibility to alter post-collision trajectories and minimize the consequences.

2.1.5 Traffic crash frequency and severity

There is a direct relationship between incidents and crashes, such that when the former occurs, there is a likelihood that crashes will occur under similar circumstances. It could be said that there is a continuum in the severity of events that occur during driving, as indicated by the so-called Hyden's Pyramid (Hydén, 1987), whose philosophy suggests a succession of events ranging from frequent and minor incidents to serious and fatal crashes that represent the highest severity. The base of the pyramid are events that pose a risk to the safety of the vehicles involved but do not disrupt their movement. From there, a series of increasingly severe phenomena follow (potential, minor, severe and fatal conflicts), although they become less frequent as the severity increases.

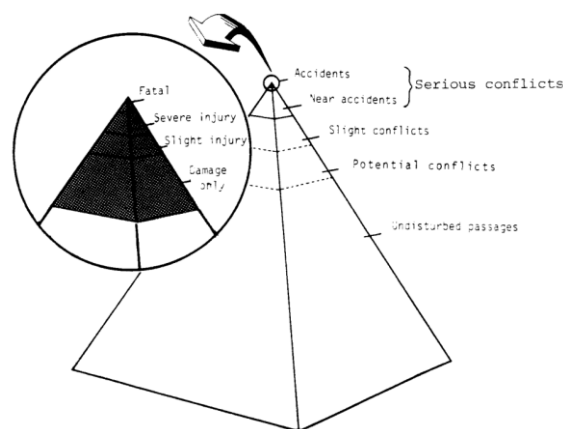


Figure 2.1. Hyden's Pyramid.

Source: Hydén (1987).

2.1.6 Traffic crash underreporting

Underreporting is a common issue of road safety analysis, which consists of not all traffic crashes being reported because of different reasons. It has been proven that less severe crashes (i.e., crashes with only slightly injured people, or property damage only crashes) show higher underreporting rates than those with more severe outcomes (Abay, 2015; A. Watson et al., 2015). This could be due to fatal crashes requiring death certificates and crashes with seriously injured requiring remarkable medical assistance and hospitalization, apart from the respective insurance requirements. Conceptually, and considering the Hyden's pyramid explained above, severe crashes are less common than slight or property damage only crashes, but these present higher reporting rates.

2.2 Road safety situation and evolution

2.2.1 Road safety situation and evolution worldwide

In 2021, roughly 1.19 million people lose their lives because of traffic crashes (World Health Organization, 2023a). This represents a decrease in total road fatalities of 5% (Figure 2.2), even though the population has increased by 13%. This reduction was observed in 108 countries. These deaths represent a huge social cost, which is estimated to be between 1% and 3% of the Gross Domestic Product (GDP), and in some cases it can reach 6% (World Health Organization, 2023a). Road deaths represented 29% of total injury-related deaths worldwide in 2019 (World Health Organization, 2022). For the general population, road traffic crashes were the 12th cause of death worldwide in 2019. Even more, for people aged 5 to 29, road traffic injuries represent the top injury-related cause of death.

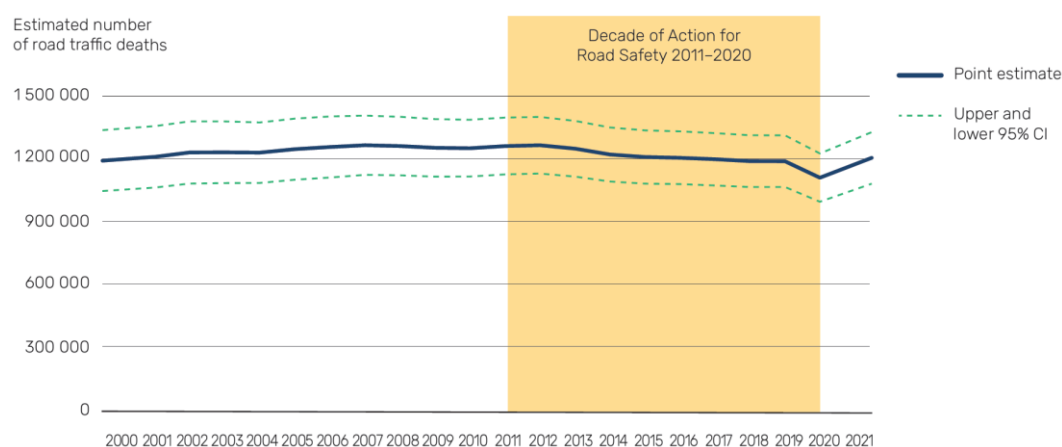


Figure 2.2. Evolution of the number of road traffic fatalities in the World, 2000-2021.

Source: World Health Organization (2023a).

Regarding the mode of transport, 4-wheel vehicles represented 25%, 2/3 powered wheelers represented 30%, bicyclists represented 5%, and pedestrians represented 21% of total roads deaths in 2021 (World Health Organization, 2023a). The distribution of crashes regarding the user type is not homogenous in the world regions. For example, 26% of road fatalities were pedestrians in Europe, while in Africa this figure was 33% for the same period.

The distribution of crashes is not homogenous in all world regions, and 92% of deaths occurred in low- and middle-income countries during 2021 (Figure 2.3). These countries represent 84% of the total population and 72% of total powered vehicles worldwide. A

key figure is that, on the other hand, high-income countries include 88% of total paved roads. In fact, in low-income countries, the risk of death is three times higher than in high-income countries, even though the first ones have less than 1% of all motor vehicles.

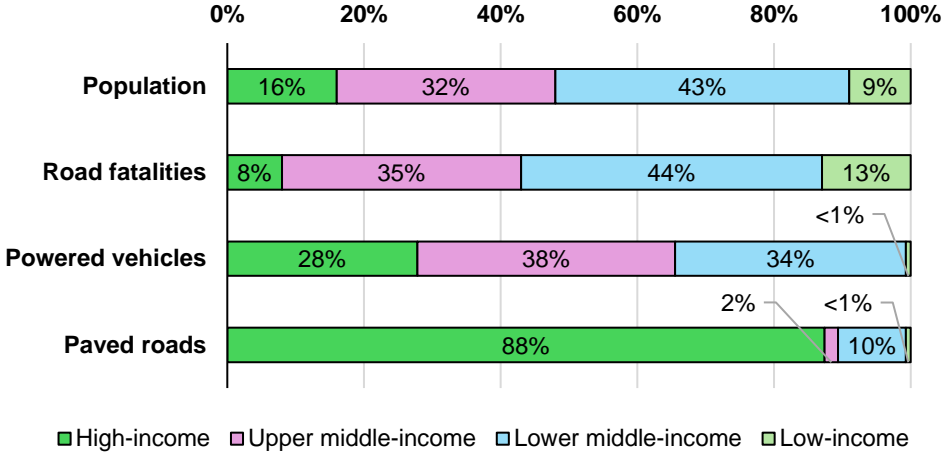


Figure 2.3. Share of population, road fatalities, motor vehicles, and paved inter-urban roads by country income level, 2021.

Source: World Health Organization (2023a)

Worldwide, the **fatality rate by population** (number of deaths by 100,000 population) has decreased from 18 in 2010 to 15 in 2021, representing a **decrease** of roughly 16% of the road traffic deaths rate. In addition, the **fatality rate by vehicle** (number of deaths by 100,000 vehicles) has decreased from 79 in 2011, to 47 in 2020, representing a decrease of 41%, even though the number of vehicles has increased by 160%(World Health Organization, 2023a).

By region, Europe showed the lowest value compared to the rest of regions in 2016, with 9.3 road deaths by 100,000 population, while the average world value was 18.2 (World Health Organization, 2018b). Between 2010 and 2021, Europe shows the highest decrease in traffic crash fatalities (36%). Other regions showing traffic death decreases were the Region of the Americas (0.1%), the South-East Asian Region (2%), and the Western Pacific Region (16%). In contrast, road traffic showed an increase in the Africa Region (17%).

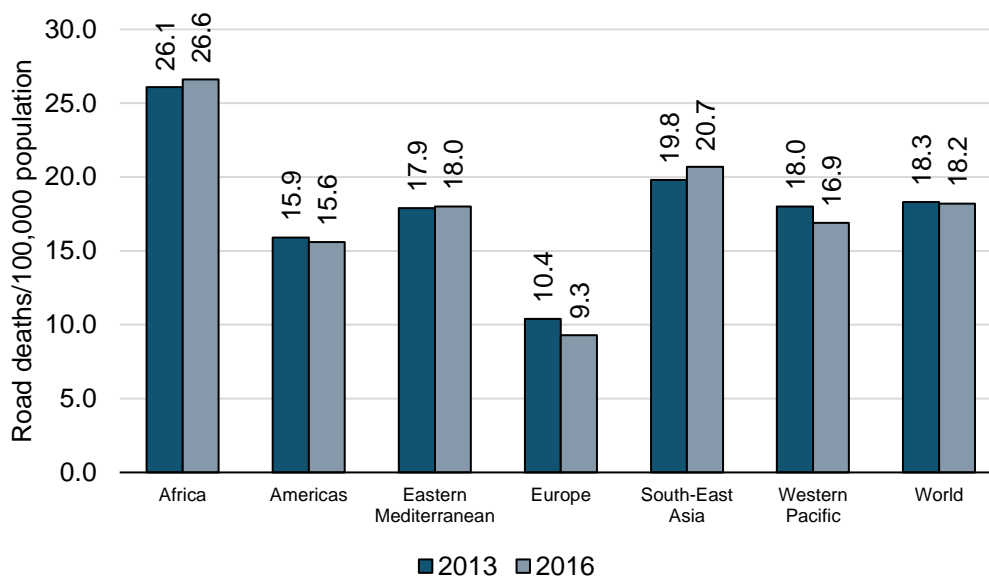


Figure 2.4. Rates of road fatalities per population by region, 2013 and 2016.

Source: World Health Organization (2018b).

2.2.2 Road safety situation and evolution in Europe

The number of road deaths in Europe during 2022 was 20,634 (Table 2.1), representing a general long-term decrease since 2012 of 22%, and a general short-term decrease since 2019 of 9%. Spain presented 1,746 road deaths in that year, having a long- and short-term decreases of 8% and 1%, respectively, below the European situation.

Table 2.1. Road fatalities by country in EU27 and EFTA, 2021-2022.

Source: Adapted from European Commission (2024b).

Country	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	LT	ST
Spain	1,902	1,680	1,688	1,689	1,810	1,830	1,806	1,755	1,370	1,533	1,746	-8%	-1%
EU	26,500	24,226	24,136	24,358	23,808	23,392	23,328	22,756	18,835	19,917	20,634	-22%	-9%

*LT = Long term change of last available year over 2012; *ST = Short term change of last available year over 2019.

Regarding road users, most road fatalities were suffered by males (77%). About the age of traffic crash victims, Table 2.2 shows the evolution of the number of road fatalities by age categories in EU27 from 2012 to 2022. The highest decrease in road fatalities since 2012 can be recorded for road users aged 15 to 17 years (39%) and 18 to 24 years (38%), while a smaller reduction occurs in age groups 65 to 74 years (1%) and 75 to 84 years (19%), and an increase in road fatalities for people aged 85 and over (27%).

Table 2.2. Evolution of the number of road fatalities by age categories in EU27, 2012-2022.

Source: European Commission (2024b).

Age Category	2012	2019	2020	2021	2022	LT	ST
< 15	672	489	384	482	450	0	0
15 - 17	703	476	412	445	429	-39%	-10%
18 - 24	3,850	2,733	2,313	2,414	2,388	-38%	-13%
25 - 49	9,704	7,650	6,346	6,666	6,620	-32%	-13%
50 - 64	5,096	4,815	4,096	4,364	4,620	-9%	-4%
65 - 74	2,629	2,710	2,303	2,384	2,612	-1%	-4%
75 - 84	2,780	2,577	2,028	2,090	2,242	-19%	-13%
85 +	897	1,160	836	941	1,138	27%	-2%
Total	26,332	22,610	18,718	19,786	20,499	-22%	-9%

*LT = Long term change of last available year over 2012; *ST = Short term change of last available year over 2019.

Regarding the location of the crashes, almost 60% of all fatalities were car occupants on motorways, where cars are the dominant mode of transport, in 2022. Conversely, on urban roads, where all modes of transport are more equally represented, the share of pedestrian fatalities among urban road fatalities is the highest (33%). Together with cyclists (15%) and powered two-wheelers (20%), the vulnerable road users constitute almost 70% of the fatalities on urban roads.

2.2.3 Road safety situation and evolution in Spain

The Directorate-General for Traffic (in Spanish: Dirección General de Tráfico, DGT) is a component of the Spanish Department of the Interior responsible for the execution of the government's road policy in the Spanish road transport network. In addition, one of the functions of DGT is the planning, development, and dissemination of statistics, indicators, and data related to traffic crashes. In other words, DGT collects and curates traffic crash data at the national level, coordinating with all the Spanish autonomous communities.

With the collected information on traffic crashes, DGT publishes each year official reports on traffic crashes in Spain, the main document is named "Main figures on road traffic crashes", published each year since 2004. In addition, each year a Statistical Yearbook of Traffic Crashes is released, containing numerous tables regarding traffic crashes and victims in Spain.

The evolution of road fatalities in Spain (Figure 2.5) shows an increasing trend from 1960 to 1989, when an historical maximum value of 9,344 road deaths was reported.

Since that year, the number of road deaths shows a decreasing trend because of the implementation of road safety policies, and the improvement of the infrastructure and vehicles. This trend of traffic crash fatalities ended in 2013, with 1,680 fatalities. From this year the registered number of fatalities has remained roughly constant. During 2023, the last complete year with available data, 1,806 people lost their lives in Spain because of a traffic crash, 60 more than in 2022 (or an increase of 3%), and 9,265 people were hospitalized because of a traffic crash, 733 more than in 2022 (or an increase of 9%). Nevertheless, the pandemic could play a role in these short-term increases in the road fatalities.

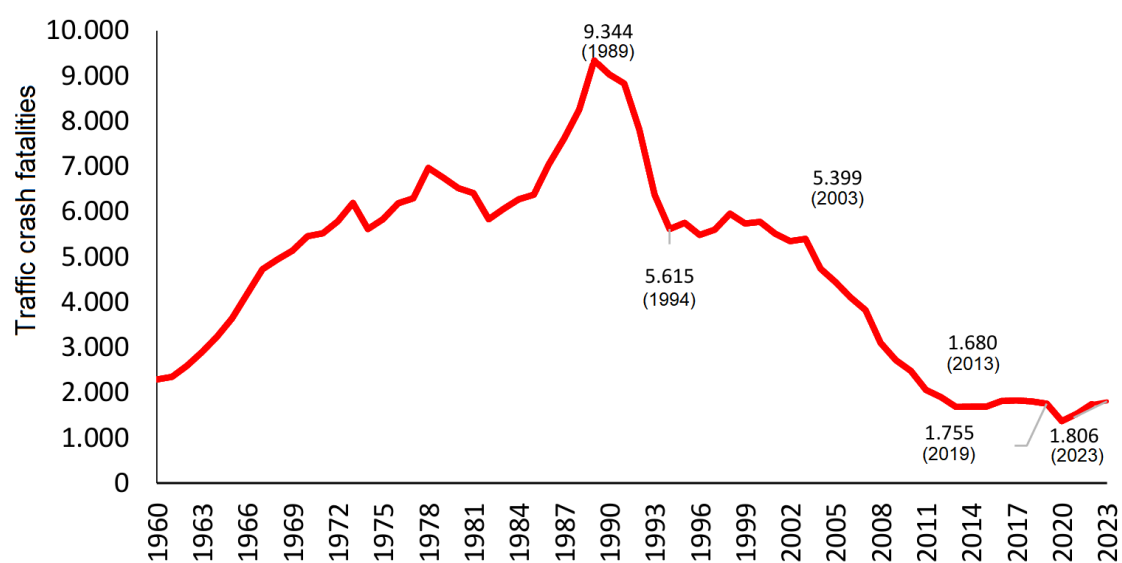


Figure 2.5. Evolution of road fatalities in Spain, 1960-2023.

Source: Adapted from Dirección General de Tráfico (2023).

Adjusting these figures by population, Spain registered 38 road deaths per million inhabitants during 2023 (Dirección General de Tráfico, 2023). This figure is below the European average (46), and Spain is in the 9th place of the safest European countries regarding traffic crashes. Notably, the number of road fatalities by population increased in one unit from 2019 (37) to 2023 (38) in Spain.

Regarding the **evolution of the number of victims considering the injury severity** in Spain, traffic crash data has distinguished between hospitalized and non-hospitalized victims since 1962. In 1965, fatalities accounted for 5%, hospitalized injuries for 26%, and non-hospitalized injuries for 68%. By 2003, the proportions shifted further, with fatalities decreasing to 3% and hospitalized injuries to 17%. In 2023, among the 135,337 traffic crash victims, 1% were fatalities, 7% were hospitalized, and 92% sustained non-hospitalized injuries. These trends indicate a significant reduction in the number and proportion of road deaths and seriously injured victims since 1960,

accompanied by an increase in the number and proportion of slightly injured victims. Consequently, the mortality rate of traffic crashes, defined as the number of fatal victims per 100 victims, has significantly decreased over the years. Specifically, it declined from 5.2 in 1993 to 1.3 in 2023, as illustrated in Figure 2.6.

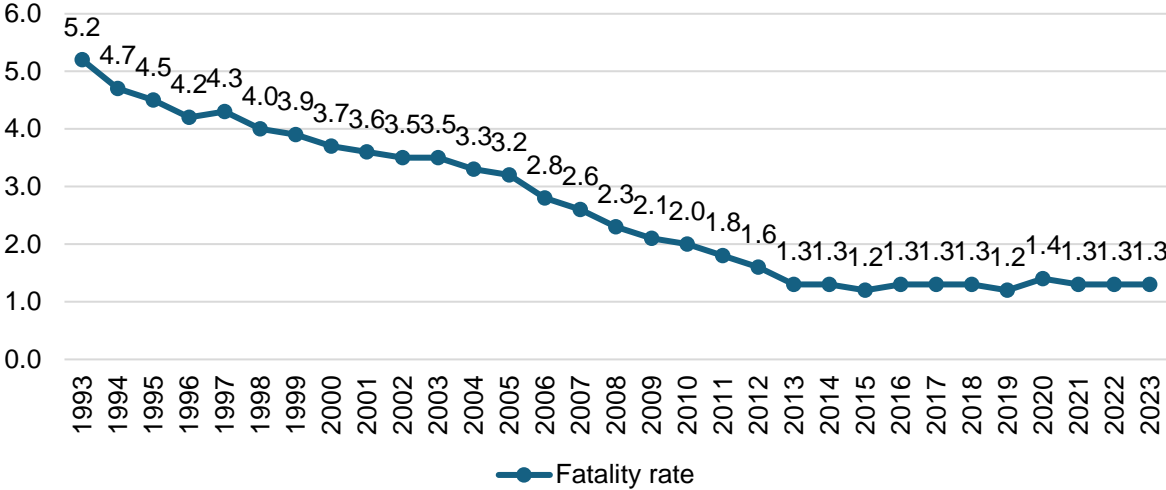


Figure 2.6. Evolution of the mortality rate in road traffic crashes in Spain, 1993-2023.
Source: Adapted from Dirección General de Tráfico (2022a, 2023).

2.3 Road safety of older pedestrians in urban roads

Once road safety generalities, the main trends in the world, Europe and Spain, and general concepts on studies with traffic crash real data have been acknowledge, in this section, the specific figures and trends on traffic crashes involving older pedestrians (i.e., pedestrians aged above 65 years old) on urban roads in Europe and Spain are presented. This is a triple-scope approach, as it is the intersection of traffic crashes involving pedestrians, older adults, and urban roads (Figure 2.7).

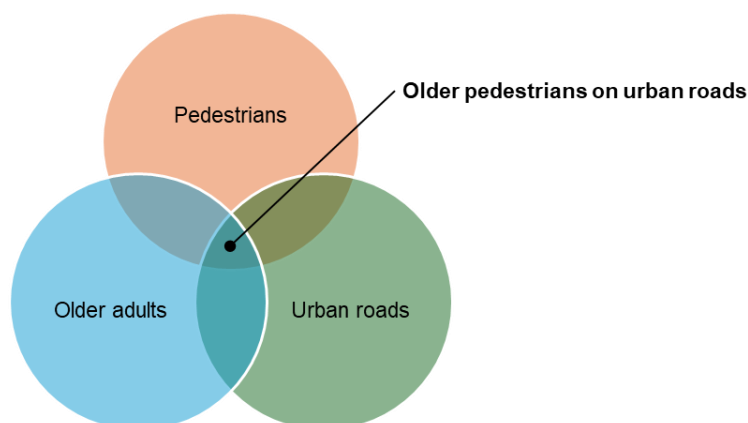


Figure 2.7. Triple scope of the analysis: older adults, pedestrians, and urban roads.

2.3.1 Older pedestrian situation in Europe

Pedestrian fatalities are a significant road safety concern in Europe, especially for older adults, who are disproportionately represented among pedestrian victims. In 2022, across the EU, there were 3,740 pedestrian fatalities, a decrease of 31% compared to 2012. As in the case of all traffic fatalities, Spain presents a smaller decrease in pedestrian fatalities since 2012 (6%) compared to Europe (31%; Table 2.3). However, despite the overall decline, people aged 65 and older accounted for 47% of all pedestrian fatalities in 2022 (Figure 2.8), although older adults represented 21.2% of the total population (European Commission, 2024c).

Table 2.3. Pedestrian fatalities in Spain and Europe, 2012-2022.

Source: Adapted from Facts & Figures: Pedestrians, 2024.

Country	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	LT*	ST*
EU	5,410	5,285	5,227	4,990	4,957	4,854	4,758	4,623	3,604	3,585	3,740	-31%	-19%
Spain	370	371	336	367	389	351	386	381	260	301	348	-6%	-9%

*LT = Long-term change of last available year over 2012; ST = Short-term change of last available year over 2019.

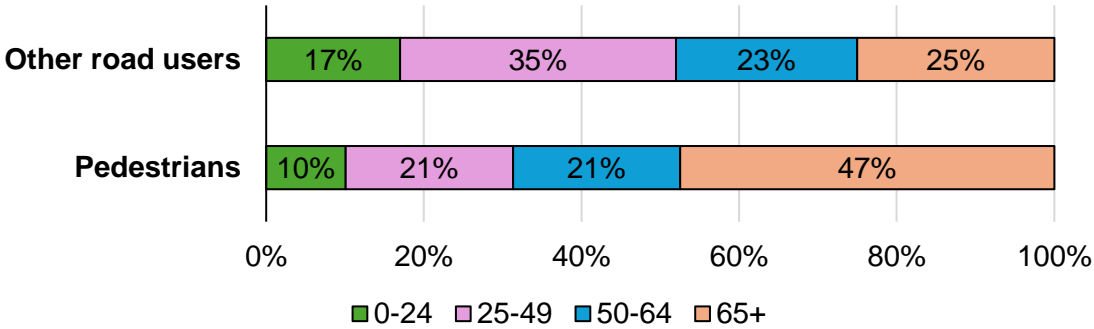


Figure 2.8. Pedestrians and other road user fatalities in traffic crashes in Europe considering the age of the victim, 2022.

Source: European Commission (2024c).

Urban roads registered 38% of total traffic crashes in 2022 in Europe (European Commission, 2024e). Nevertheless, 69% of pedestrian deaths occurred in urban areas, compared to 31% on rural roads and motorways (European Commission (2024c); see Figure 2.9). The concentration of pedestrian fatalities in urban areas is driven by the higher exposure of pedestrians in cities and the increased interaction between motor vehicles and pedestrians. Vulnerable road users (pedestrians, cyclists, motorcyclists) constitute almost 70% of all fatalities on urban roads, in which pedestrians are the most important contribution to traffic deaths on urban roads, constituting 33% of total urban road deaths (European Commission, 2024e).

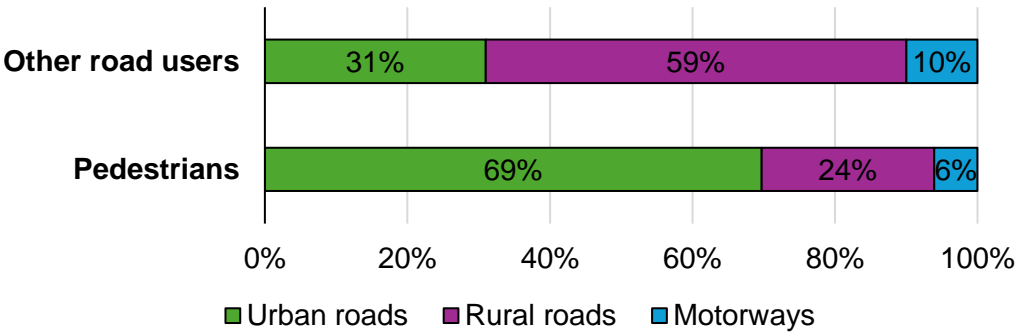


Figure 2.9. Pedestrians and other road user fatalities in traffic crashes in Europe considering the road type, 2022.

Source: European Commission (2024c).

Older adults represented 29% of total road deaths in Europe in 2022 (see Figure 2.10), a figure that has been increasing from 17% in 1992 to 29% in 2022 (European Commission, 2024f). Furthermore, older adults have the second highest mortality rate among all age groups, with an average of about 64 fatalities per million inhabitants in

road traffic within the EU27 (European Commission, 2024f). Older adults are more likely to be involved in crashes on urban roads (51%), which is notably higher than the proportion for people aged under 65 (33%). In fact, this demographic constitutes 39% of all road deaths on urban roads, and this figure is almost half (23%) for other roads (rural and motorways).

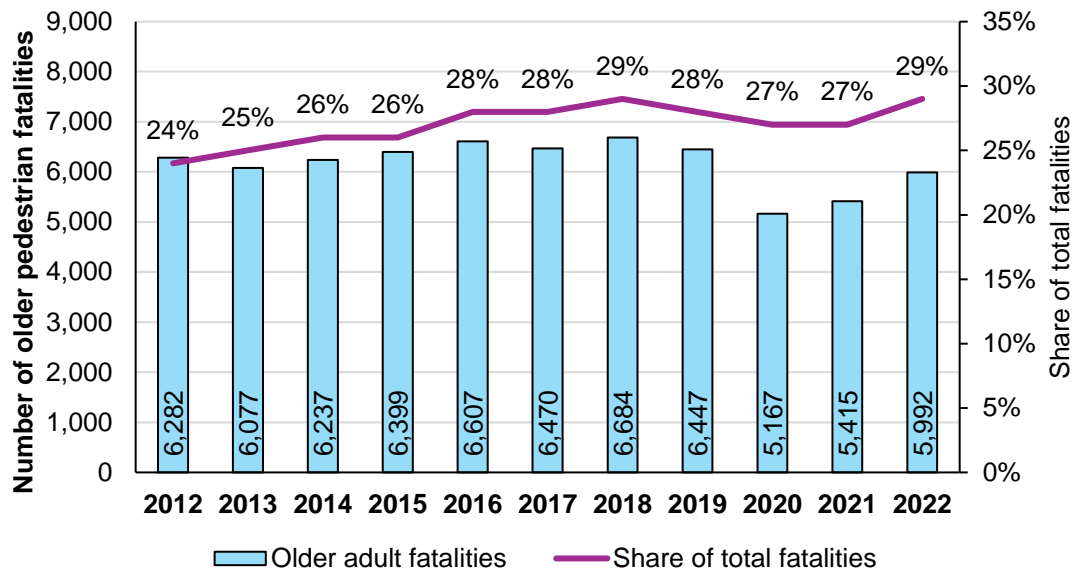


Figure 2.10. Evolution of older adult road fatalities and share in the total number of fatalities in Europe, 2012-2022.

Source: European Commission (2024d).

For older adults, the risk of fatality in traffic crashes is considerable across all transport modes. However, as pedestrians, their vulnerability is most pronounced and the high proportion of pedestrian fatalities among older adults is noticeable. In 2022, almost one third of all fatalities among older adults were pedestrians (29%, Figure 2.11). In terms of gender, females are more likely to be killed as a pedestrian than males through all age groups. The difference is most pronounced for older adults, as 43% of the female road fatalities were pedestrians, and 22% of the male road fatalities were pedestrians in 2022 (European Commission, 2024f).

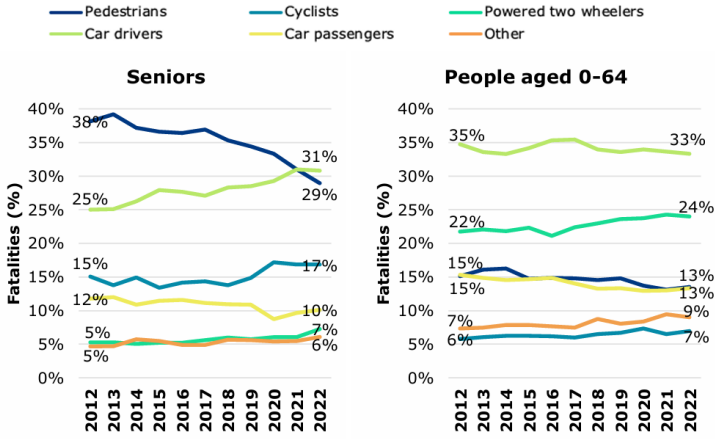


Figure 2.11. Distribution of older and non-older adult fatalities considering the transportation mode in Europe, 2022.

Source: European Commission (2024f).

Regarding pedestrian fatalities, around 1 in 2 pedestrian fatalities (47%) in 2022 were older adults (Figure 2.12), and this figure has increased since 2012 (44%). The age-related increase in pedestrian fatalities is evident when comparing different age groups: the mortality rate for older pedestrians is more than double that of adults aged 25-64 and almost five times higher than for those under 25 (European Commission, 2024c).

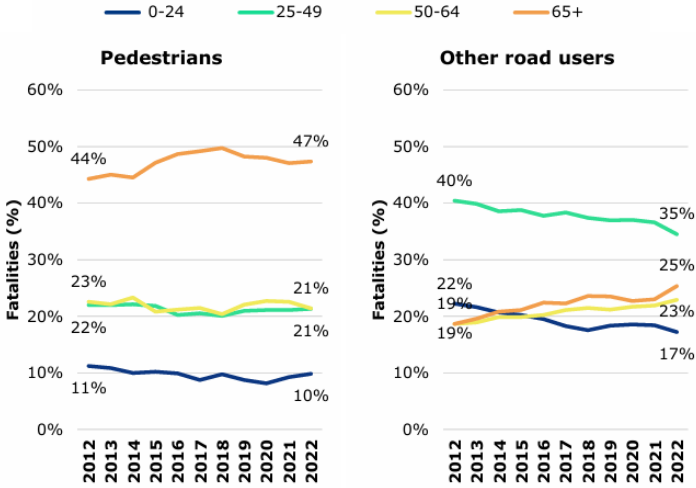


Figure 2.12. Evolution of the distribution of pedestrians and other road user fatalities considering the age of the victim in Europe, 2012-2022.

Source: European Commission (2024c).

The impact of the COVID-19 pandemic temporarily reduced the overall number of pedestrian crashes due to decreased traffic volumes, but it did not substantially alter the

long-term risks faced by older pedestrians. In fact, in some EU member states, pedestrian fatalities increased during the pandemic due to changes in traffic behaviour, such as increased speeding (European Commission, 2024c).

Older pedestrian traffic crashes on urban roads in Europe are a significant public health concern. The high fatality rates among older adults, particularly in urban areas, highlight the need for targeted road safety measures. Vulnerable road users, including pedestrians, constitute a substantial portion of urban road fatalities, and older adults are especially at risk due to their physical vulnerability. Improving urban infrastructure, enhancing pedestrian safety features, and implementing stricter speed management in cities are critical steps to protect older pedestrians and reduce fatalities across Europe.

2.3.2 Older pedestrian situation in Spain

As in the case of all traffic crashes, Spain registers a similar situation regarding older pedestrian traffic crashes to Europe. This section examines the current figures and trends of pedestrian traffic crashes involving older adults in Spain's urban areas, drawing upon data from official reports.

To provide context, traffic crash statistics from 2019 are examined next, although the trends discussed have remained consistent in subsequent years. In 2019, 1,755 individuals died because of traffic crashes. Most of these fatalities occurred on inter-urban roads (i.e., outside built-up areas), where higher vehicle speeds were a contributing factor, accounting for 1,236 deaths (70.4%). Of all fatalities, 21.7% (381) were **pedestrians**, and 64.8% (247) of these **pedestrian fatalities on urban roads**. Even though the highest number of overall traffic fatalities occurred on highways or inter-urban roads, **most pedestrian fatalities were concentrated in urban areas**.

Older adults represent a particularly vulnerable group. Figure 2.13 shows a detailed scheme as part of the exploratory analysis developed for this doctoral thesis. In total, 492 older adults lost their lives in traffic crashes in 2019, constituting 28% of all fatalities, although they make up only 19% of the population. Among them, 217 were pedestrians, representing 57.0% of all pedestrian fatalities. Notably, most of these **older pedestrians** were killed in **urban roads** (172 deaths, or 79.3% of the total older pedestrian fatalities), a proportion higher than the overall rate of pedestrian fatalities on urban roads (64.8%). Furthermore, **older adults represented a significant proportion of pedestrian deaths in urban areas, accounting for 172 out of 247 (69.6%) pedestrian fatalities in urban roads**. These statistics are illustrated in Figure 2.13.

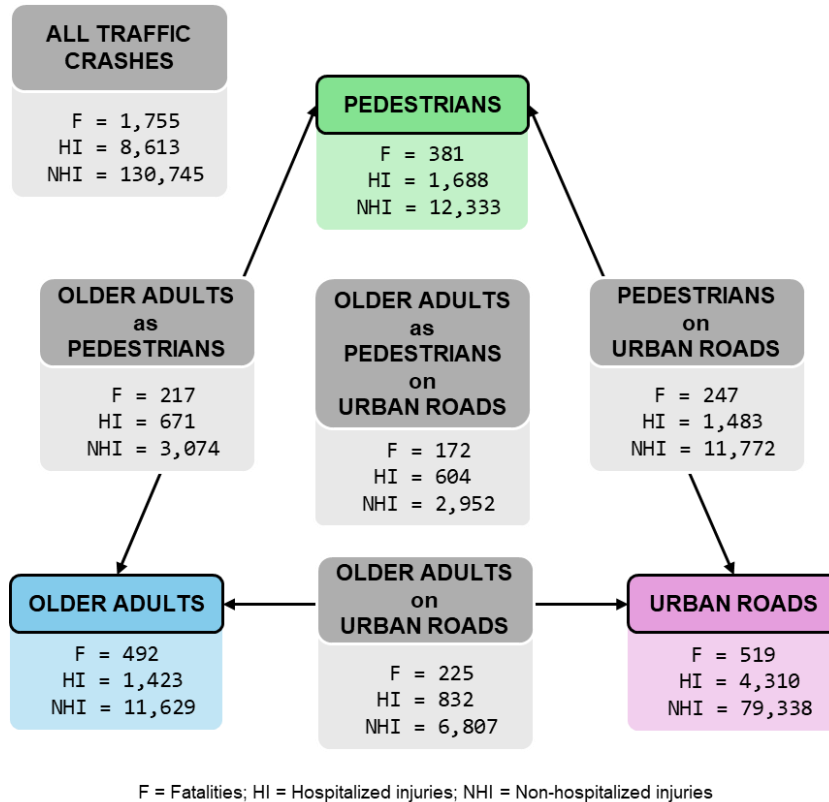


Figure 2.13. Number of fatalities, seriously and slightly injured pedestrians, older adults, and on urban roads in traffic crashes in Spain, 2019.

Among vulnerable road users (i.e., pedestrians, cyclists, and motorcyclists), pedestrians are particularly at risk in urban environments. In 2021, pedestrians accounted for 39% of all vulnerable road user fatalities across Spain. However, in urban areas, this percentage rose to 55%, reflecting the heightened risks faced by pedestrians (Figure 2.14).

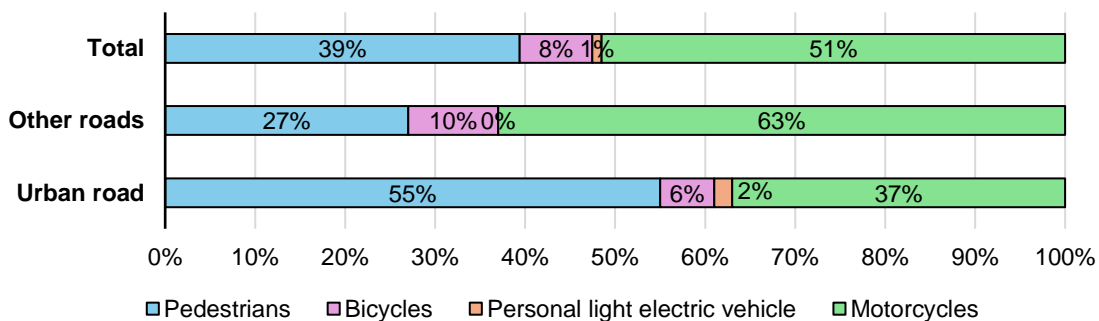


Figure 2.14. Distribution of vulnerable road user fatalities in urban and other roads considering the transport mode of the victim in Spain, 2021.

Source: Dirección General de Tráfico (2021b).

Regarding pedestrians and road type, Table 2.4 shows the number of fatally, seriously and slightly injured pedestrians in urban and other road in Spain in 2021. Most pedestrian deaths and injuries occur on urban roads (61%, in 2021), and although non-urban roads present a much higher fatality rate (15.8 vs 1.8) due to higher speeds, most pedestrian deaths, hospitalized injured and non-hospitalized injured are still registered in urban roads. The likelihood of survival may be higher in urban environments, but the frequency of fatal crashes involving pedestrians remains alarmingly high.

Table 2.4. Fatally, seriously and slightly injured pedestrians in urban and other road in Spain, 2021.

Source: Dirección General de Tráfico (2021b).

Road type	Pedestrian fatalities		Hospitalized injured pedestrians		Non-hospitalized injured pedestrians		Total #	Mortality index
	#	%	#	%	#	%		
Urban roads	183	61%	1,237	89%	8,979	95%	10,399	1.8
Other roads	118	39%	160	11%	467	5%	745	15.8
Total	301	100%	1,397	100%	9,446	100%	11,144	2.7

Regarding the **evolution of pedestrian crashes**, no significant differences in the number of killed or hospitalized pedestrians because of a traffic crash can be observed since 2012 (Figure 2.15). It is clearly shown that most of these victims are registered on urban roads, inside of built-up areas, this difference is more pronounced for the hospitalized injured pedestrians. Note that the effect of the COVID-19 pandemic is clearly shown in the figure, in which there is a notable decrease in the number of victims in 2020 and 2021. The effect of the pandemic is higher in urban areas, so the proportion of non-urban crashes increased during these two years.

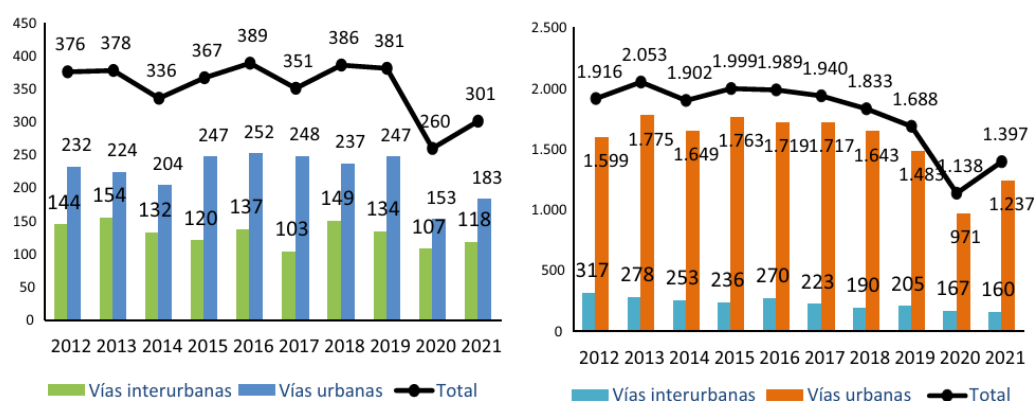


Figure 2.15. Evolution of pedestrian fatalities and seriously injured pedestrian in urban and other roads in Spain, 2012-2021.

Source: Dirección General de Tráfico (2021b).

The evolution of the proportion of crashes with pedestrians over all crashes in urban road shows no remarkable changes since 2009, since when the share of pedestrian crashes is about 20% in Spain (Table 2.5). The same applies to the share of pedestrian fatalities, seriously and slightly injured by all injured road users, with a share of pedestrians of about 20% of total fatalities, 35% of total seriously injured and 15% of all slightly injured people (Table 2.5).

Table 2.5. Traffic crashes involving pedestrians and pedestrian victims in urban roads in Spain, 2009-2018.

Source: Dirección General de Tráfico (2018).

Concept	Measure	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Traffic crashes	Pedestrians	9,080	9,112	9,543	9,548	10,721	11,190	12,797	13,303	12,877	12,642
	Total	47,462	46,329	47,149	47,690	52,222	56,423	63,198	65,641	64,740	64,407
	%	19%	20%	20%	20%	20%	20%	20%	20%	20%	20%
Fatalities	Pedestrians	269	278	222	232	224	240	247	252	248	237
	Total	584	550	457	461	450	461	519	509	499	488
	%	46%	51%	49%	50%	50%	52%	48%	49%	48%	48%
Hospitalized victims	Pedestrians	1,585	1,586	1,616	1,599	1,775	1,649	1,763	1,719	1,717	1,643
	Total	5,175	4,353	4,522	4,400	4,904	4,740	4,751	4,705	4,780	4,484
	%	31%	36%	36%	36%	36%	35%	37%	37%	36%	37%
Non-hospitalized victims	Pedestrians	7,786	7,841	8,400	8,366	9,400	9,409	11,673	12,239	11,770	11,665
	Total	56,863	56,103	56,588	57,510	63,314	68,365	76,924	79,256	77,276	76,550
	%	14%	15%	15%	15%	15%	14%	15%	15%	15%	15%

Additionally, **older adults** represent a vulnerable age group on their own, with a fatality rate of about three times higher than the general population, 3.3 deaths per 100 victims, compared to 1.1 for younger people (Table 2.6).

Table 2.6. Fatality rate of injured road users aged 65 and over, and comparison with the rest of users.

Source: Dirección General de Tráfico (2021b).

Indicator	Older adults	Rest of population	%
Fatalities	349	1,184	23%
Hospitalized Injuries	1,177	6,607	15%
Non-hospitalized Injuries	9,197	101,181	8%
Total Victims	10,723	108,972	9%
Fatality Rate	3.3	1.1	
Fatalities per Million Population	37	31	
Hospitalized Injuries per Million Population	126	174	

As mentioned previously, **pedestrian fatalities** predominantly occur on **urban roads** (Figure 2.16). However, for specific age groups, such as 15–24 and 25–34 years, a significant proportion of pedestrian deaths occur outside urban roads. Older pedestrians account for most pedestrian fatalities, and within this demographic, most deaths

occur on urban roads. Consequently, the overall trend shows that most pedestrian fatalities occur in urban roads. Yet, when disaggregated by age, pedestrians aged 15 to 64 experience more fatalities outside urban areas.

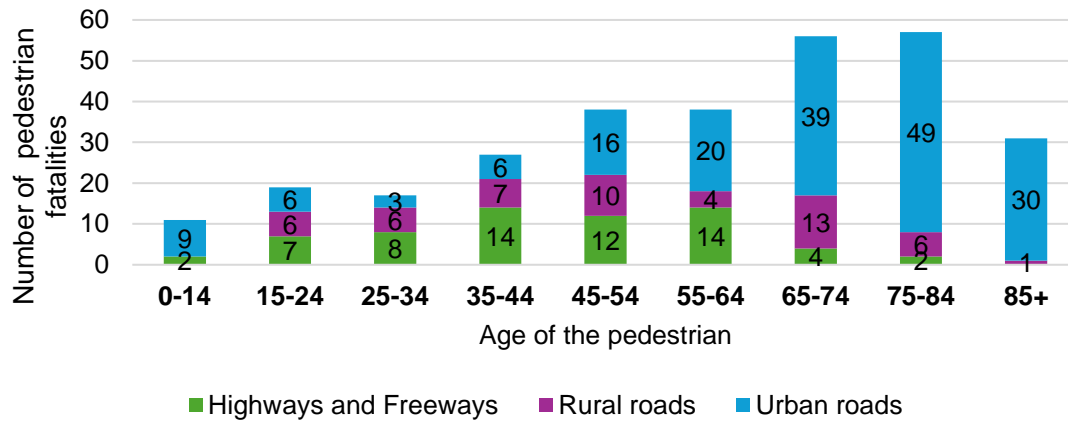


Figure 2.16. Pedestrian fatalities considering the age of the pedestrian and the road type in Spain, 2021.

Source: Dirección General de Tráfico (2021a).

A closer examination of fatalities and hospitalized victims by age group considering the transport mode shows that older adults exhibited much higher rates as pedestrians compared to other age groups, except for people under 14 (Dirección General de Tráfico, 2021b). A longitudinal look at the data from 2012, 2019, and 2021 further shows that older pedestrians consistently represent a substantial proportion of urban pedestrian fatalities (Figure 2.17). Despite advancements in road safety, this trend has remained remarkably stable, suggesting that older pedestrians have not benefited equally from general safety improvements, and that the ageing process of the population impacts road safety figures.

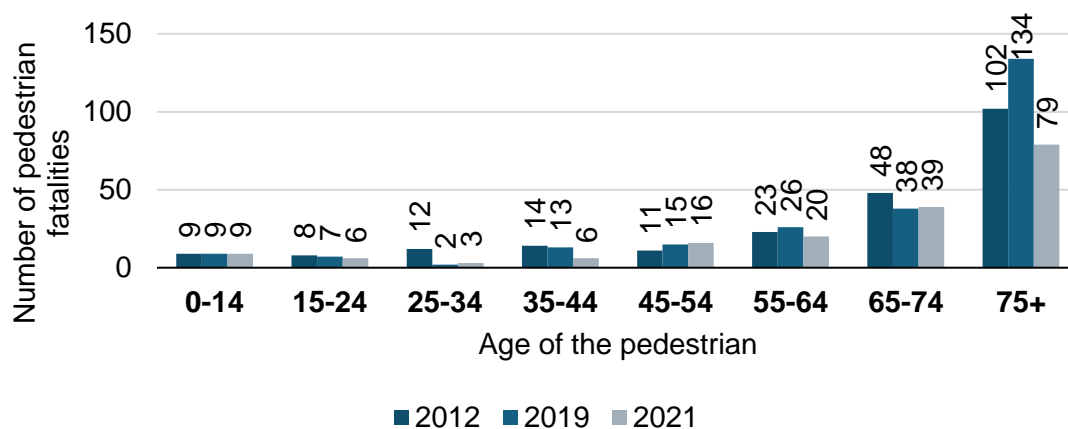


Figure 2.17. Number of pedestrian fatalities in urban roads in Spain, 2012, 2019 and 2021.

Source: Dirección General de Tráfico (2021b).

In a previous DGT's report about road safety on urban roads of 2020, it was underlined that older **pedestrians make up a significant portion of pedestrian fatalities in urban environments**, where **their risk has remained consistent for nearly a decade** year after year, totaling 70% of pedestrians killed on urban roads in 2019 (Table 2.7). The persistence of this concern, despite overall improvements in road safety, suggests that current strategies may not sufficiently address the unique vulnerabilities of older pedestrians.

Table 2.7. Older pedestrian fatalities in urban roads in Spain, 2010-2019.

Source: Dirección General de Tráfico (2020)

Pedestrian fatalities	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Older pedestrians	149	151	150	159	124	161	174	176	156	172
All pedestrians	278	222	232	224	204	247	252	248	237	247
%	54%	68%	65%	71%	61%	65%	69%	71%	66%	70%

The risks faced by older pedestrians in urban areas are further complicated by **gender differences**. Data from 2018 reveal that slightly more male pedestrians are killed on urban roads, particularly among those aged 85 and above (Dirección General de Tráfico, 2018). When adjusting to population, the differences between both genders is greater for the older pedestrians, and especially for those aged 85 and above, as male killed pedestrians above 85 show a rate of 49 per million inhabitants, while females present a rate of 23 (Figure 2.18).

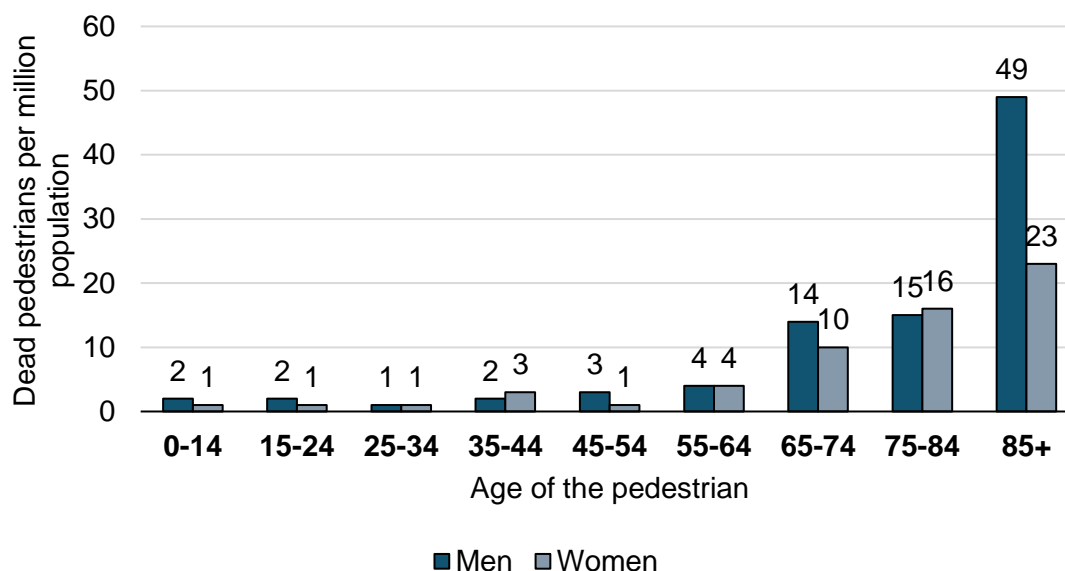


Figure 2.18. Pedestrian fatality rate per million population in urban roads in Spain, 2018.

Source: Dirección General de Tráfico (2018).

Finally, the **location of crashes within urban areas** reveals critical vulnerabilities. According to Table 2.8, most pedestrian fatalities occur on road segments (63%), rather than at intersections (36%).

Table 2.8. Pedestrians injured considering the crash location in urban roads in Spain, 2016-2018.

Source: Dirección General de Tráfico (2018).

Injured pedestrians	Urban roads	
	#	%
At a road intersection	2,076	36%
Outside of a road intersection	3,668	63%
Unknown	72	1%
Total	5,816	100%

The evidence presented underscores the critical need to address the safety of older pedestrians in urban areas. With a fatality rate significantly higher than that of the general population, older pedestrians face disproportionate risks in built-up areas. Their consistent over-representation among vulnerable road users, coupled with the persistence of high fatality rates over time, demonstrates that targeted interventions are necessary to improve their safety. Furthermore, the gendered dimension of pedestrian fatalities and the specific vulnerabilities along road segments highlight the need for a multifaceted approach to urban planning, road design, and policy measures. To reduce pedestrian fatalities in Spain, efforts must prioritize the safety of older pedestrians, who continue to face unacceptably high risks in urban environments.

In the following section, two megatrends at worldwide level that will potentially exacerbate the scenario registered for older pedestrian in the world, Europe, and particularly Spain are presented. These mega-trends represent a thread, especially for Administrations, regarding road safety of older pedestrians in urban environments.

2.3.3 Mega-trends: threads to older pedestrian urban road safety

A **mega-trend** is a broad change that affects many different concerns globally and impacts cultures. Mega-trends alter economies, have an impact on policies, and affect people's everyday lives internationally. Some notable examples include climate change, digitalization, and green economy.

Two mega-trends are especially pertinent to this doctoral thesis: **population ageing** and **urbanization**. In the upcoming decades, the proportion of older adults residing in urban areas is expected to rise sharply due to the expansion of cities and changes in

the age distribution. Understanding the intersection of these trends is crucial, as the rising number of older adults in cities will likely intensify challenges related to age-friendly urban design, especially considering older pedestrian traffic crashes in urban environments.

2.3.3.1 Population ageing

The ageing of the population is the process in which the share of older people increases. Nowadays, it is a global phenomenon, and virtually every country in the world is experiencing both growth in the number and proportion of older people. Globally, 6% of the population was above 65 in 1990, while this figure reached 9% in 2019 (United Nations, 2020b). Although Europe presents the highest recorded figures on the share of population aged 65 or over, population ageing has been fastest in Eastern and South-Eastern Asia and Latin America and the Caribbean in the last decades. By 2050, the share of older population is expected to be roughly 15% (Figure 2.19).

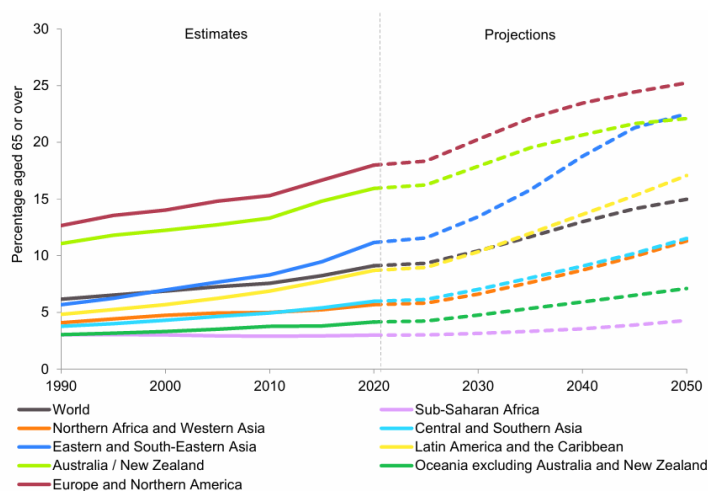


Figure 2.19. Share of population aged 65 years and over by region, 1990-2050.

Source: United Nations (2020b).

The ageing process of the population is driven by three main mechanisms (United Nations, 2023). First, the proportion of older people increases because the fertility levels decrease, reducing the proportion of younger people, which is known as ageing **“from the bottom”**. Second, the increase in life expectancy, especially in developed countries (Table 2.9), rises the number of older population, which is known as ageing **“from the top”**. The increase in lifespan is due to multiple reasons, such as advances

in medicine, economic development, social and cultural norms, urbanization and improved living conditions. The global average life expectancy at birth has increased by 7.7 years and is projected to increase by an additional 4.5 years between 2015-2020 and 2045-2050 (United Nations, 2020b). Third, when population approaching old age is larger than previous generations, their progression into older ages increases the size of the older population, which is known as the “**cohort effect**”. This substantial number of individuals entering their 60s and beyond can be attributed to earlier declines in mortality at younger ages, historical fluctuations in fertility rates, significant inflows of migrants in the past, or a combination of these factors.

Table 2.9. Life expectancy at birth and at age 65 by sex and region, 2015-2019.

Source: United Nations (2020b).

Region	Life expectancy at birth (years)				Life expectancy at age 65 (years)			
	Both sexes	Female	Male	Difference	Both sexes	Female	Male	Difference
World	72.3	74.7	69.9	4.8	17.0	18.3	15.6	2.7
Sub-Saharan Africa	60.5	62.3	58.8	3.5	12.8	13.4	12.1	1.3
Northern Africa and Western Asia	73.5	75.7	71.3	4.4	16.0	17.1	14.8	2.3
Central and Southern Asia	69.5	70.9	68.2	2.7	14.7	15.2	14.1	1.1
Eastern and South-Eastern Asia	76.3	79.0	73.7	5.3	17.2	18.9	15.5	3.4
Latin America and the Caribbean	75.2	78.5	72.0	6.5	18.2	19.5	16.7	2.8
Australia and New Zealand	83.0	85.0	81.1	3.9	21.2	22.6	19.9	2.7
Oceania, excluding Australia and New Zealand	66.3	67.8	64.9	3.0	12.6	12.9	12.3	0.6
Europe and Northern America	78.5	81.6	75.4	6.1	19.1	20.5	17.4	3.1

Despite being a “**human success story**” (United Nations, 2020b), population ageing will put increased financial pressure on old-age support systems. In countries where public transfers are high, including many in Europe and Latin America, population ageing will increase the fiscal pressure on public transfer systems, especially if patterns of taxation and benefits remain unchanged. In this regard, the old-age dependency ratio (i.e., the number of people aged 65 years or above relative to the number of people aged 20 to 64 years) is projected to grow in all regions.

Spain presented 19.6% of older adults among population in 2019, and 32.2 old-age dependency ratio (Table 2.10). By 2050, these figures will reach 36.8% and 78.4, respectively, being higher than the average values worldwide, of the continent, and of the region (Southern Europe). Spain is the only country in Europe among the 10 countries with the largest projected increase in the share of older persons between 2019 and 2050 (17.2%). By 2050, Spain will be the **third country in the world** with the highest old-age dependency ratio (78.4).

Table 2.10. Number and proportion of population aged 65 and over, and old-age dependency ratio, 2019 and 2050.

Source: Adapted from United Nations (2020b).

Region	Population aged 65 and over				Old-age dependency ratio	
	(thousands)		(%)		65+ / 20-64	
	2019	2050	2019	2050	2019	2050
World	702,935	1,548,852	9.1%	15.9%	15.9	28.4
Europe	140,410	199,895	18.8%	28.1%	31.3	53.7
Southern Europe	32,111	47,859	21.1%	35.0%	35.1	71.7
Spain	9,183	16,062	19.6%	36.8%	32.2	78.4

2.3.3.2 Urbanization

Urbanization is a complicated socioeconomic process that changes the built environment by turning previously rural regions into urban settlements and moving a population's geographical distribution from rural to urban areas. Urbanization leads to an increase in the size and population of urban settlements, as well as a higher proportion of urban people compared to rural dwellers (United Nations, 2019).

Since 1950, the global urban population has dramatically expanded, from 751 million to 4.2 billion in 2018. Over two-thirds (70%) of the world's population resided in rural areas in 1950. The number of people living in cities has been increasing more quickly than those living in rural areas since 2007 (Figure 2.20), when the worldwide urban population surpassed the global rural population for the first time in history (United Nations, 2019). The worldwide rural population is currently approximately 3.4 billion and is projected to increase slightly before declining to around 3.1 billion by 2050. In 2018, urban regions accounted for 55% of the global population, outnumbering rural areas. By 2050, metropolitan areas are expected to account for 68% of the global population, up from 30% in 1950 (United Nations, 2019).

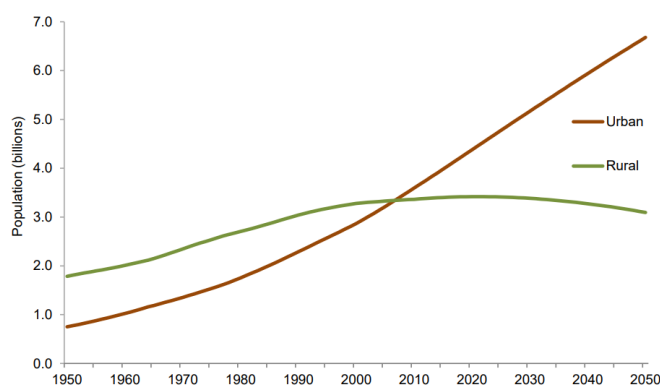
Data source: United Nations, Department of Economic and Social Affairs, Population Division (2018a). *World Urbanization Prospects 2018*.

Figure 2.20. Urban and rural population of the world, 1950-2050.

Source: United Nations (2019).

In 2018, 82% of the people in North America, 81% in Latin America and the Caribbean, 74% in Europe, and 68% in Oceania lived in cities. Asia's urbanization rate has reached over 50% and Africa's population is still primarily rural, with only 43% residing in cities. All regions are predicted to have a rise in urbanization over the next several decades, while there will be significant regional variance. By 2040, the share of urban population in Europe is predicted to rise to 80%, and by 2050, it will reach almost 85% (United Nations, 2019).

Regarding Spain, it presented 80% of urban population in 2018, and this figure is expected to grow to 83% and 88% by 2030 and 2050, respectively (Table 2.11). These figures are higher than the average worldwide, in Europe, and in the Southern Europe region. In addition, Spain has two urban agglomerations with more than 5 million inhabitants (as of 2018), which are Madrid and Barcelona. These two urban locations expect a modest 6% in 2030, compared to 2018 population (United Nations, 2019).

Table 2.11. Urban population, 1990, 2018, 2030, 2050.

Source: Adapted from United Nations (2019).

Region	Urban population (percentage of total population)			
	1990	2018	2030	2050
World	43	55	60	68
Europe	70	74	77	84
Southern Europe	65	72	75	82
Spain	75	80	83	88

About rural population, Spain is in the 21st place in the ranking of countries with the highest largest expected decline in rural population between 2018 and 2050, with a decline of 3,809 thousand inhabitants. It represents a remarkable decrease in relative terms of 41.7%, which is higher than most of the relative declines of the countries in the ranking (Table 2.12).

Table 2.12. Top 30 countries with the largest declines in rural population projected between 2018 and 2050.

Source: United Nations (2019).

Rank	Country	Projected Decline in Rural Population (thousands)	Relative Decline in Rural Population (%)
1	China	-305,515	-52.9
2	India	-110,907	-12.4
3	Indonesia	-31,746	-26.6
4	Bangladesh	-21,334	-20.2
5	United States of America	-15,735	-27.1
6	Russian Federation	-14,682	-39.9
7	Thailand	-14,664	-42.3
8	Viet Nam	-12,914	-20.9
9	Brazil	-10,697	-37.8
10	Iran (Islamic Republic of)	-7,474	-36.3
11	Italy	-7,106	-40.5
12	Turkey	-6,927	-34.0

Rank	Country	Projected Decline in Rural Population (thousands)	Relative Decline in Rural Population (%)
13	Mexico	-6,578	-25.4
14	Germany	-6,259	-33.5
15	Ukraine	-5,706	-42.3
16	Poland	-5,629	-37.0
17	Japan	-4,908	-46.0
18	South Africa	-4,614	-23.9
19	France	-4,522	-35.4
20	Myanmar	-4,420	-11.8
21	Spain	-3,809	-41.7
22	United Kingdom	-3,679	-33.3
23	Romania	-3,543	-39.3
24	Colombia	-3,404	-35.8
25	Morocco	-3,187	-23.5
26	Sri Lanka	-2,862	-16.8
27	Dem. People's Republic of Korea	-2,835	-29.0
28	Republic of Korea	-2,645	-21.9
29	Algeria	-2,593	-22.6
30	Malaysia	-2,389	-31.1

The intersection of these two demographic mega-trends, population ageing and urbanization, presents a threat to the road safety situation of older pedestrians in urban areas, especially for the Administration. As stated before, older adults are more likely to be killed or seriously injured if a traffic crash occurs. As older adult population in cities will grow for both population ageing and urbanization, a potential rising the number of injured and killed older pedestrians in traffic crashes, and the proportion of killed and seriously injured pedestrians of all pedestrian crashes could be expected.

2.4 Road safety strategies and their implications for older pedestrians

2.4.1 Sustainable Development Goals

Road safety policies in Europe and Spain are aligned with the United Nations Sustainable Development Goals (SDGs). The United Nations SDGs are a collection of 17 interlinked global goals designed to create a better and more sustainable future for all (Figure 2.21). Within the 2030 Agenda for Sustainable Development, the SDGs outline targets that span health, equality, and urban planning, establishing a roadmap for transformative change.



Figure 2.21. The 17 Sustainable Development Goals.

Specifically, for the topic of reducing traffic crashes involving older pedestrians in urban settings, several SDGs directly support this focus by advocating for safer, more inclusive, and accessible urban environments. Specific targets of goals SDG 3 (Good Health and Well-being), SDG 10 (Reduced Inequalities), and SDG 11 (Sustainable Cities and Communities) align with efforts to create safer pedestrian spaces for older adults, prioritizing infrastructure that promotes both well-being and inclusion. The following paragraphs outline how specific SDG targets directly relate to and support the thesis of improving road safety for older pedestrians in urban areas.

SDG 3: Good Health and Well-being.

- **Target 3.6:** *By 2020, halve the number of global deaths and injuries from road traffic crashes.* This target aligns directly with the focus of reducing traffic crashes involving older pedestrians, as it emphasizes minimizing preventable

injuries in road environments. Ensuring safer urban scenarios for older pedestrians directly supports Target 3.6 by reducing high-risk situations and enhancing road safety infrastructure to accommodate vulnerable groups. Efforts such as improved crosswalk design, reduced vehicle speeds in pedestrian-heavy areas, and dedicated pedestrian zones contribute to making urban settings less hazardous for older adults, thereby promoting their well-being and health.

SDG 10: Reduced Inequalities

- **Target 10.2.** *By 2030, empower and promote the social, economic and political inclusion of all, irrespective of age, sex, disability, race, ethnicity, origin, religion or economic or other status. Addressing older pedestrians' safety in urban areas contributes to this goal by fostering inclusive urban mobility.* Enhanced safety measures and infrastructure adjustments ensure that older adults can participate fully in urban life without facing disproportionate risks. By prioritizing pedestrian safety for older adults, urban spaces become more inclusive, enabling older populations to navigate their environments with confidence and autonomy, thereby supporting Target 10.2's emphasis on reducing inequalities and promoting accessibility.

SDG 11: Sustainable Cities and Communities

- **Target 11.2.** *By 2030, provide access to safe, affordable, accessible and sustainable transport systems for all, improving road safety, notably by expanding public transport, with special attention to the needs of those in vulnerable situations, women, children, persons with disabilities and older persons.* This target aligns closely with the objective of reducing traffic crashes involving older pedestrians in cities. By implementing safe crossing zones, increasing visibility, and designing streets to prioritize pedestrians, cities can create urban areas that support safe, sustainable movement for older adults. These measures not only reduce the likelihood of crashes but also contribute to sustainable urban development that accommodates the needs of an ageing population.
- **Target 11.3.** *By 2030, enhance inclusive and sustainable urbanization and capacity for participatory, integrated and sustainable human settlement planning and management in all countries.* This target supports initiatives that prioritize safe and accessible infrastructure, crucial for protecting older pedestrians in urban areas. By integrating urban planning with safety considerations for older adults, cities can create environments that are safer and more accommodating.

Target 11.3 encourages community involvement in planning, ensuring that the needs of vulnerable groups, including older adults, are reflected in urban development strategies. Such inclusive planning fosters not only physical safety for older pedestrians but also social cohesion by promoting walkable, accessible communities for all residents.

2.4.2 Second Decade of Action for Road Safety 2021-2030

The United Nations declared the decade of 2021-2030 as the Second Decade of Action for Road Safety in the resolution Improving global road safety (United Nations, 2020a). In consequence, Target 3.6 of the SDGs was updated, so traffic fatalities and seriously injured victims should be halved by 2030, considering the level of 2019 as the base. Additionally, WHO published the “Global Plan for the Decade of Action for Road Safety 2021-2030” (World Health Organization, 2021b) as a guiding document to support the implementation of the Decade of Action 2021-2030 and its objectives. These objectives were then transferred and included in the traffic crash goals of Europe, and, consequently, Spain.

2.4.3 Vision Zero in road safety

The Vision Zero concept represents a fundamental shift in the approach to road safety, aiming not merely to reduce but to eliminate road deaths and serious injuries entirely. Unlike traditional road safety strategies that often focused on mitigating the severity of traffic crashes, Vision Zero is rooted in the belief that every road death or seriously injured victim is preventable. Developed initially in Sweden, this concept has gained traction worldwide, forming a central pillar of the European Union’s road safety policy with a clear commitment to “no loss of life on the road is acceptable” (European Commission, 2020).

The European Union has firmly embedded Vision Zero within its road safety policies, demonstrating a strong commitment to drastically reducing road fatalities and severe injuries by 2050. The EU Road Safety Policy Framework 2021-2030 lays out a roadmap for achieving this goal, with a significant intermediate milestone: halving the number of road deaths and serious injuries by 2030 compared to 2020 levels (European Commission, 2020, 2022).

Central to Vision Zero is the “Safe System” approach (European Commission, 2020), which recognizes that while human errors are inevitable, fatalities and severe injuries

should not be. This approach aims to create forgiving roads (e.g. with median safety barriers to ensure that driving errors do not need to have serious consequences), where road design, vehicle safety, and speed management combine to prevent serious consequences from collisions. By considering human vulnerability, this strategy addresses several key areas:

1. **Safe Infrastructure:** Emphasizing designs that protect all road users, especially vulnerable groups such as pedestrians and cyclists.
2. **Safe Vehicles:** Promoting advanced safety features that protect not only vehicle occupants but also other road users.
3. **Safe Road Use:** Addressing behaviors such as speeding, distraction, alcohol and drugs involvement, use of safety belts, and new mobility patterns and demographic change.
4. **Post-Crash Care:** Improving emergency response systems to reduce the severity of injuries when traffic crashes do occur.

Vision Zero represents a transformative commitment by the EU to eliminate fatalities and severe injuries on the road by 2050, with a milestone of reducing such incidents by half by 2030. Through a Safe System approach, technological advancements, and targeted policies that prioritize vulnerable groups, Europe is setting a global standard in road safety. As Europe continues to advance these strategies, Vision Zero will require ongoing innovation, investment, and public awareness to adapt to the evolving challenges of mobility and road safety in an increasingly connected society.

2.4.4 Road Safety Strategy 2030 in Spain

The Spanish Road Safety Strategy 2030 aims to significantly improve road safety through a comprehensive, people-centered approach. Following European strategies, its overarching goal is to reduce fatalities and severe injuries by 50% by 2030, using 2019 data as the baseline (Dirección General de Tráfico, 2022b).

To accomplish these goals, Spain's strategy addresses various aspects of road safety through nine strategic areas, including the promotion of:

1. Trained and capable people
2. Zero tolerance towards risky behavior
3. Safe cities

4. Safer road infrastructure
5. Secure vehicles
6. Improved post-crash response
7. Enhanced data collection for risk-based management
8. Safe administrations, enterprises, and organizations
9. Integrated policies and international cooperation

The strategy underscores collaboration across governmental bodies, private sectors, and local communities, fostering a shared responsibility approach where each entity actively contributes to a safer traffic environment.

As stated previously, older adults (i.e., individuals over 65), represent a critical focus within Spain's road safety strategy due to their vulnerability in traffic environments. In 2019, older adults constituted 28% of road fatalities, although they only comprised about 19% of the population (Dirección General de Tráfico, 2022b). By 2030, the goal is to maintain the share of older adults among all fatalities to a maximum of 30%, and the share of older adults among all fatalities and severe injuries to a maximum of 20%. This means that in 2030, a maximum of 687 older adults should be severely injured or killed in urban roads. Regarding pedestrians, the objective is to halve the number of killed and severely injured individuals in urban roads by 2030. All these detailed road safety objectives are shown in Table 2.13.

Table 2.13. Objectives of the Road safety strategy 2030 in Spain.

Objective	2019	2020	2021	2022	2023	2030
All individuals						
Fatalities	1,755	137	1,533	1,746	1,806	877
Seriously injured	8,613	6,681	7,784	8,502	9,265	4,306
Older adults (65 and over)						
% of the total fatalities	28%	26%	26%	27%	26%	30%
% of the total fatalities or seriously injured	18%	16%	17%	18%	18%	20%
Fatalities or seriously injured in urban roads	1,057	703	887	1,035	1,149	687
Pedestrians						
Fatalities or seriously injured in urban roads	1,730	1,124	1,420	1,662	1,792	865

2.5 *Road safety of older pedestrians as a precursor of active and sustainable mobility*

2.5.1 Healthy and active ageing

As people live longer, there is a need to promote opportunities for health, participation and security through the course of the life of a person. That approach was firstly named “active ageing” by World Health Organization (2002), which is defined as “the process of optimizing opportunities for health, participation and security in order to enhance quality of life as people age”. Active ageing is a multi-pillar approach, including these three main pillars:

- **Health.** When the risk factors (both environmental and behavioral) for chronic diseases and functional decline are kept low while the protective factors are kept high, people will enjoy both a longer and quality life.
- **Participation.** When labor market, employment, education, health and social policies and programs support their full participation in socioeconomic, cultural and spiritual activities, according to their basic human rights, capacities, needs and preferences, people will continue to make a productive contribution to society in both paid and unpaid activities as they age.
- **Security.** When policies and programs address the social, financial and physical security needs and rights of people as they age, older people are ensured of protection, dignity and care if they are no longer able to support and protect themselves. Families and communities are supported in efforts to care for their older members.

The aim of this doctoral thesis (**reducing frequency and injury severity of older pedestrian traffic crashes**) is presented inside of the “**Health**” pillar, as the key policy proposal of “**Age-friendly, safe environments**” which consists of “preventing injuries by protecting older pedestrians in **traffic**, making **walking safe**, implementing fall prevention programs, eliminating hazards in the home and providing safety advice” (World Health Organization, 2002).

The determinants of active ageing are economic, social, physical, personal, and behavioral, as well as health and social services (World Health Organization, 2002). The functional capacity of a person declines by age. From early to adult life, a person reaches the peak of his/her functional capacity. From this point, a decline occurs to older age (Figure 2.22). Nevertheless, this decline can be higher or lower depending on multiple factors, and a life course perspective on ageing recognizes that older

people are not one homogeneous group, and that individual diversity tends to increase with age (Kalache & Kickbusch, 1997).

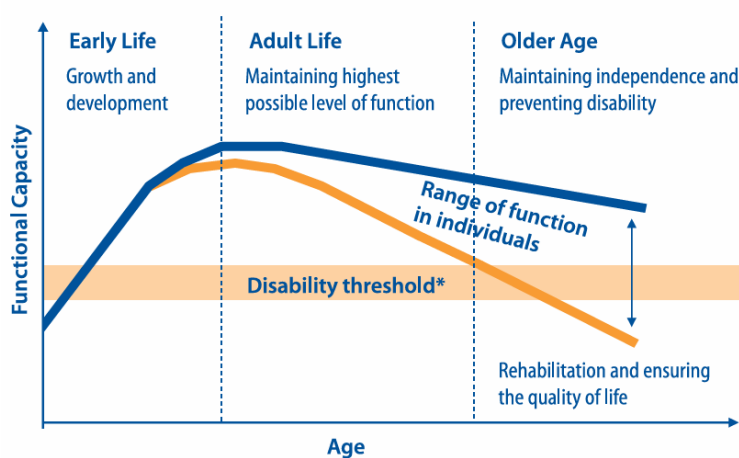


Figure 2.22. Functional capacity over the life course.

Source: World Health Organization (2002).

WHO replaced the idea of “active ageing” by “**healthy ageing**” in 2015 (World Health Organization, 2015), and it is the main concept of the **Decade of Healthy Ageing** (2021-2030), (World Health Organization, 2021a). The **definition** of healthy ageing is slightly different from the one of active ageing: “developing and maintaining the functional ability that enables well-being in older age” (World Health Organization, 2015). Functional ability is determined by the intrinsic capacity of the individual (i.e., physical, and mental capacities), the environment in which he or she lives (i.e., factors in the extrinsic world that form the context of an individual’s life), and the interactions among them. Intrinsic capacity is the composite of all the physical and mental capacities of an individual.

One of the key behaviors that influence **Healthy Ageing** is **physical activity**, which has many benefits for older adults, including increasing longevity, improving physical and mental capacities; and preventing disease. Moreover, **ageing in place** is one of the key issues for public health action, which is defined as “the ability of older people to live in their own home and community safely, independently, and comfortably, regardless of age, income or level of intrinsic capacity” (World Health Organization, 2015). Capacity declines often require older people to make transitions in their living environments, by adapting their residence or by relocating to another place. Nevertheless, older people often view their existing home or community as having the advantages of maintaining a sense of connection, security, and familiarity.

These two key features of healthy ageing (i.e., physical activity and ageing in place) are strongly related to the road safety of older adults as pedestrians, because walking is a main contributor to physical activity, and because suffering a traffic crash might force an older adult to reduce his/her physical activity, and can even force a change of residence as the older adult might necessitate cares because of the crash sequels.

There are four key areas for action on healthy ageing (World Health Organization, 2015): (1) Aligning health systems to the needs of the older populations they now serve, (2) Developing systems for providing long-term care, (3) Improving measurement, monitoring and understanding, and (4) Creating age-friendly environments. The last key area for action, **creating age-friendly environments**, is linked with the idea of **age-friendly cities** and communities.

2.5.2 Age-friendly cities

The concept of “**age-friendly cities**” emerged as a response to two global trends: population ageing and urbanization. Initiated by the WHO in 2007 and based on the explained concept of “active ageing”, the age-friendly cities movement seeks to create urban environments that promote healthy, active, and engaged ageing. The age-friendly city “encourages active ageing by optimizing opportunities for health, participation and security in order to enhance quality of life as people age” (World Health Organization, 2007). Because active ageing is a lifelong process, an age-friendly city is not just “older friendly”, but should be inclusive and accessible for people of all ages (World Health Organization, 2007, 2018a, 2023b). Age-friendly cities focus on adapting the urban environment across eight interconnected key domains identified by World Health Organization (2018a):

1. **Outdoor spaces and buildings:** Access to safe, clean, and pleasant public spaces that encourage physical activity and social interaction.
2. **Transportation:** Reliable, affordable, and accessible transport options that enable older adults to remain mobile and independent.
3. **Housing:** Diverse, accessible, and affordable housing options to allow older adults to age in place.
4. **Social participation:** Opportunities for older people to participate in social, cultural, and recreational activities.
5. **Respect and social inclusion:** Policies and attitudes that combat age discrimination and encourage intergenerational interactions.

6. **Civic participation and employment:** Opportunities for older people to stay engaged in community life, including paid or voluntary work.
7. **Communication and information:** Access to clear and relevant information, especially regarding services and amenities.
8. **Community and health services:** Accessible health and social services tailored to older adults' needs.

The WHO's Global Network for Age-friendly Cities and Communities (GNAFCC) was established in 2010 with the mission of "stimulate and enable cities, communities, and other sub-national levels of government around the world to become increasingly age-friendly" (World Health Organization, 2018c). GNAFCC now spans numerous countries and provides a framework for cities aiming to improve urban living for older residents. By 2018, the Global Network covered 39 of the 194 (20%) WHO member States. GNAFCC has a structured model for creating age-friendly environments, organized into four phases: Engage and Understand, Plan Strategically, Act and Implement, and Evaluate.

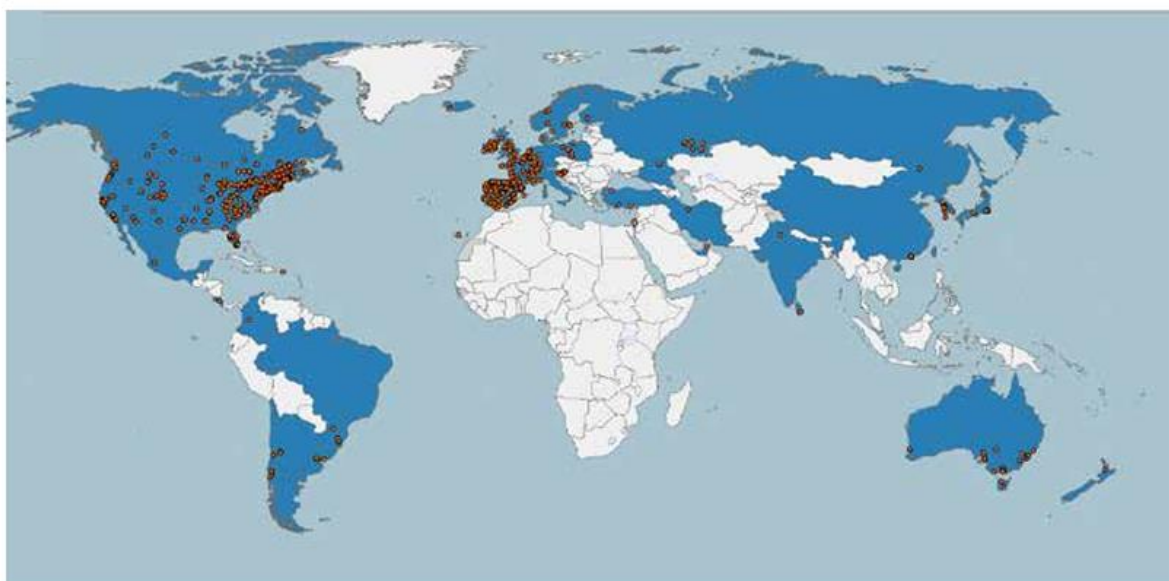


Figure 2.23. Map of cities that were members of the Global Network for Age-friendly Cities and Communities in 2018.

Source: World Health Organization (2018c).

2.5.3 Walking benefits: more than physical

Active mobility, and especially walking, offers older adults remarkable health benefits. The most straightforward benefits are in terms of physical health. Walking is a **physical activity**, and it has been related to **multiple physical benefits** (Cunningham et al., 2020; Ungvari et al., 2023), such as reduced cardiovascular risk factors, hypertension, cardiovascular and cerebrovascular diseases, type 2 diabetes, respiratory pathologies, bone health, reduced cancer risk, and even improved sleep quality and duration. In general, walking is associated with all-cause mortality. The risk of falling also decreases with the walking activity, as shown by Jansen et al. (2021) in a study on older adults aged above 70 and living at home without assistance.

Other benefits are related to cognitive functions, dementia, Alzheimer's disease, and depression, although these results are not always consistent in literature. **Mental health** is also benefited by physical activity, which mainly includes walking in the older age. **Physical activity** has been linked to reduced risk of depression, and walking reduces the symptoms of depression, and better mood state (Ungvari et al., 2023).

In addition, walking benefits are also related to a **better connection of older people with the community** (Musselwhite et al., 2015). Physical activity was related to **reduced social isolation** and **loneliness** in a survey study with 2016 data of the Health and Retirement Study (Shellito & Roldan, 2019). In addition, social isolation and loneliness were related to each other. In a study conducted in China (Liu et al., 2021), older people who walked more tended to use more the public transport and have an electric bicycle. A cross-sectional study of 173 older adults in Japan concluded that walking with someone is an effective solution to prevent to reduce loneliness among older adults (Mizuta et al., 2023).

Walking groups have demonstrated health benefits while presenting a good adherence (for a review see Hanson & Jones (2015)). Another research activity with adults aged above 50, employed group outdoor health walks in the UK (Irvine et al., 2022). This method consists of walking within a group in a natural environment, and it proved to be a suitable intervention to reduce social isolation and loneliness, as self-reported by participants. In addition, these activities were related to social well-being, in terms of sense of belonging, expanding social networks, meaningful relationships, and acting on empathy.

This **social inclusion** and **reduction of isolation** is related to **higher well-being** and **improved quality of life** (Nordbakke & Schwanen, 2014). For example, a survey with about one thousand older adults living in private households in Britain about quality of

life, revealed that having good social relationships and retaining a role in society emerged as important factors to determine the quality of life (Gabriel & Bowling, 2004).

Walking is the most common contributor to **physical activity** among older adults (Mattle et al., 2022). However, older adults are the group with the **lowest physical activity registers** (Remón et al., 2020; Whiting et al., 2021). In a survey study of 2,157 healthy older adults aged 70 and above, younger participants (70-74) showed more physical activity than older ones (above 75) (Mattle et al., 2022). The same result was found by Remón et al. (2020) in an analysis of the Eurobarometer survey 472, considering only physical activity of people above 65. This could be due to home-based being the most common contributor to physical activity (Pae & Akar, 2020), but older people, despite having a higher proportion of home-based trips, these are not school- or work-related (K. B. Watson et al., 2021), so they do not need to walk as much as younger people to fulfill their daily needs. This physical inactivity and sedentary behavior might lead to obesity (Silveira et al., 2022), which is related to several health issues such as heart disease, obstructive pulmonary disease, and diabetes.

Because of all these benefits, walking activity is related to the idea of active and healthy ageing, as it helps reduce age-related functional decline and diseases, while enhancing the connection of older adults with the community.

2.5.4 Age-related functional limitations and illnesses of pedestrians

Age-related functional limitations and illnesses present significant risks for older pedestrians. These limitations stem from natural declines in sensory, cognitive, and physical abilities, compounded by chronic illnesses and medication use. Together, they not only increase older adults' vulnerability to traffic crashes but also limit their ability to safely navigate complex traffic situations.

In terms of **sensory abilities**, **vision** and **hearing** changes are among the most pronounced challenges for older pedestrians. As adults age, **visual impairments** often manifest in various forms, including reduced contrast sensitivity, slower adaptation to light changes, and decreased ability to see in low-light conditions. Visual difficulties in pedestrians with difficulties detecting moving vehicles, judging distances, and notice potential hazards, especially in environments with fluctuating light, such as crosswalks or intersections with high glare. Visual impairment might be further affected as older adults tend to present reduced head and neck mobility (European Commission, 2015). Age-related **hearing loss** exacerbates this problem by diminishing the ability to detect higher-frequency sounds, such as those produced by approaching vehicles. Notably,

both visual and hearing impairments are related to the risk of falling, although not being the goal of this research.

Cognitive abilities also decline with age, affecting the safe navigation of urban environments. Older pedestrians often face difficulties in judging safe gaps in traffic, a task that demands quick processing of visual information and accurate assessment of vehicle speed and distance. Cognitive ageing slows down these processes (Tournier et al., 2016; Wilmut & Purcell, 2022). The cognitive burden of managing complex traffic situations is further intensified by limitations in attention and processing speed, which impact the ability to simultaneously track multiple moving vehicles, interpret signals, and assess risks (European Commission, 2024f; Fildes et al., 1994).

Physical abilities, including **balance**, **walking speed**, and **movement initiation**, also deteriorate with age, adding another layer of risk. **Balance instability** is common among older adults due to reduced muscle strength and joint stiffness, which can impair gait and increase the likelihood of falls (European Commission, 2015; Fildes et al., 1994; Oxley et al., 2004). **Reduced walking speed**, a natural consequence of ageing, is a critical limitation for older pedestrians who need more time to cross streets. Additionally, **initiating and stopping movement** becomes increasingly challenging with age. Slower reaction times and cautious gait adjustments make it difficult for older adults to respond quickly to unexpected changes, such as a vehicle turning suddenly or a pedestrian signal flashing (Fildes et al., 1994; Tournier et al., 2016; Wilmut & Purcell, 2022).

In addition, **age-related illnesses** play a crucial role in limiting older adults' capabilities. Chronic conditions, including arthritis, cardiovascular disease, and diabetes, significantly impact mobility and stamina, limiting older adults' ability to navigate urban environments effectively. For instance, arthritis can cause joint pain and stiffness, restricting movements needed to look both ways or quickly pivot in response to traffic (Oxley et al., 2004).

The use of medication, often a necessity for managing chronic conditions, introduces additional risks. Many older adults rely on medication that affect the central nervous system, such as sedatives, antihypertensives, and analgesics, which may impair coordination, cognitive function, and reaction time. For example, sedatives can cause low reaction time and diminish awareness of hazards (Oxley et al., 2004).

The **frailty** (Fried et al., 2001) that accompanies ageing significantly increases the injury risk for older pedestrians in traffic crashes. Bones weaken with age, increasing the likelihood of fractures from falls or collisions. Statistics indicate that even minor impacts can result in severe injuries, including hip fractures or head trauma, which often lead

to prolonged recovery or, in some cases, fatalities (European Commission, 2015, 2024f).

Older pedestrians face a complex set of risks stemming from sensory, cognitive, and physical limitations, compounded by age-related illnesses and medication effects. Addressing these issues through design adjustments, pedestrian-focused policies derived from traffic crash data could reduce injury rates and improve overall safety for older pedestrians in urban roads.

2.5.5 Studies on the behavior of older pedestrians

Older adults have demonstrated different behaviors when acting as pedestrians than younger people. **Self-regulation** is a crucial adaptive strategy that older pedestrians use to compensate for age-related physical and cognitive limitations, ensuring safer navigation in traffic environments (European Commission, 2015; Oxley et al., 2004; Tournier et al., 2016). Nevertheless, Oxley et al. (2004) argued that the awareness of age-related changes might not be the only reason for this self-regulation, as mature judgements about road use, lifestyle choices, and personal preferences could also play a role in the behavior.

Other studies also underline the critical role of self-regulation among older pedestrians as a key strategy to maintain safety in traffic. Phillips et al. (2013) demonstrated that older adults often adjust their pace and rely on known landmarks in unfamiliar areas, indicating an awareness of cognitive limits. Rosenbloom et al. (2016) found that older adults in higher socio-economic areas were more likely to adopt safer walking practices, highlighting the influence of awareness and available resources. Vine et al. (2012) showed that route familiarity is essential, with older pedestrians avoiding areas perceived as risky.

Due to this self-regulation, caused by an awareness of their physical limitations and an increased sensitivity to traffic risks, older pedestrians are generally **more cautious** than the rest of the population. Bernhoft & Carstensen (2008) found that older pedestrians (aged above 70) are significantly more likely than younger pedestrians (aged 40–49) to adopt careful behaviors in traffic, such as consistently using pedestrian crossings, avoiding crossing during red lights, and avoiding crossing in areas lacking traffic signals or pedestrian facilities. Older pedestrians also prioritized environmental features like sidewalks, smooth surfaces, and signalized crossings. A survey by Bailey et al. (1992) revealed similar behaviors of older adults, as this demographic tended to cross the street at designated locations and most respondents waited for the vehicle's

traffic light to turn red before crossing the street. Another survey by Rod et al. (2023) surveyed adults (above 18) residents in Australia regarding their walking frequency and perceived risk of pedestrian injury regarding traffic crashes and falling. Walking frequency activity was adversely correlated with the perceived risk of pedestrian-vehicle collision as well as the perceived risk of vulnerability to pedestrian fall events and severity. Pedestrian-vehicle collisions had a greater psychological weight than pedestrian falls, but people aged above 60 showed a lower perceived risk of vehicle collision and higher perceived risk of pedestrian fall. Additionally, a survey by Saxena (2023) regarding the crash risk perception of pedestrians aged 15-60 showed that crash risk perception negatively affects walking behavior, and that there exist significant effects of crash risk perception on perceived crossing speed.

Pedestrian behavior and walking frequency can also be influenced by **worry and anxiety**. For example, Kummeneje & Rundmo (2019) developed a survey in Norway for population aged 15 years and over on pedestrians' risk perception and worry. Results indicated that pedestrians perceived higher risk and were more worried during nighttime than in daytime. Previous involvement in traffic crash and assault had an indirect effect on worry. The perceived risk influenced worry during both nighttime and daytime, and worry influenced the outdoors walking frequency during night-time. Phillips et al. (2013) surveyed older adults to understand their experiences as pedestrians in unfamiliar urban spaces. Older adults were concerned about a set of barriers, including poor signage, confusing spaces, poor paving and sensory overload (i.e. noise and complexity of the environment) in new environments. When navigating, and distinctive buildings were more important than signage.

Regarding anxiety when crossing the street, Bailey et al. (1992) surveyed 76 citizens aged 56 and over and most of them reported that the current time to cross the street provided by the pedestrian crossing signal was sufficient, but many expressed worries about safety and feelings of anxiety, consequently increasing their walking pace when crossing the street. Infrastructure proved to be a precursor to physical activity among older adults in Wisconsin, including well-maintained sidewalks, bike paths or lanes, and traffic control; while traffic safety, not maintained or nonexistent sidewalks, crosswalks, and bike paths or lanes, discouraged physical activity (Strath et al., 2007).

These findings collectively illustrate how anxiety, driven by risk perception and environmental factors, plays a pivotal role in shaping the behavior and safety of older pedestrians. Addressing these issues through urban design improvements and supportive measures could enhance their confidence and well-being in public spaces.

Because of age-related physical and cognitive impairments and the consequent self-regulation, the **walking speed** of older adults is usually below the speed of younger

people. The most critical situation in which this fact has influence is crossing the street. A study by Romero Ortuño (2010) using data from the Survey of Health, Ageing and Retirement in Europe (SHARE) of people above 75 years old found that Spain is one of the countries with the lowest walking speed in Europe. In fact, it is the country where women older than 75 have the lowest walking speed (mean 0.40 m/s). The walking speed of older people was compared with the walking speed supposed in the design of traffic lights in the Basque country, 0.7 m/s. It was shown that 75% of men and 67% of women walked slower than that threshold of 0.7 m/s, and the reference walking speed that would contain 80% of older pedestrians was 0.25 and 0.31 m/s for men and women, respectively. Another observational study by van Hoof et al. (2020) in The Hague and Zoetermeer (The Netherlands) found that traffic lights did not allow sufficient time for older adults to cross the street before the lights turned red. Additionally, Oxley et al. (2004) in their comprehensive review on vulnerable road users emphasized that older pedestrians' slower walking speeds contribute to their difficulties in complex traffic situations, such as crossing multi-lane roads. These slower speeds, compounded by age-related sensory and cognitive challenges, often necessitate more cautious and time-consuming crossing behaviors.

Age-related limitations, and especially the reduced walking speed and cognitive declines, impact the most important action in which pedestrians and vehicles might collide, **street crossing**. A survey by Bailey et al. (1992) showed that a remarkable number of respondents avoided crossing the street during peak traffic hours and during low visibility, and about one quarter of respondents had difficulty seeing the crosswalk display. Another study by Rankavat & Tiwari (2020) examined pedestrian perceptions of safety and actual crossing behavior using a structured questionnaire in Delhi, India. Results showed that at mid-blocks, males perceived a lower risk than females, while older pedestrians perceived a higher risk than younger pedestrians. At intersections, the effect of age on the perceived risk was found insignificant.

Observational studies base their result on real crossing behaviors of older pedestrians. A study by Oxley et al. (1997) in Melbourne showed that older adults performed better on one-way roads than on two-way undivided roads, indicating that environmental complexity impacts their decision-making. Older pedestrians were engaged in riskier crossing behaviors compared to their younger counterparts in two-way undivided roads, where the gap accepted, in terms of distance, by older pedestrians was significantly higher, and older pedestrians delayed more time before making the first movement. A more recent study by Rosenbloom et al. (2016) observed crossing behavior in two Israeli cities, Tel Aviv and Beer Sheva. Older pedestrians in higher socioeconomic areas, such as Tel Aviv, adhered more to safety protocols than those in lower socioeconomic settings like Beer-Sheva. In Tel Aviv, older pedestrians keep the crossing

rules much more than the younger while in Beer Sheva older and young pedestrians behaved similarly. Similarly, Galanis et al. (2017) observed for all pedestrians more legal behavior in main arterials than in medium rate in collector streets. The lowest rate of legal behavior was shown in local streets. This indicates that higher levels of vehicle flows, and vehicle speeds prevent pedestrians from walking on the carriageways or cross the street in non-designated areas.

Simulation studies were also employed to analyze the crossing behavior of older pedestrians (Dommes et al., 2012, 2013). In the first study (Dommes et al., 2012), the crossing behavior of older and younger pedestrians in a one-way street before and after training were compared. Older pedestrians adopted smaller safety margins and made more unsafe decisions before the training program than younger pedestrians. Training improved older pedestrians' safety margins and reduced the number of unsafe decisions. Even six months after the training, the older pedestrians made more unsafe decisions when cars were approaching at high speeds than at low speeds. In a subsequent study (Dommes et al., 2013), older pedestrian crossing behavior of two-way roads was examined. Older-old pedestrians were more likely than both younger-old and young participants to make decisions that could lead to collisions with oncoming vehicles, especially with high vehicle speeds. The primary contributors to risky decisions were walking speed and time-to-arrival estimation. When crossing two-way streets, older pedestrians may overestimate their walking speed, misjudge the time-to-arrival of oncoming vehicles. These two studies emphasized the difficulties of older pedestrians to properly calculating the existing gap. Additionally, training programs can be a proper complement to infrastructure countermeasures, as training proved to be beneficial but not for all situations.

In summary, older pedestrians' street crossing behavior is shaped by a combination of age-related declines in cognitive, perceptual, and physical abilities, along with external factors like training and socioeconomic status. These elements collectively impact on their decision-making and safety, emphasizing the need for targeted interventions to improve their crossing safety.

The decision of **where and how to walk** involves various factors other than the shortest route, such as safety, convenience, and environmental features. The studies reviewed in this section offer insights into how these factors affect route choices for older pedestrians and, in some cases, all age groups. Walking activity showed to be higher among older pedestrians in neighborhoods with green spaces and more roads intersections. Additionally, higher perception of safety from traffic was observed with higher number of street intersections (Li, 2005). Micro-level route choice elements such as blind walls and litter reduce the walking activity of older adults, while front gardens,

shops, more recreational facilities, and higher perception of safety for walking increase walking activity (Borst et al., 2009; Li, 2005).

Physical safety was a main contributor to mode choice among older adults (Vine et al., 2012), which is mainly influenced by the perception of traffic and pedestrian infrastructure. The perceived physical safety can be deteriorated by the presence of shared cyclist lanes, overcrowded or busy sidewalks, poorly maintained sidewalks (Vine et al., 2012), these factors can lead older pedestrians to decide to walk on the carriageway (Corazza et al., 2017), where they are not protected from traffic. Accessibility plays a main role in walking activity and/or resistance to walking. Older pedestrians living in more accessible environments, including barrier free sidewalks, public transportation, shading, and public seating, and avoiding high slopes or stairs (Borst et al., 2009; Clarke & Gallagher, 2013; Vine et al., 2012), are expected to walk more and perceive higher safety levels. These same results of attractiveness and accessibility influencing route choice was also observed for pedestrians of all ages in Madrid (López-Lambas et al., 2021).

Apart from the reviewed literature, during the development of this thesis, a Service-Learning activity was designed as part of the Master's program in Civil Engineering. The activity, conducted during the 2021-2022 and 2022-2023 academic years, was incorporated into the elective course "Traffic Management", which is offered in the second and final year of the program. This activity involved conducting **semi-structured interviews with older adults in adult daycare centers in Madrid**. The interviews focused on their habits and behavior as pedestrians, their main road safety concerns, and potential countermeasures to enhance their safety. Further details about this activity, particularly in relation to educational innovation, can be found in a conference paper on the subject (Gálvez-Pérez, Mohíno Sanz, et al., 2023).

Although this activity was not initially intended as part of the doctoral thesis, but rather as a complement to educational innovation, the **results of the interviews** are presented here. These findings serve as **informative data** that complement the more extensive analyses presented in the primary publications.

Regarding **pedestrian habits and behavior** (Figure 2.24), almost all respondents (67%) reported that they never or rarely crossed when the pedestrian signal was red. Similarly, 60% of the respondents stated that they never or rarely crossed in areas without designated pedestrian crossings. Furthermore, about 40% mentioned that they actively avoided dangerous intersections, while the majority indicated that they rarely did so. Lastly, older pedestrians reported that they rarely made or received phone calls while walking.

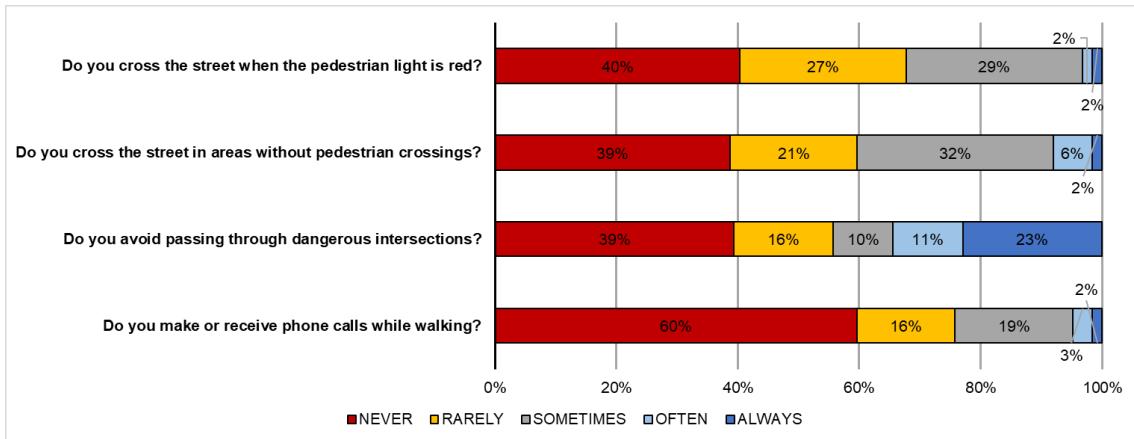


Figure 2.24. Responses of older adults to the survey regarding pedestrian habits and behavior.

Among the **main road safety concerns** of older pedestrians (Figure 2.25), the use of sidewalks by bicycles and scooters, whose usage is increasing, stood out, with almost 80% of respondents selecting this option. Other significant concerns included the condition of sidewalks and their obstruction by objects, which were marked by 63% and 46% of respondents, respectively. It is notable that, as pedestrians, older adults were more concerned with issues related to sidewalks, particularly the presence of vehicles such as scooters and bicycles, as well as the potential for falls due to poor sidewalk conditions and obstructions. Fewer respondents expressed concerns related to motorized vehicles, such as drivers speeding or issues at crossings.

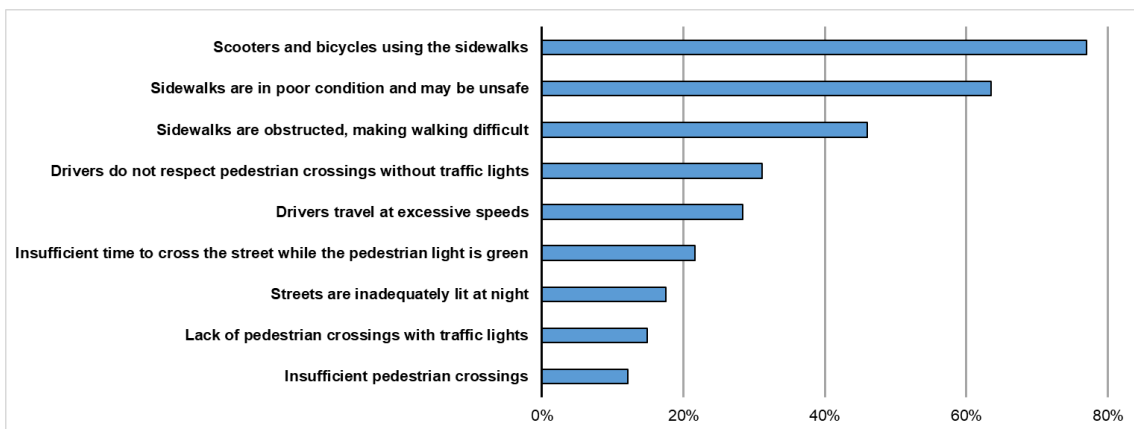


Figure 2.25. Responses of older adults to the survey regarding their main road safety concerns as pedestrians.

The most **proposed countermeasures to improve pedestrian road safety** in this pilot study (Figure 2.26) were related to sidewalks, with 80% of respondents suggesting improvements to sidewalk pavement and 50% advocating for sidewalk widening.

These responses underscore the importance older adults place on proper sidewalks, primarily to prevent falls. Although only 20% of pedestrians reported insufficient time for crossing during the green light phase as a concern, 47% proposed extending the green phase of traffic lights.

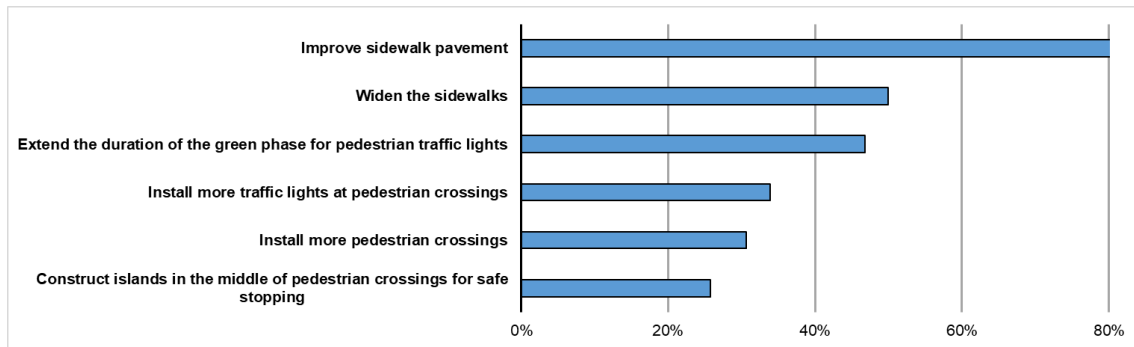


Figure 2.26. Responses of older adults to the survey regarding proposed countermeasures to improve pedestrian road safety.

2.6 *Methods and approaches in studies with traffic crash data*

2.6.1 Outcome variable: crash frequency and crash injury severity

Scientific studies utilizing road safety data from recent decades frequently employ statistical modeling. Such studies can be categorized according to the dependent variables they examine, with **crash frequency** and **crash severity** being the two most common.

Studies where **traffic crash frequency** serves as the dependent variable aim to analyze which characteristics of an area, road segment or intersection, may influence the likelihood of traffic crashes there. In contrast, studies that focus on **traffic crash injury severity**, which may be classified as binary (e.g., fatal versus non-fatal crashes) or multi-categorical (e.g., fatal, severe, minor, or property-damage-only crashes), seek to determine factors contributing to the injury severity of a crash once it has occurred. A third category of studies combines these two approaches, developing frequency models that account for both the number of crashes and their injury severity. For example, a study might include distinct frequency models for fatal and non-fatal crashes, as well as an aggregated model covering all traffic crashes.

The selection of statistical models in these studies is influenced by the nature of the dependent variable. Studies focused on traffic crash frequency typically use statistical models suited for count data, such as Poisson regression or negative binomial regression, as these models accommodate variables representing counts, including zero. In contrast, studies addressing traffic crash injury severity often apply models that allow for categorization of observations, with logistic regression (specifically binary logistic regression) being the most frequently employed. These models predict the probability of a traffic crash' severity, rather than the number of crashes in each area or road segment. The primary models used are further discussed in Section 2.6.3 Traffic crash data modeling techniques.

2.6.2 Spatial approaches in traffic crash studies

Road safety studies, particularly those focused on the frequency of traffic crashes, utilize spatial units of analysis to evaluate road safety conditions across various areas. The approach in terms of the spatial units of study is generally divided into macro-level and micro-level analyses (Ziakopoulos & Yannis, 2020).

Macro-level analysis employs areal divisions within the studied region to identify factors that influence traffic crash frequency across these divisions. In this approach, dependent and independent variables are aggregated at the level of the areal unit. Data acquisition and processing are typically more straightforward than in micro-level analysis. Additionally, the results obtained from macro-level analysis are primarily directed toward urban planning efforts, with a focus on medium- to long-term traffic crash reduction.

In this approach, spatial units can be based on either administrative divisions or ad hoc regions specifically defined by the researcher. Administrative divisions may include national or supra-national areas, such as counties in the United States, autonomous communities in Spain, or municipalities. Additionally, divisions within a municipality can also be analyzed, including smaller units such as districts, neighborhoods, postal codes, or census tracts.

The primary **advantage** of using **administrative divisions** is that socioeconomic, land use, or infrastructure data are often pre-processed by local governments and accessible through open data portals. However, utilizing these divisions may present challenges, such as **the boundary problem**. Since administrative divisions are frequently defined based on the road network (i.e., the place where traffic crashes occur), the boundary problem, or boundary effect, can arise when crashes occur near the borders of the study's areal units. In fact, when administrative divisions are smaller, a larger proportion of the road network coincides with their borders. To address this issue, researchers have applied methods such as double-counting crashes near boundaries by incorporating a buffer or band of predefined width around each spatial unit.

Using **ad hoc spatial units**, such as regular square grids or Thiessen polygons, offers certain advantages. This approach reduces the likelihood of boundary issues and allows for the use of areas with equal or nearly equal sizes, facilitating easier comparisons between zones. However, this method is relatively uncommon because data preparation, particularly data processing, can be more complex and sometimes unavailable. Specifically, raw data must be georeferenced and processed using GIS software. For example, a city could be divided into equal 100x100 m² tiles, with the total length of the road network calculated within each tile. In this case, geospatial data of the road network, including all road axes, would be required.

Micro-level studies are more suitable for analyzing and mitigating road safety issues in the short term. In these studies, dependent and independent variables are aggregated at the level of specific road elements, such as road segments or intersections. This approach allows for a detailed examination of localized factors affecting road safety, enabling targeted interventions to improve safety at high-risk locations. As a

result, these studies can identify precise risk factors and are particularly effective for short-term measures, such as traffic calming, improved signage, or adjustments to traffic signal timing. However, this approach requires more extensive data collection and processing, as variables must be recorded at finer spatial resolutions and may necessitate regular updates to reflect changes in traffic patterns and infrastructure.

Despite the added complexity in data preparation, micro-level analysis is valuable for immediate safety improvements at specific sites and can provide critical insights into the spatial distribution of road safety risks within a given area.

2.6.3 Traffic crash data modeling techniques

Data modeling techniques play a crucial role in road safety studies, especially in understanding injury severity and crash frequency. Both areas aim to uncover the relationship between various factors and the occurrence or the injury severity outcome of traffic crashes. Here, we focus on these two primary types of studies in road safety modeling, emphasizing the most common statistical modeling approaches and their relevance given the data's nature.

Injury severity studies aim to predict the level of injuries sustained in crashes, often categorized by severity levels (e.g., fatal, incapacitating, minor). The discrete, often ordinal, nature of these data requires specific modeling approaches that can handle such data structures. Traditional statistical models, such as logistic and ordered probit models, are commonly applied to address this ordinal structure, ensuring that the data's ordinal property is preserved when estimating the relationships between crash severity and influencing factors (Savolainen et al., 2011). Random parameter models, such as random parameter ordered logit or probit models, allow for varying effects across observations, which improves model accuracy in capturing this unobserved heterogeneity (Savolainen et al., 2011).

In recent years, machine learning (ML) models, particularly random forests (RF), support vector machines (SVM), k-nearest neighbors (KNN), and artificial neural networks (ANN), have gained popularity in injury severity prediction due to their flexibility and non-parametric nature, which does not require predefined relationships between independent and dependent variables (Santos et al., 2022; Silva et al., 2020). For instance, RF models have shown strong predictive performances for injury severity (Santos et al., 2022).

Crash frequency studies focus on predicting the number of crashes within a specific geographic or temporal unit, such as a road segment or intersection over a year. The

primary challenge in modeling crash frequency is dealing with count data, which are non-negative integers and often exhibit over-dispersion, a situation where the variance exceeds the mean (i.e., of the dependent variable). The Poisson regression model is a standard approach for count data, but it assumes equal mean and variance, making it unsuitable when over-dispersion is present. To address this, the negative binomial (NB) regression is widely used, as it introduces an additional parameter to account for over-dispersion, making it more flexible and robust for crash data (Lord & Mannering, 2010). Other issues that arise in crash frequency modeling include zero-inflation, where an excessive number of observations report zero crashes, especially in low-traffic areas. Zero-inflated models and hurdle models are therefore applied to capture these unique distributions accurately, ensuring better fit and more reliable parameter estimates (Lord & Mannering, 2010). Machine learning approaches have also entered crash frequency modeling, with techniques such as decision trees, and neural networks, used to predict crash counts based on a range of input features (Silva et al., 2020).

As Poisson, NB, and zero-inflated NB regressions were used in the works developed in the framework of this doctoral thesis, these are further explained in detail in Section 3.3 Statistical modeling. In terms of the contemporary classification of machine learning techniques (see Figure 2.27), the explained models are categorized under supervised learning techniques, as they operate with known dependent variables. Within this framework, the prediction of injury severity is typically treated as a classification problem, while the estimation of frequency is regarded as a regression problem.

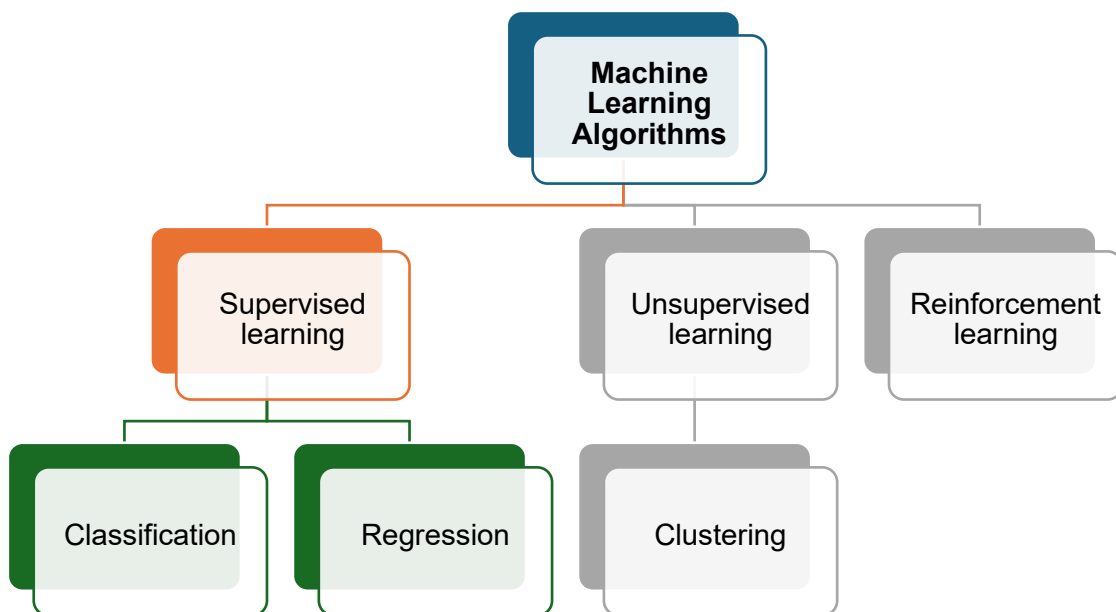


Figure 2.27. Classification of Machine learning algorithms.

2.7 *Studies on older pedestrians using traffic crash data*

During the last decades, research on the road safety of older users has been focused on older drivers (Tournier et al., 2016), even though these users do not represent a disproportion in traffic crashes (Langford et al., 2006, 2008). Nevertheless, there are investigations into older pedestrians using traffic crash data, especially exploratory analyses and research on the injury severity of these traffic crashes.

2.7.1 **Traffic crash profile or exploratory studies**

There is a first group of studies focused on exploring traffic crash data with consideration of older pedestrians. These studies do not usually include statistical modeling techniques to model the frequency or the injury severity of crashes but are more oriented to detect which patterns define older pedestrian traffic crashes, and some of them compare older and non-older pedestrian crashes.

Zegeer et al. (1993) conducted an analysis of motor vehicle crashes involving older pedestrians to understand contributing factors and recommend effective countermeasures. The study examined over 26,000 pedestrian crashes in North Carolina (1980-1990) and more than 70,000 fatal crashes nationwide (1980-1989). Results showed that while older pedestrians are slightly less likely than younger ones to be involved in crashes, they face a significantly higher risk of fatality upon being struck. Older pedestrians were overrepresented in crashes during daylight, on weekdays, and in winter. They were also more likely to be involved in crashes at intersections, especially involving turning vehicles, and in crossing wide streets. Additionally, alcohol involvement was less common among older pedestrians compared to younger groups.

Oxley et al. (1997) analyzed pedestrian traffic crashes between 1987 and 1995 at three roadside shopping precincts, adjacent to arterial roads in and around Melbourne. Most crashes occurred from 9 a.m. to 6 p.m. on weekdays for older and younger pedestrians. Slightly more younger pedestrian crashes occurred after dark than for older pedestrians. Regarding the type of crash, roughly half of crashes occurred when the pedestrian stepped off the curb for all pedestrians. Younger pedestrians were slightly more likely to suffer a collision on the far-side than the near-side of the road, but older pedestrians were much more likely to have a collision on the far-side than the near-side of the road.

Koepsell et al. (2002) employed data from six cities in Washington and California from 1995 to 1999 and analyzed street crossing locations at intersections where an older pedestrian suffered a vehicle collision. They compared the characteristics of these

locations with other control sites of similar characteristics. Results showed that the risk of pedestrian-vehicle collision was 2.1-fold greater at sites with a marked crosswalk. Additionally, results showed that there was a 3.6-fold higher risk in marked crosswalks at sites with no traffic signal or stop sign.

Abou-Raya & Elmeguid (2009) analyzed road traffic victims aged 65 and over admitted to the Emergency and Traumatology Departments of their institutions for a period of 6 months, only non-fatal injuries were considered. Results showed that most of these victims were pedestrians (57%). Regarding older pedestrian injuries, most of them were injured by falls, occurred when crossing the street and/or failing to estimate the proximity of the approaching vehicle. In addition, most injuries sustained by older pedestrians involved lower extremities.

Martin et al. (2010) studied traffic crashes in Ireland from 1998 to 2002 involving older pedestrians (aged 65 and over) and compared them with younger adults (aged 18 to 64). Mortality was more than doubled for older pedestrians compared to younger adults. About the traffic crash profile, older pedestrian crashes occurred in daylight and in good weather conditions.

Das et al. (2019) analyzed older pedestrian fatal crashes in the U.S. from 2014 to 2016 using the empirical Bayes (EB) geometric mean method to identify high-risk factors. The study categorized older pedestrians by age groups (65–69, 70–74, 75–79, and 79+) and examined various crash variables, including pedestrian position, gender, lighting conditions, and crash type. The analysis revealed that older pedestrians, particularly women over 79, were overrepresented in backing vehicle crashes. Male pedestrians aged 65–69 experienced a higher frequency of nighttime segment-related crashes, especially those involving expressway crossings. Additionally, failure to yield at intersections and inadequate street lighting were associated with older pedestrian traffic crashes.

O'Hern et al. (2015) analyzed traffic crashes collisions involving older pedestrians in Victoria, Australia, from 2003 to 2012 and compared them with those involving younger adult pedestrians. The analysis focused on two age groups: 65-74 years and 75+ years. Variables such as pedestrian age, gender, location, injury severity, and collision characteristics (e.g., road geometry, speed zones) were analyzed. The study found that older adults, especially those over 75, experienced higher rates of fatalities and serious injuries compared to younger cohorts. Collisions involving older pedestrians were more frequent at intersections, driveways, and in parking lots. The findings emphasized the increased vulnerability of older adults due to frailty and slower mobility, with a higher likelihood of sustaining severe injuries even in low-speed environments.

S. Kim & Ulfarsson (2019) used single pedestrian traffic crashes in the U.S. from 2012 and 2013 to analyze specific factors differentiating older and non-older pedestrian traffic crashes. The employed model was a logistic regression with a random effect to capture the regional heterogeneity. Results showed that older pedestrians are more likely to be involved in a traffic crash with higher driver age, during street crossing, if the vehicle was turning, if the vehicle is a minivan or a SUV (Sport utility vehicle), and in parking or driveways. On the other hand, pedestrian crashes involving older pedestrians are less common than those involving younger pedestrians in dark (with or without light) conditions, and during summer.

Table 2.14. Reviewed literature on older pedestrian traffic crash profile or exploratory studies.

Study	Analyzed groups	Statistical model	Main results
Zegeer et al. (1993)	Older pedestrians (65+) Non-older pedestrians	-	More fatality for older pedestrians Older pedestrian crashes are more common: daylight, weekday, winter, intersection, wide street, turning vehicle
Oxley et al. (1997)	Older pedestrians (65+) Non-older pedestrians	-	Older pedestrian crashes more common: daylight, weekday, far-side
Koepsell et al. (2002)	Older pedestrians (65+)	Logistic regression	Older pedestrian crashes are common: marked crosswalks, marked crosswalks without signalization
Abou-Raya & Elmeguid (2009)	Older pedestrians (65+)	-	Most pedestrian fall Lower extremity most common injury Pedestrians fail to estimate proximity approaching vehicle
Martin et al. (2010)	Older pedestrians (65+) Non-older pedestrians	-	Higher mortality Older pedestrian crashes are more common: during daylight, good visibility and good weather
Das et al. (2019)	Older pedestrians (65+) fatal crashes	Empirical Bayes	High crash potential for older pedestrians: Failure to yield at intersections Dark with poor street lighting Backing vehicle-related crashes for female pedestrians segment-related crashes at night and expressway crossing at night for male pedestrians
O'Hern et al. (2015)	Older pedestrians (65+) Non-older pedestrians	-	Higher fatality rate Older pedestrian crashes are more common: at intersections, driveways, parking lots, while crossing Higher average hospital stay length for older adults
S. Kim & Ulfarsson (2019)	Older pedestrians (65+) Non-older pedestrians (18-60)	Logistic regression	Older pedestrian crashes are more common: higher driver age, during street crossing, turning vehicle, minivan/SUV, at parking/driveway Older pedestrian crashes are less common: during dark hours, during summer

The reviewed literature along with the used statistical model, analyzed groups of pedestrians and main results are shown in Table 2.14. From the reviewed literature, the most repeated result is that the higher fatality rate (number of fatalities by population) and mortality (number of fatalities by 100 victims) of older pedestrians are higher for older pedestrians than for the rest of the population (Martin et al., 2010; O'Hern et al., 2015; Oxley et al., 1997; Zegeer et al., 1993). Other characteristics that differentiate

older pedestrian traffic crashes from the rest of pedestrian crashes are daylight conditions (S. Kim & Ulfarsson, 2019; Martin et al., 2010; Oxley et al., 1997; Zegeer et al., 1993), weekday (Oxley et al., 1997; Zegeer et al., 1993), winter season (Zegeer et al., 1993), good weather and visibility (Martin et al., 2010). Older pedestrian crashes are less common than other pedestrian crashes during summer season (S. Kim & Ulfarsson, 2019). Regarding the vehicle involved in the crash, older pedestrians are more likely to be hit by a minivan or SUV (S. Kim & Ulfarsson, 2019), and when the vehicle is turning (S. Kim & Ulfarsson, 2019; Zegeer et al., 1993). Regarding the location of the crash, older pedestrians are more involved in crashes than the rest of the population at intersections, during street crossing, in wide streets, in the far-side of the street crossing, and in parking/driveways (Abou-Raya & Elmeguid, 2009; S. Kim & Ulfarsson, 2019; O'Hern et al., 2015; Oxley et al., 1997; Zegeer et al., 1993). About intersections, Koepsell et al. (2002) found greater risk for older pedestrians in marked crosswalks, especially if there are no traffic lights or stop signals. The most common locations of the injury were lower extremities, head/neck, and abdominal areas, and the hospital stay length was greater for older adults (Abou-Raya & Elmeguid, 2009; O'Hern et al., 2015).

2.7.2 Traffic crash injury severity studies

Other specialized studies on older pedestrians deal with the injury severity of traffic crashes. Given that a traffic crash occurred, these studies employ statistical models to show which factors influence the injury severity of this crash. In comparison with crashes with a focus on traffic crash frequency, which are explained in the subsequent section, there are remarkably more studies into injury severity of older pedestrian traffic crashes.

In the reviewed literature specialized on the injury severity of older pedestrians, higher values of both the age of the pedestrian (Guo et al., 2021; Kitali et al., 2017; Laković et al., 2023; Wang et al., 2017) and the posted/driver speed (Guo et al., 2021; Kitali et al., 2017; Lalika et al., 2022; Wang et al., 2017) showed to be related to more severe traffic crashes. Additionally, crashes involving alcohol (Kitali et al., 2017; Noh et al., 2018), vehicle moving straight (Kitali et al., 2017; Lalika et al., 2022), heavy vehicle (Noh et al., 2018; Wang et al., 2017), and male driver (Lalika et al., 2022; Wang et al., 2017). Conversely, crashes were expected to be less severe during daytime (Guo et al., 2021; Lalika et al., 2022; Wang et al., 2017), at intersection (Guo et al., 2021; Noh et al., 2018), and with a higher age of the driver (Lalika et al., 2022).

Table 2.15. Reviewed literature on older pedestrian traffic crash injury severity.

Study	Analyzed groups	Statistical model	Main results
Wang et al. (2017)	Older pedestrians (65+)	Ordered probit regression	More severe crashes: higher age of the pedestrian, posted/driver speed, heavy vehicle, male driver Less severe crashes: daytime
Kitali et al. (2017)	Older pedestrians (65+)	Logit, probit, cloglog regressions	More severe crashes: higher age of the pedestrian, posted/driver speed, alcohol, vehicle moving straight
Noh et al. (2018)	Older pedestrians (65+)	Binary logit regression	More severe crashes: alcohol, heavy vehicle Less severe crashes: at intersection
Guo et al. (2021)	Older pedestrians (65+)	XGBoost	More severe crashes: higher age of the pedestrian, posted/driver speed Less severe crashes: daytime, at intersection
Lalika et al. (2022)	Older pedestrians (65+)	Bayesian logistic regression	More severe crashes: posted/driver speed, vehicle moving straight, male driver Less severe crashes: daytime, higher age of the driver
Laković et al. (2023)	Older pedestrians (65+)	Ordinal logistic regression	More severe crashes: higher age of the pedestrian

The reviewed scientific articles on injury severity in older pedestrian traffic crashes do not use separate models for older pedestrians and the rest of the population (Table 2.15). This lack of distinction is significant because it prevents meaningful comparisons. Without separate models, it is impossible to determine whether older pedestrians exhibit unique patterns or if the factors influencing their injury severity are the same as those affecting younger pedestrians. Furthermore, these studies fail to consider environmental variables related to the crashes, such as socioeconomic factors, land use, or infrastructure characteristics, and are limited to independent features present in the traffic crash database.

2.7.3 Traffic crash frequency studies

Frequency studies analyze which factors make areas, streets or junctions more likely to register a higher number of traffic crashes. These studies are scarce in comparison to the rest of studies on older pedestrian traffic crashes using traffic crash data, maybe because the data collection process, which should include gathering data on socioeconomics, land use, and infrastructure, is intricate and data might not be available for all cases.

Two main scientific contributions are included in the group of traffic crash frequency studies at the **macro-level**. First, Dumbaugh & Zhang (2013) used traffic crash data from San Antonio-Bexar County (Texas, U.S.) from 2003 to 2007 to analyze traffic crashes at urban locations involving older drivers and older pedestrians and cyclists.

In this study, older people were selected with the threshold of 75 years old. Two negative binomial models were fitted to predict the number of crashes involving older pedestrians and cyclists and killed and severely injured crashes involving older pedestrians and cyclists per block group. Results showed that the number of older adults, arterial roads and big box stores increased crashes involving older pedestrians, while dense networks of lower-speed roads reduced them.

Other study published by Lee et al. (2020) investigated older pedestrian traffic crashes in Seoul (Korea) at the census block group level with data from 2015 to 2017. They built different negative binomial models for all and older pedestrian crashes, considering the injury severity of those crashes and the area's economic wealth. Results showed that four-way intersections and crosswalks impacted the number of pedestrian crashes and the severity of older pedestrians only in low-priced-housing areas, implying that pedestrian measures may be unevenly distributed across the city.

In addition to this studies at the macro-level, two articles were identified to be traffic crash frequency studies at the **micro-level**. First, a study focused on **street intersections** by D. Kim (2019) examined older pedestrian traffic collisions at intersections in Los Angeles County, focusing on the impact of intersection characteristics, ambient conditions, and land use patterns. Using data from 2015 to 2017, the study employed a multinomial logistic regression model to differentiate factors contributing to older pedestrian collisions compared to younger pedestrians. Results indicated that raised medians, three-way intersections, parks and recreational land, and street trees make a positive contribution to the safety of older pedestrians, while bus stops increased the likelihood of an intersection to have a high concentration of older pedestrian collisions. In contrast, intersections with crosswalks or colored crosswalks contribute to safer scenarios for younger pedestrians.

A second study focused on the **street segment level** by Lv et al. (2021) investigated older pedestrian safety in Changning District (Shanghai) from 2013 to 2015, using Poisson and geographically weighted Poisson regression (GWPR) models. The study analyzed the influence of the built environment, exposure variables, and community features on older pedestrian collisions. The Poisson model indicated that green space and sidewalks proportions reduced the number of older pedestrian crashes. On the other hand, a higher number of road junctions, roads around nursing homes, schools, bus stops, metro stations, traditional markets, and supermarkets were hazardous for older pedestrians. Results from the GWPR model highlighted spatial variation, showing that green space decreased crash risk only in non-congested areas.

Table 2.16. Reviewed literature on older pedestrian traffic crash frequency.

Study	Spatial unit of study	Analyzed groups	Statistical model	Main results
Dumbaugh & Zhang (2013)	Neighborhood (macro-level)	Older pedestrians and cyclists (75+)	Negative binomial regression	More crashes: arterial roads, big box stores Less crashes: low speed network
Lee et al. (2020)	Census block group (macro-level)	Older pedestrians (65+) All pedestrians	Negative binomial regression	More crashes: crosswalks, BRT station, 4-leg intersections, hospitals, pubs
D. Kim (2019)	Intersection (micro-level)	Older pedestrians (65+) Non-older pedestrians	Multinomial logistic regression	More crashes: bus stops Less crashes: raised median, 3-way intersections, parks, trees
Lv et al. (2021)	Street segment (micro-level)	Older pedestrians (65+)	Poisson regression and geographically weighted Poisson regression	More crashes: junctions, schools, bus stops, metro stations, traditional supermarkets Less crashes: green space, sidewalks

From the reviewed literature (Table 2.16), it can be summarized that areas with higher arterial roads, big box stores and public transport points (i.e., bus stops and metro stations) are riskier for older pedestrians. In contrast, areas with lower speed road networks are safer for them. Regarding road junctions, the existence of raised median, 3-way intersection, parks and trees seem to reduce the likelihood of an intersection to have a high concentration of older pedestrian collisions, while the presence of bus stops increases that likelihood. Regarding road segments, higher proportions of green space and sidewalk reduce the number of older pedestrian crashes, and more road junctions make segments riskier.

Although these results are interesting and focused on the features of the surroundings of each traffic crash, which implies remarkable data acquisition and data processing efforts, some of the used methodologies do not allow us to gain sufficient insight into the problem. For example, the works by Dumbaugh & Zhang (2013) and Lv et al. (2021) develops only models for older pedestrian traffic crashes, and not for the rest of the pedestrian crashes. Similarly, the work by Lee et al. (2020) develops separate models for older pedestrian and all pedestrian traffic crashes. This approach allows comparison between the two samples, but one sample is part of the other, potentially leading to effects on the older pedestrian model disappearing in the general model if the effect is only significant for older pedestrians. Even worse, contrary effects on older and non-older populations could cause a variable to simply be not significant in the “all” pedestrian model, resembling the Simpson’s paradox (Blyth, 1972).

2.7.4 Main countermeasure proposals to improve older pedestrian road safety

General countermeasures to improve older pedestrian road safety typically include behavior and education programs, and infrastructure improvements, with measures aimed at speed reduction, separation of vehicular and non-vehicular traffic, simplification of traffic situations (European Commission, 2015; Oxley et al., 2004). Specifically, the commonly proposed countermeasures are as follows:

- **Behavior and education programs** should be aimed at educational campaigns to:
 - Promote the use of specialized clothing.
 - Increase awareness of age-related cognitive and sensory declines.
 - Inform and train pedestrians about the proper search behavior on the road or related to novel traffic rules, situations and road layouts in urban and rural areas.

- Measures aimed at vehicle **speed reduction**
 - Reducing the average speed of motorized traffic at locations with high pedestrian flows.
 - Setting appropriate speed limits in high-risk areas

- **Separation of vehicular and non-vehicular traffic** may also benefit pedestrian road safety
 - Provision of a pedestrian footpath network that is: comfortable, safe, and well maintained
 - Provision of barriers or guardrails to limit pedestrian access to the road at unsafe mid-block.
 - Provision of crossing facilities in high-risk areas.

- **Simplification of traffic situations**
 - Adjusting the traffic signals to allow for the slower walking speed of older pedestrians in urban areas.

- Lighting at intersections and pedestrian crossings.
- Installation of conspicuous 'give way to pedestrian' signals.

2.7.5 Scientific gap in the literature and needs

The body of literature addressing the behaviors, limitations, and perceptions of older adults as pedestrians is significantly broader than that utilizing traffic crash data. Furthermore, most studies analyzing the frequency or severity of injuries among older pedestrians through statistical modeling do not typically develop distinct models for comparing older and non-older pedestrian traffic crashes. Instead, they predominantly focus on older pedestrian traffic crashes.

As a result, there is a clear need for further research on older pedestrian traffic crashes, particularly regarding their frequency, as the existing studies are more limited in this area. Additionally, such research should involve the development of distinct statistical models to identify the factors influencing the frequency and injury severity of traffic crashes separately for older and non-older pedestrians. This approach would enable the analysis of both common patterns shared by the two age groups and the unique patterns specific to crashes involving older pedestrians. Furthermore, this framework is essential for testing whether certain factors simultaneously influence the road safety of older and non-older pedestrians in opposing ways. Under such circumstances, countermeasures designed to improve the road safety of older pedestrians could inadvertently compromise the safety of other pedestrian groups.

Rather than relying solely on the variables already available in the traffic crash database, this research should include variables concerning the surrounding conditions of traffic crashes, such as socioeconomic factors, land use characteristics, and, particularly, infrastructure. These additions are crucial for gaining deeper insights into the road safety challenges faced by older pedestrians and for proposing effective countermeasures to address these issues.

Older pedestrian fatalities are currently concentrated on urban roads, and this number may potentially rise due to demographic mega-trends such as population ageing and urbanization. As a result, urban scenarios (i.e., areas within built-up environments) should be prioritized in such analyses. However, conducting road safety analyses in urban settings presents unique challenges compared to studies focused on interurban roads. Factors such as limited data on vehicle flows, the absence of comprehensive information on street infrastructure, and the inherent complexity of urban environments significantly complicate this type of analysis.

In addition, the covid-19 pandemic situation could have altered mobility patterns of older adults, and, consequently, factors related to the traffic crashes they suffer as pedestrians. Furthermore, covid-19 has boosted the use of micromobility vehicles, such as bicycles and electric scooters, even after the pandemic ended. In consequence, there is a need to study the impact of the pandemic on the road safety of older pedestrians in urban areas in terms of frequency, injury severity, and crash-related factors.

The following section outlines the materials and methods employed in the development of this doctoral thesis. The section includes details on data sources, ad hoc designed systematic algorithms for data processing, statistical modeling techniques, and the chronology of publications in indexed journals.

3

Materials and Methods

The third chapter of this thesis examines the materials and methods used to investigate traffic crashes involving older adults in urban areas of Spain. The study primarily focuses on exploring the factors linked to the surroundings of these crashes, particularly the influence of socioeconomic, land use, and infrastructural variables on the frequency and injury severity of traffic crashes involving older pedestrians. This research addresses a significant gap in literature specifically, the lack of detailed understanding of how these contextual factors contribute to the vulnerability of older adults in urban settings.

To bridge this gap, a comprehensive dataset was required. The traffic crash database provided by the Directorate-General for Traffic (DGT) was selected as the primary data source. The DGT database is an official and detailed repository of traffic incidents across Spain, offering extensive information about the characteristics of crashes, vehicles, and individuals involved. However, despite its richness in certain aspects, the database often lacks critical variables necessary for analyzing the contextual factors influencing crash occurrences. Furthermore, the location of traffic crashes is reported as postal addresses, necessitating an additional process to geolocate these events.

To address these limitations, this chapter outlines the comprehensive data processing methods developed to enhance the DGT database. The integration of additional data sources was essential to construct a more complete depiction of the context in which these crashes occurred. This integration involved significant data wrangling and geospatial processing, particularly for geocoding crash locations—a major undertaking—and linking them to relevant environmental characteristics. A substantial portion of this chapter is devoted to describing the geolocation of traffic crashes. A custom geocoding process was developed to accurately determine the locations of crashes in Madrid.

The chapter begins by describing the content and structure of the DGT traffic crash database (**Section 3.1**). It then proceeds to detail the procedures for database enhancement (**Section 3.2**), including the definition of variables, the acquisition of supplementary data to fill the gaps in the DGT database, and the subsequent data

processing. Finally, the statistical modeling techniques used (**Section 3.3**) and the application of the methodology to the developed publications (**Section 3.4**) are presented.

In summary, Chapter 3 presents a comprehensive methodology for investigating the influence of factors on traffic crashes involving older adults. By addressing the limitations of the DGT database through extensive data processing and the integration of supplementary data sources, this research provides valuable new insights into the road safety risks faced by older pedestrians in urban environments.

3.1 Spanish traffic crash database

The primary data source for this thesis was the traffic crash database of the Directorate-General for Traffic (DGT). Traffic crashes in Spain are systematically recorded at the national, regional, and local levels by law enforcement agencies using the ARENA System, which serves as the official registry. ARENA stands for ‘Accidentes: REcogida de iNformación y Análisis’ or ‘Traffic crashes: data gathering and analysis.’ The DGT oversees the National Registry of Traffic Crashes Victims, responsible for collecting essential data used in the analysis of traffic crashes and informing road safety policy decisions.

Traffic officers are responsible for completing on-site traffic crash reports, which include detailed information about the victims, traffic crash circumstances, vehicles involved, and individuals affected. In the event of serious or fatal traffic crashes, data transmission occurs rapidly, with a one-month maximum period for complete submission. These records are further supplemented by additional sources, such as hospital data and forensic reports, to ensure data quality and accuracy.

The traffic crash database contains variables associated with the following categories:

- **Crashes:** Includes the date, time, location (province, municipality), type of road, weather conditions, and road surface conditions.
- **Vehicles:** Covers the type, any prior anomalies, maneuvers, insurance details, and vehicle inspection (ITV) records.
- **Individuals:** Includes details such as age, gender, nationality, use of seatbelt or helmet, injury severity level sustained, results of alcohol/drug tests, and driving licenses.

3.1.1 Definitions applicable to the Spanish traffic crashes

Key definitions from Spanish traffic law are presented here, such as the distinction between “fatal traffic crashes”, “serious traffic crashes”, and the specific categorization of injuries sustained by victims. A complete understanding of these terms is essential to analyzing the patterns of crashes and injuries accurately. Important definitions regarding the traffic crashes recorded in Spain are included in the “Orden INT/2223/20-14, de 27 de octubre, por la que se regula la comunicación de la información al Registro Nacional de Víctimas de Accidentes de Tráfico”. These definitions are summarized here for better understanding of the data used:

A **traffic crash with victims** refers to an incident occurring on roads or areas governed by traffic and road safety legislation, where one or more individuals are injured or killed. Traffic crashes must involve at least one moving vehicle, which may include trams, trains, or other rail vehicles, provided the incident occurs on public roads. Multi-vehicle collisions are treated as single traffic crashes. This definition excludes traffic crashes resulting from natural causes (e.g., natural deaths), suicides, attempted suicides, homicides, or intentional injuries.

A **vehicle** is considered **involved in a traffic crash** under the following conditions:

- It collides with another vehicle, pedestrian, animal, or obstacle.
- It causes injury or death to its driver or passengers, or results in property damage, even without a collision.
- It is stopped or parked dangerously and contributes to the traffic crash.
- The behavior of the driver or passenger influences the traffic crash, irrespective of a collision.
- The driver or passenger is struck while entering, exiting, or after falling from the vehicle.

At the same time, a **persons involved in a traffic crash** includes vehicle occupants and pedestrians whose behavior or presence contributes to the traffic crash. Definitions include:

- **Driver:** The individual operating the vehicle at the time of the traffic crash, including those supervising learner drivers.

- **Passenger:** An individual inside or on the vehicle or struck while entering or exiting the vehicle.
- **Pedestrian:** A person not inside the vehicle involved. This includes those pushing or walking with bicycles or mopeds, individuals in wheelchairs or on skates, people repairing or pushing a vehicle, or individuals struck after leaving a vehicle and walking away. Pedestrians also encompass those guiding or riding on animals.

Among people involved in traffic crashes, victims can be classified according to the severity of injuries as follows:

- **Fatality:** A person who dies because of a traffic crash, either at the scene or within thirty days after the crash, based on mortality databases. Cases of natural death or suicide are excluded.
- **Seriously injured:** A person who is injured in a traffic crash and requires hospitalization of more than 24 hours.
- **Slightly injured:** A person who receives medical treatment without prolonged hospitalization and is not classified as deceased or seriously injured.

Considering the previous victim severity categorization, traffic crashes can be classified into three groups:

- **Traffic crash with victims:** This refers to a crash involving at least one injured or deceased person.
- **Fatal traffic crash:** A crash where at least one person has died.
- **Serious traffic crash:** Classified as serious if, although there are no fatalities, at least one person is injured and requires hospitalization for more than twenty-four hours.

In this context, it is important to distinguish between the concepts of severity of the victims and severity of the crash, always bearing in mind that severity is linked to the fatal result or the time of hospitalization, and not directly linked to the severity of the injuries of the victim. Any crash data filtering and analysis should relate the traffic crash variables with the information of the victims involved, as the severity of the victims will condition the severity of the crash.

3.1.2 Communication of traffic crashes

The procedure for reporting traffic crash data depends on the severity of the incident:

- **Fatal and serious crashes:** For crashes resulting in fatality or an injury requiring hospitalization for more than 24 hours, the data classified as urgent must be submitted within 24 hours of the traffic crash. The remaining information may be submitted incrementally; however, the full report must be completed within one month of the traffic crash date.
- **Other crashes:** For crashes that do not involve fatalities or serious injuries, the urgent data must be submitted within 10 calendar days from the traffic crash date. The complete report should be submitted within one month, allowing for incremental updates as more information becomes available.

There are three main possibilities for traffic officers to include the recorded traffic crashes in the database in the case of the current ARENA2 system, which are recorded in the Order INT/2223/2014:

- A. Use of the ARENA2 Application, which includes the following functionalities:**
 - Input of rapid supply data, which serves as a preliminary version of the complete form.
 - Progressive updating of data with each new piece of information available, maintaining historical versions.
 - Control measures to prevent the duplication of traffic crash records.
 - Selection of vehicle and individual data from the DGT's central records, minimizing errors.
 - Data quality control to ensure consistency and completeness of the forms.
 - Information queries with filters to visualize traffic crashes by location and date, allowing the download of data tables for analysis.
- B. Web Service.** Local or regional administrations that do not use ARENA2 can submit their data via a web service, which meets the same quality and communication time requirements.
- C. Paper Forms.** In municipalities where it is not possible to use ARENA2 or the web service, paper forms are allowed for a period of three years. These forms

are sent to the Provincial Traffic Headquarters, where they are then digitized and entered ARENA2.

The trend in Spain is the massive use of ARENA 2 Application, as it avoids the loss of traffic crash information. The accident GPS coordinates is beginning to be recorded by some Spanish local Administrations and this fact will ease many road safety territorial analysis in the future.

3.1.3 Data collection systems: ARENA and ARENA2

ARENA and ARENA2 systems serve as essential tools for collecting and analyzing traffic crash data in Spain. Introduced in 2005, the ARENA system aimed to unify various traffic crash data services into a single platform. Officially named 'Traffic crashes: Information Collection and Analysis,' (*Accidentes: REcodiga de Información y Análisis*, in Spanish) it was designed to streamline traffic crash data management by enabling the rapid collection of information from police reports for subsequent analysis.

On January 1, 2014, ARENA was succeeded by ARENA2, which introduced significant advancements. These improvements included enhanced data traceability and the ability to access web services from the DGT for the automatic inclusion of vehicle and driver identification from their respective records. This development improved data accuracy, reduced human error, and facilitated the integration of information from multiple sources.

Although these systems primarily focus on data collection, they do not perform the final statistical analysis of the data, as the information in ARENA2 remains subject to modification until verified. Nevertheless, they have played a crucial role in consolidating a national database, thus enhancing tracking and statistical analysis of traffic crashes and informing road safety decision-making.

3.1.4 Database structure

As the traffic crash collection system used by the DGT has undergone two different versions, ARENA and ARENA2, the data schema formats also changed over time. The ARENA schema was operational from 2006 until 2016, although some municipalities transitioned to the new ARENA2 format starting in 2014. While the differences between these two schemas are subtle, they are nonetheless significant.

The ARENA data scheme comprised three tables (crashes, individuals, and vehicles), whereas ARENA2 expanded this data structure to include five tables (crashes, pedestrians, drivers, passengers, and vehicles). The scheme of both database formats is shown in Figure 3.1. Furthermore, the ARENA2 schema was specifically designed to minimize data loss and errors during data collection. Despite these improvements, both schemes support similar workflows, although ARENA2 allows for more comprehensive data collection.

Database format	Tables				
ARENA	Crashes	Individuals			Vehicles
ARENA 2		Drivers	Passengers	Pedestrians	

Figure 3.1. Scheme of ARENA and ARENA 2 database formats.

The **ARENA database** is provided as text files with a 'txt' extension, and it contains three tables that can be connected to each other:

1. The crashes table contains one entry per traffic crash.
2. The individuals table lists everyone involved in a traffic crash, whether they are drivers, passengers, or pedestrians.
3. The vehicles table includes a record for each vehicle involved in a traffic crash.

Both the individuals and vehicle tables may contain multiple entries per traffic crash. For example, a single crash might involve two vehicles or both a vehicle and a pedestrian.

The **ARENA2** database is provided as files with a 'csv' extension, and it contains five tables that can be connected to each other:

1. The crashes table contains one entry per traffic crash.
2. The pedestrians table includes each pedestrian involved in a traffic crash.

3. The drivers table includes each driver involved in a traffic crash.
4. The passengers table includes each vehicle passenger involved in a traffic crash.
5. The vehicles table includes a record for each vehicle involved in a traffic crash.

Consequently, each of these three tables (pedestrians, drivers, and occupants) may contain multiple records for the same traffic crash. However, it is possible that some crashes may not have corresponding records, as not all user types are necessarily present in every crash.

3.1.5 Relational databases

The DGT traffic crash database is structured as a **relational database**. This type of database is organized into various tables to optimize storage efficiency and facilitate user queries. The tables are related to each other with unique identifiers, known as keys.

In the case of the ARENA structure, the main table is the 'crashes' table, which rows can be uniquely identified using a single **identifier** (ID) for each crash. This table is related with the 'individuals' table with a 'one-to-many' relationship, as for one crash, more than one individual can be present. The rows can be uniquely identified using simultaneously the **ID** of the **crash**, the **ID** of the **vehicle**, and the **position inside of the vehicle**. Finally, the 'vehicles' table is related to the 'crashes' table also with a 'one-to-many' relationships, as for each crash we can have more than one vehicle. The rows of the 'vehicles' table are uniquely identified with the **ID of the crash**, and the **ID of the vehicle**. The scheme of the ARENA database is shown in Figure 3.2.

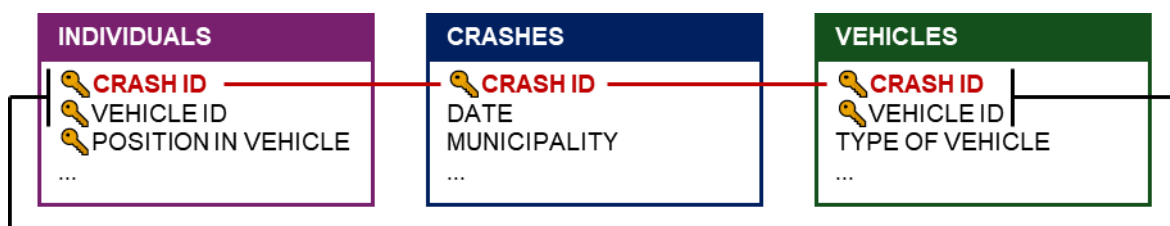


Figure 3.2. Scheme of the ARENA Database.

In the case of the **ARENA2 structure**, the 'crashes' and 'vehicles' tables present the same guidelines as in the ARENA structure. The 'pedestrians', 'drivers', and 'passengers' tables are related with the 'crashes' table with 'one-to-many' relationships; and with the 'vehicles' tables with 'one-to-one' relationship in the case of the 'drivers' table,

and with 'one-to-many' relationship in the case of the 'passengers' table. The scheme of the ARENA2 database is shown in Figure 3.3.

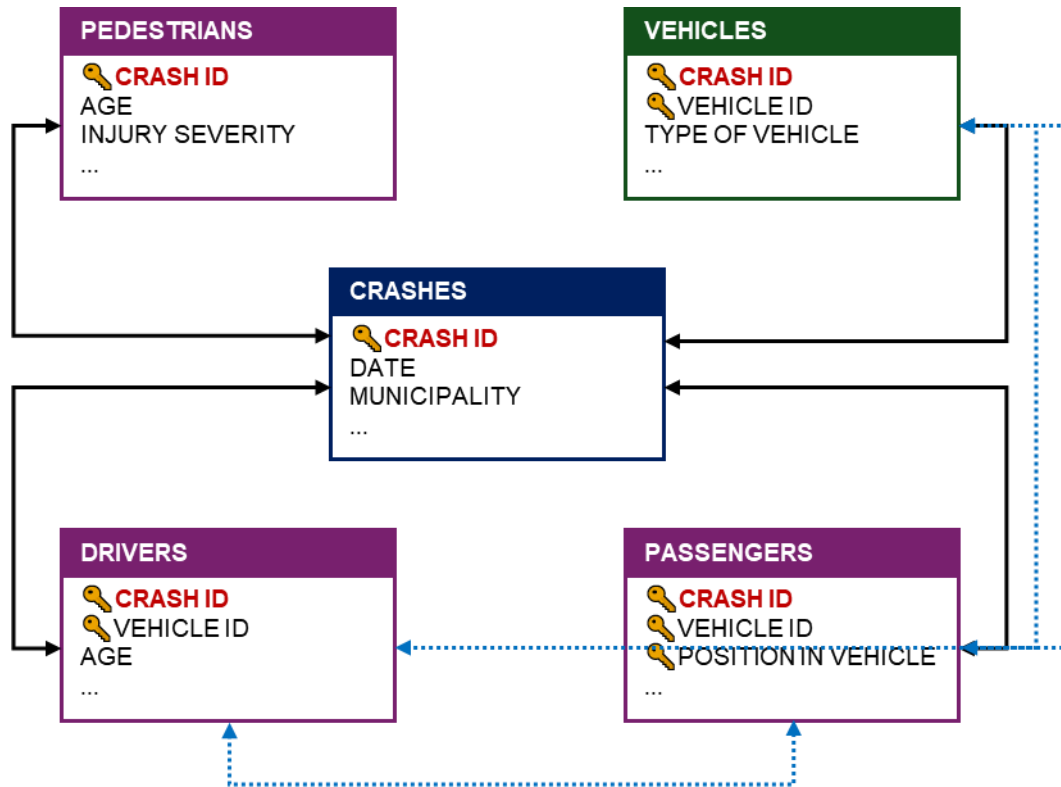


Figure 3.3. Scheme of the ARENA2 Database.

3.1.6 Database queries

Relational databases often require joining multiple tables to perform queries involving data from more than one table. While this process can be implemented using widely available software like Microsoft Excel, it typically necessitates modifying the original dataset, which could lead to errors and be highly time-consuming. In contrast, employing more specialized procedures, such as those involving SQL, is preferable. Structured Query Language (SQL) is a domain-specific language designed for managing databases, particularly relational databases.

SQL allows users to perform queries without altering the original database and executes them efficiently. Its primary functions include joining tables, selecting specific columns, and summarizing data by grouping it based on columns. The principal methods for joining two tables are inner, full, left, and right joins (Figure 3.4). SQL can be employed in specific software such as Microsoft Access or PostgreSQL; however,

these platforms are not capable of conducting statistical analyses on the resulting datasets. Therefore, result tables need to be exported to different software for analysis. In this thesis, SQL techniques were implemented in R using the 'dplyr' package, which enabled direct data import into R and allowed for complex queries, such as generating a table of vehicle-pedestrian collisions that includes the age of the pedestrians. Later, the treated data could be directly analyzed using the same software, eliminating the need for external tools for statistical analysis.

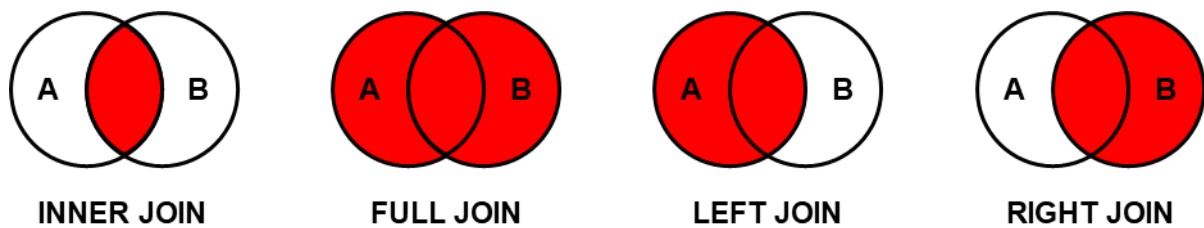


Figure 3.4. Common SQL join types applicable to the road safety official database.

3.1.7 Database deficiencies

Although the DGT's traffic crash database is a rich and valuable resource for road safety research, it suffers from several notable deficiencies that complicate the analysis, particularly when the focus is on understanding the factors related to the surroundings of crashes involving older pedestrians. One of the most significant limitations is the absence of precise geographic coordinates for each crash. Instead of providing exact latitude and longitude data, the database typically reports crash locations using postal addresses. This reliance on postal addresses introduces several challenges, including inconsistencies in address formatting, the potential for missing or incomplete address data, and difficulties in geocoding these addresses accurately to pinpoint the precise crash locations.

In addition to the absence of geographic coordinates, the database lacks critical variables that describe the surrounding environment of each crash. Variables related to socioeconomic, land use and infrastructure features are not included in the DGT database. These contextual factors are crucial for understanding the underlying causes of crashes, especially for vulnerable populations like older pedestrians. Their absence necessitates the integration of supplementary data from external sources, such as the National Statistics Institute (INE), OpenStreetMap (OSM), and the National Geographic Institute (IGN), to build a more comprehensive analysis of crash risk.

These deficiencies significantly limit the DGT database's usefulness as a standalone resource for examining the influence of environmental and contextual factors on traffic crashes. As a result, considerable effort must be invested in data processing and the integration of additional data sources to overcome these gaps and enable a more detailed spatial analysis of crash patterns and contributing factors.

3.2 *Database elaboration for the research*

This section covers all the issues concerning database elaboration considering the needs of this research. During the development of this doctoral thesis, database elaboration, which includes data acquisition, and data wrangling, has been a major task, consuming remarkable time and effort. In the next paragraphs, general approaches are shown for reproducibility purposes. Finally, the application of the general procedure to the four published articles is described.

3.2.1 Variable definition

3.2.1.1 **Dependent variable**

The dependent variable of road safety studies is commonly the frequency of crashes within a zone or road element, or the injury severity of each crash.

If the dependent variable is the frequency of traffic crashes, it typically represents the number of traffic crashes occurring within a defined spatial unit, which can be at the macro- or micro-level. This count-based nature often lends itself to statistical models such as Poisson or negative binomial regression, depending on the presence of over-dispersion. These models are used for analyzing what makes that in a spatial unit of analysis more crashes occur than in others.

If the dependent variable is the injury severity of each crash, it is usually categorized into different levels, such as minor injuries, serious injuries, or fatalities. In such cases, ordered logit or probit models are commonly employed to account for the ordinal nature of the severity levels. Alternatively, when focusing on binary outcomes, such as major versus minor crashes, logistic regression models are often used. Given that a crash occurred, these models help identify factors associated with more severe outcomes.

3.2.1.2 **Independent variables**

3.2.1.2.1 *Socioeconomic variables*

In this doctoral thesis, socioeconomic variables related to the environment of older pedestrian-vehicle collisions were examined. The variables considered included population by age group, population density, and average income per person. Although these variables are critical for analyzing road safety, they are not directly available in the DGT's traffic crash database. However, socioeconomic data can be sourced from

official databases such as the National Statistics Institute (INE) for the entire Spanish territory or from municipal open data portals for studies focused on specific cities. This information is generally provided in terms of administrative divisions (e.g., municipalities, districts, or neighborhoods), meaning that the level of detail for these variables cannot exceed the granularity of the administrative division used.

3.2.1.2.2 Land use variables

Land use variables were also included in this study. As in the case of socioeconomic variables, this information is not available in the traffic crash database. Land use variables include principal land use (i.e., primary, secondary, tertiary, residential) proportion, land use mix index (e.g., entropy index, or HHI index), and points of interest (POI) of different nature. These variables can be retrieved from multiple data sources, such as OpenStreetMap (OSM) for worldwide data, National Geographic Institute (IGN) for Spain data, or municipal open data portals for studies focused on specific cities. Note that land use presence data is often provided as geographic area data, while POI data, because of its own nature, is often provided as geographic point data.

3.2.1.2.3 Infrastructure variables

Regarding infrastructure variables, these are also not available in the traffic crash dataset. Instead, these were retrieved from OSM and IGN in most cases. This group of variables includes street devices, such as trash containers or trees, public transport nodes, such as bus stops and metro stations, proportions of type of streets (e.g., major, or minor roads), and sidewalk surface. These elements are vector geographic data that can be either points, lines or polygons. The data processing of these variables is considerably different if the study is at the macro- or micro-level.

3.2.2 Spatial unit of analysis

In spatial analysis, the choice of the unit of study plays a fundamental role in the structuring and interpretation of the data. The unit of study refers to the spatial or non-spatial entity to which all variables, both dependent and independent, are associated and analyzed. It is crucial to define these units appropriately, as they serve as the foundation for aggregating, modeling, and interpreting the data in the context of the research question. Units of study can broadly be categorized into macro and micro levels, depending on the scale and granularity of the spatial entities being analyzed.

In both macro- and micro-level, the dependent and independent variables are expressed and processed in relation to each spatial unit of analysis. This ensures that the data is structured coherently, facilitating accurate spatial analysis and the interpretation of results. Whether focusing on broad trends at the macro level or detailed patterns at the micro level, selecting the appropriate unit of study is crucial for achieving the research objectives and for the proper application of statistical and spatial models in the analysis of urban mobility and road safety.

3.2.2.1 Macro-level units

Macro units of study typically refer to larger, aggregated geographic areas. These may be administrative divisions, such as cities, municipalities, neighborhoods, or even larger areas such as regions and provinces. Alternatively, macro units could also be regular polygons, such as grid cells of equal size, that are used to discretize the study area into manageable units. For instance, a study might focus on comparing road safety outcomes across different neighborhoods or administrative districts, where each unit of study represents an entire district. In this case, variables such as the number of pedestrian crashes, average age of the population, or the level of road infrastructure investment would be aggregated and analyzed at the macro-level. Three examples of macro-level units of analysis are shown in Figure 3.5.

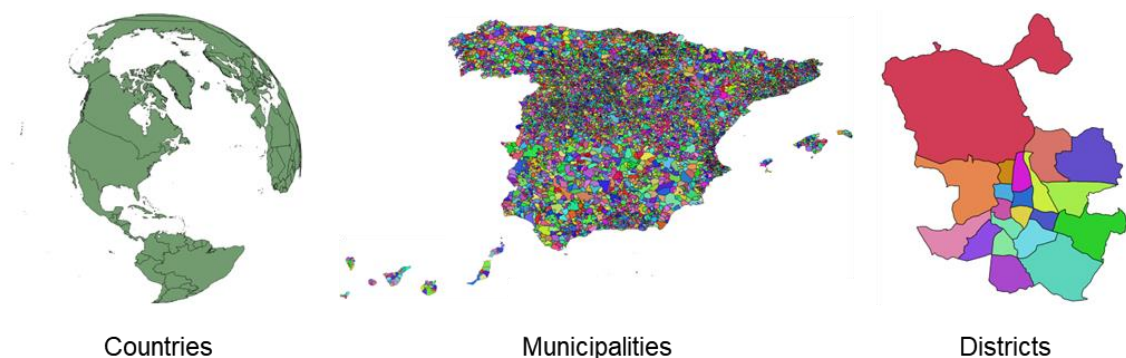


Figure 3.5. Examples of macro-level spatial units of analysis.

3.2.2.2 Micro-level units

Micro units of study, on the other hand, refer to more detailed and smaller-scale spatial entities, such as road segments, intersections, or junctions. These micro units provide a finer resolution for analyzing spatial phenomena and are particularly useful when

studying localized effects or high-resolution spatial patterns. For example, in pedestrian safety research, each road segment or intersection could represent a micro unit, with variables such as the number of traffic crashes, pedestrian flow, visibility, and the presence of crosswalks or traffic signals being measured and analyzed for each unit. Micro units allow for a more precise understanding of how specific locations contribute to overall trends and can be critical for targeted interventions, such as identifying hazardous junctions or road segments where traffic crashes are more likely to occur.

3.2.2.3 Traffic crash level unit

An alternative method is to select the traffic crash as the unit of study, instead of any other spatial unit of analysis. This approach would consider each crash as an observation, and the variables of the surrounding of each crash would be associated with it. This procedure can be done by assigning for each crash the characteristics of an administrative division where it is located. Another approach, more computationally demanding but more precise, would be to assign the characteristics of the surroundings of the crash, considering the surroundings of the crash as a buffer from the crash with a determined radius. This unit of analysis is more oriented towards injury severity studies, in which the dependent variable is different for each crash, regardless of the district or street segment where they were registered.

3.2.3 Data acquisition

3.2.3.1 Data formats

The data used in this thesis can be categorized into two primary formats: tabular and geographic. Both formats are essential for handling different aspects of the analysis, and understanding their structure and use is crucial for accurate data manipulation and interpretation.

Tabular data

Tabular data refers to data that is organized in rows and columns, often presented in a spreadsheet-like structure. Each row represents a unique observation, while each column represents a variable or attribute related to those observations. This format is widely used for storing and processing data in many scientific fields due to its simplicity

and ease of manipulation. Common formats for tabular data include CSV (Comma-Separated Values), Excel spreadsheets, and relational databases. This format is especially useful for non-spatial attributes and is typically used as input for statistical models, including regression analysis and machine learning algorithms. Sometimes tabular data refers to a geographical space (e.g., related to administrative divisions, or even includes the specific coordinates of point). In those cases, tabular data can be converted or considered geographic data.

Geographic data

Geographic data, on the other hand, includes spatial information that describes the location of features on the earth's surface. It combines locational information (such as coordinates) with attribute data to represent geographic phenomena. Geographic data can be further categorized into vector and raster formats. Vector data includes points, lines, and polygons, which are used to represent discrete features such as cities, roads, or regions. Raster data, on the other hand, represents continuous phenomena, such as elevation or temperature, using a grid of cells, each containing a value. In this thesis, vector data was mainly used. Formats for storing geographic data include shapefiles, GeoJSON, and TIFF, among others.

3.2.3.2 Data sources

As introduced in the previous section, multiple data sources were used during the development of this doctoral thesis. In this point, the main data sources considered along with their spatial extension, data formats, variables are presented (Table 3.1).

Table 3.1. Main data sources used in the development of this doctoral thesis.

Data source	Spatial extension	Data formats	Key variables
INE	National (Spain)	Tabular	Socioeconomics
IGN	National (Spain)	Geographic	Land use and infrastructure
Madrid's Open Data Portal	Municipal (Madrid)	Tabular and geographic	Socioeconomics, land use and infrastructure
OSM	Global (Worldwide)	Geographic	Land use and infrastructure

3.2.4 Data processing

3.2.4.1 Software tools

For the data processing in this research, two main software tools were utilized: R and ArcMap. These tools were selected due to their powerful data manipulation capabilities, statistical functionalities, and geographic analysis features.

- **R** (R Core Team, 2013) is an open-source programming language and software environment commonly used for statistical computing and data visualization. It provides a comprehensive ecosystem for handling various data structures, performing statistical tests, and creating advanced graphical outputs. In this research, R was primarily used for data wrangling, statistical analysis, and model fitting. The most important R packages used in the workflow are:
 - ***tidyverse***: This collection of packages facilitated efficient data manipulation, transformation, and visualization through functions from *dplyr*, *ggplot2*, *tidyr*, and others.
 - ***glm***: The Generalized Linear Models package was used to fit various logistic and Poisson regression models.
 - ***MASS***: This package was instrumental for fitting negative binomial models, which were crucial for dealing with overdispersed count data. It also provided tools for model selection through functions such as *stepAIC*.
 - ***pscf***: This package was used for fitting Hurdle models, including Hurdle negative binomial models. These models were important for handling data with many zeros, which could not be properly modeled using standard count regression approaches.
 - ***INLA***: The Integrated Nested Laplace Approximation (INLA) package was essential for fitting Bayesian models.
 - ***sf***: The *sf* package, used for handling simple features, played an important role in processing and analyzing spatial data in a more modern and efficient way than using software other than R.
- **ArcMap** (Esri Inc., 2020) is a geographic information system (GIS) software developed by Esri. It is designed for working with maps and spatial data, providing tools for spatial analysis, data visualization, and cartography. ArcMap was used in this research for geoprocessing tasks, including spatial data management, visualization of traffic crash locations, and the creation of maps.

3.2.4.2 Extraction of vehicle-older pedestrian collisions

The extraction of vehicle-older pedestrian collisions was a critical step in the development of this doctoral thesis. Based on the processes outlined in previous sections, data was extracted from the Spanish traffic crash database provided by DGT using the R software. To accurately extract crashes involving vehicle-pedestrian collisions where older pedestrians were involved, several procedures were necessary. This section will explain these procedures without delving into technical specifics (e.g., variable names and values). The code in R for the extraction is shown in Annex 1. The datasets from ARENA and ARENA2, referenced throughout, are presented in Table 3.2 for clarity.

Table 3.2. Dataset terminology in the traffic crash database extraction.

TABLE	ARENA	ARENA2
Crashes	CRASHES_TABLE	CRASHES_TABLE
Individuals	INDIVIDUALS_TABLE	
Pedestrians		PEDESTRIANS_TABLE
Drivers		DRIVERS_TABLE
Passengers		PASSENGERS_TABLE
Vehicles	VEHICLES_TABLE	VEHICLES_TABLE

3.2.4.2.1 Main functions used

In the next paragraphs, the process to extract the data regarding older pedestrian traffic crashes is explained. The code for the data extraction process was written in R, but it could be adapted to any other programming language. For reproducibility, the main functions used were:

- `group_by()` takes an existing table and converts it into a grouped table where operations are performed "by group".
- `summarize()` creates a new data frame. It returns one row for each combination of grouping variables.
- `mutate()` creates new columns that are functions of existing variables. It can also modify and delete columns.
- `select()` selects variables in a data frame.
- `left_join()` adds columns from table y to table x, matching observations based on the keys, while keeping all observations in x.

- `coalesce()` finds the first non-missing value at each position given a set of vectors.

3.2.4.2.2 Vehicle-pedestrian collisions

Initially, vehicle-pedestrian collisions were extracted from the comprehensive traffic crash database. This operation was conducted identically across both the ARENA and ARENA2 databases. The input table for this process was the crashes table, labeled `CRASHES_TABLE`. The extraction was performed using the following process (Figure 3.6):

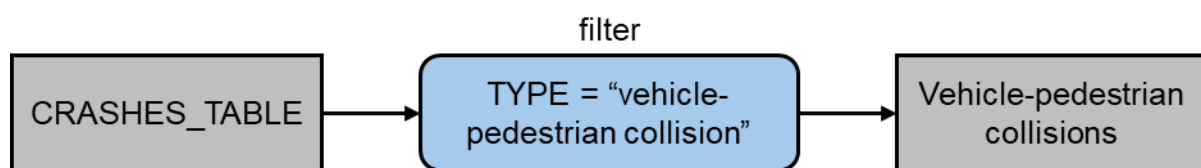


Figure 3.6. Process to filter vehicle-pedestrian collisions from the traffic crash database.

3.2.4.2.3 Single vehicle-pedestrian collisions

Multiple vehicles or pedestrians might be involved in a vehicle-pedestrian collision, although it is not the most common case. In this specific doctoral thesis, considering these traffic crashes might be misleading, as if two pedestrians were involved in a crash one can be older and the other non-older. Even more, maybe the older pedestrian was injured and the non-older was not, or vice versa, so maybe we should only consider the injured as being involved in the crash. There are other cases where we find no records for the pedestrian and/or for the driver of the vehicle, or the vehicle itself. In those cases, it is not possible to determine the age or the injury severity of the pedestrian and/or driver.

Considering only single vehicle-pedestrian collisions, which only include 1 pedestrian, and 1 vehicle involved, makes it easier to relate traffic crash characteristics with pedestrian, driver, and vehicle characteristics. In other cases, when relating the crashes table to the others, the output would contain more rows than the original traffic crashes table, as one-to-many joints would be present. In addition, data is easier to be processed and interpreted considering only single vehicle-pedestrian collisions, as the number of crashes would be equal to the number of pedestrians.

The process starts with creating a table containing, for each vehicle-pedestrian collision, the number of pedestrians and vehicles involved. Later, the number of pedestrians and vehicles is included in the crashes table. Then, the number of pedestrians and

vehicles involved in each crash is attached for each crash in the crashes table. Finally, crashes with one pedestrian and one vehicle can be identified in that table. This process is performed slightly differently when dealing with ARENA and ARENA2 databases.

In ARENA database, the process requires the individuals table (*INDIVIDUALS_TABLE*) and the vehicles table (*VEHICLES_TABLE*). With this data we need to filter those individuals involved in vehicle-pedestrian collisions, and those who are pedestrians. Later, we would count how many “rows” are there for each crash, as one row correspond to one pedestrian. The same is carried out for the vehicles table. The complete process is illustrated in Figure 3.7.

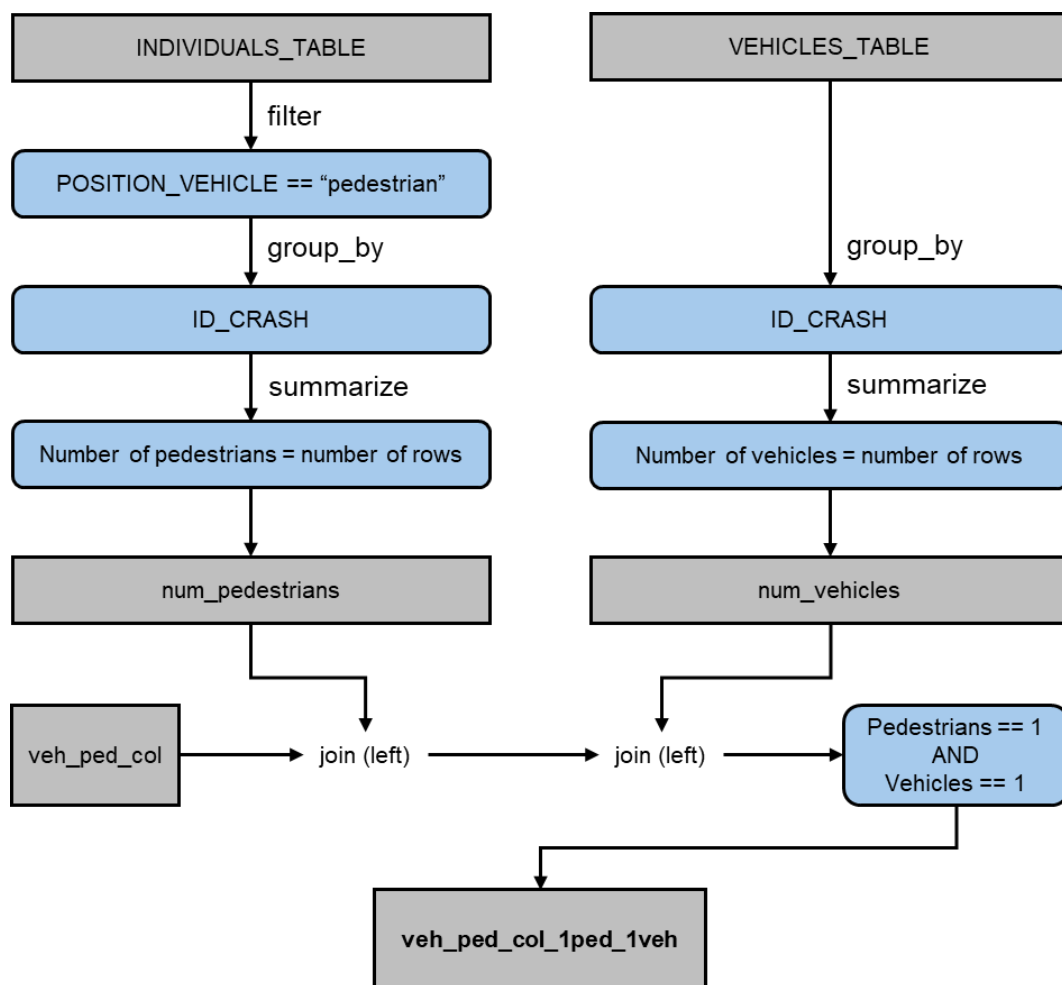


Figure 3.7. Process to filter single vehicle single pedestrian collisions from the traffic crash database.

In turn with the ARENA2 database, the process is slightly different, as information on the pedestrians is already in a separate table, so we will need the pedestrians table (*PEDESTRIANS_TABLE*) and the vehicles table (*VEHICLES_TABLE*).

3.2.4.2.4 Vehicle-pedestrian collisions considering pedestrian's age

Once single vehicle-pedestrian collisions are identified, the age of each pedestrian can be directly assigned or related to its corresponding traffic crash without having duplicate records. Later, the crashes table can be divided depending on the pedestrian being older or not. This process is performed slightly differently in ARENA and ARENA2. For the ARENA case, both the table containing single vehicle-pedestrians collision created above (*veh_ped_col_1ped_1veh*) and individuals table (*INDIVIDUALS_TABLE*) are used as illustrated in Figure 3.8.

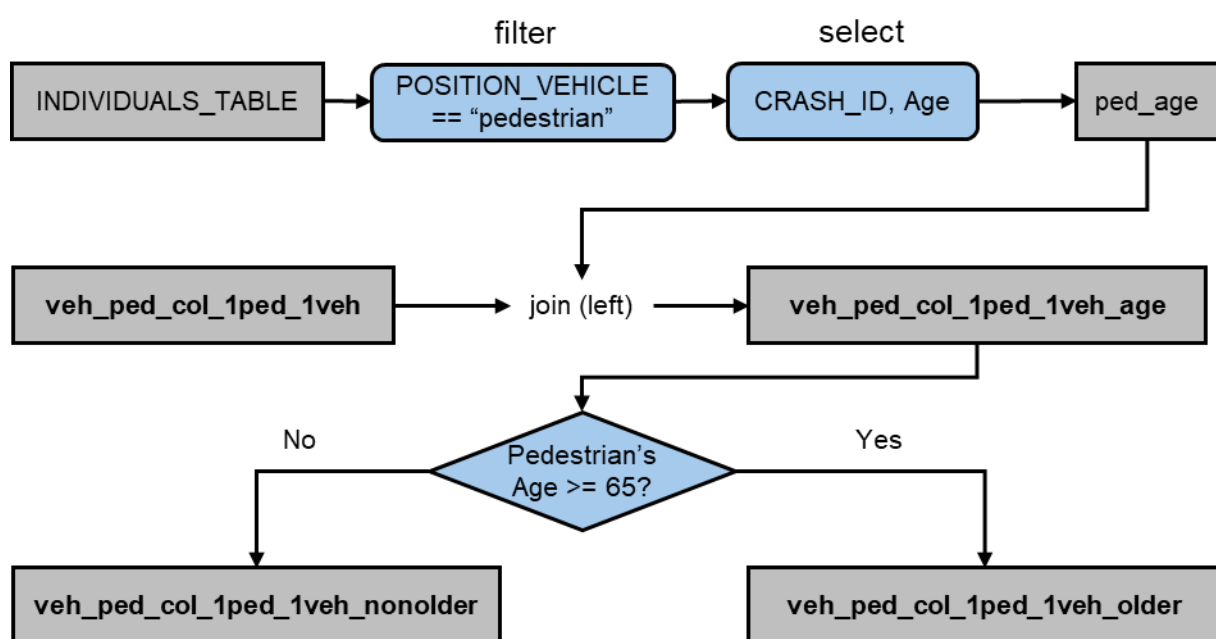


Figure 3.8. Process to filter singles vehicle single pedestrian traffic crashes involving older and non-older adults from the traffic crash database.

3.2.4.3 Geolocation of traffic crashes

Geocoding traffic crashes was a major challenge in this doctoral thesis. The geolocation of crashes was provided through postal addresses in the DGT traffic crash database. Converting these addresses into coordinates is a process known as geocoding. Since the thesis placed considerable emphasis on the characteristics of the surroundings of traffic crashes, accurately locating the crashes was essential.

In principle, geocoding addresses should not pose a significant challenge if the number of addresses is limited or if third-party paid software is used. However, this was not the case in this study, as more than ten thousand locations needed to be geocoded for the Madrid case study. Additionally, no third-party paid software was available.

Consequently, an ad hoc process was developed to geocode any address in the city of Madrid. This approach also has the advantage of enabling rapid geocoding of future crashes.

In the Spanish traffic crash database, crashes are located either by the street name and number if they occurred on a road segment, or by the names of two streets if they occurred on a road junction. Relying solely on this information, without precise geographical coordinates, presents several challenges. First, street names may change over time, making it difficult to accurately identify streets in the present day. Second, the ARENA structured database, which is provided as fixed-width text files, allows only 26 characters for storing addresses, which is sometimes insufficient. Third, the exact crash locations are not recorded, meaning the specific side of the road or lane remains unknown for road segment crashes, and the specific leg of the junction is unclear for crashes at road junctions. Lastly, maintaining the database demands considerable effort to ensure the correct spelling of street names. If this process fails for any record, some crashes may not be properly geocoded.

In Spain, street names typically consist of three elements: the type of street (e.g., avenue, street, alley), an optional particle (e.g., of, of the), and the name of the street itself. Considering this, and that traffic crash addresses might be written in a different way than the official street name (e.g., without the particle, or with a different particle), an ad hoc process was developed to geocode a set of addresses, and it was applied specifically to the Madrid case study. The process is described in the following paragraphs for this case, but it can be adapted for use in other studies.

The required materials are as follows:

- A set of addresses to be geocoded.
- A database containing all possible addresses in the city or region, along with their geographic coordinates. This database is often accessible through official open data portals.
- A database of all road junctions in the city or region, including their coordinates. While this database may not always be available, it can be created using GIS software and the geometry of the region's roads.
- A database of road name changes covering at least the period during which the data to be geocoded was collected.

Addresses are geocoded differently depending on whether they consist of a street name and number or two street names (i.e., junctions). Therefore, both methods are presented separately in the following sections.

3.2.4.3.1 Geocoding of traffic crashes at street segments

The geocoding process for addresses consisting of a street name and number is outlined in the following paragraphs and illustrated in Figure 3.9.

A. Data preparation:

1. In the first stage, the addresses to be geocoded are separated into street type, particle (if any), and street name, and number. For a few examples, see Table 3.3.
2. In the second stage, a set of combinations of the street type, particle, and street name is created: the complete street name (street type + particle + name), street name without the particle (type + name), street name without the type (particle + name), and street name without both the type and particle (name).
3. In the third stage, the street name is updated to the most current name if a name change has occurred, using the record of street name changes.
4. Finally, a set of combinations of the street type, particle, and street name is created in the database of all geocoded addresses: street type + particle + name, type + name, particle + name, name.
5. As a result, two tables are obtained. One of them contains the cleaned and updated set of addresses to be geocoded, with columns for street type, particle, street name, three columns with the combinations of them, and number. The second table consists of the geocoded addresses, including columns for street type, particle, street name, three columns with the combinations of them, and coordinates.

Table 3.3. Examples of the parts of a street name considered in the geoprocessing of addresses.

Complete street name	Street type	Particle	Name
Avenida de la Reina Victoria	Avenida	De la	Reina Victoria
Paseo de la Castellana	Paseo	De la	Castellana
Calle Mayor	Calle		Mayor

B. Matching the address. A series of SQL joins is performed to find matches between the table of the addresses to be geocoded and the table containing all postal addresses with coordinates. When matches are found after a join, they are stored separately, and those records are removed from the dataset for subsequent joins.

In addition, if duplicate matches are found for a single address, the duplicates are removed, and the record is reintroduced into the list of addresses to be geolocated.

1. Join considering the complete street name, and the number.
2. Join considering the street name without the particle, and the number.
3. Join considering the street name without the type, and the number.
4. Join considering the street name without both the type and the particle, and the number.

C. Combining the Results. The tables resulting from the four joins are combined into a single table that contains all the matched or found addresses.

D. Final Geocoding Step. Finally, the geocoded addresses are introduced into the input table of traffic crash addresses to be geocoded.

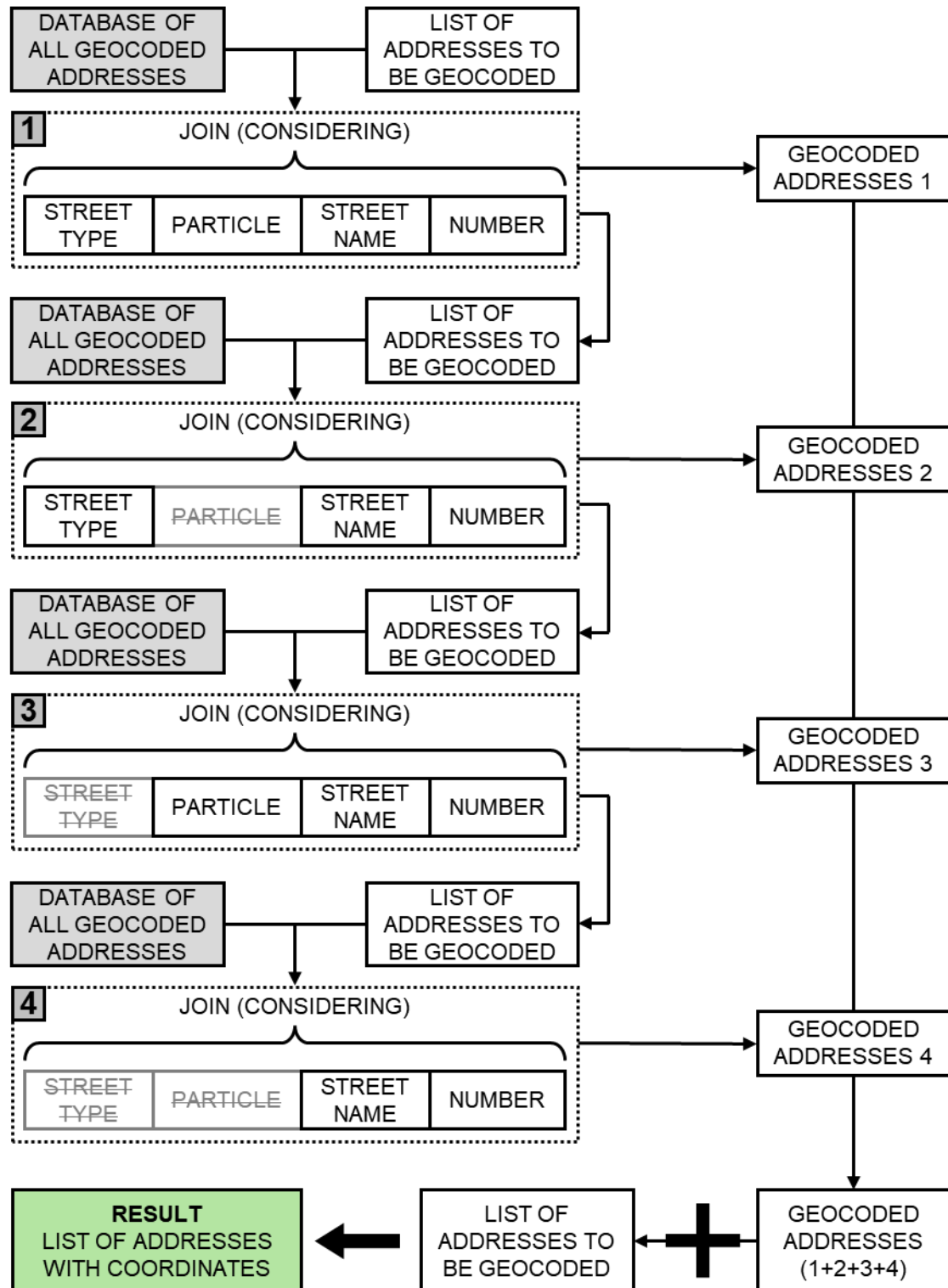


Figure 3.9. Developed process for geocoding of traffic crashes at street segments.

3.2.4.3.2 Geocoding of traffic crashes at street junctions

The geocoding process for addresses consisting of two street names (junction) was conducted as follows.

A. Data preparation:

1. In the first stage, both street denominations to be geocoded are separated into street type, particle (if any), street name, and number.
2. In the second stage, a set of combinations of the street type, particle, and street name is created for both streets: the complete street name (street type + particle + name), street name without the particle (type + name), street name without the type (particle + name), and street name without both the type and particle (name). The different combinations are then joined considering all possible combinations, in which the first of both streets use alphabetical order (i.e., Street A is the first one considering alphabetical order, and Street B the other).
3. In the third stage, street names are updated to the most current name if a name change has occurred, using the record of street name changes.
4. Finally, a set of combinations of the street type, particle, and street name is created for both streets in the database of all geocoded addresses: street type + particle + name, type + name, particle + name, name. The different combinations are then joined considering all possible combinations, in which the first of both streets is set considering alphabetically order.
5. As a result, two tables are obtained. One of them contains the cleaned and updated set of addresses to be geocoded, with columns for street type, particle, street name, three columns with the combinations of them, and number. The second table consists of the geocoded addresses, including columns for street type, particle, street name, three columns with the combinations of them, and coordinates.

B. Matching the street junction. A series of all possible SQL joins is performed to find matches between the table of the addresses to be geocoded and the table containing all street junctions with coordinates. When matches are found after a join, they are stored separately, and those records are removed from the dataset for subsequent joins. In addition, if duplicate matches are found for a single address, the duplicates are removed, and the record is reintroduced into the list of addresses to be geolocated.

C. Combining the Results. The tables resulting from the joins are combined into a single table that contains all the matched or found addresses.

D. Final Geocoding Step. Finally, the geocoded addresses are introduced into the input table of addresses to be geocoded.

3.2.4.3.3 Final step

Once the crash locations at street junctions and street segments have been geopositioned, both tables (i.e., addresses of street and number, and of two streets) are joined and related to the identifiers of the traffic crashes. As a result, we have a table with the identifiers, the address, and the coordinates (i.e., two separate variables, X and Y) of the crashes. The crashes unable to geocode present either unknown coordinates (*NA*) or coordinates set to zero (0, 0).

3.2.4.4 Systematic processing of variables with respect to spatial units

Data processing involves multiple operations with a single purpose: to have a final database suitable for data exploratory analysis and data statistical modeling. This section outlines all the necessary process to process the raw data. As these processes might be different depending on the unit of study and the nature of the source data, all the encountered situations during the redaction of this doctoral thesis are presented. The processes explained here can be implemented using multiple software, but in the elaboration of these thesis ArcGIS and R software were used. The author would like to suggest, if possible, using only R, as we can process data and perform GIS operations all within the same program.

To summarize and guide the reader, all the possible combinations of units of study and type of data are presented in Table 3.4. Note that, as stated before, data can be tabular or geographic, but in this section, we suppose that tabular data shows geographic information, so it can be converted to geographic data. Hence, in this section we suppose that the data is geographic.

Table 3.4. Data processing in terms of the spatial unit of study and the nature of the data.

Unit of study	Areal data	Line data	Point data
Macro-level Administrative division	Directly assign Or Intersect ↓ Area	Intersect ↓ Length	Intersect ↓ Count

Unit of study	Areal data	Line data	Point data
Micro-level Road segment Road junction	Centroid of the unit ↓ Assign value	Assign ↓ Count	Assign ↓ Count
Crash-level Point	Crash location ↓ Assign value	Crash location ↓ Create buffer	Crash location ↓ Create buffer

3.2.4.4.1 Systematic processing of variables in macro-level units of study

Macro-level studies show considerably easier data acquisition and processing than micro-level studies.

Regarding **data associated with areas**, administrative divisions were used as the unit of analysis in the two developed works at the macro-level. If the spatial unit of analysis is not an administrative division but other type of area or polygon, the calculations are analogous. With this configuration, working with areal or polygonal data may have two possibilities in function of the availability of data from official sources:

- **Data is also provided regarding administrative divisions.** This is the most straightforward case. The data should be proportioned in the same administrative divisions of the unit of study, or in smaller ones, but not in bigger ones. For instance, if the unit of study are city districts, data could be related to city districts, neighborhoods, or census tracts, but not related to the whole city. If the data are in the same administrative divisions as the unit of study, the data is simply related to each unit of analysis considering the administrative divisions (Figure 3.10). If not, the data should be aggregated in each unit of study.

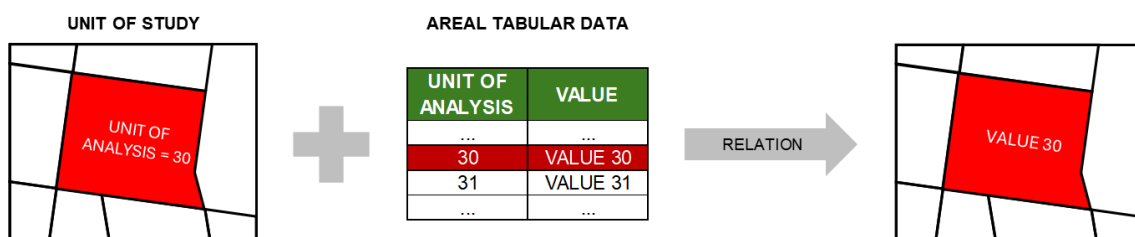


Figure 3.10 Workflow with areal units of study and areal tabular input data.

- **Data is not provided regarding administrative divisions.** If data is not provided regarding any administrative division, polygonal data should be aggregated according to the unit of analysis, using GIS operations, as this data is geographic. In this case, the research can face two options:

- **Option 1. Only the amount of area inside of each unit of study wants to be studied.** As a first stage, data should be spatially intersected with the units of study (Figure 3.11). Second, the areas of the entities inside each region should be aggregated (Equation 3.1).

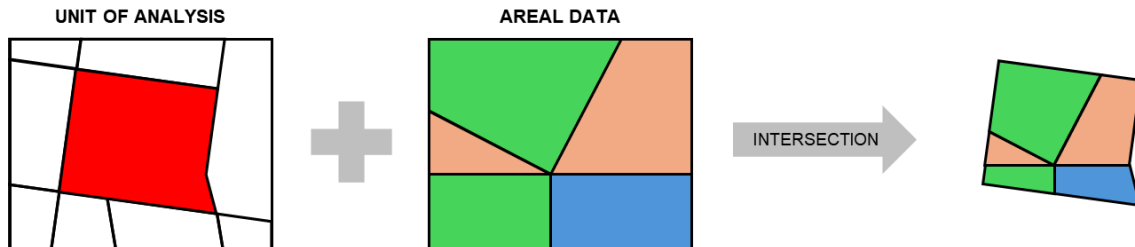


Figure 3.11 Workflow with areal units of study and areal input data.

$$A(E_i^k) = \sum_j A(E_{i,j}^k) \quad (3.1)$$

Where $A(x)$ is the surface of x , $E_{i,j}^k$ is the intersection between the units of study i and the element j , of type k .

- **Option 2. The proportion of areas of a certain type over others inside of each unit of study wants to be studied.** As a first stage, data should be spatially intersected with the unit of study. Secondly, the areas of the entities inside each region should be aggregated considering the “type” variable (Equation 3.1). Third, the proportion of area of a certain type is calculated as shown in Equation 3.2.

$$\%A(E_i^k) = \frac{\sum_j A(E_{i,j}^k)}{\sum_i A(E_{i,j}^k)} \quad (3.2)$$

Where $A(x)$ is the surface of x , $E_{i,j}^k$ is the intersection between the units of study i and the element j , of type k .

Regarding **geographic line data**, the common variables that we would need would be total length of the lines, total length by type, and proportion of length by type inside

each administrative division. For example, the total length of main roads within a city district. These procedures are analogous to the areal data case, and, like previous procedures, the geographic line data should be intersected with the units of study (Figure 3.12); and, as a second task, the length of the lines inside each administrative division can be aggregated (Equations 3.3 and 3.4).

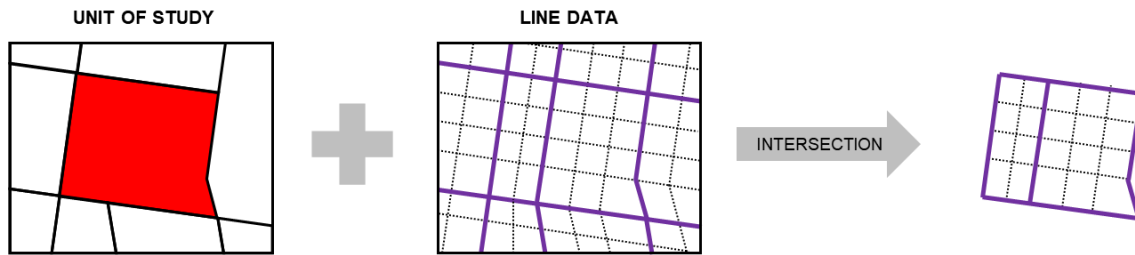


Figure 3.12 Workflow with areal units of study and linear input data.

$$L(E_i^k) = \sum_j L(E_{i,j}^k) \quad (3.3)$$

$$\%L(E_i^k) = \frac{\sum_j L(E_{i,j}^k)}{\sum_i L(E_{i,j}^k)} \quad (3.4)$$

Where $L(x)$ is the length of x , $E_{i,j}^k$ is the intersection between the units of study i and the element j , of type k .

Regarding **geographic point data**, the common variable would be count of points inside of each administrative division. The process would involve assigning each data point to each administrative division (Figure 3.13), and then counting how many points there are for each administrative division.

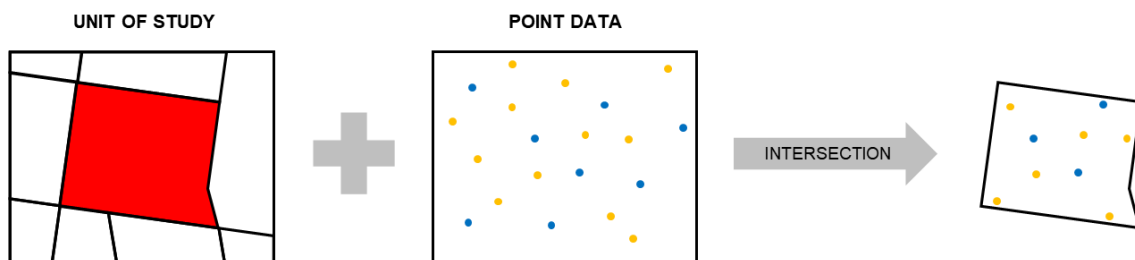


Figure 3.13 Workflow with areal units of study and point input data.

Normalization of independent variables is needed to make different areas comparable between each other. In this case, normalization was performed in terms of the total length of the road network inside each unit of analysis. This was preferred to normalization over the area, as some units of analysis might present a wide area, but a concentrated road network, such is the case, for instance, of suburban city districts in Madrid (Spain). Normalization was only applied to point data, such as POIs or street devices. For example, the number of bus stops per street kilometer within a city district.

3.2.4.4.2 Systematic processing of variables in micro-level linear units of study

Data processing is more challenging in micro-level studies using linear units of study, as it is the data acquisition, as normally aggregated data is easy to access but the reliability and availability of micro-level data is not so common. In this case, the common unit of study is the road or street segment.

Regarding **spatial data associated with areas**, in this doctoral thesis, the value of the entity of the areal data was assigned to each spatial unit of analysis (e.g., street segment) considering the entity in which the centroid of each spatial unit of analysis is located (Figure 3.14).

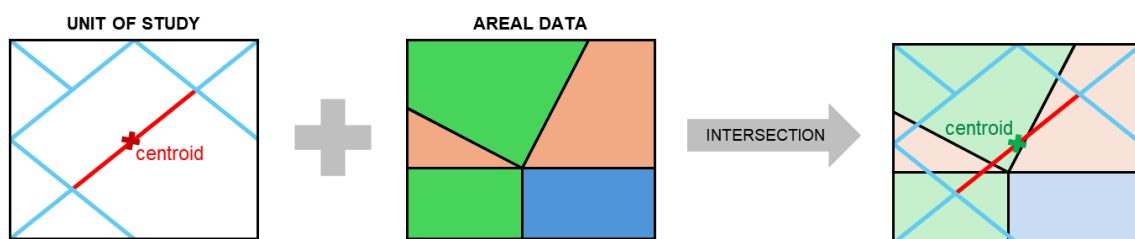


Figure 3.14 Workflow with linear units of study and areal input data.

Regarding **spatial data associated with lines**, the nearest spatial unit of study to each line can be obtained. Later, the total length of data for each spatial unit of study can be calculated as the sum of the lines which nearest spatial unit of study is that one spatial unit of study.

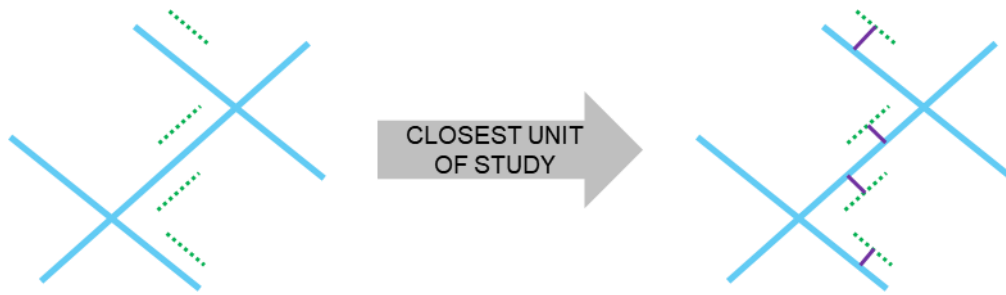


Figure 3.15 Workflow with linear units of study and line input data.

Regarding **spatial data associated with points**, an analogous procedure can be carried out. For each data point, the nearest linear spatial unit of study (Figure 3.16), the distance between them and the angle of the orthogonal line between the point and the unit of study can be calculated. After this, considering the angle, the total number of points in each road segment and the number of points on each side of the road segment can be processed (an example is shown in Figure 3.17).

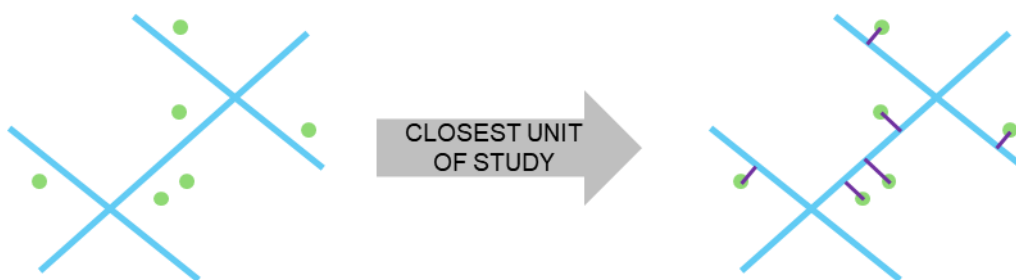


Figure 3.16 Workflow with linear units of study and point input data.

POINT ID	SEGMENT ID	DISTANCE (m)	ANGLE
1	34	19.2	109.8
2	34	17.4	26.1
3	56	40.0	61.1
4	56	6.9	-47.1
5	56	8.9	-65.9
6	56	19.9	-13.9
7	79	30.1	-55.6
8	79	27.1	213.5

SEGMENT ID	NUMBER OF POINTS	SIDE A	SIDE B
34	2	2	0
56	4	1	3
79	2	1	1

Figure 3.17 Example of data processing for linear units of study and point input data.

3.2.4.4.3 Systematic processing of variables in micro-level point units of study

In this case, the common unit of study is the road junction, or the traffic crash itself, representing a point in space. One can perform similar operations as the explained for the micro-level line units or create a buffer around each spatial unit of analysis. In the

latter case, the problem could be reduced to what was explained for macro-level units of study.

Regarding **spatial data associated with areas**, the value of the entity of the areal data was assigned to each spatial unit of analysis considering the entity in which each spatial unit of analysis (i.e., a point) is located, as in the case of micro-level with line spatial units of analysis.

Regarding **spatial data associated with lines**, a buffer from the units of study can be drawn and the lines inside each buffer would be assigned to each unit of analysis. For example, it can be calculated the proportion of main roads around a street junction. With this procedure, the same case as for line data in macro-level study can be used, as now the buffers function as how it does the administrative divisions in those studies.

Regarding **spatial data associated with points**, a buffer can be as well created for each unit of study, so the method explained for the macro-level would be applicable. However, in this study we tried to relate each point to only one road segment, and creating buffers would potentially generate duplicates for points if the buffers were big enough to intersect each other. Instead, the distance of each data point to each unit of study can be calculated. Later, and with the possibility of adding a threshold, each data point is related to the nearest unit of study. If the unit of study is the road segment (i.e., lines) the amount of data points on each side of the road can be easily computed separately considering the sign of the angle of the lines that joins the points with the streets.

3.3 *Statistical modeling*

Statistical modeling is essential for analyzing complex data and making informed decisions based on empirical evidence. In the context of traffic safety, statistical models help in understanding the factors that contribute to the frequency and severity of crashes involving older pedestrians. This section discusses various statistical models used to predict or infer the number and severity of older pedestrian traffic crashes, providing valuable insights for policymakers to enhance street safety.

3.3.1 Traffic crash frequency modeling

The models described in this section have been used in the development of this doctoral thesis to predict or infer the number of older pedestrian traffic crashes in each unit of study. These models help policymakers to assess which factors make streets riskier, in this case, in terms of the number of crashes suffered by older pedestrians.

3.3.1.1 Poisson regression model

Poisson regression is part of the generalized linear models (GLM). A GLM is a generalization of the ordinary linear regression, which relates the linear regression model with the outcome or dependent variable using a link function. This regression assumes that the dependent variable (Y_i) has a Poisson distribution and assumes the logarithm of the expected value as a linear combination of the independent variables with unknown parameters. That is why it is commonly known as the log-linear model. In the case of Poisson regression, the logarithmic function is the link function, and the Poisson distribution is the probability function of the outcome.

Poisson regression is used to model count data (i.e., non-negative integers), so it has been typically applied to study the frequency of traffic crashes, as an area or road might present no or more than 0 crashes.

$$Y_i | \lambda_i \sim \text{Poisson}(\lambda_i) \tag{3.5}$$

Where $\lambda_i > 0$ is the rate parameter.

Given an aleatory variable Y_i with a Poisson distribution, the probability mass function (PDF) for a given value $Y_i = y_i$ can be expressed as follows:

$$P(Y_i = y_i | \lambda_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} \quad (3.6)$$

The link function between the independent variables and the expected outcome Y_i of the dependent variable can be expressed as follows:

$$E[Y_i | x] = \lambda_i = e^{\beta x_i} = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki}) \quad (3.7)$$

Where λ_i is the rate parameter, β is the vector of the parameters of the model, and x_i is the vector of independent variable for the observation i .

In consequence, the Poisson regression model for the i^{th} observation can be expressed as follows:

$$P(Y_i = y_i | x_i, \beta) = \frac{e^{-\exp(\beta x_i)} \exp(\beta x_i)^{y_i}}{y_i!} \quad (3.8)$$

Where λ_i is the rate parameter, β is the vector of the parameters of the model, and x_i is the vector of independent variable for the observation i .

The Poisson distribution is easy to implement and interpret. However, it assumes that the mean and variance values of the expected outcome are equal:

$$E[Y_i] = \lambda_i = \text{Var}[Y_i] \quad (3.9)$$

Where $E[Y_i]$ is the expected value (mean) of Y_i , and $\text{Var}[Y_i]$ is the variance of Y_i .

Count data (e.g., frequency of traffic crashes) might exhibit two possible characteristics that make other regression models more appropriate:

- **Overdispersion.** As stated above, the Poisson distribution assumes that its mean value is equal to its variance. Nevertheless, data might have a variance higher than the mean value. This phenomenon is known as overdispersion. In such cases, other distributions such as the negative binomial distribution is preferred.
- **Excess of zero.** If the outcome variable presents a high proportion of zeros, higher than should be expected considering the parameter of the used probability mass function, other regressions models, such as the hurdle or zero-inflated models should be used.

3.3.1.2 Negative Binomial regression model

The negative binomial (NB) distribution is a discrete probability distribution that models the number of failures in a sequence of independent and identically distributed Bernoulli processes before a definite number of successes occurs. NB regression is also known as Poisson-gamma regression, as NB is a Poisson-gamma mixture model. When count data exhibits overdispersion (i.e., the variance exceeds the mean), NB regression is preferred. Note that when we refer to NB, we are referring to the standard 'NB2' regression (Hilbe, 2011, 2014).

In contrast to the Poisson model, the variance of the NB model is influenced by two parameters: the dispersion parameter (α) and the mean (μ), as outlined in Equation 3.10.

$$Var[Y] = \mu(1 + \alpha\mu) = \mu + \alpha\mu^2 \quad (3.10)$$

Where $Var[Y]$ is the variance of the dependent variable Y , μ is the mean of the distribution, and α is the dispersion parameter, which quantifies the degree of overdispersion in the data.

When the dispersion parameter (α) approaches zero, the variance equals the mean ($Var[Y] = \mu$), and the NB model converges to the Poisson model. Some statistical

software uses the inverse of α , denoted as θ , to parameterize the model. In terms of θ , the variance is expressed as:

$$\theta = \frac{1}{\alpha} \Rightarrow \text{Var}[Y] = \mu \left(1 + \frac{\mu}{\theta}\right) = \mu + \frac{\mu^2}{\theta} \quad (3.11)$$

The negative binomial regression model assumes that the dependent variable Y_i follows a negative binomial distribution:

$$Y_i \sim NB(\mu_i) \quad (3.12)$$

Where Y_i is the count outcome for the i^{th} observation, and μ_i is the mean parameter for the i^{th} observation, which depends on the independent variables.

The probability mass function (PMF) of the observation Y_i being equal to y_i , knowing the mean value (μ_i), and the overdispersion parameter (α) is shown in Equation 3.13.

$$P(Y_i = y_i | \mu_i, \alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(\alpha^{-1})y_i!} \left(\frac{1}{1 + \alpha\mu_i}\right)^{\alpha^{-1}} \left(\frac{\alpha\mu_i}{1 + \alpha\mu_i}\right)^{y_i} \quad (3.13)$$

Where $P(Y_i = y_i | \mu_i, \alpha)$ is the probability that $Y_i = y_i$, given μ_i and α , $\Gamma(\cdot)$ is the gamma function, α is the dispersion parameter, and μ_i is the mean parameter for observation i . This formulation allows the model to account for overdispersion in the count data by introducing the dispersion parameter α , providing a better fit when the variance exceeds the mean.

The mean parameter μ_i is linked to the independent variables through a log-linear relationship, like the Poisson regression model:

$$E[Y_i | x_i] = \mu_i = e^{\beta x_i} = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki}) \quad (3.14)$$

Where $E[Y_i | x_i]$ is the expected value of Y_i given the independent variables x_i , $\beta = (\beta_0, \beta_1, \dots, \beta_k)$ is the vector of regression coefficients, and $x_i = (1, X_{1i}, X_{2i}, \dots, X_{ki})$ is the vector of independent variables for observation i , including an intercept term.

3.3.1.3 Hurdle Negative Binomial regression model

The hurdle negative binomial (HNB) regression model is a two-stage model that accounts for two data generation processes simultaneously: one generating zero observations, and other generating non-zero observations.

The first or zero component of the model (f_1) is the process generating zero observations. A binary logistic regression is used for this purpose, and its probability mass function, Bernoulli distribution, is shown in Equation 3.15.

$$f_1(z_i) = p_i^{z_i}(1 - p_i)^{1-z_i} \quad (3.15)$$

Where z_i is an indicator variable such that $z_i = 1$ if $Y_i = 0$ and $z_i = 0$ if $Y_i > 0$, and p_i is the probability that $Y_i = 0$ for observation i , which can be modeled as:

$$\log\left(\frac{p_i}{1 - p_i}\right) = \gamma_0 + \gamma_1 W_{1i} + \gamma_2 W_{2i} + \dots + \gamma_k W_{ki} \quad (3.16)$$

Where $\gamma = (\gamma_0, \gamma_1, \gamma_2, \dots, \gamma_k)$ are the coefficients for the zero component model, and $W_i = (1, W_{1i}, W_{2i}, \dots, W_{ki})$ are the independent variables for the zero component, which could be the same or different from those in the count component.

The second or count component of the model (f_2) is the process generating the non-zero observations. In this case, a negative binomial regression is used to account for overdispersed count data, which PMF is shown in Equation 3.17.

$$f_2(z_i) = P(Y_i = y_i | \mu_i, \alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(\alpha^{-1})y_i!} \left(\frac{1}{1 + \alpha\mu_i}\right)^{\alpha^{-1}} \left(\frac{\alpha\mu_i}{1 + \alpha\mu_i}\right)^{y_i} \quad (3.17)$$

where Y_i refers to the outcome value of the i^{th} observation, $\Gamma(\cdot)$ is the gamma function, α is the dispersion parameter and μ_i is the mean value of the count feature.

These two components are considered simultaneously in the HNB probability mass function, which is shown in Equation 3.18.

$$P(Y_i = y_i) = \begin{cases} f_1(0) = p_i, & \text{if } y_i = 0 \\ \frac{1 - f_1(0)}{1 - f_2(0)} \cdot f_2(y_i), & \text{if } y_i > 0 \end{cases} \quad (3.18)$$

The term $1 - f_2(0)$ in the denominator for the case $y_i > 0$ serves to truncate the negative binomial PMF (Equation 3.13), as that PMF considers null and non-null outcomes, but the $y_i > 0$ branch should only consider non-null outcomes. Considering the values of $f_1(0)$ and $f_2(0)$, HNB's PMF is shown in Equation 3.19.

$$P(Y_i = y_i) = \begin{cases} p_i, & \text{if } y_i = 0 \\ \left(\frac{1 - p_i}{1 - \left(\frac{1}{1 + \alpha\mu_i}\right)^{\alpha-1}} \right) f_2(y_i), & \text{if } y_i > 0 \end{cases} \quad (3.19)$$

3.3.2 Traffic crash injury severity modeling

Injury severity modeling aims to understand and predict the factors that influence the severity of injuries resulting from traffic crashes. By identifying these factors, interventions can be designed to reduce the likelihood of severe injuries.

3.3.2.1 Binary logistic regression

Binary logistic regression is used when the dependent variable is binary (i.e., it takes on two possible outcomes, such as “serious injury” vs. “non-serious injury”). This regression model estimates the probability that a certain event occurs (e.g., a severe injury) given a set of independent variables.

$$p_i = P(Y_i = 1|x_i) = \frac{e^{\beta x_i}}{1 + e^{\beta x_i}} = \frac{1}{1 + e^{-\beta x_i}} \quad (3.20)$$

Where p_i is the probability of $Y_i = 1$, $\beta = (\beta_0, \beta_1, \dots, \beta_k)$ is the vector of regression coefficients, and $x_i = (1, X_{1i}, X_{2i}, \dots, X_{ki})$ is the vector of independent variables for observation i , including an intercept term.

Alternatively, the logistic regression model can be expressed in terms of the log-odds:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta x_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} \quad (3.21)$$

Where $\log\left(\frac{p_i}{1-p_i}\right)$ is the log-odds of a severe injury occurring, $\beta = (\beta_0, \beta_1, \dots, \beta_k)$ is the vector of regression coefficients, and $x_i = (1, X_{1i}, X_{2i}, \dots, X_{ki})$ is the vector of independent variables for observation i , including an intercept term.

Logistic regression coefficients are estimated using the maximum likelihood estimation (MLE) method, which finds the parameter values that maximize the likelihood function, given the observed data. This model is useful in injury severity modeling as it helps identify significant predictors of severe injuries among older pedestrians, allowing policymakers to implement targeted interventions.

In addition, this model was used in Article D: The effect of COVID-19 on older pedestrian road safety: A holistic analysis through pandemic phases (Gálvez-Pérez et al.,

2025), to distinguish between pedestrian crashes occurring before (class 0) and crashes occurring after (class 1) the COVID-19 pandemic.

3.3.2.2 Random forest

Machine learning (ML) approaches, as exposed in Section 2.6.3 Traffic crash data modeling techniques, have been used to analyze injury severity of traffic crashes. In this doctoral thesis, the random forest (RF, Breiman (2001)) algorithm was used to classify traffic crashes to complement the predictive capacity of logistic regression, which is restricted to linear relationships between predictors and outcome.

RF is a supervised ML technique that operates as an ensemble of decision trees. Each decision tree, another ML algorithm, classifies data by applying a series of rules. This process involves sequentially splitting the data at each node based on one or more features, starting from the root node and continuing until a leaf node is reached, which provides the final classification output. In classification tasks, the RF predicts outcomes by aggregating the majority vote of its decision trees.

RF is capable of capturing non-linear relationships between variables, which often allows it to outperform traditional statistical models, such as logistic regression. However, RF models are less interpretable due to the large number of decision trees and the complexity of the splitting rules at each node. The relative importance of variables in the classification process can be assessed using two metrics (Nembrini et al., 2018): mean decrease impurity (MDI) and mean decrease accuracy (MDA). Variables with higher values of MDI or MDA are considered more influential in the classification. While MDI is computed during model training, MDA is calculated after training. Since MDI may be biased toward variables with multiple categories (Nembrini et al., 2018), this study primarily uses MDA as the variable importance measure, although both metrics are included in the results.

3.3.3 Bayesian inference

Bayesian inference combines information of the model's parameters (prior probability) and a likelihood function of the model for the observed data to derive the posterior probability distribution using the Bayes' theorem (Equation 3.22).

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (3.22)$$

Where $P(A|B)$ the probability of event A occurring given that B is true, $P(B|A)$ the probability of event B occurring given that A is true, and $P(A)$ and $P(B)$ are the probabilities of observing A and B, respectively.

In Bayesian approach the following elements are used:

- The model. A model with defined parameters θ that explains the data generation process of the dependent variable (X).

$$X \sim p(x|\theta) \quad (3.23)$$

- The prior distribution. The distribution of the parameters that describes prior knowledge about the parameter of the model without information of the observed data.

$$\theta \sim p(\theta|\alpha) \quad (3.24)$$

- The sampling distribution or likelihood.

$$p(X|\theta) \equiv L(\theta|X) \quad (3.25)$$

- The marginal likelihood or the evidence. The distribution of the observed data marginalized over the parameters.

$$p(X|\alpha) = \int p(X|\theta)p(\theta|\alpha)d\theta \quad (3.26)$$

- The posterior distribution. The distribution of the parameters considering the observed data and the prior knowledge, determined using the Bayes' rule.

$$p(\theta|X, \alpha) = \frac{p(X|\theta)p(\theta|\alpha)}{p(X|\alpha)} \propto p(X|\theta)p(\theta|\alpha) \quad (3.27)$$

In these equations, X is the data sample, θ are the parameters of the distribution of the data point, α are the hyperparameters of the parameter distribution. A visual representation of the prior, likelihood and posterior distributions is shown in Figure 3.18.

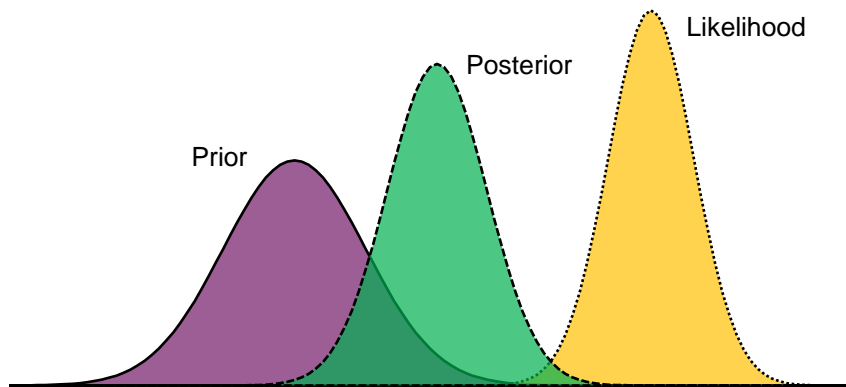


Figure 3.18. Representation of the prior, likelihood and posterior distributions used in Bayesian statistics.

The posterior distribution $p(\theta|X, \alpha)$ often does not have a closed form, so approximation techniques should be used. Historically, the most used technique to approximate the posterior distributions of the parameters is the Markov Chain Monte Carlo (MCMC), using algorithms such as Metropolis-Hastings and Gibbs sampling. These algorithms can be time- and computationally consuming if the structure of the model is intricate or the dataset is large.

In contrast, Integrated Nested Laplace Approximations (INLA) is a method for Bayesian inference based on Laplace's method, which was proposed by Rue et al. (2009). Its main advantage is that it can provide approximations in a considerably shorter time compared to the use of MCMC, even when the model structure is complex (Rue et al., 2009, 2017).

3.3.4 Model assessment

3.3.4.1 Akaike information criterion (AIC)

The Akaike information criterion (AIC) is an estimator of the predicting error or the quality of the statistical model (Akaike, 1987). This measure can be used to compare models used on the same data, and lower AIC values represent better regression model performance. The AIC does not only consider the maximized value of the likelihood of the model (\hat{L}), but also the number of independent variables used in the model (k). A model with more independent variables is expected to perform better than another model with fewer variables, so the number of variables penalizes the AIC value.

$$AIC = 2k - 2 \log\{p(y|\theta)\} = 2k - 2 \log(\hat{L}) \quad (3.28)$$

3.3.4.2 Deviance information criterion (DIC)

The deviance information criterion (DIC) is a hierarchical modeling generalization of the AIC. It is commonly used to assess the performance of Bayesian models. As AIC, it considers the goodness-of-fit and the model complexity, and lower DIC values indicate better model performance. The deviance $D(\theta)$ is a measure of goodness-of-fit and is defined as:

$$D(\theta) = -2 \log(p(y|\theta)) + C \quad (3.29)$$

Where y is the data, θ are the unknown parameters of the model, $p(y|\theta)$ is the likelihood function, and C is a constant that depends only on the data and can be ignored when comparing models.

The DIC is calculated as:

$$DIC = p_D + \overline{D(\theta)} = D(\bar{\theta}) + 2p_D \quad (3.30)$$

Where p_D is the effective number of parameters, $\overline{D(\theta)}$ is the posterior mean deviance, and $D(\bar{\theta})$ is the deviance of posterior means.

The expression to compute the number of parameters (p_D) can be obtained from the previous equations:

$$p_D = \overline{D(\theta)} - D(\bar{\theta}) \quad (3.31)$$

3.3.5 Two independent models: older and non-older pedestrian traffic crashes

The primary objective of this thesis is to identify the variables that contribute to an increase in traffic crashes involving older pedestrians. However, it is also essential to consider the factors influencing traffic crashes among other pedestrians for two key reasons. First, variables affecting older pedestrian crashes may similarly impact the broader pedestrian population, necessitating countermeasures that benefit all groups. Second, some variables may exert opposing effects on different age groups, potentially increasing crashes among older pedestrians while decreasing them for the rest of pedestrians, or vice versa. In such cases, interventions designed to enhance road safety for one group could inadvertently compromise safety for the other.

Consequently, this doctoral thesis employs independent statistical models to analyze traffic crashes involving older and non-older pedestrians, using a consistent set of independent variables (Figure 3.19). The results of are then compared to identify:

- Variables influencing the number of crashes affecting both age groups, and
- Variables specifically influencing the number of crashes involving older pedestrians.

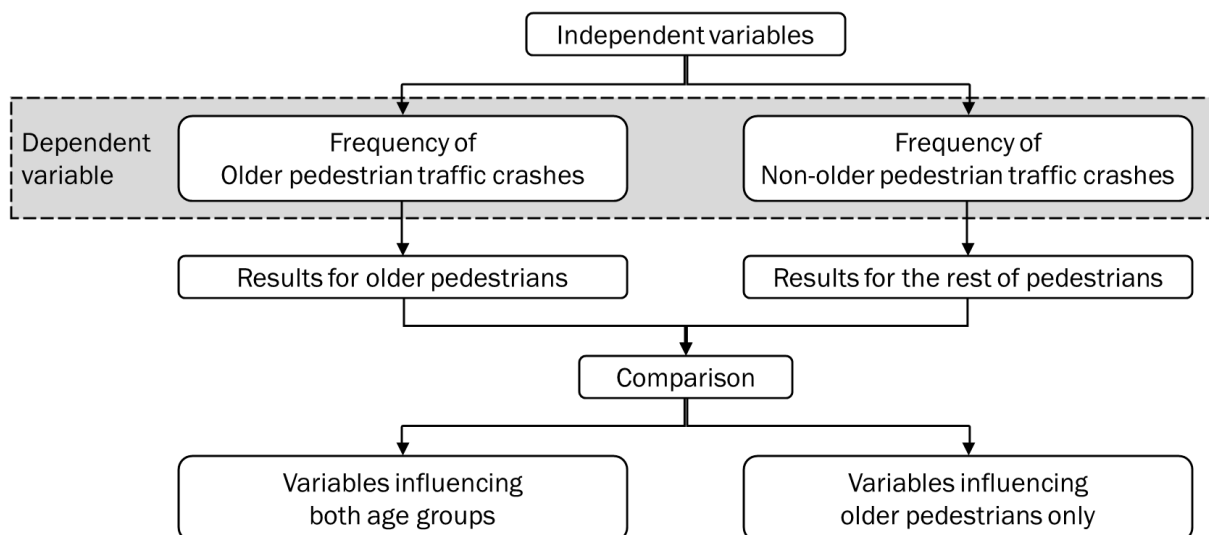


Figure 3.19 Conceptual configuration of statistical modeling in the doctoral thesis.

3.4 Application of the methodology to the publications

Once the generalities on data formats, types of variables, and data processing techniques have been acknowledged, the different variables included in each of the developed works are presented. Also, the chronology of these publications and the reasons for one leading to the next one are presented. For better understanding, the main features of the developed publications are shown in Table 3.5.

Table 3.5 Main features of the developed publications.

Feature	Article A	Article B	Article C	Article D
Case study	Madrid city	Spain	Madrid city	Spain and Madrid city
Spatial unit of study	City district	Municipality	Street segment	Country and traffic crash
Dependent variable	Number of older pedestrian traffic crashes per district	Number of older pedestrian traffic crashes per municipality	Number of older pedestrian traffic crashes per street segment	Number of older pedestrian traffic crashes per month Moment of the crash (pre- or post-pandemic)
Nature of the data	Count data with overdispersion	Count data with overdispersion and excess of zero	Count data	Count data Categorical data
Statistical model	Negative binomial regression	Hurdle negative binomial regression	(Bayesian) Poisson regression	Poisson regression Binary logistic regression

3.4.1 Chronology of publications

Given the previously exposed challenges related to materials and methods, particularly the geolocation of traffic crashes and the lack of socioeconomic, land use, and infrastructure variables in the traffic crash database, this section provides a detailed explanation of the decisions that led from one scientific publication to the next. For better understanding, this process is shown in Figure 3.20.

Initially, a macroscopic analysis at the district level, with Madrid as the case study, was conducted (**Article A**). The primary objectives of this first article were to establish the theoretical foundation of the doctoral thesis, familiarize the author with the methodology, including statistical models, and present initial findings regarding the differences between traffic crashes involving older and non-older pedestrians. Additionally, a tailored process was developed to accurately geolocate the traffic crashes recorded in the DGT's database. The required data for the geocoding process is typically supplied by municipal authorities and may require further data cleaning. Consequently, the process was specifically designed for the city of Madrid, as expanding the process of geocoding postal addresses to the entire country was deemed unfeasible within the scope of this doctoral thesis.

The main limitation of this paper lies in the use of city districts as the spatial unit of analysis. Using districts limits the assessment of certain variables, as they may have similar or identical values across districts within the same city (for example, the motorization rate), preventing a meaningful analysis of their impact on road safety. Moreover, although geocoordinates are unavailable in the traffic crash database, the specific municipality where each crash occurred is recorded, as well as the zone designation (whether the crash happened inside or outside built-up areas). These limitations, identified in the previous study, were transformed into an opportunity for further research. Consequently, a subsequent study was designed to examine road safety for older pedestrians at the municipal macroscopic level.

As a result, a second study was conducted (**Article B**), focusing on a macroscopic analysis at the municipal level, with Spain as the case study. The primary challenge of this analysis was the presence of an excess of zero values in the data (i.e., most municipalities presented no crashes for the studied period), making a hurdle negative binomial regression more appropriate than a standard negative binomial regression. Additionally, significant efforts were made in data acquisition and processing to ensure enough independent variables for each municipality. Since hurdle negative binomial models are complex to interpret due to their two components (sub-models) that must be analyzed concurrently, the relative importance of the independent variables was assessed, alongside performance indicators specifically designed for each variable.

These two scientific contributions fall within the category of macro-level studies. These types of studies are more suitable for mid- and long-term urban planning than for proposing short-term infrastructure measures, which is the focus of micro-level studies. However, analyzing road safety at the micro-level requires highly detailed data acquisition, typically at the road segment level, which is often unavailable, particularly regarding infrastructure. Additionally, street segments need traffic flow data, such as Annual Average Daily Traffic (AADT), an essential measure of exposure to traffic crash risk. Inter-urban traffic crash studies are sometimes easier to conduct than those in urban areas (i.e., the aim of this doctoral thesis) because AADT is generally available for most road segments, and road characteristics, such as lane and shoulder widths, are either available or can be measured using high-resolution orthophotos. This process is more challenging for urban studies, as buildings or street trees often obstruct the view of streets in these images.

As a third contribution, a micro-level study focusing on street segments in Madrid was conducted (**Article C**), utilizing the geolocation process developed in the first article. Significant efforts were made to clean the available data and automate the processing of variables that were not directly accessible. Street segments were selected instead

of street junctions, as most pedestrian traffic crashes, and those involving older pedestrians, occur on street segments. Only road segments with available AADT data were included in the analysis. Nevertheless, the study encompassed road segments from all city areas and across various street types. This final article regarding the frequency of older pedestrian crashes provided valuable insights into the specific infrastructure factors that increase the risk for older pedestrians on certain streets.

After these three publications regarding contributing factors to the frequency of older pedestrian traffic crashes, a fourth publication regarding the effects of the COVID-19 pandemic on older pedestrian traffic crashes in terms of frequency, injury severity and crash-related factors was developed (**Article D**). This publication used all the collected knowledge and developed tools of the three previous publications, including the geolocating procedure and the systematic data processing tools. Both macro- and micro-level analyses used in the previous publications were also employed in two separate stages of the methodology.

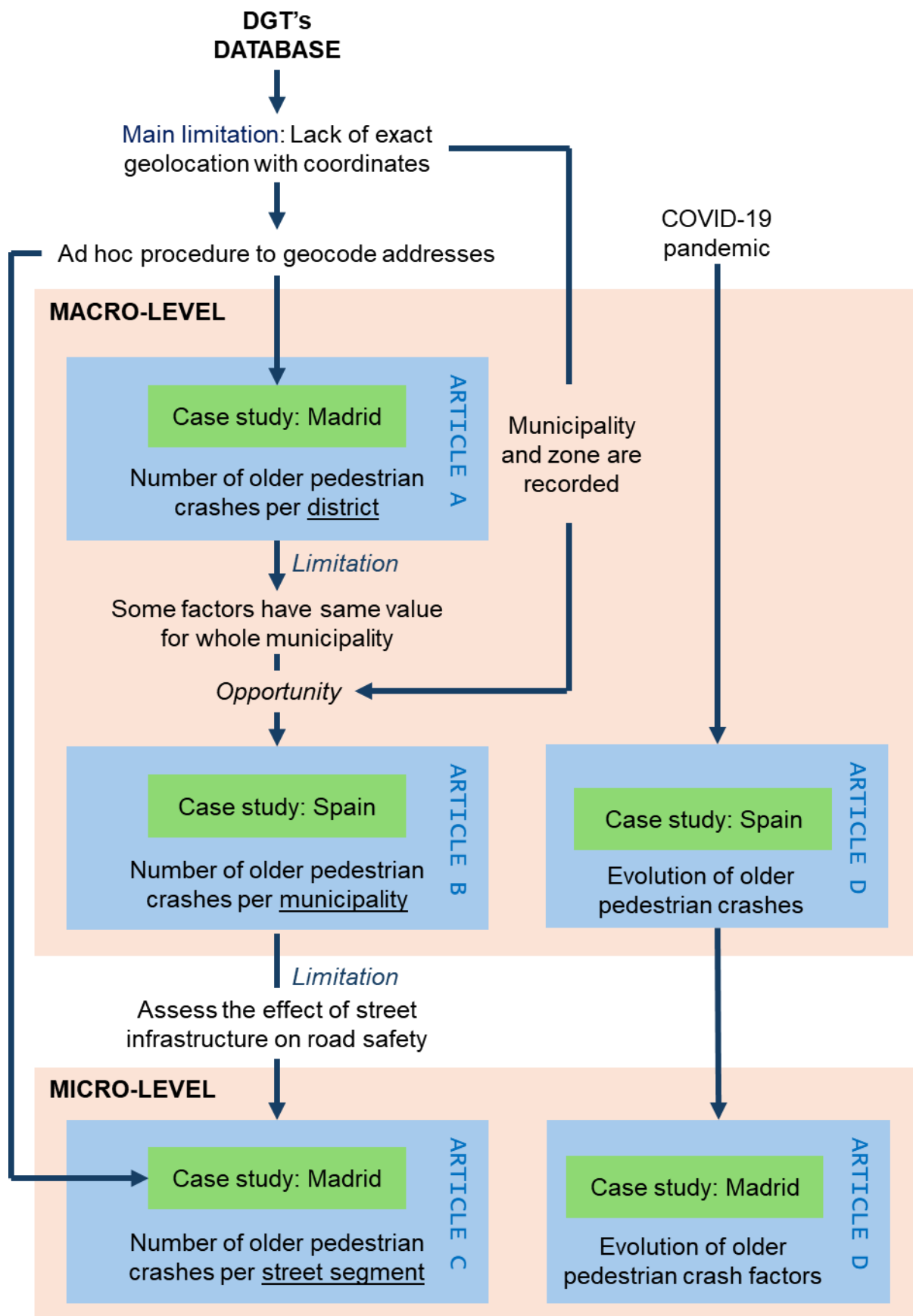


Figure 3.20. Chronology of methodological decisions across scientific publications.

3.4.2 Article A Macro-level analysis using District as the spatial unit of study

The first publication is a macro-level study with Madrid districts as the units of study, which is included in the Section 4.2 Article A: The Influence of Built Environment Factors on Elderly Pedestrian Road Safety in Cities: The Experience of Madrid. In this study, the dependent variables were the number of older and non-older pedestrian traffic crashes for each city district.

The **socioeconomic variables** were straightforward to obtain; population density and average income data were directly accessed from the Madrid City Council's database. Since these datasets were already aligned with the city's administrative divisions, no further processing was necessary.

Land use variables were mainly extracted from the urban plan of Madrid (Plan General de Ordenación Urbana, or PGOUM) from 1997. This information is geographical data consisting of polygons with different land uses. The proportion of residential, main street and green area surfaces over the total surface of each city district were obtained. Additionally, the number of points of interest (POIs) were obtained from the business census of Madrid, which provides the geographical coordinates (a point) for each business in the city.

As with the other articles in this study, the **infrastructure** variables required extensive data processing. A critical metric for deriving these variables was the total length of streets within each district. This measurement was determined by processing geographic data on Madrid's Road axes, intersecting these axes with the geometry of the city's districts, and calculating the total length of roads within each district. Based on this foundational value, additional variables were calculated. For instance, sidewalk density in each district was determined as the ratio of the sidewalk area to the total roadway length, thereby reflecting the average sidewalk width.

Other variables were geographical point based. The number of intersections per kilometer was calculated by dividing the total number of intersections within a district by its total roadway length. The intersection locations were extracted directly from the road axis geometry. Moreover, data from the Madrid City Council's open data portal were processed in a similar manner. These data included metrics such as the number of traffic signs, traffic lights, metro stations, and bus stops per kilometer of roadway.

3.4.3 Article B: Macro-level analysis using Municipalities as the spatial unit of study

The second publication is a macro-level study with Spanish municipalities as the units of study, which is included in the Section 4.3 Article B: Analysis of the elderly pedestrian traffic accidents in urban scenarios: the case of the Spanish municipalities. In this study, the dependent variables are the number of older and non-older pedestrian traffic crashes for each municipality.

It is important to emphasize that while the study units analyzed in this article are Spanish municipalities, only their urbanized areas were considered, with non-urbanized areas being excluded.

Socioeconomic variables were obtained and processed following the methodology employed in the preceding study. Specifically, data were sourced from the Spanish National Statistics Institute (INE), including metrics such as total population, population aged over 65, population aged under 65, population density, femininity index, the proportion of residents aged over 65, and the proportion of primary residences. These variables were directly linked to municipalities under the assumption that population and housing are predominantly located within their urbanized areas. Additionally, data from the Directorate-General for Traffic (DGT) were utilized to analyze the total number of registered vehicles, vehicle types, and the average vehicle age.

Geographic data on **land use** in Spain, sourced from the National Geographic Institute, were processed to include only urbanized areas within each municipality. These data cover the entire national territory. For each municipality, the proportion of different land use types relative to the total urbanized area was calculated based on this refined dataset. A land use mix indicator was also derived from these proportions to quantify the diversity of land use patterns. Additionally, data on points of interest across Spain, obtained from OpenStreetMap, were incorporated to complement the analysis and provide a broader perspective on land use patterns.

Two types of **infrastructure variables** were identified: lengths and point counts. Using data obtained from OpenStreetMap, road networks were extracted for the entire national territory. These networks were subsequently clipped to include only roads located within the urbanized areas of municipalities, following a methodology analogous to the one used for land use data. After refinement, various road attributes were quantified in proportional terms. Specifically, the proportion of each road type (i.e., motorways, main roads, residential streets, and paths) relative to the total road length within each municipality was calculated. Additionally, the proportion of two-way roads in relation to total road length was determined. Furthermore, the analysis incorporated

point-based infrastructure variables normalized by road length. These included the number of intersections, pedestrian crossings, traffic lights, and bus stops per unit of road length.

3.4.4 Article C: Micro-level analysis using Street segments as the spatial unit of study

The third publication is a micro-level analysis with streets segments of the city of Madrid as the units of analysis, which is included in the Section 4.4 Article C: Age-Friendly Urban Design for Older Pedestrian Road Safety: A Street Segment Level Analysis in Madrid. In this study, the dependent variables are the number of older and non-older pedestrian traffic crashes in each street segment.

The third article developed as part of this doctoral dissertation required significantly more effort in data collection and processing than the earlier articles. As previously mentioned, this type of research demands far more detailed data compared to those employed in macro-level studies. In this context, geographic data is considerably more advantageous than tabular data because they can be processed to refer to specific street segments or intersections, allowing for more precise and localized analyses.

However, certain variables used in the study were necessarily tabular, as they are commonly referenced and collected based on various administrative divisions. Consequently, the most detailed administrative divisions available were utilized. This was the case for the **socioeconomic variables** employed in the study, such as population density and per capita income. To obtain these variables, data corresponding to the different census tracts of the city of Madrid were utilized. Each street segment was associated with the socioeconomic variables corresponding to the census tract where its centroid was located.

For **land use variables**, points of interest of various types were sourced from OpenStreetMap data. Three main categories of points of interest (POIs) were considered: public, health and education, and leisure. A systematic process was employed in which, for each point of interest, the closest road segment was identified, and the distance between them, as well as the angle formed by the line connecting the point of interest to the road segment relative to the north, were both calculated. Based on the type of point of interest, the number of points located on each road segment was determined (as shown in). Furthermore, using the angle information, the number of points located on each side of the street was calculated for each type of point of interest. This data was then used to develop a distribution index for points of interest. The index

assigns a value of zero when there are no points of interest on the street or when the number of points is equal on both sides of the street. Conversely, it reaches its maximum value of one when all points of interest are located on a single side of the street.

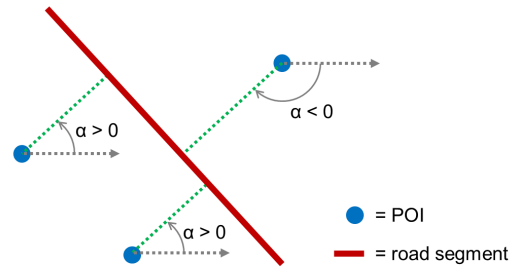


Figure 3.21. Process to assign each point to a street segment.

Source: Gálvez-Pérez et al. (2024a).

Significant emphasis was placed on the **infrastructure variables** employed in this article. As this is a micro-level study, these variables pertain to individual road segments rather than encompassing larger areas of the city. This approach enables a high level of detail, particularly regarding infrastructure. Nevertheless, certain information is often unavailable at the required level of granularity or may not exist altogether. Despite these limitations, as in the other articles, considerable effort was made to standardize the processes of data acquisition and processing to enhance the reproducibility of the study. The infrastructure variables were categorized into three groups, considering the primary challenges encountered by older adults in their role as pedestrians. Collected infrastructure variables are shown in Figure 3.22.

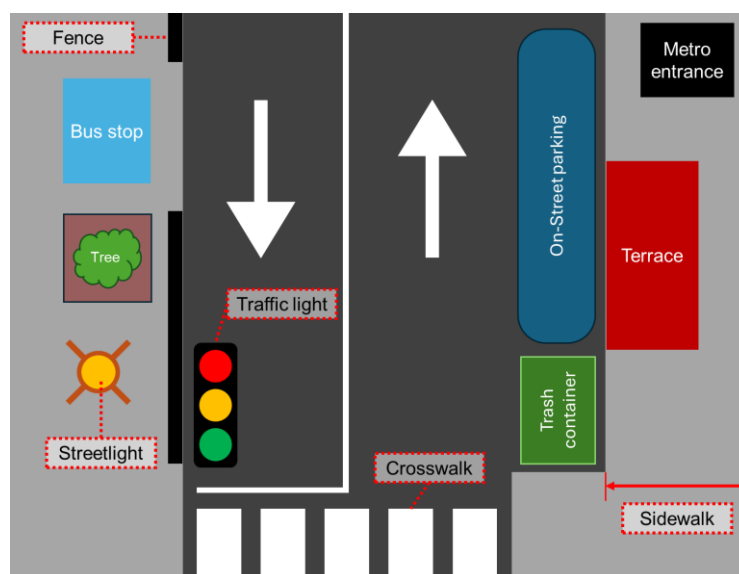


Figure 3.22. Infrastructure variables considered in the micro-level analysis.

Source: Gálvez-Pérez et al. (2024a).

A subgroup of infrastructure variables was first identified to reflect the **ease with which pedestrians can navigate sidewalks**. This group included two variables related to sidewalk width. For each road segment, the widths of sidewalks on both sides of the street were manually measured due to the absence of sufficiently precise data for systematic processing, ensuring the reliability of the results. The sum of these values was then calculated to represent the average width of the street segment. Also, a sidewalk disparity index was processed. This index assigns a value of one if a sidewalk exists on only one side of the street and a value of zero if the sidewalk widths are identical on both sides.

The presence of trees on sidewalks was manually measured due to the lack of reliable open data on the city's tree coverage. In contrast, the average slope of each street was derived from the digital terrain model provided by the Instituto Geográfico Nacional (IGN), which features a 2-meter grid resolution. The remaining variables in this subgroup are binary and pertain to the presence of various elements on sidewalks. These variables were processed similarly to points of interest, with each point assigned to the nearest road segment. Examples of these variables include the presence of restaurant terraces, streetlights, bus stops, and metro stations.

The second subgroup of **infrastructure** variables pertains to the **ease with which pedestrians can cross the street**. This subgroup includes the number of lanes on the street, the number of pedestrian crossings per unit length of the street, the presence of fences designed to prevent pedestrian crossings (either along sidewalks or in the median), and the presence of traffic lights. Due to reasons similar to those previously mentioned, the first three variables, number of lanes, pedestrian crossings, and pedestrian fences, were manually measured to ensure data accuracy and reliability.

The third subgroup of **infrastructure** variables pertains to the **ease with which pedestrians can detect vehicle traffic**. This subgroup includes the classification of the street type (i.e., one-way street, two-way street without a median, or two-way street with a median) the presence of surface parking in areas adjacent to sidewalks, the level of complexity in the road segment's contour, and the presence of garbage containers along the segment.

The first three variables were measured manually. Notably, the level of complexity of a road's outline depends on the types of intersections present at its endpoints. To address this, intersections were classified into three categories: three-way, multi-way, or roundabout. Because of difficulties in neck rotation of older pedestrians, 3-legged intersections were supposed to provide an advantage for one side of the street (i.e., the side of the street in front of the third leg), as shown in Figure 3.23. The road outline complexity variable was then generated based on the criteria, outlined in Table 3.6.

This method enabled the incorporation of intersection types at the ends of road segments into the model without explicitly prioritizing one type over another. Conversely, the presence of trash containers was processed in a manner similar to points of interest, as their geographic locations were available.

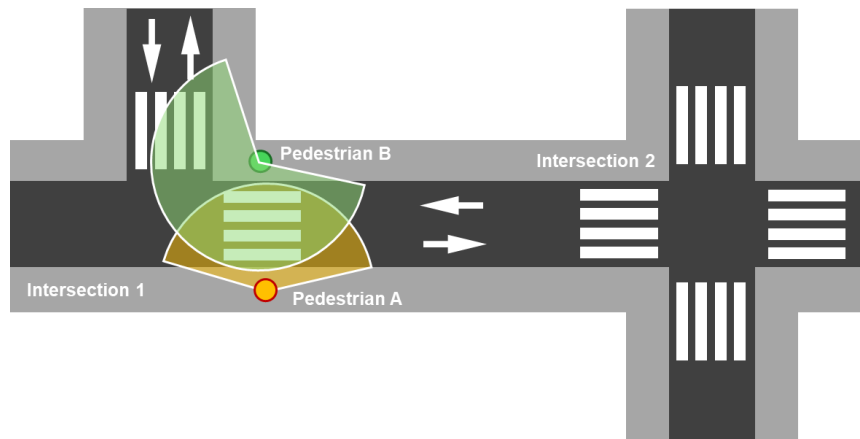


Figure 3.23. Example of street segment with contour complexity B, and pedestrians' field of view.

Source: Gálvez-Pérez et al. (2024a).

Table 3.6. Combinations of possible types of intersections at street edges and definition of level of complexity of the road segment contour.

Source: Gálvez-Pérez et al. (2024a).

Intersection 1	Intersection 2	
	3-legs	+3-legs or Roundabout
3-legs	A	B
+3-legs or Roundabout	B	C

3.4.5 Article D: Impact of COVID-19 pandemic on older pedestrian road safety

Article D is about the impact of covid-19 pandemic on older pedestrian road safety, in terms of frequency of crashes, injury severity of crashes and crash-related factors before and after the pandemic, considering traffic crash data until 2022. Unlike the three previous publications, Article D does not deal with frequency models of traffic crashes suffered by older pedestrians. Article D is included in Section 4.5 Article D: The effect of COVID-19 on older pedestrian road safety: A holistic analysis through pandemic phases.

This publication used two case studies. In a first stage of the methodology, Spain was employed as the case study. The effects of the pandemic in terms of frequency and

injury severity on traffic crashes involving older adults in urban scenarios were analyzed. In a second stage, the city of Madrid was used as the case study to investigate if the pandemic has modified crash-related factors of crashes involving older adults.

Regarding the first stage of the methodology, injured pedestrians in traffic crashes in urban roads in Spain were obtained from the traffic crash database for the years 2016 to 2022. With this data, the number of injured pedestrians by age (i.e., older and non-older) and month was obtained. Each month (i.e., January 2016 to December 2022) was identified with a phase of the pandemic (i.e., pre-pandemic, lockdown, new normality and after vaccination). Finally, the number of injured pedestrians was modeled with multiple Poisson regressions, separating older and non-older pedestrians and total, killed, seriously injured and slightly injured pedestrians. This national analysis was complemented with an analysis of the injury severity and age of the injured pedestrian through the pandemic phases in comparison with the pre-pandemic situation employing chi-squared tests.

About the second stage of the methodology, injured pedestrians in traffic crashes in urban roads in Madrid were obtained from the traffic crash database for the years 2018-2019 and 2021-2022. These events were geolocated using the process developed in Article A. Prior to statistical data modeling, a hot spot analysis of the traffic crashes before and after the pandemic was performed to analyze if these hotspots have been altered. For this analysis, the city of Madrid was divided using a square grid of 500 x 500 m squares. The number of older and non-older injured pedestrians pre- and post-pandemic inside of each square were calculated. Finally, for each case (older and non-older pedestrians, and pre- and post-pandemic) the Getis-Ord G_i^* statistic was calculated for each cell to detect statistically significant hot spots of older and non-older pedestrian traffic crashes.

For data statistical modeling, a database with features for each crash was built. For each injured pedestrian, crash variables present in the traffic crash database were obtained and re-categorized if necessary: weekday, time of the day, lighting condition, weather condition, crash location, pedestrian gender.

Apart from this variables, in a deeper effort to detect changes in crash-related factors, the characteristics of the surroundings of each crash were obtained using the data processing techniques developed throughout this doctoral thesis. Data on socioeconomics, land use and infrastructure of the surroundings of each crash were collected.

First, socioeconomic features included the population density and ageing rate of the census tract where the crash was registered, in a process similar to the one developed in Article C. Second, land use variables included the number of points of interest per

kilometer and the tertiary, residential land use proportions and Herfindahl-Hirschman index (HHI) land use mix index (Equation 3.32) within a buffer of radius 400 m (a 5-minute walk) around each crash. The process is analogous to the one developed in Articles A and B.

$$HHI_i = \sum_{j=1}^k (\%LU_{i,j})^2 \quad (3.32)$$

Where HHI_i is the HHI land use mix index around the i^{th} traffic crash, $\%LU_{i,j}$ is the proportion of the j^{th} land use around the i^{th} traffic crash, and k is the number of land uses included around the i^{th} traffic crash.

Third, infrastructure variables represent a high proportion of the collected independent variables. As in the land use variables case, infrastructure variables were collected from around a buffer of 400 m around each crash (Figure 3.24). Infrastructure variables processed were proportion of major roads, mean width of roads and sidewalks, bus stops per kilometer, metro stations per kilometer, traffic lights per kilometer, crosswalks per kilometer, and length of bike lanes per kilometer.

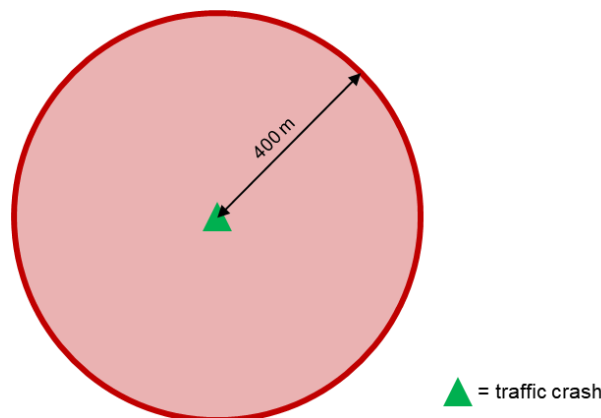


Figure 3.24. Scheme of the buffer around each traffic crash.

The proportion of major roads were calculated as in Article B, using data from OpenStreetMap. The mean widths of sidewalks and roads around each crash were obtained as the area of roads and sidewalks divided by total road length in each buffer. Finally, bus stops, metro, traffic lights and crosswalks by kilometer were obtained by counting

the number of these devices inside each buffer and dividing these values by the total road length in each buffer. These techniques have been already used in Articles A and B, but in this case instead of considering administrative divisions, the buffer around each crash would be the unit of study.

In summary, for each pedestrian traffic crash the period (the dependent variable, pre- or post-pandemic), the age of the pedestrian (older or non-older adult) and the independent variables were registered. Finally, the probability of a crash to be registered during the post-pandemic period was modeled through a binary logistic regression. With this model, the objective was to analyze crash-related variables that have been modified after the pandemic compared to the pre-pandemic period. As in the rest of published articles, two separate and independent models were built: one for older pedestrians and the other for non-older pedestrians.

4

Publications

4.1 Summary of publications

Article A

Title	The Influence of Built Environment Factors on Elderly Pedestrian Road Safety in Cities: The Experience of Madrid
Authors	Daniel Gálvez-Pérez, Begoña Guirao, Armando Ortuño, Luis Picado-Santos
Journal	International Journal of Environmental Research and Public Health (ISSN: 1660-4601)
Issue, page(s)	19(4), 2280
Publication date	17 February 2022
DOI	https://doi.org/10.3390/ijerph19042280
SJR category	Public Health, Environmental and Occupational Health
Quartile	Q2

Article B

Title	Analysis of the elderly pedestrian traffic accidents in urban scenarios: the case of the Spanish municipalities
Authors	Daniel Gálvez-Pérez, Begoña Guirao, Armando Ortuño
Journal	International Journal of Injury Control and Safety Promotion (ISSN: 1745-7300)
Issue, page(s)	31(3), 376-395
Publication date	22 April 2024
DOI	https://doi.org/10.1080/17457300.2024.2335482
JCR category	Public, Environmental & Occupational Health
Quartile	Q2

Article C

Title	Age-Friendly Urban Design for Older Pedestrian Road Safety: A Street Segment Level Analysis in Madrid
Authors	Daniel Gálvez-Pérez, Begoña Guirao, Armando Ortuño
Journal	Sustainability (ISSN: 2071-1050)
Issue, page(s)	16(19), 8298
Publication date	24 September 2024
DOI	https://doi.org/10.3390/su16198298
JCR category	Environmental Studies
Quartile	Q2

Article D

Title	The effect of COVID-19 on older pedestrian road safety: a holistic analysis through pandemic phases
Authors	Daniel Gálvez-Pérez, Begoña Guirao, David del-Villar-Juez
Journal	Journal of Transport and Health (ISSN: 2214-1413)
Issue, page(s)	41, 101986
Publication date	17 January 2025
DOI	https://doi.org/10.1016/j.jth.2025.101986
JCR category	Public, Environmental & Occupational Health
Quartile	Q2

4.2 *Article A: The Influence of Built Environment Factors on Elderly Pedestrian Road Safety in Cities: The Experience of Madrid*



Article

The Influence of Built Environment Factors on Elderly Pedestrian Road Safety in Cities: The Experience of Madrid

Daniel Gálvez-Pérez ^{1,*}, Begoña Guirao ¹, Armando Ortuño ² and Luis Picado-Santos ³

¹ Departamento de Ingeniería del Transporte, Territorio y Urbanismo, Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos, Universidad Politécnica de Madrid, Calle del Profesor Aranguren, 3, 28040 Madrid, Spain; begona.guirao@upm.es

² Escuela Politécnica Superior, Universidad de Alicante, 03690 San Vicente del Raspeig, Spain; arorpa@ua.es

³ CERIS, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal; luispicadosantos@tecnico.ulisboa.pt

* Correspondence: daniel.galvezp@upm.es

Abstract: With the progressive ageing of the population, the study of the relations between road safety and elderly users is becoming increasingly relevant. Although the decline of pedestrian skills in the elderly has been widely studied in the literature, few studies have been devoted to the contributing built environmental factors of the elderly pedestrian collisions, such as the sidewalk density, the presence of traffic lights, or even some indicator related to land use or the socioeconomic features of the urban fabric. This paper contributes to the limited literature on elderly pedestrian safety by applying a negative binomial regression to a set of built environmental variables to study the occurrence of accidents involving elderly and younger (non-elderly) pedestrians in Madrid (Spain) between 2006 and 2018. The model considers a selection of built environmental factors per city district, linked to land use, infrastructure, and socioeconomic indicators. Results have highlighted that the elderly pedestrian collisions could be avoided with the existence of a wider sidewalk in the district and a greater traffic lights density. Unlike younger pedestrian accidents, these accidents are much more favored in ageing districts with higher traffic flows.

Keywords: elderly pedestrians; road safety; road traffic collisions; accident analysis; built environment; street design



Citation: Gálvez-Pérez, D.; Guirao, B.; Ortuño, A.; Picado-Santos, L. The Influence of Built Environment Factors on Elderly Pedestrian Road Safety in Cities: The Experience of Madrid. *Int. J. Environ. Res. Public Health* **2022**, *19*, 2280. <https://doi.org/10.3390/ijerph19042280>

Academic Editors: Francisco Alonso and Cristina Esteban

Received: 18 January 2022

Accepted: 14 February 2022

Published: 17 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

An increase in life expectancy has led to a generalized phenomenon of population ageing. The United Nations forecasts that global life expectancy will increase from 70 years in 2015 to 83 years by 2095 [1]. This demographic change means, unavoidably, more elderly are actively part of mobility and road traffic, and this fact is translated into an increasing number of elderly road fatalities and severe injuries. Sustainable cities should also be inclusive for the elderly pedestrians, and mobility (mainly walking) is a fundamental prerequisite for the well-being of older persons [2–4]. Promoting walking for elderly pedestrians in sustainable cities implies the analysis of the urban built environment variables that affect the occurrence of elderly pedestrian accidents, because in most of the cases, policymakers and other professionals (e.g., urban planners, urban designers, and architects) could implement actions on the built environment variables.

Although elderly mobility has more advantages than drawbacks, in terms of road safety, population ageing shows an indirect negative effect as the number of elderly road fatalities and severe injuries is increasing [5]. For the coming years, forecasts at European level are not better: while at the moment one road traffic fatality out of five is aged 65 or over, it is expected that by 2050 one road traffic fatality out of three will be an older person if the risk rates of older people and younger age groups decline at the same pace [5]. Crashes involving users over 65 years old will rise to alarming figures if nationwide policy

actions are not taken to improve safety. Moreover, elderly pedestrians and cyclists are the weakest group as, within the entire European population, the elderly make up 39% of all pedestrian fatalities and 40% of all pedal cyclist fatalities compared to 18% and 19% of all car driver and passenger fatalities [5]. In OECD countries, persons 65 years and older represent 13% to 20% of the population, but they make up more than 50% of pedestrian fatalities [6]. Urban areas are especially risky scenarios for the elderly road safety as there are remarkably more elderly fatalities (55%) than there are middle-aged fatalities (33%) in urban areas.

The higher road accident risk of the elderly has been often linked to the reduction in physical and mental faculties with advancing age, which leads to inappropriate (and unexpected) behavior in elderly pedestrians and drivers. A large body of research [7–9] has dealt with specific physical and mental limitations of the elderly as road users, paying special attention to drivers [10,11]. However, while fatal accidents involving elderly drivers are still very few [12,13], elderly pedestrian fatalities in urban scenarios are statistically a matter of concern. Physical and mental limitations of elderly pedestrian are usually related to a poorer vision and hearing, lower walking speeds, longer reaction time, reduced ability to make head and neck movements, or less muscle agility. These physical and mental limitations can be exacerbated by age-related illnesses and certain chronic medication. Comorbidity (having more than one illness) is also more usual among the elderly population and is also linked to a higher crash risk [5]. Due to their physical and mental state, the elderly group also registers greater fatality rates [14].

In the literature, surveys and questionnaires to elderly road users have helped to investigate elderly pedestrians' and drivers' perception on road safety [8,15]. Older pedestrians and drivers can compensate with "self-regulation" actions for age-related functional declines. "Self-regulation" means that the individual is aware, acknowledges, and has insight of their functional impairments, being conscious of his own declining capacities and adapting his behavior to their limitations (e.g., avoiding complex traffic situations or only crossing at formal pedestrian crossings). Due to their mental maturity and "self-regulation", in general, older road users are more cautious compared to younger age groups [8].

Literature on older pedestrians' perception of their own declining capacity suggests that they have good awareness when there is clear feedback from the urban built environment. The urban built environment is very important for older pedestrians as inadequate infrastructures and unfamiliar environments can increase such anxieties and, in some cases, also reduce mobility [16,17]. For example, the fear of falling is a common fear among elderly pedestrians, and, through surveys, poorly maintained pavements were identified as the most important risk factor in their outdoor environment [18,19]. The built and physical environment can be a strong determinant of mobility, but in most studies the analysis of the built environment and walking infrastructure is reduced to subjective perceptions declared by the elderly by surveys and rarely backed by pedestrian collision data and the analysis of the accident location in the city.

In fact, little is known about the contributing built environmental factors and injury outcomes of the elderly pedestrian collisions. Apart from road safety education of elderly pedestrians, the implementation of actions on the urban built environment (adapting sidewalks, street and intersections signaling and lighting, urban car speed reduction, etc.) is a key issue to reduce elderly pedestrian fatalities and severe injuries. Therefore, investigating the contributory urban built environment factors to collision and injury risk is a first step in the development of appropriate road safety strategies and countermeasures.

This paper contributes to the limited existing literature by studying the built environment contributory factors to elderly pedestrian collisions in the city of Madrid (Spain) for the period 2006–2018. Madrid, with a population of more than 3 million inhabitants, is the capital and biggest Spanish city. Spain has one of the lowest road fatality rates per million of inhabitants (the seventh in 2019) among European countries, but population ageing and elderly pedestrian fatalities are starting to become a matter of concern. In 2019, 19% of the Spanish population was aged over 65 and 70% of pedestrian fatalities in

urban roads were individuals over 65 years old [20]. Methodology was based on a negative binomial regression, applied to a database made of built environmental factors and urban accidents involving one vehicle and one elderly or younger (non-elderly) person, thus capturing the effect of the built environmental factor on the accident occurrence per city district (administrative division). Previous works developed in Madrid, at a lower level of data disaggregation, by Gálvez-Pérez, Guirao, and Ortuño [21] already showed the importance of the built environmental variables. However, some key ad hoc variables were missing in the study, and the modelling results obtained for elderly pedestrian collisions were not compared with the analysis of the accident experienced by the rest of younger pedestrian. For that reason, this research is not only an extension of the cited study but also a comparative analysis among two age groups with a larger number of studied built environment features, obtaining a better holistic approach of the problem.

The paper is divided into the following sections: Section 1 contains the introduction; Section 2 presents the state of the art on pedestrian road safety and elderly pedestrians as a study group; Section 3 gives a detailed description of Madrid as a case study. Section 4 explains the model structure, application to the database, and discusses the most important results affecting the process. Finally, Section 5 presents the conclusions and future research lines drawn from this research.

2. State of the Art on Road Safety of Pedestrians and Elderly Pedestrians

The literature has demonstrated that active mobility is more intricate than a logistics optimization problem [22–25], reflecting that walking (like cycling) is complex behavior resulting from the interaction between individuals, groups, and their environment. While frequent car routes for drivers are limited to the road network and route choice is more dependent on generalized cost optimization, pedestrians' routes include all the city streets' networks, and route choice does not necessarily follow the shortest itinerary in distance or time or that more efficient in terms of energy or cost, as it usually depends on other route features [26]. As a consequence, pedestrian mobility analysis requires integrated perspectives and approaches from different disciplines such as urbanism, psychology, engineering, ecology, and physical health.

The complexity of the pedestrian mobility analysis is also extended to the study of the pedestrians' road safety. Pedestrian fatalities and injuries due to road safety accidents are some of the limited negative externalities derived from walking in our cities. The location of the accident can be considered a built environment factor [27]. According to Stocker et al. [27], road safety risk in the built environment can be studied at regional level (population density, land use mix, urban sprawl, socioeconomic variables, etc.) and at local level (pedestrian infrastructure and roadway design). With this approach, conditions of the built environment include also factors such as traffic speeds, traffic flows, and visibility, and the pedestrian accident risk is a mixture of the level of risk provided by the built environment and the type of pedestrian (elderly, young children and young adults, disable pedestrians, intoxicated pedestrians, gender roles, etc.).

Tiwari [28] has studied the relationship between the evolution of pedestrian safety research and the measures to reduce these type of accidents for the last 120 years, taking as reference year the date of the first reported pedestrian fatality in 1899 [29]. Although the epidemiology of pedestrian crashes through available accident databases has been continuously developed, findings of Tiwari [28] describe an initial special focus on the analysis of pedestrian crossing behavior, and research study designs based primarily on field observations, complemented with pedestrian interviews, state preference studies and risk perception at various facilities. Street crossings were identified from the beginning as risky locations for pedestrians because there is always a time limitation to making a decision to cross the street: the time elapsing between the pedestrians first observation of the oncoming vehicle and the arrival of the vehicle at the crossing. As a consequence, gap acceptance has been one of the first variables used in pedestrian field observation, and later, preferences of route choice and location for crossing roads were analyzed. These primary

field observations at road crossings have migrated in the last decades towards the use of video cameras, image processing [30], and multi-agent simulation systems [31] to analyze pedestrian behavior in general. The first age group of pedestrians investigated in detail was children [32], in order to better understand their cognitive skills and limitations to avoid road accidents.

Tiwari [28] pointed out two main results from the research progress on pedestrian road safety. The first one deals with individual gap acceptance levels, which are quite dependent of the width of the crossing points, but can be strongly mitigated by individual capabilities. Indirectly, this result accepts “self-regulation” of older pedestrians when their capabilities decline (e.g., older people, who walk slower, will take longer gaps). The second result is related to children road safety, giving priority to the improvement of the built environment over traditional pedestrian traffic education (especially for children below the age of 10 years). Despite the large body of research on pedestrian behavior and risk exposure, pedestrian crashes have not decreased at the desirable levels in both motorized and less-motorized countries. A new research approach is required to increase road safety of pedestrians, and there is a consensus on built environment principles that can lead to safer cities for pedestrians [33]. In this regard, the concern on designing a better built environment for certain population groups (such as the elderly) in urban scenarios has increased, and the idea of creating “age-friendly” cities has been promoted by institutions such as the World Health Organization [34].

Pedestrian and roadway infrastructure variables are key variables of the built environment at local level [27], and the literature shows a broad inventory of infrastructure features affecting pedestrian road safety. In the recent literature, some authors [18,35,36] have analyzed pedestrian road safety in relation to urban road type and traffic flows. The works developed by Galanis et al. [36], unlike in the case of Corazza et al. [18] and Demasi et al. [35], are more focused on pedestrian behavior in relation to urban road type and traffic flows. The main results by Galanis et al. [36], focusing on pedestrian illegal behavior, underlined that a low level of motorized traffic flow in combination with maintenance and mobility problems in pedestrian infrastructure incites pedestrians to walk outside the sidewalk (in the street) and underestimate their safety issues. In the studies developed by Corazza et al. [18], the pavement of the sidewalk has been directly related to pedestrian road safety, as distressed or too narrow sidewalks induce pedestrians to walk outside the sidewalks and on the carriageways, which is very unsafe. Demasi et al. [35] proposed a methodology to estimate the level of road safety for vulnerable users (pedestrians, cyclists, and motorcyclists) of each section of a street and the hazard index (with infrastructure variables) of the overall branch. Research developed recently by Kim [37] represents a milestone in the study of elderly pedestrian collisions, comparing this group of accidents with younger pedestrian collisions in a specific type of road section, the intersection level. This work has proved the influence of land use, ambient conditions, and intersection characteristics on pedestrian safety and showed differences for both age groups. Specifically, three-way intersections, raised medians, street trees and park and recreational land use were found to have a positive effect on the elderly pedestrians’ safety, and the number of bus stops increased the chance to have an elderly pedestrian collision. Moreover, Kim [37] demonstrated that some measures implemented to reduce pedestrian collisions may favor an age group against others. For example, according to Kim [37], intersections with crosswalks or colored crosswalks do not contribute to elderly pedestrians’ safety, but rather to the safety of younger pedestrians. Muhan Lv et al. [38] analyzed the occurrence of vehicle–elderly pedestrian collisions in relation to the characteristics of the built environment at the road segment level (microlevel) in a district of Shanghai. They considered the number of elderly pedestrian accidents in each segment as dependent variables, taxi flow and elderly population as exposure variables, and built environment features, extracted from online databases and image processing, as independent variables. The authors built both Poisson and GWPR models to assess the effect of built environment on the occurrence of elderly pedestrian collisions and found that road segments near schools, supermarkets, traditional

markets, bus stops, and metro stations were more dangerous for elderly pedestrians. Moreover, it was found, through the GWPR model, that the influence of the built environment features varied throughout the studied district. For instance, green space may improve elderly pedestrian road safety only in noncongested environment.

As described in the mentioned research works, pedestrian road safety studies that have analyzed the infrastructure built environment as a safety-contributing factor rarely single out the elderly group. These studies focused mainly on pedestrian crossings [15,39,40] and sidewalk pavement state [18]. The majority of elderly pedestrian fatalities occur in crashes in which the elderly pedestrian is hit by another vehicle, and this scenario is overrepresented in such accidents in which they initiated a crossing maneuver [5] (European Commission, 2015). In relation to signalized pedestrian facilities, Koepsell et al. [39] found that a high rate of older-pedestrian crashes still occurs at signalized (or marked) pedestrian facilities, and, specifically, signalized crosswalks generate a 2.1 times increased crash risk for older pedestrians (even after controlling methodologically for confounding factors such as site characteristics and pedestrian and vehicle volumes). In general, elderly pedestrians tend to believe that the rest of road users obey traffic regulations, which might give them too much confidence in these rules, and this fact increases their crash risk [8]. Additionally, older pedestrians tend to wear dark clothes, which can reduce their visibility for drivers at intersections [8].

In relation to the sidewalk pavement state, uneven surfaces are a matter of concern especially for older pedestrians because this population group has a higher risk of falling, stumbling, or stripping while walking compared with younger adults. For the elderly, maintaining postural stability and balance is more difficult [8], and some authors have shown [41] that the rate of single-pedestrian accidents is significantly higher than the rate of any other accident type for older road users, but injuries in this type of accident tend to be less severe [8].

Built environmental factors at regional level [27] include also neighborhood design and land use and socioeconomic features of the districts where the elderly pedestrians live. Many authors have studied the relationship between pedestrian collisions and macrolevel built environmental factors [42–45]. Wedagama et al. [42] found that pedestrian casualties in the city center zone are particularly associated with an increase in retail and community land use during working hours. In the city center zone, out of working hours, an increase in retail land use (almost certainly clubs and bars) is also associated with an increase in pedestrian casualties. Wier et al. [43] considered, as a built environmental factor to explain pedestrian crashes, the proportion of people living in poverty and the proportion aged 65 and older. Ukkusuri et al. [44] found that tracts of land (districts or neighborhoods) with a greater proportion of industrial, commercial, and open land use types have greater likelihood for collisions while tracts with a higher fraction of residential land use have significantly lower likelihood of pedestrian collisions. Moreover, census districts that have a greater number of transit stops and schools are more likely to have greater pedestrian crashes. Some authors have even demonstrated that some built environmental factors contribute to pedestrian collisions more than road infrastructure conditions do. Aparidian and Smirnov [45] found that socioeconomic neighborhood features are more significant to predict the occurrence of pedestrian collisions than is traffic exposure, measured in Vehicles Miles Travelled (VMT). Furthermore, they suggested that a proper policy to improve road safety in urban areas would be periodic renovation and building of easy-to-adapt environments, because traffic volume or land use changes may deteriorate the road safety of a certain territory. Sugie Lee, Junho Yoon, and Ayoung Woo [46] studied Seoul pedestrian safety in relation to pedestrian age and the price of housing in the neighborhoods on a macro-level scale. They built 12 separate negative binomial regression models to calculate the number of collisions in a neighborhood, disaggregating these crashes by the age of the pedestrian (elderly pedestrians and all pedestrians), the severity of the injury of the pedestrian (total crashes, KSI, and slight), and the price of the housing in the neighborhood (high- and low-price housing). They found that traffic regulators, such as crosswalks and

four-way intersections, had a substantial impact on the occurrence of elderly crashes in areas with low housing prices. It was suggested that this result was caused by an uneven distribution of special pedestrian road safety measures. In addition, the authors remarked on the importance of collaborating in a multidisciplinary work group (engineer and non-engineer specialists) to address elderly pedestrian road safety issues through improvements in built environments and educational programs for elderly pedestrians and drivers.

This paper helps to shed light on the influence of built environmental factors on elderly pedestrian road safety. Although road safety education is still an important tool to avoid accidents, Tiwari [28], based on a literature review on pedestrian road safety, pointed out that decreasing fatalities and injuries is only possible if we take focused and targeted actions on the built environmental factors. Until now, the most recommended countermeasures [5] to protect elderly pedestrians from vehicle crashes are focused on reducing interactions between pedestrians and other road user types, and reducing the average speed of motorized traffic at locations with high pedestrian flow. There is also a need to implement measures that improve the conspicuity of pedestrians for drivers and specifically improve the perception of elderly pedestrians about other road users. These latter actions need a better understanding of elderly pedestrian crashes, in order to make decisions on the changes needed in the built environment.

These results lead us to the scientific need to complement studies based on road safety perceptions of elderly pedestrian and the mobility of elderly pedestrians with the analysis of the accident location in the urban road network. The next section describes the city of Madrid (Spain) as a case study for this research. An exploratory analysis of the elderly pedestrian accidents database (2006–2018) is presented per city district, as well of a selection of street built environment variables and socioeconomic factors.

3. Materials and Methods

3.1. Madrid Case Study

Spain, together with Japan, Finland, Sweden, Greece, Italy, and Germany, is among the countries in the world with a larger number of older adults. In 2019, approximately 19.3% of Spanish people were over the age of 65, and almost 6% were over 80 years old [47]; this figure could reach 40% in 2060 [48]. The city of Madrid offers a good case study for this research as it is the capital of one of the most rapidly ageing nations in the world and it also has a high proportion of elderly residents (19%). Large cities offer opportunities for better management of dedicated aging resources, and Madrid has almost 3.5 million inhabitants and is administratively divided into 21 heterogeneous districts, which are further subdivided into 131 neighborhoods (barrios). Figure 1 shows the ageing rate per city district in Madrid, revealing a higher ageing rate for central districts. In this case study, there is a great heterogeneity in terms of administrative district surface, and bigger districts are located in the periphery of the city, involving a lower residential proportion, and a lower population density and aging rate. In terms of number of inhabitants, although districts located in the periphery have low density, due to their high surface area, their population (number of inhabitants) is higher than that of central districts.

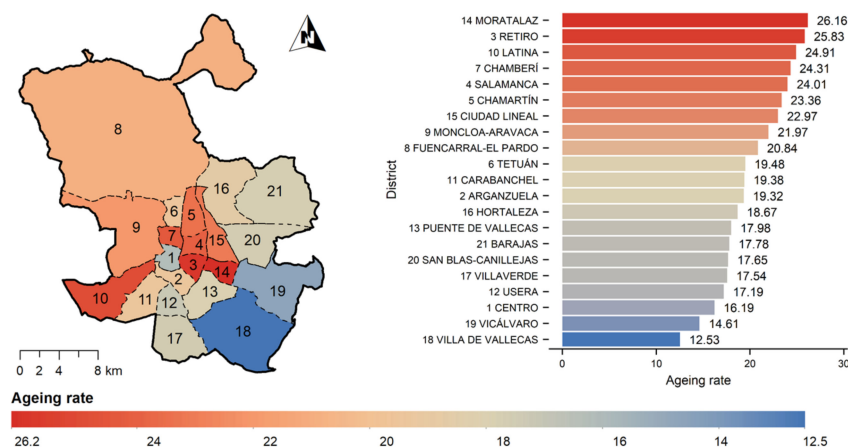


Figure 1. Madrid ageing rates for city districts (2018).

Madrid also provides a good road safety database for the pedestrian collisions study. The quality of the available accident database determines, to a large degree, the success and approach of any research on road safety. Compared to other national databases, the Spanish database is sufficiently consolidated [49,50]. The pedestrian collision data used in this research were extracted from the Spanish Accident Statistics Database and consist of accidents on Madrid city streets involving a single vehicle and a pedestrian, during a period of 13 years (2006–2018). According to the literature [5], the majority of elderly pedestrian fatalities occur in crashes in which the elderly pedestrian is hit by another vehicle. There is a higher rate of single-pedestrian accidents with older roads, but injuries in this type of accident tend to be less severe [8] with lower rates of fractures (intracranial and other injuries) compared to those who are involved in pedestrian–vehicle events [8,41]. Moreover, in the Madrid accident database, the profile of pedestrian accident involving “1 person and 1 pedestrian” was predominant in the sample (89% of vehicle–pedestrian collisions).

One of the weakest points of the Spanish accident database is the absence of traffic exposure data (traffic flow), street road layout, and traffic signaling information associated to the accident location. Moreover, the Spanish Accident Statistic Database does not supply the accident location with GPS coordinates (as in the US), using instead the kilometric point on the road of an interurban road or the closest number of a street (the name of two streets in case of accidents located in intersections) in urban scenarios, leading to further data-processing problems. Collecting these variables is very laborious, but is the only way to obtain a holistic approach for a road safety study. Furthermore, this information, if obtained, is very valuable for all stages of this line of research.

In this study, comparison between elderly pedestrian accidents and non-elderly pedestrian accidents is necessary to analyze the variables that specifically affect elderly pedestrian. Consequently, two dependent variables have been considered: the number of vehicle–elderly pedestrian collisions and the number of vehicle–non-elderly pedestrian collisions in each Madrid district. With this criteria, 20,236 records of vehicle–pedestrian collisions were filtered. Later, records of people involved in these accidents were studied to acquire collisions that have a configuration of “1 person and 1 vehicle”. Thus, 18,118 (89% of vehicle–pedestrian collisions) records were filtered to be studied. At this point, a homogenous and tidy dataset was already built, containing basic information about vehicle–pedestrian collisions in Madrid city: date, time, and type of the accident and age, gender, and injury level of the pedestrian and the driver. The last step of this procedure was to divide the ad hoc created database into two separated databases, considering the age of the pedestrian: one for pedestrians that were 65 years or older and one for the rest of the records. The result was two separate groups of data with 4663 vehicle–elderly pedestrian collisions and 13,455 vehicle–non-elderly pedestrian collisions. An initial assessment of these data highlights the fragility of the elderly age group over the rest of the population, because

this age group accounts for 51% of total fatalities and 34% of total serious injuries, but only accounts for 26% of total accidents of the studied type.

Geolocation of the accidents is a key issue in this research, as accidents will be assigned to the Madrid city district they occurred in. A logical process was developed to geolocate these accidents from the available alphanumeric information about their position (i.e., street name and number or two street names if the location is a junction). It was possible to geo-locate 93% of the vehicle–pedestrian collisions. As a result, Figure 2 shows the number of collisions suffered by elderly pedestrians, non-elderly pedestrians, and all pedestrians per Madrid district for the period 2006–2018. A larger number of accidents are in the central districts in comparison to the periphery. Moreover, the figure shows that the area (surface) of districts is quite heterogeneous, and this heterogeneity is also common for other district variables, such as their road network, and other built environment indicators. Moreover, this heterogeneity is also maintained, to a greater or lesser extent, inside each district.

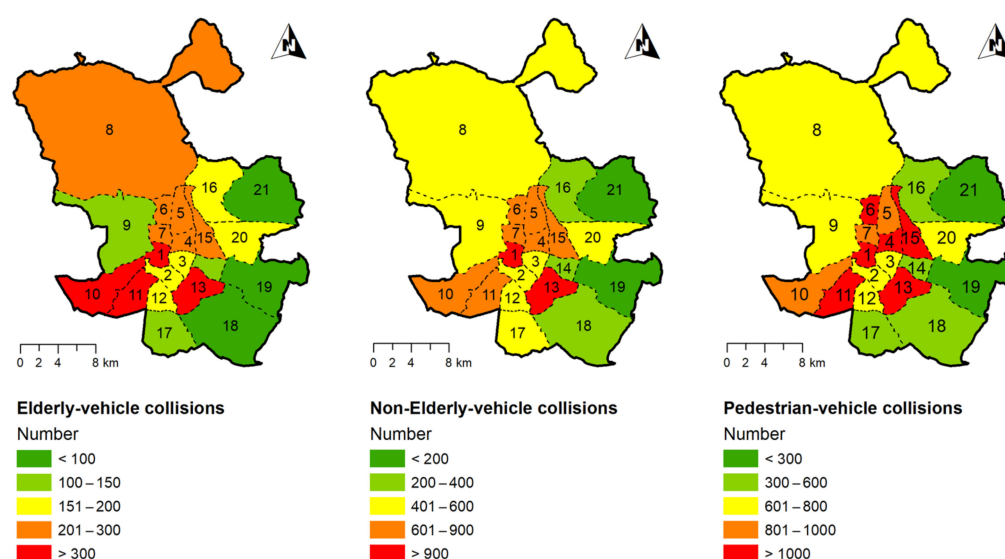


Figure 2. Number of collisions suffered by elderly pedestrians, non-elderly pedestrians, and all pedestrians per Madrid district for the period 2006–2018.

To assess the impact of the built environment features on elderly pedestrian road safety and compare it to impact on the group of non-elderly pedestrians, data regarding Madrid districts were gathered. According to the most used built environment variables in the literature [42–44], different groups of variables were collected and processed. These variables can be classified as follows: (i) socioeconomic variables, (ii) land use variables, and (iii) infrastructure features of each district. The variables of each group were selected considering the literature analysis and the availability of the data. It was especially difficult to select infrastructure variables, as there are no exhaustive inventories on Madrid city street features (e.g., width of the streets and existence and width of the sidewalks). As exposure variables, the Annual Average Daily Traffic (AADT) and total street length of each district were used. The pedestrian flow was not considered because of the lack of these data in official databases. The exposure variables are expected to contribute to the occurrence of vehicle–pedestrian collisions. Due to heterogeneity of districts, the variables were normalized to compare districts with different surfaces and street lengths. There is a big heterogeneity between the surfaces of the districts, and some of them—especially the peripheral ones—include high-capacity street (highways) and non-urban areas. As a consequence, the surface of the district seems unsuitable to normalize all the independent variables. Land use variables that deal with surfaces will be normalized by the district surface, but district surface does not provide information about the “useful area” of the districts, meaning the area where pedestrians have real interactions with the motorized traffic. Nevertheless, total street length of each district is a fair normalizing value for

infrastructure variables because most of these indicators are distributed along the streets (e.g., signals, traffic lights, and street junctions).

The first group of gathered data, socioeconomic indicators (i), consists of the number of inhabitants (elderly, non-elderly, and total), population density, ageing rate, and average annual income per household in each district. The number of inhabitants of each district per year was available directly at Madrid City Council in five-year age groups. Hence, the number of the elderly, non-elderly, and total inhabitants was calculated for each district and year. These variables were normalized by total street length of the district to represent the level of exposure of each age group. Population density was also available at Madrid City Council for each district and year. Moreover, as the data about the number of inhabitants of each district were separated into age groups, the ageing rate (the proportion of inhabitants older than 65 years over the total number of inhabitants of each district) was calculated. The average annual income per household was accessible through the Madrid Council Statistical Portal for the years 2013 to 2017 and in the INE for 2015 to 2018. With respect to the rest of the years, the value of the closest available year was used.

The second group (ii), land use variables, consists of the number of Points of Interest (POIs) of a different nature per street kilometer and the proportion of a group of nominal land use indicators—residential, green area, and main street proportion. The number of POIs per kilometer in each district was considered invariant over the studied period. It was obtained through the Businesses Census published at Madrid Council Open Data Portal (MCOOP), which includes leisure, retail, education centers and hospitals, among others. Finally, the number of these points was divided by the total street length of each district. Nominal land use was obtained directly from MCOOP as a polygon-shaped file compatible with GIS software, containing information from the “Plan General de Ordenación Urbana de Madrid (PGOUM)”. For this study, residential and green area proportions of the total district surface were selected and considered invariable over time. Commercial use was dismissed as the variance between districts and its absolute value were almost null. Green area is defined by the PGOUM (1997) [51] as the land that provides one of these services: (i) neighborhood green area, garden area of small or medium surface, (ii) district park, gardens with a medium or large surface where people can develop multiple activities, (iii) urban park, similar to district parks but with historical or functional singularities, and (iv) metropolitan park, forest areas that offer different cultural and recreational activities that are integrated in the natural environment. The “main street” proportion was estimated according to the definition of “main street”, defined by [51] as the public road that enables the mobility and accessibility between districts due to its functional conditions, design features, traffic flow, or associated activities. Usually, main streets are arterial streets at urban level that are wider, more signalized, more illuminated, and with a higher vehicle traffic volume than the rest of the road network. The main street proportion was obtained using the nominal land use data as the ratio of main street surface over the total street surface.

The third group (iii), infrastructure, consists of those variables linked to the street type directly available from the official and open databases, such as total street length, sidewalk density (district total sidewalk surface per district street length), street junctions per kilometer, signaling per kilometer, traffic lights per kilometer, and public transport presence—bus stops and metro stations per kilometer. Total street length and street junctions per kilometer were extracted from the National Center for Geographic Information (CNIG) road axis. “Road axis” includes roads that can be used by vehicles and pedestrians simultaneously (e.g., excluding highways). Sidewalk density was obtained as the surface of sidewalk over the total street length of each district. Signaling density and traffic light density were obtained as the total number of road signals and traffic lights in a district, with their location available through the MCOOP, over the total street length. Bus stops and Metro station locations and opening/closing dates were available through the Madrid Regional Transport Consortium (CRTM) database.

Among infrastructure variables, AADT was also considered as an indicator of the infrastructure use. The AADT was only available for a group of street segments in the MCODP, so AADT was estimated for each district as the median of the traffic flow registered in the traffic gauging station located in a district. AADT is an important variable to include in this research as it shows the level of accident risk exposure of each street. This macroscopic research, at district level, forces the estimation of a medium AADT per district in the city, which can be a rude estimation mainly in those districts with high differences between traffic flows in their streets.

The result of this process was an ad hoc-designed database that contains different built environment indicators of Madrid as a case study. In order to show district heterogeneity, this information can be displayed on a map to study the spatial distribution of the variables. Figure 3 shows graphically a subset of the gathered variables on a Madrid districts map. The total street length is higher in bigger peripheral districts, mainly in Fuencarral-El Pardo (8), Moncloa-Aravaca (9), Hortaleza (16), Villa de Vallecas (18), and Puente de Vallecas (13). In terms of traffic, AADT is greater in central districts such as Retiro (3), Salamanca (4), and Chamberí (7), where there is a higher concentration of POIs. Nevertheless, in this study AADT was obtained as the median value of a group of measuring points provided by Madrid Council. Hence, this variable is only an average measure of a district traffic flow as gauging stations are usually located in main streets and not in narrower residential ones. In relation to population density, values are larger in the city center districts, those that have a smaller surface area, especially in Chamberí (7), Tetuán (6), Salamanca (4), and Centro (1), which are the four most centrally located districts. Elderly and non-elderly inhabitants per total street length have a similar behavior, being maximum in Chamberí (7), Retiro (3), Salamanca (4), and Moratalaz (14) for the elderly pedestrians, and in Chamberí (7), Arganzuela (2), Salamanca (4), and Tetuán (6) for the non-elderly pedestrians. If we focus on land use variables, POIs per kilometer have the same distribution as that of population density. This variable also represents an attraction motivation for inhabitants of other districts to visit, being this indicator higher in Centro (1), Chamberí (7), Arganzuela (2), Salamanca (4), and Tetuán (6). Sidewalk density is higher in those districts with wider sidewalks or sidewalks all along the streets. In fact, the sidewalk density is an average indirect measurement of the sidewalk width. The lack of sidewalks in some streets reduces the value of this indicator even though the existent ones are wider than those in other districts. Figure 3 shows that this variable does not follow a spatial pattern of distribution among districts as happens with previous variables. However, sidewalk is higher in districts located in a central north–south band, especially the suburban districts, being remarkably higher in Moratalaz (14), Chamberí (7), and Carabanchel (11). The indicator “junctions per kilometer” is higher in districts with intricate road networks. This fact applies mainly in central and south districts, especially in Centro (1), Tetuán (6), Usera (12), and Carabanchel (11). Finally, “bus stops per kilometer” is higher in central districts, as these are attraction areas and, as a consequence, the public transport network is denser. This indicator is also related with population density and inhabitants per kilometer variables, being maximum in Moratalaz (14), Chamberí (7), Arganzuela (2), and Salamanca (4).

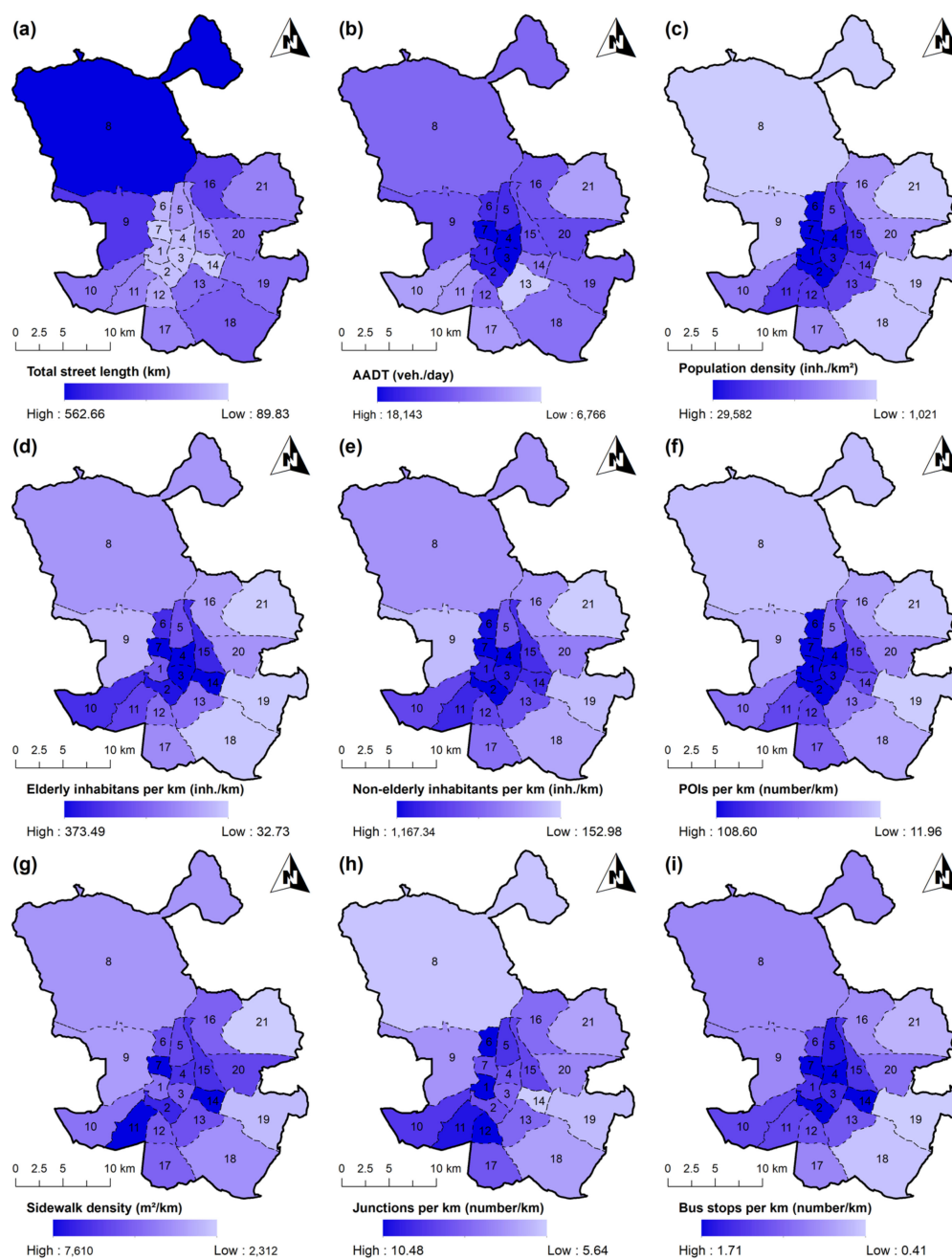


Figure 3. Spatial distribution of built environment variables in Madrid case study in 2018: (a) total street length, (b) AADT, (c) population density, (d) elderly inhabitants per kilometer, (e) Non-elderly inhabitants per kilometer, (f) POIs per kilometer, (g) sidewalk density, (h) junctions per kilometer, and (i) bus stops per kilometer.

3.2. Methodology

The main objective of this paper was to assess the influence of the built environment of a city on the occurrence of vehicle–pedestrian collisions considering the ageing of the population. This analysis was developed at the macroscopic level, using Madrid districts as the spatial unit of reference. District level was selected with the aim to use mainly the data available in official databases that did not require a big amount of postprocessing operations, and the lowest level of detail found in these data was the district level. Once a database containing the dependent (number of accidents) and independent (built environment features) variables was obtained, two regression models were formulated and compared, using all the studied contribution factors: one model for vehicle–pedestrian collisions

where the pedestrian was 65 or older, and other for the rest of vehicle–pedestrian crashes. The purpose was to evaluate the statistical significance and sign of each variable used in the models and study the differences between them. For this objective, the statistical regression model employed was the negative binomial regression, a widely used model on random discrete events with overdispersion. In this section, the construction of the two negative binomial statistical models is described, using the ad hoc-designed datasets explained in the previous section.

The Negative Binomial (NB) distribution is a very frequently employed regression model in studies on the occurrence of road crashes of different natures, including studies at a macroscopic level [52]. The NB is a distribution derived from the Poisson gamma distribution [53], and it was employed because it can operate with data with overdispersion. This feature cannot be found in a Poisson regression model, as mean and variance are identical. Since overdispersion was found in the occurrence of pedestrian–vehicle collisions by district dataset, the NB model was appropriate to build the statistical models to be analyzed. The Probability Density Function (PDF) of the NB distribution is as follows:

$$P(Y = y_i) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(\alpha^{-1})y_i!} \left(\frac{\alpha \cdot \mu_i}{1 + \alpha \cdot \mu_i} \right)^{y_i} \left(\frac{1}{1 + \alpha \cdot \mu_i} \right)^{\alpha^{-1}} \quad (1)$$

where $P(Y = y_i)$ is the probability of Y resulting in y_i , μ_i is the projected number of crashes, y_i is the number of vehicle–pedestrian collisions at the district I , and α is the dispersion parameter.

The expected number of crashes can be calculated using the following equation:

$$\mu_i = \exp \left(\beta_0 + \sum_{j=1}^n \beta'_j x_{ij} \right) \quad (2)$$

where μ_i is the projected number of crashes, β_0 is the intercept of the model, β_j are the estimated parameters of the variables, and x_{ij} are the independent and known variables at the district i .

As accident occurrence is a random event and the studied independent variables of Madrid districts do not vary greatly, the input of the negative binomial model had to be treated due to the variance of the number of vehicle–pedestrian crashes in consecutive years of the study. With this approach, the Sliding Window Method (SWM) was used in the data. The SWM consists of “moving” a virtual window of a specified width over a tidy database to obtain the result of a certain function (e.g., sum, average, or median value) applied to a variable in the range covered by the window [54,55]. A five-year window was used in this investigation, and, as the period of study was 13 years, nine complete windows were created for each city district (see Figure 4). As a consequence, nine records of each district (21 in total) were available, summing up a total of 189 observations. On the one hand, the number of vehicle–pedestrian collisions (the dependent variable) in each five-year window was calculated as the sum of the events that occurred during that period. Thus, the output variable of the model is “number of vehicle–pedestrian collisions in 5 years in the district”. On the other hand, for the independent variables (socioeconomic, land use, and infrastructure), the used function was the average value for each five-year period.

At this point, dependent and independent variables were already collected and processed. Table 1 shows the main statistics associated to the built environmental variables classified in three groups (socioeconomic, land use, and infrastructure). Thus, the Negative binomial regression model was applied, and two models, using all the studied variables, were formulated in order to compare results: one model for vehicle–pedestrian collisions where the pedestrian was 65 or older, and other for the rest of vehicle–pedestrian crashes. In both models, two exposure variables were used: total street length and AADT (Equation (3)).

$$\text{Acc} = L^{\beta_1} \cdot \text{AADT}^{\beta_2} \cdot \exp(\beta_0 + \beta_3 \cdot \text{InhExp} + \beta_4 \cdot \text{Pop D} + \beta_5 \cdot \text{AI} + \beta_6 \cdot \text{POIs} + \beta_7 \cdot \text{R Prop} + \beta_8 \cdot \text{G Prop} + \beta_9 \cdot \text{MS Prop} + \beta_{10} \cdot \text{Swk D} + \beta_{11} \cdot \text{Junct} + \beta_{12} \cdot \text{Signals} + \beta_{13} \cdot \text{TLights} + \beta_{14} \cdot \text{Metro} + \beta_{15} \cdot \text{Bus}) \quad (3)$$

where “Acc” is the number of vehicle–pedestrian collisions, where pedestrians are elderly or non-elderly depending on the model, during a five-year window in a district, “L” is the total district street length (km), “AADT” is the average AADT per district (veh./day), “InhExp” is the number of inhabitants of the studied age group (elderly or non-elderly, depending on the model) per kilometer (inhabitants/km), “Pop D” is population density (inhabitants/km²), “AI” is average annual income per household (€), “POIs” is POIs per street kilometer (points/km), “R Prop” is residential proportion (% of the district surface), “G Prop” is green area proportion (% of the district surface), “MS Prop” is main street proportion (% of the street surface over the district surface), “Swk D” is sidewalk density (m²/km), “Junction” is the number of street junctions per street kilometer (junctions/km), “Signals” is the number of signals per street kilometer (signals/km), “Tlights” is the number of traffic lights per street kilometer (traffic lights/km), “Metro” is the number of metro stations per street kilometer (metro stations/km), and “Bus” is the number of bus stops per street kilometer (bus stops/km).

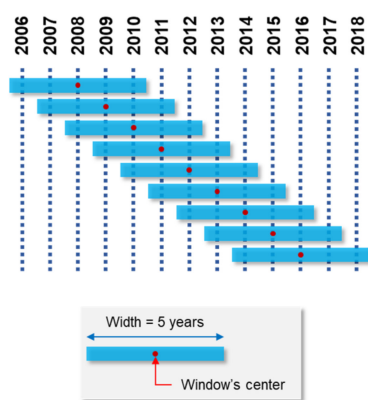


Figure 4. Sliding window method applied to Madrid case study in the period 2006–2018.

Table 1. Main statistics of the built environment variables considered in the research by type (land use, socioeconomic, and infrastructure).

Variable	Unit	Min.	Max.	Mean	Median	σ (SD)
Exposure						
Total street length (L)	km	89.83	562.66	243.54	252.82	123.28
AADT (AADT)	veh./day	6766.00	18,143.00	11,299.00	11,052.00	2727.46
Socioeconomics						
Elderly inhabitants per km (InhExp)	inh./km	18.80	378.02	157.24	157.22	92.35
Non-elderly inhabitants per km (InhExp)	inh./km	148.00	1298.70	622.70	642.00	298.14
Population density (Pop D)	inh./km ²	880.00	32,227.00	14,263.00	15,822.00	9726.10
Average income per household (AI)	€/district	23,517.00	70,735.00	38,456.00	35,532.00	10,992.18
Land use						
POIs per km (POIs)	points/km	11.96	108.60	49.14	43.09	28.17
Residential proportion (R Prop)	% Surface	0.03	0.48	0.27	0.28	0.14
Green area proportion (G Prop)	% Surface	0.00	0.41	0.08	0.05	0.09
Main street proportion (MS Prop)	% Surface	0.28	0.68	0.46	0.46	0.10
Infrastructure						
Sidewalk density (Swk D)	m ² /km	2312.00	7610.00	4702.00	4588.00	1370.84
Junctions per km (Junct)	junctions/km	5.64	10.48	7.59	7.68	1.27
Signals per km (Signals)	signals/km	14.47	83.05	42.95	40.41	16.35
Traffic lights per km (TLights)	lights/km	1.62	21.02	9.16	8.10	4.83
Metro stations per km (Metro)	stations/km	0.00	0.36	0.08	0.04	0.09
Bus stops per km (Bus)	stops/km	0.26	1.71	1.01	1.05	0.39

4. Modelling Results

Table 2 shows the results of the modelling, the level of significance of each independent variable, the Akaike Information Criterion (AIC) and the log-likelihood of the models. Differences between both models have been studied through the statistical significance (*p*-value) of the independent variables and the sign (positive or negative) of the parameters.

Table 2. Results of the Negative Binomial (NB) regression models.

Variable	Model 1: Elderly Pedestrians				Model 1 and Model 2 Comparison		Model 2: Other Pedestrians			
	Estimate	S. Error	z Value		Sign	p-Value	Estimate	S. Error	z Value	
Intercept	-1.49×10^{-1}	1.21	-12.321	***	=	<	-8.15	9.50×10^{-1}	-8.574	***
Exposure										
log (Total street length)	2.00	1.10×10^{-1}	18.199	***	=	<	1.55	9.92×10^{-2}	15.659	***
log (AADT)	5.73×10^{-1}	1.14×10^{-1}	5.032	***	=	<	1.26×10^{-1}	8.08×10^{-2}	1.563	
Socioeconomics										
Inhabitants per km	2.74×10^{-3}	3.20×10^{-4}	8.572	***	=	<	5.80×10^{-5}	2.01×10^{-4}	0.289	
Population density	6.01×10^{-5}	7.48×10^{-6}	8.036	***	=	>	8.33×10^{-5}	7.07×10^{-6}	11.787	***
Average income	-2.42×10^{-5}	2.50×10^{-6}	-9.661	***	=	<	-1.66×10^{-5}	1.92×10^{-6}	-8.624	***
Land use										
POIs per km	2.42×10^{-3}	3.49×10^{-3}	0.693		≠	>	-1.20×10^{-2}	3.42×10^{-3}	-3.500	***
Residential proportion	9.26×10^{-1}	5.43×10^{-1}	1.705	.	≠	>	-1.66	4.47×10^{-1}	-3.705	***
Green area proportion	5.98×10^{-2}	1.54×10^{-1}	0.388		=	>	8.20×10^{-1}	1.44×10^{-1}	5.691	***
Main street proportion	2.04	4.59×10^{-1}	4.436	***	=	>	2.04	3.58×10^{-1}	5.692	***
Infrastructure										
Sidewalk density	-8.26×10^{-5}	1.57×10^{-5}	-5.276	***	=	<	-1.39×10^{-5}	1.38×10^{-5}	-1.010	
Junctions per km	8.15×10^{-2}	2.65×10^{-2}	3.078	**	=	>	2.18×10^{-1}	2.40×10^{-2}	9.078	***
Signals per km	-1.42×10^{-3}	2.92×10^{-3}	-0.488		≠	>	1.17×10^{-2}	2.67×10^{-3}	4.389	***
Traffic lights per km	-2.50×10^{-2}	8.66×10^{-3}	-2.889	**	≠	<	7.73×10^{-3}	8.68×10^{-3}	0.891	
Metro stations per km	3.87×10^{-1}	4.20×10^{-1}	0.920		=	>	2.65	3.55×10^{-1}	7.465	***
Bus stops per km	1.42	1.20×10^{-1}	11.820	***	=	<	9.41×10^{-1}	1.28×10^{-1}	7.367	***
Number of observations (n)	189						189			
AIC	1444.2						1713.6			
Log-likelihood	-705.1						-839.8			

Significance codes (p-value): '***' for 0.1%, '**' for 1%, '*' for 5%, and '.' for 10%.

Two exposure or control variables were considered: total street length and AADT of the district. The first one, total street length, is significant in both models and has been included in the models to distinguish two districts with the rest of the variables being equal. As expected, this indicator has a negative effect on road safety, as higher total street length involves more locations at which to suffer an accident. Despite that, this macroscopic analysis does not consider if the streets “could not be crossed”, understanding the formal allowance of crossing a certain road through the existence of zebra crossing and the real allowance of crossing a certain road based on the width, traffic flow, and type of the road. Ring roads linked to big cities are usually located at peripheral districts, and this regional approach does not allow a deeper level of analysis. This way, each street of a certain district is thought to have the same accident risk exposure, regardless of the differences between them in terms of infrastructure design, and this fact is crucial in heterogeneous districts. The research also did not consider if streets have at least a sidewalk on one of its sides. The second exposure variable, AADT, shows that elderly pedestrian collisions are much more conditioned by AADT than younger-pedestrian collisions are, and it has a positive impact (positive sign) in both models, as it is a measure of accident risk exposure. This result is consistent, since the elderly have a longer reaction time, and a high traffic volume involves the reduction of time between vehicles (i.e., the spatial and temporal window gap to cross a street). Nevertheless, in this study, AADT is estimated as an average value per district, and the study does not consider if this traffic is registered in roads that are not “walkable” (clearly with an urban topology, provided with suitable crossing areas).

Regarding socioeconomic variables, both models show that population density has a positive and significant effect on the number of collisions, with the significance of the ageing rate being higher in model 2. Population density collects how crowded a district is, considering the average height of buildings, and usually it represents a measure of pedestrian accident risk exposure. The number of elderly inhabitants per total street length has a positive effect on the number of elderly pedestrian collisions (model 1), while the number of non-elderly inhabitants per kilometer is not statistically significant for non-elderly casualties (model 2). This fact is consistent with the idea that elderly pedestrians are likely to walk within their neighborhood/district of residence while younger people usually move towards crowded and central areas. Annual average income is an indicator of the economic status of a district, and as it increases, the likelihood of the number of collisions is reduced. Both models reflect this effect, and this variable is significant in both models.

In relation to land use variables, model 2 shows a high effect of POIs per street length, residential proportion, and green area proportion on the occurrence of non-elderly pedestrian collisions. In that model, more residential proportion means less risk for non-elderly pedestrians. In relation to Points of Interest (POIs) per street length, a higher density means less risky areas for non-elderly pedestrians. This result was not a priori expected but deals with the idea that a higher density of POIs indicates more crowded streets where pedestrians walk slower and are more alert to their surroundings. POIs per street length, residential proportion, and green area proportion are statistically insignificant for the occurrence of elderly pedestrian collisions (model 1). Finally, main street proportion is an indicator of the importance of the streets within a certain district. Usually, these streets are wider and have higher traffic flows operating at higher speeds compared with the rest of streets. A higher proportion of these streets represents a higher level of accident risk exposure for pedestrians in both models (regardless of their age).

Concerning infrastructure variables, both models show a high significance of junction density and bus stop density. Sidewalk density has a negative effect on the number of accidents, and this effect is only statistically significant in the case of elderly pedestrians. Sidewalk density is an indicator of the amount of sidewalk available for the pedestrian in relation to total street length. In other words, it represents an average width of the sidewalk, being narrower in central districts and wider in the newer districts. Elderly pedestrians are more sensitive to the state and the width of the sidewalk, as this population group has

a higher risk of falling, stumbling, or tripping while walking compared with younger adults. For the elderly, maintaining postural stability and balance is more difficult [8], and usually too narrow sidewalks incite pedestrians to walk outside the sidewalk [18]. This is a riskier task for the elderly, as inadequate infrastructures can increase elderly pedestrians' anxieties on the perception of the built environment [16,17]. Results also indicate that a higher number of street junctions per kilometer favors the number of collisions in one district for both elderly and younger pedestrians, as this variable represents how complex and intricate a street network is and, in consequence, how many hazards a pedestrian or driver can find for the same total street length. The density of traffic lights significantly improves road safety for elderly pedestrians, while it is statistically insignificant for the rest of pedestrians. The presence of traffic lights implies a lower speed of the traffic flows as vehicles are forced to stop when traffic lights are red. The statistical insignificance of this variable in the non-elderly pedestrian model suggests a higher level of awareness of the elderly. In relation with public transport, bus stop density is an indicator of riskier districts for both elderly and non-elderly pedestrians. The density of bus stops has a positive effect on the number of accidents, as it represents a possible higher pedestrian flow, the presence of large vehicles in the streets, and the existence of critical spots where the elderly can mislead when getting off the bus. Metro station density, which is a measure of generation and attraction of pedestrian trips of a certain zone, has also a positive effect, but it is only statistically significant for non-elderly pedestrians (model 2), reinforcing the idea that non-elderly pedestrians are mobile over wider zones than the surrounding of the district of residence of the pedestrian and that this age group moves towards central areas. In summary, among all the infrastructure variables analyzed, there are two that significantly affect the road safety of the elderly pedestrian: the average sidewalk width and traffic lights per kilometer.

In conclusion, the comparison of the two models reveals that districts with higher total street length, population density, "main street" proportion, street junctions per kilometer, and bus stops per kilometer and lower average income are riskier for pedestrians in general, regardless their age. In particular, the occurrence of elderly pedestrian collisions (model 1), unlike younger pedestrian accidents, is favored by traffic flow (AADT) levels and the population of the age group per kilometer, while these variables are statistically not significant in the occurrence of non-elderly pedestrian casualties. In turn, accident occurrence is less likely with a higher sidewalk density and more traffic lights per kilometer. Non-elderly pedestrian collisions (model 2) are favored by greater signals density and metro stations per kilometer, and are less affected by AADT levels.

The results of this study lead us to think that the general actions aimed at reducing the road accident rate of pedestrians are not going to be enough to drastically decrease the number of fatal and severe elderly pedestrian collisions. There is a scientific consensus on the need to improve street connectivity for all road users, for well-designed sidewalks in all streets, tighter intersection turning radii, bicycle facilities, and the use of two-lane streets whenever possible. However, the results of this paper show that neighborhoods with a higher number of elderly inhabitants per kilometer are going to need specific actions, apart from those previously mentioned. These actions should not only concentrate on "over-marking" (and "over-signaling") crosswalks and intersections in these districts or reinforcing the use of "traffic calming" measures. Pedestrian facilities near transit stations also require special attention in those ageing districts as the interaction with traffic flows is higher and the perception and reaction times of pedestrians increase with the physical deterioration derived from ageing. In conclusion, a special package of countermeasures is needed for the elderly if policymakers aim to achieve the design of more sustainable "age-friendly" cities.

5. Conclusions and Discussion

The current ageing process of the population is a matter of concern in almost every scientific field of knowledge. This paper assessed this issue from the urban road safety

perspective, considering vehicle–pedestrian collisions in Madrid city as a case study. Elderly pedestrians show a higher intrinsic fragility when they suffer an accident due to their physical and mental limitations caused by age-related illnesses and certain chronic medication. These limitations are also connected to a higher crash risk. Previous literature has shown behavior peculiarities of the elderly as pedestrians, such as acting more cautiously because of their maturity or avoiding complex traffic situations. The study of this age group, particularly acting as pedestrians, will be suitable to prepare the built environmental factors and adapt our cities for the future with a higher rate of aged population.

In this research, data per Madrid district were gathered to build two negative binomial regressions in order to estimate the number of vehicle–pedestrian (elderly and non-elderly) collisions considering built environment indicators. For this purpose, an ad hoc dataset containing variables at Madrid district level per year was constructed and treated using the sliding window method with five-year width windows. Mainly primary data were gathered, available directly from official databases that did not require a big amount of postprocessing operations. The dependent variable of the models was the number of vehicle–pedestrian collisions per district during each five-year sliding window for the total studied period, 2006–2018.

The main results show that elderly pedestrian collisions are favored by traffic flow (AADT) levels and the population of the age group per kilometer, while these variables are statistically not significant in the occurrence of non-elderly pedestrian casualties. On the other hand, elderly collisions are less likely with a higher sidewalk density and more traffic lights per kilometer. These results show that countermeasures based only on improving infrastructure (signaling of crosswalks and intersections, implementing crossing islands, and well-designed sidewalks) will help to reduce pedestrian collisions but will not be enough to reduce specifically elderly pedestrian collisions. As the level of road safety risk exposure of the residence district is dependent on population density and AADT as determinant factors, additional actions will have to be implemented specifically in districts with an aged population. Most of them should have the objective to improve the conspicuity of elderly pedestrians for drivers and specifically improve the perception of elderly pedestrians about other road users in these districts. For example, apart from reducing vehicle flow speeds, narrowing street lanes, and widening sidewalks, new traffic signs with elderly pedestrian symbols could be implemented.

Despite the results obtained in this investigation, new research lines emerge, as improvements in the methodology can be achieved. First, the sample size was conditioned by the level of detail of the information available at official data sources but could be widened in the future. Moreover, the study at district level may lead to confusing results because this spatial unit is not found to be homogenous itself, and this problem could be solved by studying the phenomenon using a smaller spatial unit, such as neighborhoods, or using regular grids. Consequently, a more complex and ad hoc data collection process would be needed to feed the new territorial divisions. If using regular grids, instead of neighborhoods, the size of the cells of the grid would be unknown a priori. Thus, a comparative analysis between a set of possible sizes would also be required. Finally, the elderly pedestrian group could be compared to other groups with physical limitations, such as disabled people, to analyze potential common measures to improve road safety of both groups in urban areas.

A second research line can be oriented to study groups of street segments with the same typology and traffic flow. This line represents a major challenge in the already used methodology, and it is a common tool in the study of accident black spots (high-accident-concentration sections) in interurban roads. It is considered to be a correct alternative, as it will be possible to find and classify street segments based on built environment features and the existence or absence of vehicle–pedestrian collisions. The street segments approach would need a deeper definition of variables with an ad hoc measurement of infrastructure indicators (e.g., the sidewalks width). Despite the opening of new research lines, one of the main challenges of the study of urban accidents, the geolocation of the accidents, was

properly solved in this paper by developing an ad hoc procedure. The validation of this procedure in this Madrid case study will be valuable in further investigations, regardless of the research line followed.

The need to obtain a more complete dataset on infrastructure variables is a common issue in future research lines and, as this study has demonstrated, is also a key point (among the rest of built environment variables) to achieve a holistic understanding of elderly pedestrian collisions. Transport policymakers can take actions on the city infrastructure, and the final goal of this research is to define scientifically a set of rules and recommendations to build or redesign sustainable cities considering the ageing process of the population, thus reducing the risk of elderly pedestrian collisions.

Author Contributions: Conceptualization, D.G.-P. and B.G.; methodology, D.G.-P.; software, D.G.-P.; validation, B.G., A.O. and L.P.-S.; formal analysis, D.G.-P., B.G., A.O. and L.P.-S.; investigation, D.G.-P., B.G., A.O. and L.P.-S.; resources, D.G.-P.; data curation, D.G.-P.; writing—original draft preparation, D.G.-P. and B.G.; writing—review and editing, B.G., A.O. and L.P.-S.; visualization, D.G.-P. and A.O.; supervision, B.G., A.O. and L.P.-S.; project administration, B.G. All authors have read and agreed to the published version of the manuscript.

Funding: Daniel Gálvez-Pérez is developing his doctoral thesis while he enjoys a grant from to the Universidad Politécnica de Madrid through the “Programa Propio de I + D + I 2020: Ayudas para Contratos Predoctorales”.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors would like to thank the Dirección General de Tráfico (DGT) for providing the traffic accidents data to develop this study.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

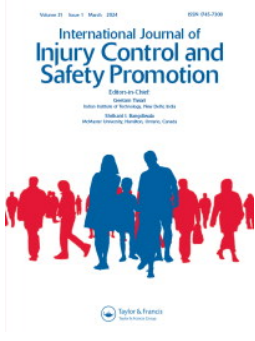
References

1. Population Division—Desa. *World Population Prospects—The 2015 Revision: Key Findings and Advance Tables*; United Nations: New York, NY, USA, 2015.
2. Dugan, E.; Lee, C.M. Biopsychosocial Risk Factors for Driving Cessation: Findings from the Health and Retirement Study. *J. Aging Health* **2013**, *25*, 1313–1328. [[CrossRef](#)] [[PubMed](#)]
3. Musselwhite, C.; Holland, C.; Walker, I. The Role of Transport and Mobility in the Health of Older People. *J. Transp. Health* **2015**, *2*, 1–4. [[CrossRef](#)]
4. Nordbakke, S.; Schwanen, T. Well-Being and Mobility: A Theoretical Framework and Literature Review Focusing on Older People. *Mobilities* **2014**, *9*, 104–129. [[CrossRef](#)]
5. European Commission. *ElderSafe-Risks and Countermeasures for Road Traffic of Elderly in Europe*; No. MOVE/C4/2014-244; European Commission-Directorate-General for Mobility and Transport (DG-MOVE): Brussels, Belgium, 2015.
6. ITF. *Pedestrian Safety, Urban Space and Health*; OECD Publishing: Paris, France, 2012. [[CrossRef](#)]
7. Dunbar, G.; Holland, C.A.; Maylor, E.A. *Older Pedestrians: A Critical Review of the Literature*; Department of Transport: London, UK, 2004.
8. Oxley, J.; Corben, B.; Fildes, B.; O’Hare, M.; Rothengatter, T. Older Vulnerable Road Users—Measures to Reduce Crash and Injury Risk. *Monash Univ. Accid. Res. Cent. Rep.* **2004**, *218*, 162.
9. Palamara, P.; Broughton, M. An Investigation of Pedestrian Crashes at Traffic Intersections in the Perth Central Business. *Regulation* **2013**, *5*, 169–175.
10. Charlton, J.; Oxley, J.; Koppel, S.; Aburumman, M.; Hua, P.; Williams, T. *Older Drivers: A Review of the Literature, Programs and Practices*; Monash University: Melbourne, Australia, 2017.
11. Charlton, J.L.; Oxley, J.; Fildes, B.; Oxley, P.; Newstead, S.; Koppel, S.; O’Hare, M. Characteristics of Older Drivers Who Adopt Self-Regulatory Driving Behaviours. *Transp. Res. Part F Traffic Psychol. Behav.* **2006**, *9*, 363–373. [[CrossRef](#)]
12. Langford, J.; Bohensky, M.; Koppel, S.; Newstead, S. Do Older Drivers Pose a Risk to Other Road Users? *Traffic Inj. Prev.* **2008**, *9*, 181–189. [[CrossRef](#)]

13. Langford, J.; Methorst, R.; Hakamies-Blomqvist, L. Older Drivers Do Not Have a High Crash Risk—A Replication of Low Mileage Bias. *Accid. Anal. Prev.* **2006**, *38*, 574–578. [[CrossRef](#)]
14. Henary, B.Y.; Ivarsson, J.; Crandall, J.R. The Influence of Age on the Morbidity and Mortality of Pedestrian Victims. *Traffic Inj. Prev.* **2006**, *7*, 182–190. [[CrossRef](#)]
15. Bernhoft, I.M.; Carstensen, G. Preferences and Behaviour of Pedestrians and Cyclists by Age and Gender. *Transp. Res. Part F Traffic Psychol. Behav.* **2008**, *11*, 83–95. [[CrossRef](#)]
16. Clarke, P.; Gallagher, N.A. Optimizing Mobility in Later Life: The Role of the Urban Built Environment for Older Adults Aging in Place. *J. Urban Health* **2013**, *90*, 997–1009. [[CrossRef](#)]
17. Phillips, J.; Walford, N.; Hockey, A.; Foreman, N.; Lewis, M. Older People and Outdoor Environments: Pedestrian Anxieties and Barriers in the Use of Familiar and Unfamiliar Spaces. *Geoforum* **2013**, *47*, 113–124. [[CrossRef](#)]
18. Corazza, M.V.; Di Mascio, P.; Moretti, L. Management of Sidewalk Maintenance to Improve Walking Comfort for Senior Citizens. *WIT Trans. Built Environ.* **2017**, *176*, 195–206.
19. Ståhl, A.; Carlsson, G.; Hovbrandt, P.; Iwarsson, S. “Let’s Go for a Walk!”: Identification and Prioritisation of Accessibility and Safety Measures Involving Elderly People in a Residential Area. *Eur. J. Ageing* **2008**, *5*, 265–273. [[CrossRef](#)]
20. Dirección General de Tráfico (DGT). *Las Principales Cifras de La Siniestralidad Vial. España 2019*; Dirección General de Tráfico: Madrid, Spain, 2020.
21. Gálvez-Pérez, D.; Guirao, B.; Ortuño, A. Road Safety of Elderly Pedestrians in the Urban Context: An Approach Based on Infrastructure and Socioeconomic Variables. *Transp. Res. Procedia* **2021**, *58*, 254–261. [[CrossRef](#)]
22. Cervero, R.; Sarmiento, O.L.; Jacoby, E.; Gomez, L.F.; Neiman, A. Influences of Built Environments on Walking and Cycling: Lessons from Bogotá. *Int. J. Sustain. Transp.* **2009**, *3*, 203–226. [[CrossRef](#)]
23. Khan, M.; Kockelman, K.M.; Xiong, X. Models for Anticipating Non-Motorized Travel Choices, and the Role of the Built Environment. *Transp. Policy* **2014**, *35*, 117–126. [[CrossRef](#)]
24. Orellana, D.; Guerrero, M.L. Exploring the Influence of Road Network Structure on the Spatial Behaviour of Cyclists Using Crowdsourced Data. *Environ. Plan. B Urban Anal. City Sci.* **2019**, *46*, 1314–1330. [[CrossRef](#)]
25. Talavera-García, R.; Soria-Lara, J.A. Q-PLOS, Developing an Alternative Walking Index. A Method Based on Urban Design Quality. *Cities* **2015**, *45*, 7–17.
26. Sevtsuk, A.; Basu, R.; Li, X.; Kalvo, R. A Big Data Approach to Understanding Pedestrian Route Choice Preferences: Evidence from San Francisco. *Travel Behav. Soc.* **2021**, *25*, 41–51. [[CrossRef](#)]
27. Stoker, P.; Garfinkel-Castro, A.; Khayesi, M.; Odero, W.; Mwangi, M.N.; Peden, M.; Ewing, R. Pedestrian Safety and the Built Environment: A Review of the Risk Factors. *J. Plan. Lit.* **2015**, *30*, 377–392. [[CrossRef](#)]
28. Tiwari, G. Progress in Pedestrian Safety Research. *Int. J. Inj. Control Saf. Promot.* **2020**, *27*, 35–43. [[CrossRef](#)]
29. Culver, G. Death and the Car: On (Auto) Mobility, Violence, and Injustice. *ACME Int. J. Crit. Geogr.* **2018**, *17*, 144–170.
30. Jackson, S.; Miranda-Moreno, L.F.; St-Aubin, P.; Saunier, N. Flexible, Mobile Video Camera System and Open Source Video Analysis Software for Road Safety and Behavioral Analysis. *Transp. Res. Rec.* **2013**, *2365*, 90–98. [[CrossRef](#)]
31. Teknomo, K. Application of Microscopic Pedestrian Simulation Model. *Transp. Res. Part F Traffic Psychol. Behav.* **2006**, *9*, 15–27. [[CrossRef](#)]
32. Sandels, S. Young Children in Traffic. *Br. J. Educ. Psychol.* **1970**, *40*, 111–116. [[CrossRef](#)]
33. Clifton, K.J.; Burnier, C.V.; Akar, G. Severity of Injury Resulting from Pedestrian–Vehicle Crashes: What Can We Learn from Examining the Built Environment? *Transp. Res. Part D Transp. Environ.* **2009**, *14*, 425–436. [[CrossRef](#)]
34. World Health Organization. *Global Age-Friendly Cities: A Guide: World Health Organization*; World Health Organization: Geneva, Switzerland, 2007.
35. Demasi, F.; Loprencipe, G.; Moretti, L. Road Safety Analysis of Urban Roads: Case Study of an Italian Municipality. *Safety* **2018**, *4*, 58. [[CrossRef](#)]
36. Galanis, A.; Botzoris, G.; Eliou, N. Pedestrian Road Safety in Relation to Urban Road Type and Traffic Flow. *Transp. Res. Procedia* **2017**, *24*, 220–227. [[CrossRef](#)]
37. Kim, D. The Transportation Safety of Elderly Pedestrians: Modeling Contributing Factors to Elderly Pedestrian Collisions. *Accid. Anal. Prev.* **2019**, *131*, 268–274. [[CrossRef](#)]
38. Lv, M.; Wang, N.; Yao, S.; Wu, J.; Fang, L. Towards Healthy Aging: Influence of the Built Environment on Elderly Pedestrian Safety at the Micro-Level. *Int. J. Environ. Res. Public Health* **2021**, *18*, 9534. [[CrossRef](#)] [[PubMed](#)]
39. Koepsell, T.; McCloskey, L.; Wolf, M.; Moudon, A.V.; Buchner, D.; Kraus, J.; Patterson, M. Crosswalk Markings and the Risk of Pedestrian–Motor Vehicle Collisions in Older Pedestrians. *JAMA* **2002**, *288*, 2136–2143. [[CrossRef](#)] [[PubMed](#)]
40. Kubitzki, J.; Janitzek, T. *Safety and Mobility of Older Road Users*; Allianz Deutschland AG: Munich, Germany, 2009.
41. Feypell, V.; Methorst, R.; Hughes, T. Non-Motor Pedestrian Accidents: A Hidden Issue, Results of the ITF/OECD Report on Pedestrian Safety, Urban Space and Health. Available online: https://www.ictct.net/wp-content/uploads/23-Hague-2010/ictct_document_nr_740.pdf (accessed on 1 September 2021).
42. Wedagama, D.M.P.; Bird, R.N.; Metcalfe, A.V. The Influence of Urban Land-Use on Non-Motorised Transport Casualties. *Accid. Anal. Prev.* **2006**, *38*, 1049–1057. [[CrossRef](#)] [[PubMed](#)]
43. Wier, M.; Weintraub, J.; Humphreys, E.H.; Seto, E.; Bhatia, R. An Area-Level Model of Vehicle-Pedestrian Injury Collisions with Implications for Land Use and Transportation Planning. *Accid. Anal. Prev.* **2009**, *41*, 137–145. [[CrossRef](#)]

44. Ukkusuri, S.; Miranda-Moreno, L.F.; Ramadurai, G.; Isa-Tavarez, J. The Role of Built Environment on Pedestrian Crash Frequency. *Saf. Sci.* **2012**, *50*, 1141–1151. [[CrossRef](#)]
45. Apardian, R.E.; Smirnov, O. An Analysis of Pedestrian Crashes Using a Spatial Count Data Model. *Pap. Reg. Sci.* **2020**, *99*, 1317–1338. [[CrossRef](#)]
46. Lee, S.; Yoon, J.; Woo, A. Does Elderly Safety Matter? Associations between Built Environments and Pedestrian Crashes in Seoul, Korea. *Accid. Anal. Prev.* **2020**, *144*, 105621.
47. INE. INEbase/Demography and Population. 2019. Available online: <https://www.ine.es/dyngs//INEbase/en/listaoperaciones.htm> (accessed on 1 September 2021).
48. Schmid, K. *World Population Ageing 2019*; Department of Economic and Social Affairs Population Division: New York, NY, USA, 2020.
49. Casado-Sanz, N.; Guirao, B.; Gálvez-Pérez, D. Population Ageing and Rural Road Accidents: Analysis of Accident Severity in Traffic Crashes with Older Pedestrians on Spanish Crosstown Roads. *Res. Transp. Bus. Manag.* **2019**, *30*, 100377. [[CrossRef](#)]
50. Casado-Sanz, N.; Guirao, B.; Lara Galera, A.; Attard, M. Investigating the Risk Factors Associated with the Severity of the Pedestrians Injured on Spanish Crosstown Roads. *Sustainability* **2019**, *11*, 5194. [[CrossRef](#)]
51. Ayuntamiento de Madrid. *Plan General de Ordenación Urbana de Madrid. PGOU-M-97*; Ayuntamiento de Madrid: Madrid, Spain, 1997.
52. Ziakopoulos, A.; Yannis, G. A Review of Spatial Approaches in Road Safety. *Accid. Anal. Prev.* **2020**, *135*, 105323. [[CrossRef](#)]
53. Hilbe, J.M. *Negative Binomial Regression*; Cambridge University Press: Cambridge, UK, 2011; ISBN 0521198151.
54. Mozaffari, L.; Mozaffari, A.; Azad, N.L. Vehicle Speed Prediction via a Sliding-Window Time Series Analysis and an Evolutionary Least Learning Machine: A Case Study on San Francisco Urban Roads. *Eng. Sci. Technol. Int. J.* **2015**, *18*, 150–162. [[CrossRef](#)]
55. Fitzpatrick, K.; Avelar, R.E.; Turner, S.M. *Identification of High Pedestrian Crash Locations*; Technical Report HWA-HRT-17-108; U.S. Department of Transportation, Federal Highway Administration: McLean, VA, USA, 2018.

4.3 *Article B: Analysis of the elderly pedestrian traffic accidents in urban scenarios: the case of the Spanish municipalities*



Analysis of the elderly pedestrian traffic accidents in urban scenarios: the case of the Spanish municipalities

Daniel Gálvez-Pérez, Begoña Guirao & Armando Ortuño

To cite this article: Daniel Gálvez-Pérez, Begoña Guirao & Armando Ortuño (22 Apr 2024): Analysis of the elderly pedestrian traffic accidents in urban scenarios: the case of the Spanish municipalities, International Journal of Injury Control and Safety Promotion, DOI: [10.1080/17457300.2024.2335482](https://doi.org/10.1080/17457300.2024.2335482)

To link to this article: <https://doi.org/10.1080/17457300.2024.2335482>



Published online: 22 Apr 2024.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Analysis of the elderly pedestrian traffic accidents in urban scenarios: the case of the Spanish municipalities

Daniel Gálvez-Pérez^a, Begoña Guirao^a  and Armando Ortuño^b

^aIngeniería del Transporte, Territorio y Urbanismo, Universidad Politécnica de Madrid, Madrid, Spain; ^bIngeniería Civil, Universidad de Alicante, Alicante, Spain

ABSTRACT

As the elderly population grows, there is a greater concern for their safety on the roads. This is particularly important for elderly pedestrians who are more vulnerable to accidents. In Spain, one of the most aged countries in the world, the elderly accounted for 70% of all pedestrian deaths in 2019. In this study, the focus was on analysing the occurrence of elderly pedestrian-vehicle collisions in Spanish municipalities and how it is related to the built environment. The study used the hurdle negative binomial model to analyse the number of elderly and non-elderly pedestrian accidents per municipality in 2016-2019. The exploratory analysis showed that cities above 50,000 inhabitants were safer for the elderly, and larger provincial capitals had lower elderly pedestrian traffic accident rates. The occurrence of all pedestrian traffic accidents was linked to the socio-demographic features. For elderly pedestrians, land use was found to be influential, with a lower proportion of land covered by manufacturing and service activities linked to a smaller number of accidents. Results showed that improving road safety for older pedestrians may not necessarily compromise the situation for the rest of population. Hence, policymakers should focus on infrastructure improvements adapted to the needs of elderly pedestrians.

ARTICLE HISTORY

Received 24 April 2023
Revised 4 January 2024
Accepted 23 March 2024

KEYWORDS

Elderly pedestrians; road safety; road traffic collisions; accident analysis; built environment

Introduction

The ageing process of the population has a direct impact on road safety statistics. In Europe, 17% of people who suffered a road accident in 1992 and died were over the age of 65, and this figure reached 29% in 2018. Elderly road users represent 50% and 46% of pedestrians and cyclists killed, respectively (European Commission, 2021b). Regarding the location of the elderly pedestrian fatalities, 53% of them were on urban roads, while 42% and 5% were in rural roads and motorway, respectively (European Commission, 2021a). Particularly in Spain, one of the most aged countries in the world, the proportion of people over 65 years old was 19% in 2019, but the elderly pedestrian fatalities on urban roads represented almost 70%, over the total pedestrian fatalities (DGT, 2020). A traffic crash occurs at a geographical location, often characterized by specific built environment features. Built environment is the physical surroundings of the traffic crash. Understanding the correlation between the crash site and its associated built environment is a key issue to understand why elderly pedestrians are leading the figures of the urban road traffic victims.

The analysis of accidents from a territorial approach is not new within the road safety literature, as the location of the accident within the territory is crucial to understand built environmental factors affecting road safety. Basically, two types of studies can be found in the literature: 'macro-scale level' and 'microscale level' studies (Ziakopoulos & Yannis, 2020). The first one analyses traffic accidents by

their data aggregation in territorial units, while the 'microscale level' is reserved, in terms of road safety, to the accident analysis at homogeneous sections of infrastructure level (intersections, street crossing, or street sections).

The ageing process of the population and the fragility of the elderly create the need to study the road safety of elderly pedestrians at territorial level (both at macro-level and micro-level) with special attention. Previous macro-level studies have analysed the relationship between road safety and built environment at various spatial units of analysis, including country, regional, municipal, and intra-municipal levels. **Country studies**, usually conducted by international organizations, aim to compare road safety indicators across countries. On the other hand, there are several **regional studies** developed by researchers (Aguero-Valverde, 2013; Castro-Nuño et al., 2018), focused on large administrative divisions. These studies have limitations, such as combining urban and interurban accidents in the same group, and the potential high impact that the larger cities have on results, as they register remarkable proportions of both traffic accidents and population size. **Municipal studies** (Cabrera-Arnau & Bishop, 2021; Lee et al., 2018; Santos & Carvalho, 2019) use municipal boundaries (cities/towns) as spatial unit of analysis and mitigate some of the issues of regional studies, although researchers often focus on the largest cities within country and do not pay attention to road safety situation in small cities. **Intra-municipal studies** (Aparidian & Smirnov, 2020; Cai et al., 2016; Yang et al., 2021, Gálvez-Pérez

et al., 2022) use spatial units of analysis smaller than a city, including administrative (census tracts, districts, neighbourhoods, etc.) or non-administrative units (regular grids). These studies offer high levels of detail but are limited in scope, often examining only a single city; and, in consequence, introducing bias in the data as the studied city may present very similar features in all the intra-municipal units, like public transport presence. Moreover, this methodology would not be feasible for examining road safety throughout an entire nation due to data availability at the required level of detail and geographical extension. Given the exposed limitations of the different levels of analysis and the availability of data, this study focuses on the **municipal level** and uses all **Spanish municipalities**, not just big cities, as the territorial units of study. In the literature (Cabrera-Arnau et al., 2020; Quistberg et al., 2022), studies at municipal level have shown that big cities seem to be safer than small cities; and rural cities seem to be riskier than metropolitan cities. These findings should be tested specifically for the occurrence of elderly pedestrian traffic accidents.

The objective of this paper is to evaluate the role of urban built environment on the frequency of elderly pedestrian crashes (versus non-elderly pedestrian crashes) at municipal level. Two hurdle negative binomial (HNB) models were applied to a database including 8,028 municipalities (ranging from small to large cities) as case study and aim to compare elderly pedestrian (EP model) versus non-elderly pedestrian (NEP model) accidents. The database was created by extracting pedestrian traffic accidents from the Spanish database for the period 2016-2019. Accidents involving one pedestrian and one vehicle (45,098 accidents) were used in this investigation, with 28.4% of these accidents suffered by elderly pedestrians. Regarding built environment features for each municipality, a set of socio-demographic, land use and infrastructure were gathered from different multiple data sources through a systematic ad hoc procedure. After a preliminary analysis, 15 independent variables were chosen to develop an exhaustive model selection procedure by computing all possible configurations of these independent variables to build different HNB regression models, summing 32,768 EP models and 32,768 NEP models. This process allowed us to identify which independent variables were good predictors, regardless of the presence of other covariates in the model. The impact of each independent variable in the EP and NEP models were evaluated through a multi-criteria analysis considering the parameter sign, p-value and the goodness-of-fit of each model configuration.

The paper is divided into the following sections: Section 1 contains the introduction; Section 2 presents the state of the art on territorial analysis of road safety at different scale levels; Section 3 gives a detailed description of the urban pedestrian accidents distributed among the Spanish municipalities as case study. Section 4 explains the database elaboration and analysis, statistical model selection and structure, and modelling results. Finally, Section 5 presents the conclusions and future research lines drawn from this research.

State of the art

There is a consensus on the importance of the **traffic crashes location** as a variable affecting the occurrence and severity of accidents, which can be considered as a built environment factor for road safety (Stoker et al., 2015). The location of the pedestrian accidents is associated to urban scenarios and the built environment variables can be considered at different macro-scale territorial levels (country, region, city or city district levels) or at micro-scale level (similar sections or junctions of street infrastructure). There is extensive literature on road safety macroscale studies (Ziakopoulos & Yannis, 2020) at spatial units which can range from a county level to a small city administrative division or a traffic analysis zone (TAZ). In this paper, a literature review on macroscale road safety studies have been analysed in order to examine their main results and the built environment factors considered.

Regarding pedestrian macroscale studies **at country level**, they are usually developed by international organizations with the aim of comparing the road safety situation of different countries (International Transport Forum (ITF), 2020; World Health Organization (WHO), 2018). In contrast, there are many studies on road safety at **regional (provincial, county, etc.) level** developed by academics, analysing different traffic accident data including all types of accidents (Castillo-Manzano et al., 2020; Castro-Nuño et al., 2018; Castro-Nuño & Arévalo-Quijada, 2018) or accidents by level of injury (Aguero-Valverde, 2013; Li et al., 2013), but only a few of them are focused pedestrian crashes (Lee et al., 2019). These studies normally use aggregated available socio-demographic and infrastructure data, as the existent level of detail of each spatial unit is very small to assess the effect on road safety. Population density is one of built environmental variables most frequently used to analyse road safety in this type of regional analysis. For example, the study developed by Castro-Nuño et al. (2018) proved that population density, health services improvements together with accessibility and sustainability of urban transportation were key factors for traffic crashes occurrence at province level. Results obtained by Castro-Nuño & Arévalo-Quijada (2018) showed that safer regions also depended on higher urban population, higher services concentration and more advanced road networks and transport systems. Tourism has also been revealed as a factor affecting road safety and some authors (Castillo-Manzano et al., 2020) have found a correlation between foreign tourism and increased traffic accidents in Spanish provinces. Higher volumes of foreign overnight stays were linked to higher rates of both urban and interurban traffic accidents.

From an accident severity approach, Aguero-Valverde (2013) analysed regional level crashes by injury severity in Costa Rica at canton macro-scale (urban and interurban scale), finding a correlation between all crashes and vehicle-kilometres and urban population. The population age groups were significant factors affecting fatal and injury collisions (0-14 years group) and injury and property damage only collisions (15-24 group, and over 65 years group). Likewise, in the U.S., Li et al. (2013) studied fatal crashes in

California at county level, obtaining that freeway mileage, urban traffic, road density, and median household income had a negative effect on the occurrence of fatal accidents, while population density and youth population showed a positive effect.

In conclusion, although these regional studies do not focus only on pedestrians, there are some variables, such as **population density, population age distribution, median household income, and public transport system**, which are revealed as key factors that indirectly may also affect to the occurrence of elderly pedestrian accidents. These studies have some limitations that should be considered, as these analyses normally focus on traffic accidents regardless of their location, so accidents on urban and interurban roads are treated equally. In addition, as large regions are studied, accidents in both large cities and small towns are grouped together and analysed as one unit, and their features are analysed following the same approach. This fact means that, although heterogeneity across areas could be evaluated, the influence of the largest cities within the study territorial units may significantly impact the analysis results.

Down to a lower territorial unit, **like the municipal or metropolitan area level**, literature show many studies to evaluate all type of traffic collisions (Cabrera-Arnau et al., 2020; Cabrera-Arnau & Bishop, 2021), fatal crashes (Quistberg et al., 2022) and pedestrian crashes (Lee, Abdel-Aty, Huang, et al., 2019; Lee et al., 2018; Santos & Carvalheira, 2019). Cabrera-Arnau et al. (2020) found that urban areas (identified with towns with more than 10,000 inhabitants) have more slight and serious accidents, while fatal accidents are more common in rural areas (less than 10,000 inhabitants) in England and Wales. Their study showed that areas with higher population size are prone to crashes, but these locations have a higher probability of less severe accidents. The authors stated that this result could be caused by an increase in population size leading to an increase in traffic congestion and drivers' stress level. In a subsequent study, Cabrera-Arnau & Bishop (2021) studied all road accidents of 1 year of 300 European cities over 50,000 inhabitants and found no significant change in the number of crashes per person among these cities. However, this rate was remarkably higher than the national average in some cases, which could be due to omitted variables in the study, such as traffic volume, traffic congestion and proximity to a port. In Portugal, Santos & Carvalheira (2019) developed a pedestrian road accident index for Portuguese municipalities, based on population growth rate and pedestrian crashes. The index showed that municipalities with higher population growth rates, generally located in more consolidated urban areas and along the coastline were safer for pedestrians, while less consolidated areas resulted less safe and needed more urgent road safety interventions.

Out of the European context, and in relation to pedestrian accident analysis at metropolitan area level, Lee et al. (2018) found a positive relationship between fatal pedestrian and bicycle crashes in U.S. metropolitan areas and the expected pedestrian trips, the proportions of black and Hispanic householders, and below poverty level. A negative relationship was found with the proportion of open space.

Employing similar data, Lee, Abdel-Aty, Huang, et al. (2019) proposed a methodology for calculating walking exposure and pedestrian road safety, finding that the best measure of walking exposure was walking hours, which was influenced by population, climate, and vehicle usage. Higher pedestrian fatalities were related to higher walking hours, mid-elderly population (64-75), minority ethnicity, and lower land occupation of tertiary industry. Regarding road traffic deaths, Quistberg et al. (2022) studied the relationship between mortal road crashes and the built environment in 366 Latin American cities (over 100,000 inhabitants), finding that accident mortality was higher in cities with isolated urban development and lower in cities with higher population density, gross domestic product per capita, higher intersection density, and mass transit systems. Authors defined city 'isolation' as the measured surface-weighted mean of the nearest neighbour distance between built-up patches. They recommended avoiding disconnected development and promoting walkable streets and public transportation to reduce road fatalities.

According to the studies at municipal level, big cities seem to be safer than small cities, and rural cities seem to be riskier than metropolitan cities for all kind of road users, but the contribution of population size and rural areas to specifically elderly pedestrian accidents remains still unexplored. The municipal-level approach offers the advantage of being able to examine different towns and cities, allowing researchers to evaluate the influence of factors that are prevalent throughout a municipality. However, in the revised literature, special attention is devoted to big cities (which frequently share characteristics such as public transportation, well-developed urban planning, and strong presence of tertiary economic activities), while the road safety behaviour of the smallest towns remains unexplored.

The study of road accident accidents at **intra-municipal level is the subsequent level of analysis in this state of the art**. There is a significant body of literature utilizing smaller territorial units than a city to study exclusively pedestrian road safety, with two main approaches: using the city administrative territorial divisions or designing ad hoc regular grids. In the first approach, spatial units such as the census tracts, census block groups, traffic analysis zones (TAZs), neighbourhoods and districts of a single city have been used, as socio-demographic data are usually provided based on administrative criteria (Aparidian & Smirnov, 2020; Cai et al., 2016; Casares Blanco et al., 2019; Gálvez-Pérez et al., 2022; Lee et al., 2020; Marshall & Ferenchak, 2017; Yang et al., 2021).

Using census tracts as the spatial unit of analysis, Aparidian & Smirnov (2020) found in Toledo (Ohio, U.S.) higher pedestrian accident occurrence in census tracts with more percentages of black residents and total population, and lower crash frequencies in areas with higher home ownership, populations over 65 and under 5, and higher median house values. Using a smaller unit of analysis, the census block groups, Lee et al. (2020) studied elderly pedestrian accidents in Seoul at the census block group level and found that traffic regulators had a significant effect on elderly pedestrian accidents in low-priced housing areas. They

suggested that this result was due to an uneven distribution of road safety measures for pedestrians. Cai et al. (2016) and Yang et al. (2021) used traffic analysis zones (TAZs) to predict pedestrian and bicycle accidents in Florida (U.S.) and traffic accidents in general in Suzhou Industrial Park (China), respectively, observing the impact of land use, education and service institutions, and traffic lights on accidents.

Marshall & Ferenchak (2017) studied the fatality rate of vehicle occupant and pedestrian/bicyclist across the U.S. using the zip code as the spatial unit of analysis. They analysed the value of these two indicators depending on the location of the accident (urban or rural), population density, household income and race/ethnicity. Results showed that higher fatality rates were found in rural and sparse areas for vehicle occupants, while this figure was similar for pedestrians and bicyclists in rural and urban areas and in dense and sparse areas. In Spain, Casares Blanco et al. (2019) used custom non-regular cells, similar to city neighbourhoods, in Benidorm (Spain) and found that mixed land uses improved overall traffic safety, while higher tourist accommodations were linked to safer pedestrian areas. More recently, Gálvez-Pérez et al. (2022) analysed vehicle-elderly pedestrian collisions in Madrid at the city district level, finding that AADT, traffic lights, sidewalk surface, and elderly population per street length, showed a more statistically significant impact on elderly pedestrian collisions (compared to the rest of pedestrian collisions).

Following the second approach, authors such as Asadi et al. (2022), Cai et al. (2017), Kim et al. (2006), and Xie et al. (2017) have analysed traffic accidents at an intra-municipal level using custom regular spatial units, rather than administrative divisions. For instance, Asadi et al. (2022) analysed vehicle-bicycle and vehicle-vehicle accidents causing property damage, fatal, and severe injuries using 100×100 m cells in several Dutch municipalities, *via* a hurdle negative binomial regression model to handle data with excess of zeros (cells with zero accidents). Although this second approach could provide the smallest territorial unit, this methodology is not so widely used, as it requires more available official disaggregated data and more data processing.

In conclusion, the intra-municipal level approach shows some disadvantages related to the heterogeneity of the studied features, and research should focus on sufficiently heterogeneous intra-municipal units. The study of a single municipality, even if it is highly diverse, can introduce bias in the data, as the city can present fixed characteristics of the population, land use, or infrastructure (e.g. the existence of public transport, as normally a city presents this feature in all its zones). In addition, there might be some features only available at the municipal level, such as the number of registered vehicles; and if the study only includes one municipality, their effects cannot be analysed, as there is only one value available. This intra-municipal territorial level is practically non-approachable when results of all the territorial extension of a country are expected.

This paper aims to analyse the road safety data of elderly pedestrians and their relationship with the characteristics of the accident location at the municipal level of detail,

considering that the municipality belong to the same country. Despite the fact that many studies have explored the impact of macroscale built environment features on urban road accidents, **relatively few have focused specifically on pedestrian accidents, and even fewer on elderly pedestrian accidents.** Elderly pedestrian studies at country, regional and municipal level are practically non-existent, and the only pedestrian studies at municipal spatial level reveal the important role played by the variable of population density, population age distribution, median household income, and public transport system. Furthermore, if the municipal level of detail is to be used, a large set of municipalities should be studied to have enough sample size and remarkably different locations not to introduce bias in the analysis, and to avoid studying only large cities. In this paper, all the Spanish municipalities were included in the case study.

Next section shows elderly pedestrian accidents data in Spain (2016-2019) and their distribution among all Spanish municipalities. As population size and density have been revealed in the literature being two key variables to influence road safety, next chapter also includes an index of collisions per population density based on the population size. In addition, in order to analyse the behaviour of the cities playing an administrative role, a ranking of the elderly pedestrian road safety in the Spanish provincial capitals is also presented in the subsequent section.

Case study and database: Spanish municipalities

Spain has been selected as the case study as it is one of the most aged countries in the world, together with Japan, Italy, Finland, Portugal, or Greece (Population Reference Bureau PRB, 2019). In 2021, 19.8% of the population was older than 65 in Spain (Instituto Nacional de Estadística INE, 2022); and this statistic is expected to reach 36.8% in 2050 (United Nations, Department of Economic and Social Affairs, Population Division, 2020). This country is the only European country among the world top 10 list with the highest expected increase in the ageing rate between 2019 and 2050 (United Nations, Department of Economic and Social Affairs, Population Division, 2020). The ageing process of the population, similarly to other top ageing countries, has had a direct impact on road safety statistics and these accident official figures already showed that 70% of the total pedestrian fatalities on urban Spanish roads in 2019 were elderly.

In order to study which urban features make a city safer for elderly pedestrians, a database of urban pedestrian accidents (elderly and non-elderly victims) per Spanish municipality has been built using and filtering the official Spanish accident base (provided by DGT). The Spanish database of traffic accidents is a relational database that contains information on accidents (location, date), pedestrians, drivers, and passengers involved, and vehicles. The municipality where each accident was registered is a variable collected in the database (location), so a traffic accident frequency analysis can be developed using the municipality as the spatial unit of study. To avoid data homogenization processes, in this study only data provided with the last data structure design of the

DGT database have been used. Consequently, the study period included 4 years and ranged from 2016 to 2019, the last year available suitable for the study, as 2020 was not a representative year for mobility due to the COVID-19 pandemic.

The aggregated data show that in Spain 51,068 vehicle-pedestrian collisions were registered on **urban roads** between 2016 and 2019, and almost 90% of them (45,098) had a 'single vehicle-pedestrian configuration' (1 vehicle and 1 pedestrian involved). This configuration was the only one considered in this study, as it is the most common and at the same time allows to assign the age of the pedestrian victim to each collision in order to group the accidents of elderly pedestrians per municipality. According to this data consideration, Spanish elderly pedestrians (65 years or older) have suffered, for this period, 28.4% of the accidents (see Table 1). The total number of victims, considering the pedestrian who suffered the collision as the victim, coincides with the number of accidents, 45,098 victims, and these victims were grouped, according to the severity into fatalities (1.9%), seriously injured (12.0%), slightly injured (85.1%), and unharmed (1.0%). Data shows that although older pedestrians account for 28.4% of accidents, they constitute 71.1% and 41.3% of total deaths and serious injuries, respectively, which reveals the importance of studying this age group. As the previous figures represent aggregated data for all Spanish municipalities, it is important to analyse the role that the main urban features can play in the road safety of elderly pedestrians.

Among these urban features, the literature review has shown that demographic features like **population size** and **population density** are key performance indicators to analyse the occurrence of urban road accidents. Lower crash rates (all types of road accidents) are associated with higher population densities and highly consolidated urban development (Castro-Nuño et al., 2018; Quistberg et al., 2022). As larger cities are usually associated with higher population densities and more consolidated urban areas, highly populated cities are supposed to be safer from a road safety criterion, although this hypothesis has not been widely assessed for pedestrian road safety and even less for elderly pedestrians. For this reason, as part of the exploratory analysis of our case study, especial attention has been paid to demographic variables and their link to elderly pedestrian collisions.

Spain had a heterogeneous spatial distribution of the population (47 million inhabitants in total), allocated in 8,131 different municipalities in 2019. Most of these municipalities (5,877) have less than 2,000 inhabitants, summing up 3 million inhabitants (5.7% of the total population). There are 1,501 (18%) municipalities, with a size that ranges from 2,000 to 10,000 inhabitants, almost 7 million (14.6%) inhabitants in total. Finally, there are 753 (9%) cities with more than 10,000 inhabitants that contain 79.7% of the

population of the country, which means that almost 80% of the country's population lives in cities with more than 10,000 inhabitants. The relation between ageing rates and population size in Spain is proportionally inverse (see Figure 1), and small municipalities are the most aged on average. However, the distribution of the absolute value of the elderly population shows that there are more elderly inhabitants in the two biggest cities, Madrid and Barcelona, than in the 5,877 smallest municipalities. In those small towns, the ageing ratio is 28.5%, while in large cities (above 500,000 inhabitants) it is 18.5% (Pérez Díaz et al. 2020). The study of the ageing rate distribution indicates the need of normalizing the number of elderly pedestrian accidents by the elderly population instead of the total population, as the proportion of elderly inhabitants is not constant.

As part of this analysis, a comparison between the number of pedestrian collisions in groups of cities, according to their population density as a control variable, is necessary. At this point, a database with the number of elderly, non-elderly and total pedestrian accidents in each municipality was already created. With the population size and total surface of each municipality, an average population density for each group of municipalities by population size was calculated as the total population over the total surface of each set of locations. The indicator of the number of elderly, non-elderly and total pedestrian traffic accidents over the average population density and the average ageing rate of each group of municipalities, considering the total population size, is shown in Figure 1. Larger cities seem to be less risky for pedestrians in general, and this result is consistent with previous literature (Castro-Nuño et al., 2018; Quistberg et al., 2022). The elderly pedestrian accident indicator remains constant for municipalities from 101 to 20,000 inhabitants, while the rest of pedestrian collision indicators have an increasing tendency in that range. The following group of cities, ranging from 20,001 to 50,000 inhabitants, is the most dangerous for all pedestrians, and the subsequent groups of municipalities with more than 50,000 inhabitants register a decreasing tendency of the indicators for elderly, non-elderly and all pedestrian accidents. Consequently, cities with more than 500,000 inhabitants are less risky for all pedestrians in Spain, although other additional variables, indirectly linked to population size, should be considered (like the presence of urban public transport or the condition of 'province capital'). For example, in Spain, municipalities with more than 50,000 inhabitants are required to have **public urban** passenger transportation services (de España, 1985), which facilitates the transport of people throughout the city, reducing the risk of suffering an accident as a pedestrian (Quistberg et al., 2022). This effect is also captured by the figures of elderly pedestrian road safety, making cities greater than 50,000 inhabitants safer for

Table 1. Annual evolution of the number of vehicle-pedestrian collisions in Spain between 2016 and 2019.

Age of the pedestrian	2016	2017	2018	2019	Total
Elderly (65+)	3,316	3,191	3,072	3,210	12,789
Non-elderly	8,030	7,553	7,537	7,627	30,747
Unknown	369	491	373	329	1,562
Total	11,715	11,235	10,982	11,166	45,098

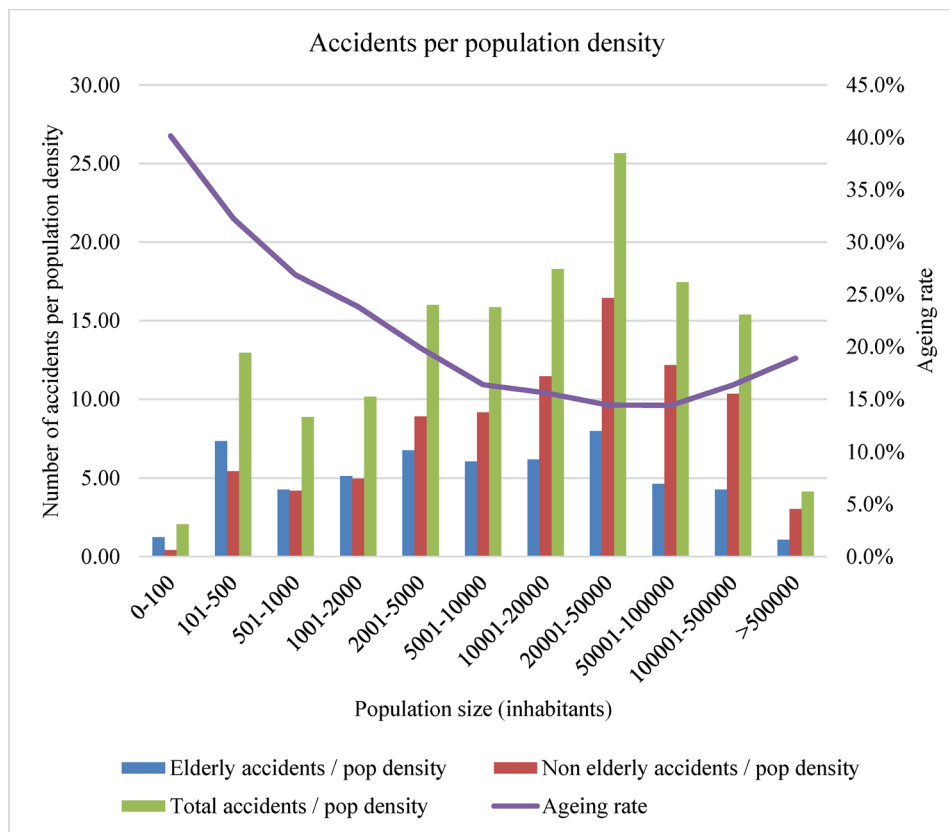


Figure 1. Number of elderly, non-elderly and total pedestrian traffic accidents over population density considering the population size of the municipality in Spain between 2016 and 2019.

them. **Figure 1** also shows the evolution of the ageing rate with the increase of population size, revealing that smaller rural (under 50,000 inhabitants) municipalities register higher ageing rates.

Provincial capitals are usually large cities with a developed and consolidated urban scheme that concentrate the Public Administrative headquarters of the province, being Spain administratively divided into 17 regions and 50 provinces. These 50 cities have more facilities and services than non-capital municipalities, and their main economic activity is focused on the services economic sector. A ranking of these provincial capital cities can be useful to assess road safety for elderly pedestrians in those consolidated areas, and to achieve this objective, the number of elderly pedestrian accidents was normalized by 10,000 elderly inhabitants, which was used as an exposure variable to address size disparities between these cities and compare them (see **Table 2**). These cities vary in population size, ranging from 34,670 (Teruel) to 3,186,620 (Madrid) inhabitants and the indicator of 'elderly pedestrian accidents per 10,000 elderly inhabitants per year' varies from 0.46 (Ciudad Real) to 13.24 (León) in the province capitals for the study period. In **Table 2**, these cities have been grouped by size: cities with over than 500,000 inhabitants, cities between 500,000 and 250,000 inhabitants, cities between 250,000 and 100,000, towns between 50,000 and 100,000, and towns under 50,000. Cities at the top of the ranking (15 riskier cities for elderly pedestrians) are concentrated in the groups between 50,000

and 250,000 inhabitants, except for Palma (5th) and Barcelona (13th), which are two cities with considerable tourist activity. Although being the most populated cities in Spain, Madrid and Barcelona are in the 25th and 13th positions, respectively. The largest cities, with more than half million inhabitants, except for Barcelona, are in the 50% of the safest provincial capital cities for elderly pedestrians. These facts indicate that larger capital cities use to be safer for the elderly, with some exceptions in which tourism could play an important role (increasing normal road traffic flows and the functioning of the city). Middle-sized provincial capitals could be riskier than small and large ones because small ones have less traffic volume and large ones have better and denser road infrastructure and denser public transport. Hence, capitals from 50,000 to 250,000 inhabitants may present an unfair balance between the volume of vehicles and the road infrastructure/public transport.

This section began by describing the urban pedestrian traffic accidents in Spain considering the age of the pedestrian and assessing the effects of population and population density on elderly road safety. The following step of this investigation involves collecting more built environment features apart from these two demographic variables, and modelling the number of elderly and non-elderly accidents in each municipality with a comparison of the results of both models. Next section describes the data and methods used in this research, including the elaboration and analysis of the database, the data modelling, and the results.

Table 2. Ranking of Spain province capitals in relation to their elderly pedestrian road safety indicator (elderly pedestrian accidents per 10,000 elderly inhabitants) between 2016 and 2019.

Population size group (inhabitants)	Ranking (position)	Municipality	Total population (inhabitants)	Ageing rate (%)	Elderly pedestrian accidents (2016-2019)	Elderly Road Safety Indicator*
>500,000	13	Barcelona	1,601,985	20.64%	1054	7.97
>500,000	25	Madrid	3,186,620	19.30%	1409	5.73
>500,000	29	València	790,755	18.53%	314	5.36
>500,000	31	Sevilla	696,320	17.11%	249	5.23
>500,000	33	Zaragoza	672,975	18.28%	253	5.14
>500,000	42	Málaga	559,700	15.21%	119	3.49
250,001-500,000	5	Palma	400,370	14.00%	234	10.43
250,001-500,000	22	Murcia	436,480	13.98%	149	6.10
250,001-500,000	28	Valladolid	309,920	21.01%	144	5.53
250,001-500,000	38	Córdoba	327,240	16.34%	91	4.25
250,001-500,000	46	Las Palmas de Gran Canaria	380,320	15.35%	44	1.88
250,001-500,000	47	Bilbao	349,305	22.06%	49	1.59
250,001-500,000	49	Alacant	328,105	16.70%	21	0.96
100,001-250,000	1	León	130,780	23.39%	162	13.24
100,001-250,000	3	Albacete	171,020	13.81%	110	11.64
100,001-250,000	4	Logroño	152,070	16.86%	110	10.72
100,001-250,000	6	Vitoria-Gasteiz	239,000	17.52%	165	9.85
100,001-250,000	8	Badajoz	150,660	13.98%	80	9.50
100,001-250,000	9	Burgos	177,245	18.27%	116	8.95
100,001-250,000	10	Tarragona	133,040	15.66%	73	8.76
100,001-250,000	15	Donostia-San Sebastián	184,425	21.15%	118	7.56
100,001-250,000	16	Granada	239,940	18.41%	128	7.25
100,001-250,000	18	A Coruña	244,320	21.80%	144	6.76
100,001-250,000	20	Oviedo	223,475	18.94%	105	6.20
100,001-250,000	21	Pamplona/Iruña	193,930	19.35%	93	6.20
100,001-250,000	23	Ourense	106,280	22.34%	57	6.00
100,001-250,000	24	Lleida	136,660	15.58%	49	5.76
100,001-250,000	30	Almería	188,970	14.07%	56	5.27
100,001-250,000	32	Santa Cruz de Tenerife	204,055	16.59%	70	5.17
100,001-250,000	34	Cádiz	123,700	19.66%	48	4.93
100,001-250,000	35	Salamanca	150,690	23.54%	67	4.72
100,001-250,000	40	Jaén	116,120	15.39%	28	3.92
100,001-250,000	43	Huelva	147,275	15.28%	26	2.89
100,001-250,000	45	Castelló de la Plana	175,835	14.95%	21	2.00
100,001-250,000	48	Santander	177,010	20.77%	18	1.22
50,001-100,000	2	Girona	95,675	13.63%	64	12.27
50,001-100,000	7	Segovia	54,510	19.47%	41	9.66
50,001-100,000	11	Zamora	64,820	20.66%	45	8.40
50,001-100,000	12	Pontevedra	82,010	17.17%	46	8.17
50,001-100,000	14	Lugo	97,470	19.59%	59	7.73
50,001-100,000	17	Ávila	58,700	15.56%	26	7.12
50,001-100,000	19	Huesca	51,170	16.97%	23	6.62
50,001-100,000	26	Palencia	79,475	18.98%	34	5.63
50,001-100,000	36	Cáceres	95,010	14.22%	25	4.62
50,001-100,000	37	Cuenca	55,780	15.59%	16	4.60
50,001-100,000	39	Guadalajara	83,710	14.39%	20	4.15
50,001-100,000	41	Toledo	83,085	13.63%	16	3.53
50,001-100,000	50	Ciudad Real	74,880	14.48%	2	0.46
20,001-50,000	27	Soria	39,845	18.02%	16	5.57
20,001-50,000	44	Teruel	34,670	15.91%	5	2.27

Data and methods

Database elaboration and analysis

Apart from population size and population density (analysed in the previous section), other built environment variables have shown to influence the occurrence of traffic accidents at different spatial units of analysis. At this stage, the independent variables were divided into the 3 groups of built environment features proposed by Ukkusuri et al. (2012): (i) socio-demographic, (ii) land use, and (iii) infrastructure variables. The variables in this study should be collected through a systematic process due to the large number of municipalities included in the research. It should be noted

that the quality and scale of the data from larger municipalities may differ from smaller ones. In this regard, homogeneous databases are preferable to avoid further data processing and potential data loss. In this analysis, only data from the urbanized zones of municipalities were used to include only features of zones where the vast majority of studied urban accidents took place. In order to work with only urbanized zones, an ad hoc and systematic process was designed to select these areas for more than 8,000 municipalities. This ad hoc process included GIS algorithms using the ArcMap software (Esri Inc, 2020) and further data wrangling in the R software (R Core Team, 2022); it was applied to land use and infrastructure data, as

socio-demographic features and urban road accidents are supposed to be already concentrated in the urbanized areas within municipal boundaries.

In the literature, socio-demographic factors (i) were found to be important in predicting the frequency of accidents in an area, such as population density (Castro-Nuño et al., 2018; Quistberg et al., 2022), total population (Aparidian & Smirnov, 2020; Cabrera-Arnau & Bishop, 2021; Castro-Nuño & Arévalo-Quijada, 2018), and the proportion of population considering age groups and ethnicity (Aparidian & Smirnov, 2020; Lee et al., 2018; 2019). In this investigation, these characteristics were collected together with other demographic and socioeconomic municipal data available that affect pedestrian road safety.

The collected socio-demographic variables include population size, population density (inhabitants per urbanized surface), ageing rate (proportion of inhabitants over the age of 65), femininity index (number of females over number of males), main houses ratio (number of main houses over the total number of houses, including secondary and empty houses), average age of the vehicles, motorization rate (number of registered vehicles over 1,000 inhabitants) and the proportion of vehicle types (cars and motorcycles over all vehicles). Data showed that, in general, larger cities show a higher proportion of women, a higher proportion of first residences, and more registered motorcycles than smaller ones, which are characterized by having higher motorization rates and older vehicles. These facts indicate the differences between the Spanish municipalities considering their population size, as highly populated cities have more stable population during the year (main house rate), and newer vehicles. This last fact could be caused as higher GDP per capita is found in those cities, so the population might renew their vehicles more often than in smaller ones. In contrast, public transport, and some services (e.g. hospitals or stores) are not present in every of the smaller municipalities, and people are supposed to have a private vehicle in order to move to another municipality to fill their necessities. That fact might cause the existence of higher motorization rates in smaller municipalities. Note that the motorization rate is not a measure of vehicle traffic volume, but a measure of the necessity of the inhabitants of owning a vehicle. This necessity may be due to the lack of points of interest in their municipality, the necessity of covering long distances, or the lack of public transport services.

In relation to land use indicators (ii), a set of variables have given adequate results in previous urban accidents predictive models, such as the proportion of industry land use (Lee et al., 2019), open spaces (Gálvez-Pérez et al., 2022; Lee et al., 2018), and points of interest (Castro-Nuño & Arévalo-Quijada, 2018; Gálvez-Pérez et al., 2022; Yang et al., 2021)

Land use variables (ii) selected for this study are the proportions of main land uses (primary, secondary, tertiary, and residential) of the total urbanized surface, the land use mix index of entropy (ENT) and the number of points of interest (POIs) per total length of the road network. Land use proportions were calculated as shown in Equation 1. Primary land use includes agriculture, forestry, mining, and fishing

areas. Secondary land use is filled with industrial and manufacturing activities. Commercial, community, and cultural services, such as retail, hotels, and restaurants, compose the tertiary land use and residential land use is covered by housing areas. The ENT land use mix index is an indicator used in previous literature that reflects the balanced distribution of land uses in an area (Song et al., 2013). The ENT index can take values between 0 and 1, representing 0 the minimum balanced distribution possible and 1 the situation in which all land use types are equally existent, and it was calculated according to Equation 2. According to Asadi et al. (2022), higher land use mix values at regular grid intra-municipal level lead to safer zones, as inhabitants should move less to reach their destiny locations; but the effect of this variable might be different, as in our study other land use mix is used and it is calculated at municipal scale. Other land use variable included in this study is the number of points of interest (POIs) per km, strongly linked to pedestrian and vehicle flows (higher POIs means higher pedestrian and vehicle flows) and, according to the literature (Gálvez-Pérez et al., 2022; Yang et al., 2021), this variable was shown to increase the number of accidents at the city district and traffic analysis zone levels. Figure 2 shows the proportion of primary, secondary, tertiary, and residential land uses, and the land use mix index in the urbanized areas of Spanish municipalities by population size. Primary land use presence is higher in less populated municipalities, and it decreases in more populated areas, where secondary, tertiary, and especially residential land uses are more present. The land use mix entropy index is higher in less populated municipalities, where land uses are more homogeneously present than in bigger cities, where most of the surface is covered by residential and tertiary areas. Additionally, more POIs per km are found in the largest cities.

$$p_{i,j} = \frac{\sum_{m=1}^{m_{i,j}} A_{i,j,m}}{\sum_{n=1}^{k_i} \sum_{m=1}^{m_{i,n}} A_{i,n,m}}, \quad (1)$$

where $p_{i,j}$ is the proportion of surface of land use type j in the municipality i , $A_{i,n,m}$ is the surface of the land use section m of type n in the municipality i , $m_{i,n}$ is the total number of sections of the type n in i and k_i is the total number of land use types present in the municipality i .

$$ENT_i = - \frac{\sum_{j=1}^{k_i} (p_{i,j} \cdot \ln(p_{i,j}))}{\ln(k_i)}, \quad (2)$$

where ENT_i is the entropy land-use mix index in the area i , $p_{i,j}$ is the proportion of land use j in the area i , and k_i is the total number of land use types in the municipality i .

In relation to other built environment independent variables, previous research has shown that some features of the road infrastructure (iii) can affect urban road safety, and some examples are the density of intersections (Gálvez-Pérez et al., 2022; Quistberg et al., 2022), presence of mass transit

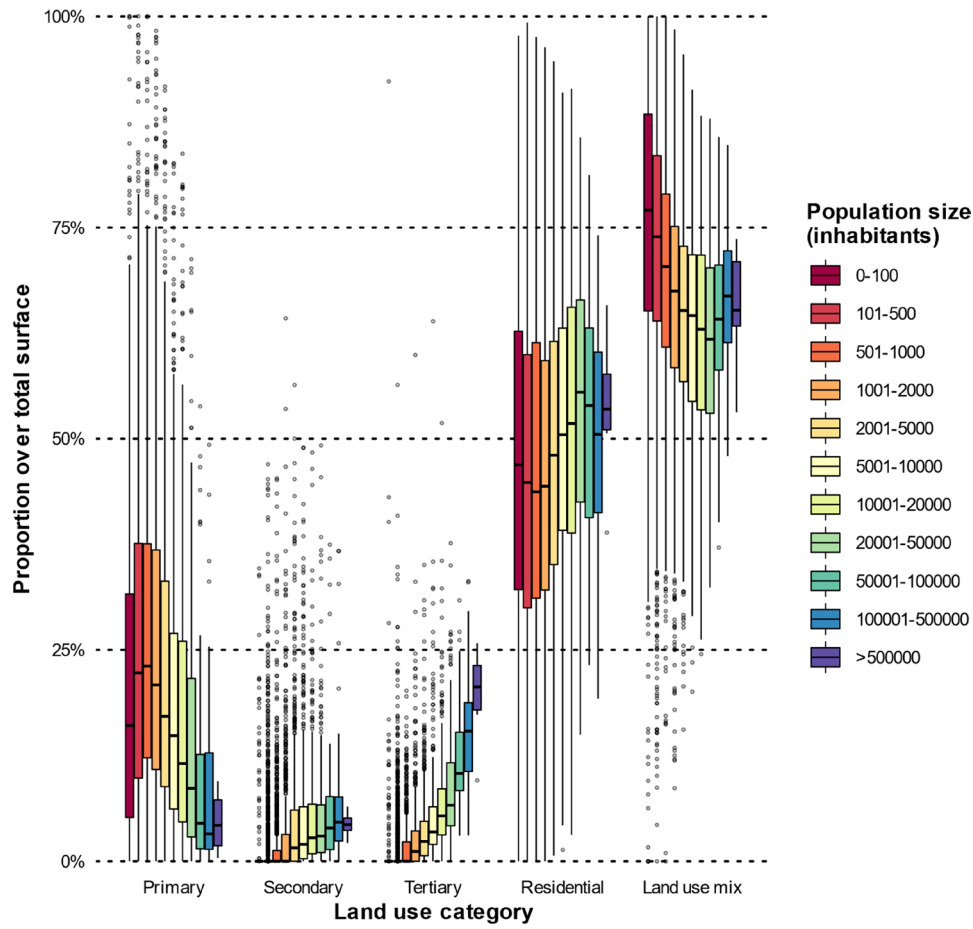


Figure 2. Proportion of primary, secondary, tertiary, and residential land use, and land use mix index in Spanish municipalities urbanized areas by population size of the municipality.

(Castro-Nuño et al., 2018; Quistberg et al., 2022) or traffic lights (Gálvez-Pérez et al., 2022; Yang et al., 2021).

The availability of a reliable urban road infrastructure database usually conditions the selection of these variables to be used in accident modelling. In our research, OpenStreetMap (OSM) offered a minimum of road infrastructure variables covering the Spanish municipalities, so the infrastructure variables (iii) were obtained after processing data extracted from OSM. These data include the total length of roads in each municipality, proportion of the roads by type (motorways, major and residential roads, and paths; calculated according to Equation 3), two-way road proportion, road intersections per km, pedestrian crossings per km, traffic light per km and bus stops per km. Due to the large set of road types present in the OSM data, the roads were re-categorized into 4 groups: motorways, major roads (trunk, primary, secondary, and tertiary), residential roads (residential, living streets, and unclassified), and paths (paths and tracks).

Smaller municipalities are usually crossed or surrounded by major roads that are part of longer interurban routes, and these are usually classified by OSM as trunk, primary, secondary or tertiary roads, and the rest of the road network are residential streets and paths. Larger cities generally have a more varied road network, where we can find that major avenues and streets are identified as primary,

secondary, or tertiary roads in the OSM data, depending on their 'relative importance'. These cities also show higher proportions of interurban roads, lower proportions of residential roads and paths, and a lower proportion of two-way roads than smaller municipalities. In terms of the collected road elements per km figures, larger cities have more bus stops, pedestrian crossing points and traffic lights per km, while smaller municipalities have more road intersections per km.

$$p_{i,j} = \frac{\sum_{m=1}^{m_{i,j}} l_{i,j,m}}{\sum_{n=1}^{k_i} \sum_{m=1}^{m_{i,n}} l_{i,n,m}}, \quad (3)$$

where $p_{i,j}$ is the proportion of length of road type j in the municipality i , $l_{i,n,m}$ is the length of the road segment m of type n in the municipality i , $m_{i,n}$ is the total number of road segments of the type n in i and k_i is the total number of road types present in the municipality i .

Once the data collection and the variables definitions have been accomplished, the outcome of this process is a database that includes the number of vehicle-pedestrian collisions (elderly pedestrians and non-elderly pedestrians) as well as the socio-demographic, land use, and infrastructure variables for each municipality during the period 2016-2019. Table 3 summarizes the variables collected and data sources

used in this study. Municipalities with null inhabitants, registered vehicles, or road infrastructure, were removed from the dataset, summing up a sample size (n) of 8,028 municipalities. Table 4 shows the main statistics of the dependent and independent variables included in the study.

Data modelling

Two statistical models to analyse the elderly pedestrian accidents per municipality should be built: one to predict the elderly (model EP) and other for the non-elderly (model NEP) pedestrian accidents. Comparison between EP and NEP models will help to understand which factors influence elderly pedestrian accidents but not necessarily the rest of pedestrians. The structure of the models was determined by the nature of the dependent variable (the number of pedestrian accidents in each municipality for the period 2016-2019). This variable only takes positive values (count data) that show an excess of zero occurrence (there is a vast number of municipalities where no accidents were registered) and also presents over-dispersion (the variance is remarkably higher than the mean).

Using count data models could produce inaccurate findings if there are more occurrences of 'zero' than would be predicted by the model's probability mass function

(PMF). This is usually known as having an excess of zeros (Hilbe, 2014). In the case study, there are 6,350 (79%) municipalities with null vehicle-pedestrian collisions during the studied period (Table 5). The proportion of municipalities that registered at least 1 accident (regardless of the pedestrian age) is positively related to the population size of the municipality (see Table 5). Most of municipalities with less than 10,000 inhabitants registered zero elderly and non-elderly pedestrian accidents for the study period. This fact could be due to the limitation of the study period (4 years), the randomness of the accident, and a lower exposure to crash risk in the smaller cities, because the traffic and pedestrian flows are supposed to be smaller.

If we use the Poisson distribution to model the number of traffic accidents suffered by pedestrians, whose mean value in the case study is 5.62 accidents by municipality, we expect 0.4% of zero occurrence based on the Poisson's PMF (Equation 4), but this result is remarkably lower than the real data. This excess of municipalities with zero accidents, according to Table 6, is common for both samples, the sample of elderly and non-elderly accidents.

$$f(k=0; \lambda=5.62) = \frac{\lambda^k e^{-\lambda}}{k!} = \frac{5.62^0 e^{-5.62}}{0!} = 0.004 = 0.4\%, \quad (4)$$

Table 3. Description of accident and municipal variables.

Variable	Unit	Definition	Source
Vehicle-pedestrian collisions			DGT (2022)
Elderly pedestrian collisions	#	Number of elderly ped. accidents in the municipality	
Non-elderly pedestrian collisions	#	Number of non-elderly ped. accidents in the municipality	
Socio-demographic			INE (2013)
Total population	Inhabitants	Total number of inhabitants of the municipality	
Elderly population	Inhabitants	Total number of elderly inhabitants of the municipality	
Non-elderly population	Inhabitants	Total number of non-elderly inhabitants of the municipality	
Population density	Inh/ha	Number of inhabitants by municipality surface	
Femineity index	women/men	Number of women by number of men ratio	
Ageing rate	%	Number of people aged 65+ by total population	
Main houses proportion	%	Main houses by total number of houses ratio	
Number of registered vehicles	#	Number of registered vehicles in the municipality	DGT (2022)
Vehicle type proportion	%	Proportion of * of total number of vehicles	
Cars proportion	%	Cars	
Motorcycles proportion	%	Motorcycles	
Average age of the vehicles	Years	Average age of vehicles registered in the municipality	
Land Use			IGN (2022), Geofabrik (2022)
Primary use	%	Primary land use surface by total urbanized surface	
Secondary use	%	Secondary land use surface by total surface	
Tertiary use	%	Tertiary land use surface by total surface	
Residential use	%	Residential land use surface by total surface	
Entropy mix index	index [0-1]	Entropy land use mix index	
POIs per road length	#/km	Number of points of interest (POI) per total road length	
Infrastructure			Geofabrik (2022)
Total length of roads	km	Total length of roads of all types	
Road type proportion	%	Proportion of * of the total length of roads	
Motorway proportion	%	Motorways	
Major roads proportion	%	Trunk, primary, secondary, and tertiary roads	
Residential roads proportion	%	Residential, unclassified roads and living streets	
Paths proportion	%	Paths and tracks	
Two-way road proportion	%	Proportion of two-way roads of the total road length	
Road intersections per road length	#/km	Number of road intersections by total road length	
Crossings per road length	#/km	Number of pedestrian crossing points by total road length	
Traffic light per road length	#/km	Number of traffic lights by total road length	
Bus stops per road length	#/km	Number of bus stops by total road length	
Auxiliary information			IGN (2022)
Municipal boundaries	-	Boundaries of the Spanish municipalities	
Settlement boundaries	-	Boundaries of the urbanized zones	

Table 4. Variable descriptors of the dependent and independent variables used in the models (min, max, 1st and 3rd quartiles, mean, and standard deviation).

Variable	Min.	Max.	1 st Qu.	3 rd Qu.	Mean	S.D.
Dependent variables						
Elderly pedestrian accidents	0.00	1,409	0.00	0.00	1.60	21.88
Non-elderly pedestrian accidents	0.00	4,098	0.00	0.00	3.84	61.82
Socio-demographic						
Elderly	5.00	615,145.00	60.00	511.20	986.90	8,841.54
Non-elderly	5.00	2,571,475.00	120.00	1,980.00	4,812.00	38,059.91
Population density	0.55	1,179.64	6.94	26.91	22.92	34.17
Femininity index	0.00	5.00	0.84	1.00	0.91	0.15
Main houses rate	0.08	1.00	0.41	0.70	0.55	0.19
Ageing rate	0.03	0.89	0.18	0.36	0.28	0.13
Motorization rate	246.60	119,450.50	779.20	993.80	1,000.80	2,527.37
Average age of the vehicles	1.62	18.82	12.00	13.61	12.75	1.30
Cars proportion	0.15	0.96	0.58	0.69	0.63	0.08
Motorcycle proportion	0.00	0.36	0.05	0.09	0.07	0.03
Land use						
Primary proportion	0.00	1.00	0.08	0.35	0.23	0.19
Secondary proportion	0.00	0.64	0.00	0.02	0.03	0.06
Tertiary proportion	0.00	0.92	0.00	0.03	0.02	0.05
Residential proportion	0.00	0.99	0.33	0.62	0.47	0.20
Land use mix (ENT)	0.00	1.00	0.60	0.79	0.69	0.16
POIs per km	0.00	40.47	0.24	1.70	1.41	2.26
Infrastructure						
Motorway road proportion	0.00	0.38	0.00	0.00	0.00	0.01
Major road proportion	0.00	0.83	0.08	0.19	0.14	0.09
Residential road proportion	0.00	1.00	0.64	0.82	0.73	0.14
Path road proportion	0.00	1.00	0.05	0.18	0.13	0.12
Two-way proportion	0.14	1.00	0.93	1.00	0.93	0.13
Road intersections per km	0.00	29.29	8.16	11.79	10.21	3.10
Crossing per km	0.00	14.65	0.00	0.44	0.57	1.45
Traffic signals per km	0.00	13.54	0.00	0.00	0.02	0.21
Bus stops per km	0.00	3.75	0.00	0.08	0.08	0.20

Table 5. Number of municipalities with at least one vehicle-pedestrian (V-P) accident by population size group in Spain (2016-2019) and age of pedestrian.

Population size (inhabitants)	No. of municipalities	No. of municipalities with at least 1 V-P accident (%)	No. of municipalities with at least 1 V-EP* accident (%)	No. of municipalities with at least 1 V-NEP** accident (%)
0 to 100	1,101	5 (0.5%)	3 (0.3%)	1 (0.1%)
101 to 500	2,676	79 (3.0%)	46 (1.7%)	33 (1.2%)
501 to 1,000	1,046	103 (9.8%)	53 (5.1%)	48 (4.6%)
1,001 to 2,000	907	153 (16.8%)	85 (9.4%)	79 (8.7%)
2,001 to 5,000	998	342 (34.3%)	188 (18.8%)	224 (22.4%)
5,001 to 10,000	548	349 (63.7%)	222 (40.5%)	273 (49.8%)
10,001 to 20,000	359	283 (78.8%)	229 (63.8%)	263 (73.3%)
20,001 to 50,000	249	222 (89.2%)	206 (82.7%)	214 (85.9%)
50,001 to 100,000	83	81 (97.6%)	80 (96.4%)	79 (95.2%)
100,001 to 500,000	55	55 (100.0%)	55 (100.0%)	55 (100.0%)
> 500,000	6	6 (100.0%)	6 (100.0%)	6 (100.0%)
Total	8,028	1,678 (20.9%)	1,173 (14.6%)	1,275 (15.9%)

Table 6. Expected number of municipalities with zero (0) pedestrian traffic accidents based on the Poisson probability mass function in Spain for the period 2016-2019.

Vehicle collisions with	Mean number of traffic accidents by municipality	Number of municipalities with zero (0) registered pedestrian traffic accidents	
		Observed	Expected*
Elderly pedestrians	1.60	6,855 (85.4%)	1,621 (20.2%)
Non-elderly pedestrians	3.84	6,753 (84.1%)	173 (2.1%)
All pedestrians	5.62	6,350 (79.1%)	29 (0.4%)

*Expected number based on the Poisson probability mass function (Equation 4).

where λ is the average number of events (pedestrian traffic accidents per municipality) and k is the number of occurrences.

In our case study, the excess of zero occurrence could be due to two reasons: accident **under-reporting** or **accident**

randomness. On the one hand, traffic accident under-reporting is a common practice in Spain, especially when accident victim is slightly injured (which is the case in most accidents); and, on the other hand, traffic accidents are random events that might not be registered during the study period. Despite analysing a 4-year period in our study, we observed a lack of accidents in most municipalities, especially the smaller and less populated ones. We should not consider this fact as a fixed result regardless of the study period. In those locations with null vehicle-pedestrian accidents observations, these events may occur in a different period of study (e.g. a wider period) and we can assume that these municipalities would count with at least one accident if the study period was wider. Municipalities with zero accidents throughout many years are conceptually possible: for instance, due to a very low exposure to the accident risk. Another possibility of registering null accidents during the

sampling period is the case of areas with higher exposure to accident risk but with infrastructure and city features that make them safer. In summary, the excess of zeros might be caused by accidents that were not registered or by accidents that did not occur but would occur during a different sampling period.

To overcome this situation, adjusted models such as ‘zero-inflated’ models (Cai et al., 2016; Pew et al., 2020) and ‘hurdle’ models (Asadi et al., 2022; Cai et al., 2016) have been used in the literature. The main difference between them is that the first ones are mixture models that assume the existence of **structural and sampling zeros**, and the latest ones are two-stage models that only account for **sampling zeros**. The structural zeros approach is based on the consideration that there are zero observations that will always remain zero, regardless of the sampling period. In contrast, sampling zeros are observed zeros due to the sampling process and that would be different from zero if the sampling period was different.

In our analysis, the hypothesis of structural zeros would be associated with municipalities with null accidents (even if the study period was infinite), while the consideration of having only sampling zeros would be associated with areas with zero registered accidents in 2016-2019, but that would have accidents if the period were different. This latter assumption seems more suitable, because every sampled municipality has shared roads for vehicles and pedestrians and not null traffic volumes, which can be inferred as the population is not null, and that means that a traffic accident may always occur, although the exposure risk is low. In other words, the risk of having an accident for a pedestrian is not zero, and if the study period was infinite, every municipality would account for at least 1 accident, and no structural zeros would be found. As a consequence, the hurdle model, which was firstly proposed by Cragg (1971) and further discussed by Mullahy (1986), was chosen over the zero-inflated model in this paper.

Furthermore, apart from an excess of zeros, our crash occurrence data shows over-dispersion, and the model should be capable of managing over-dispersed data. In this regard, using a Negative Binomial model is a fair solution. Therefore, the collected data will be analysed by implementing a **hurdle negative binomial (HNB) regression model**. The HNB regression model is a two-stage model that contains two model components or sub-models. First, a classification model (**zero component, ZC**), which in this case is a **binomial logistic regression model**, classifies municipalities into two groups: those with zero accidents and those with at least 1 accident. Second, a frequency model (**count component, CC**), a **zero-truncated negative binomial regression model** in this case, provides the number of accidents for the municipalities of the second group. The zero component is trained with the complete dataset (all municipalities) and the count component only with no-null accident data (municipalities with at least 1 accident). When the features of a municipality are inserted into the HNB model, the first sub-model estimates whether it will have zero or more than zero pedestrian accidents for the next 4-year

period. If the municipality falls under the first option, the forecast would be zero accidents during the next period, and, consequently, the second sub-model would not be applied. On the other hand, if the municipality is classified in the second group, the second sub-model would be applied, and the number of accidents would be estimated.

The probability density function (PDF) of the HNB model was structured according to Equation 5. The PDF of the count component f_2 of the model, which is a negative binomial regression, is shown in Equation 6. We implemented the HNB model in the R software using the package ‘pscl’ published by Zeileis et al. (2008).

$$P(Y_i = y_i) = \begin{cases} f_1(0) = p_i, & y_i = 0 \\ \frac{1 - f_1(0)}{1 - f_2(0)} \cdot f_2(y_i), & y_i > 0 \end{cases} \quad (5)$$

where f_1 and f_2 are the probability functions of the first sub-model (f_1), which filters out null accident observations through a logistic classification model, and the second sub-model (f_2), which is the count component, respectively. p_i is the probability of having a null count response ($0 < p_i < 1$). Note that the term $1 - f_2(0)$ in the denominator considers the truncation of the count component of the model, as the probability of f_2 considers null observations, whereas the $y_i > 0$ section should only consider non-null responses.

$$f_2(y_i; \mu_i, \alpha) = P_{NB}(y_i; \mu_i, \alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(\alpha^{-1}) y_i!} \left(\frac{\alpha \dots}{1 + \alpha \mu_i} \right)^{y_i} (1 + \alpha \mu_i)^{-\alpha^{-1}} \quad (6)$$

where y_i refers to the number of vehicle-pedestrian collisions in the municipality i , $\Gamma(\cdot)$ is the gamma function, α is the dispersion parameter and μ_i is the mean value of the count feature.

The **zero component** of the model follows the logit function, and the probability of having zero accidents can be estimated according to Equation 7.

$$\text{logit}(p_i) = \log\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1 X_{i1} + \dots + \beta_m X_{im} = X_i^T \beta \quad (7)$$

where p_i is the probability of having zero accidents in the municipality i , X_i^T is the vector of predictors and β is the vector of parameters.

The **count component** of the model follows the negative binomial formulation, and the expected number of crashes in municipalities with no-null data can be estimated using Equation 8.

$$\log(\mu_i) = \gamma_0 + \gamma_1 X_{i1} + \dots + \gamma_m X_{im} = X_i^T \gamma, \quad (8)$$

where μ_i is the number of crashes in the municipality i , X_i^T is the vector of predictors and γ is the vector of parameters.

With the HNB regression model, an exhaustive model selection procedure was developed to assess the behaviour

of the built environment variables and their main contributions to the elderly pedestrian road safety. This process involved computing all possible configurations of the independent variables to build different HNB regression models to predict both elderly (EP model) and non-elderly (NEP model) pedestrian crashes. As the number of combinations of variables grows exponentially with the number of tested independent variables, a preliminary analysis was performed to reduce the number of tested variables and the consequent computing cost. This preliminary analysis started by identifying and removing the independent features that showed linear correlation with each other. After that, a set of HNB models to predict EP and NEP accidents was built, including the full model (all independent variables) and other models using stepwise algorithms (James et al., 2013). Variables that did not show a statistically significant effect on the response variable were also removed from the final set of independent variables. In total, 15 variables, were selected for the final data modelling, which formed 32,767 ($2^{15} - 1$) possible variable combinations. Even though the HNB is theoretically a fair model for the case study, the HNB model was tested against the Negative Binomial (NB) model prior to data modelling. A NB and a HNB model were built for both EP and NEP datasets with the final set of 15 independent variables. The HNB model showed to be a better option against the NB model based on the AIC, with values of 8,296.78 vs 8,350.37 for the EP case and 9,554.24 vs 9,617.78 for the NEP case.

The structure of the tested model configurations (MCs) was designed to always include the exposure variable (number of elderly and non-elderly inhabitants for the EP and NEP models, respectively), an intercept, and at least 1 of the 15 independent variables. Moreover, the same set of explanatory variables was used in each MC to build both sub-models (count and zero components). The 'empty' model, that included only the intercept and the exposure variable, was also tested, summing 32,768 possible MCs in total. The MCs were tested to model the two groups of accidents being studied: elderly (EP model) and non-elderly (NEP model) pedestrian traffic accidents. In other words, 32,768 EP models and 32,768 NEP models were built, each one composed of the count (CC) and zero (ZC) model components. The result of this process was a dataset containing for each MC:

- The 4 parameters of the independent variables (one for each component of each model; EP-CC, EP-ZC, NEP-CC and NEP-ZC) and their significance levels (p-value).
- The goodness-of-fit measure of the model through the following indicators: Akaike information criterion (AIC) and the Akaike weight (defined according to Equation 9).

With this information, and considering all MCs, the minimum, maximum, mean, and standard deviation values of the variable parameters were computed. Likewise, the proportion of times the parameter is positive (PP)

and the proportion of times each independent variable is significant at the alpha levels of 10% (PV10), 5% (PV5), and 1% (PV1) were calculated (Table 7). A consistent impact of a variable on road safety, either positively or negatively, regardless of the presence of other factors in the regression model, suggests a strong relationship with road safety; and the significance of that variable should be further examined. Additionally, the Akaike weight (w_i) is a measure of evidence of the model i to be the best model among the K tested MCs. It is a normalized measure of the likelihood of a model being the best model of a set of K models based on a goodness-of-fit measure (AIC in this case), and the sum of the Akaike weights of the K models is equal to 1 (Anderson & Burnham, 2004). The Akaike weight can be used to assess the model, but also to measure the relative importance of the independent variables. If a predictor appears in all models that have high Akaike weights, that indicates that it is an important variable to predict the outcome variable. Therefore, the relative importance of a variable is the sum of the Akaike weights of the models that contain that covariate. The number of times each variable appears should be balanced in the set of regression models (Anderson & Burnham, 2004; Murray & Conner, 2009), which is accomplished in our case, as every independent variable appears in the same number of tested models, because all possible MCs were computed.

$$w_i = \frac{\exp\left(-\frac{1}{2}\Delta(AIC)_i\right)}{\sum_{k=1}^K \exp\left(-\frac{1}{2}\Delta(AIC)_k\right)} = \frac{\exp\left(-\frac{1}{2}(AIC_i - \min(AIC))\right)}{\sum_{k=1}^K \exp\left(-\frac{1}{2}(AIC_k - \min(AIC))\right)}, \quad (9)$$

where w_i is the Akaike weight of the model i , Δ_i is the increment of the AIC of the model i , and K is the number of computed regression models.

A multi-criteria analysis was performed to evaluate which predictors of EP collisions were the best. For this purpose, a set of indexes were defined for each independent variable in each model (EP and NEP) using the processed data: (i) sign, (ii) p-value and (iii) Akaike weight. The values of all these indexes ranges from 0 to 1, representing 1 the best possible value and 0 the worst.

- The **sign index** (i) measures the determination of the sign of the variable parameter, being 0 is the worst situation (where half of the calculated parameters are positive, and the other half has negative values) and being 1 the best scenario (where all variable parameters have the same sign positive or negative).
- The **p-value index** (ii) was defined as the PV10 value described above.
- The **Akaike weight index** (iii) was designed as the relative importance of the variable based on the Akaike weight.

Each independent variable has 5 indexes: 2 for CC (sign and p-value), 2 for ZC (sign and p-value), and the Akaike

Table 7. Assessment of the behaviour of each independent variable (parameter, p-value and Akaike weight) in the EP and NEP models.

Variable	No. of models where the variable appears	Model EP: Elderly pedestrian accidents						Model NEP: Non-elderly pedestrian accidents												
		Parameter (β)			P-value			Parameter (β)			P-value									
		Min	Max	Mean	SD	> 0	α = 10%	α = 5%	α = 1%	Σ w _i	Min	Max	Mean	SD	> 0	α = 10%	α = 5%	α = 1%	Σ w _i	
Count component	32,768	-11.31	-5.67	-8.75	1.07	0.00	1.00	1.00	1.00	1.00	1.00	-14.36	-8.29	-11.44	1.04	0.00	1.00	1.00	1.00	1.00
Intercept																				
Socio-demographic																				
log(inhabitants)	32,768	1.24	1.34	1.28	0.02	1.00	1.00	1.00	1.00	1.00	1.37	1.50	1.42	0.02	1.00	1.00	1.00	1.00	1.00	1.00
Population density	16,384	-0.00	0.00	-0.00	0.00	0.11	0.04	0.00	0.47	0.00	0.00	0.00	0.00	0.00	0.00	0.85	0.69	0.31	0.91	0.91
Main house rate	16,384	-0.48	0.28	-0.14	0.15	0.19	0.01	0.00	0.21	-0.71	0.08	-0.33	0.15	0.01	0.33	0.18	0.03	0.79	0.79	0.79
Average age of vehicles	16,384	-0.18	0.01	-0.08	0.04	0.01	0.59	0.54	0.48	-0.17	-0.01	-0.09	0.03	0.00	0.78	0.69	0.48	0.93	0.93	0.93
Motorization rate	16,384	-0.00	0.00	-0.00	0.00	0.05	0.02	0.00	0.38	0.00	0.00	0.00	0.00	0.86	0.09	0.00	0.00	0.56	0.56	0.56
Cars proportion	16,384	-1.39	1.51	0.03	0.81	0.50	0.21	0.06	0.26	-1.20	1.55	-0.05	0.68	0.47	0.14	0.02	0.00	0.23	0.23	0.23
Motorcycle proportion	16,384	3.05	6.16	4.54	0.61	1.00	1.00	1.00	1.00	2.23	5.62	3.73	0.57	1.00	1.00	1.00	0.94	0.99	0.99	0.99
Land use																				
Secondary proportion	16,384	0.74	1.63	1.23	0.17	1.00	1.00	0.96	0.64	-0.28	0.79	0.26	0.20	0.91	0.01	0.00	0.00	0.43	0.43	0.43
Tertiary proportion	16,384	0.55	1.64	1.16	0.22	1.00	0.87	0.64	0.90	0.50	1.93	1.17	0.27	1.00	0.82	0.63	0.24	0.35	0.35	0.35
Land use ENT	16,384	-0.12	0.98	0.46	0.24	0.96	0.43	0.31	0.76	-0.14	0.94	0.42	0.18	0.99	0.38	0.21	0.03	1.00	1.00	1.00
POIs per km	16,384	-0.01	0.01	0.00	0.00	0.62	0.00	0.00	0.93	0.00	0.02	0.01	0.00	0.91	0.02	0.00	0.00	0.76	0.76	0.76
Infrastructure																				
Residential road proportion	16,384	-0.59	0.61	0.04	0.20	0.64	0.00	0.00	0.76	-1.28	-0.07	-0.65	0.20	0.00	0.46	0.21	0.00	0.96	0.96	0.96
Major road proportion	16,384	-0.66	0.53	-0.09	0.18	0.27	0.00	0.00	0.18	-0.87	0.85	0.12	0.41	0.59	0.02	0.00	0.00	0.32	0.32	0.32
Motorway proportion	16,384	-5.19	-0.97	-3.19	0.78	0.00	0.62	0.36	0.51	-4.72	-0.57	-2.74	0.83	0.00	0.51	0.28	0.00	0.79	0.79	0.79
Road intersections per km	16,384	-0.00	0.05	0.02	0.01	0.99	0.08	0.01	0.62	0.01	0.07	0.04	0.01	1.00	0.71	0.49	0.11	0.64	0.64	0.64
km																				
Traffic lights per km	16,384	-0.09	0.03	-0.03	0.03	0.12	0.00	0.00	0.96	0.03	0.15	0.07	0.02	1.00	0.06	0.02	0.00	0.74	0.74	0.74
log(theta)	32,768	0.67	0.77	0.72	0.02	1.00	1.00	1.00	1.00	0.39	0.48	0.44	0.02	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Zero component																				
Intercept	32,768	-12.85	-5.49	-9.71	1.36	0.00	1.00	1.00	1.00	-13.80	-5.66	-10.65	1.50	0.00	1.00	1.00	1.00	1.00	1.00	1.00
Socio-demographic																				
log(inhabitants)	32,768	1.33	1.52	1.43	0.03	1.00	1.00	1.00	1.00	1.25	1.47	1.36	0.04	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Population density	16,384	-0.00	0.00	-0.00	0.00	0.01	0.07	0.01	0.47	-0.00	-0.00	-0.00	0.00	0.00	0.85	0.53	0.03	0.91	0.91	0.91
Main house rate	16,384	-0.41	0.56	0.03	0.20	0.53	0.00	0.00	0.21	-0.87	0.05	-0.43	0.18	0.01	0.31	0.13	0.00	0.79	0.79	0.79
Average age of vehicles	16,384	-0.23	-0.02	-0.12	0.04	0.00	0.89	0.80	0.48	-0.23	-0.01	-0.11	0.04	0.00	0.78	0.69	0.49	0.93	0.93	0.93
Motorization rate	16,384	-0.00	0.00	-0.00	0.00	0.00	0.99	0.91	0.38	-0.00	-0.00	-0.00	0.00	0.00	0.50	0.50	0.47	0.56	0.56	0.56
Cars proportion	16,384	-1.60	0.39	-0.65	0.41	0.06	0.16	0.06	0.26	-2.03	0.17	-0.90	0.47	0.02	0.31	0.19	0.03	0.23	0.23	0.23
Motorcycle proportion	16,384	-0.31	5.24	2.99	1.08	1.00	0.68	0.56	1.00	0.76	6.21	4.23	0.96	1.00	0.93	0.87	0.69	0.99	0.99	0.99
Land use																				
Secondary proportion	16,384	1.33	2.61	2.07	0.28	1.00	1.00	1.00	0.95	0.38	2.24	1.38	0.43	1.00	0.74	0.58	0.45	0.43	0.43	0.43
Tertiary proportion	16,384	0.95	3.43	2.24	0.46	1.00	0.94	0.83	0.48	-0.35	2.38	1.08	0.52	0.98	0.22	0.08	0.00	0.35	0.35	0.35
Land use ENT	16,384	0.52	1.52	1.05	0.20	1.00	0.99	0.97	0.76	0.91	1.96	1.47	0.18	1.00	1.00	1.00	1.00	1.00	1.00	1.00
POIs per km	16,384	-0.07	-0.01	-0.04	0.01	0.00	0.86	0.71	0.41	-0.06	-0.01	-0.04	0.01	0.00	0.73	0.52	0.16	0.76	0.76	0.76
Infrastructure																				
Residential road proportion	16,384	-1.62	-0.56	-1.03	0.18	0.00	0.93	0.80	0.76	-2.53	-0.94	-1.63	0.41	0.00	1.00	1.00	0.99	0.96	0.96	0.96
Major road proportion	16,384	-0.76	1.18	0.29	0.55	0.56	0.16	0.02	0.18	-2.22	0.81	-0.56	0.97	0.50	0.43	0.30	0.04	0.32	0.32	0.32
Motorway proportion	16,384	-2.94	3.46	0.00	1.12	0.49	0.00	0.00	0.51	2.25	7.82	5.01	1.10	1.00	0.79	0.54	0.15	0.79	0.79	0.79
Road intersections per km	16,384	0.01	0.06	0.03	0.01	1.00	0.49	0.22	0.62	-0.02	0.04	0.01	0.01	0.79	0.01	0.00	0.00	0.64	0.64	0.64
Traffic lights per km	16,384	0.76	1.67	1.20	0.18	1.00	1.00	0.87	0.96	0.33	1.08	0.66	0.15	1.00	0.19	0.01	0.00	0.74	0.74	0.74

weight index. As the same variables were used in both MC components, the Akaike weight index is equal for each variable between the sub-models and was only counted once. The multi-criteria analysis was based on 3 performance scores per variable, calculated as follows: performance scores for count and zero components were obtained by multiplying the indexes of the respective components, and the overall score was obtained by multiplying all 5 indexes of a variable. Table 8 displays the overall performance score of each variable in the multi-criteria analysis and Figure 3 illustrates the performance scores of each independent variable for count and zero components and Akaike weight index. Furthermore, this methodology was used to find the best model, in terms of the lowest AIC value, being all the variables statistically significant at a 10% level. In this regard, the best HNB models for the EP and NEP crashes are shown in Table 9.

Modelling results

This section provides the interpretation and comparison of the statistical modelling results regarding the frequency of vehicle-pedestrian collisions involving elderly pedestrians (model EP) and non-elderly pedestrians (model NEP). The results can be interpreted in two steps due to the inherent characteristics of the statistical model that was used, which is composed of two model components: count and zero. On the one hand, the **zero component (ZC)** estimates the likelihood of having at least one accident. On the other hand,

the **count component (CC)** predicts the frequency of accidents in municipalities where there has been at least one accident during the study period. In this regard, the exposed performance scores (count, zero and overall; Table 8 and Figure 3) were used to interpret the results.

The exposure variable, the number of inhabitants, is a strong predictor for both models and in both model components (CC and ZC), as the parameter is positive and statistically significant in every tested model. Regarding **socio-demographic features**, the average age of the vehicles and the motorcycle proportion are significant variables at alpha 10% in both model components of the EP and NEP models in most of the tested variable combinations, for both age groups, having older vehicles is related to safer municipalities, while a higher proportion of motorcycles indicates riskier areas. Additionally, the occurrence (CC) of non-elderly pedestrian traffic accidents occurrence is affected by the population density of the municipality, and higher population densities imply higher accident frequencies. The occurrence of elderly pedestrian traffic accidents seems not to be strongly linked to this feature, which could be inferred from Figure 1, as the number of these accidents by population density is roughly constant in the dataset. The motorization rate is a good predictor in the ZC of the EP model; more registered vehicles per inhabitant means safer municipalities. This fact might be due to smaller municipalities usually presenting higher motorization rates, and private vehicles are used mainly to move between municipalities and not inside the area itself, being interurban accidents not studied in this paper. The main house

Table 8. Indexes and overall performance scores of the multi-criteria analysis to assess the behaviour of the independent variables in the modelling of elderly and non-elderly pedestrian accidents.

Variable	Model EP Elderly pedestrian traffic accidents						Model NEP Non-elderly pedestrian traffic accidents					
	Count component indexes		Zero component indexes		Akaike weight index	Overall score	Count component indexes		Zero component indexes		Akaike weight index	Overall score
	Sign	PV10	Sign	PV10			Sign	PV10	Sign	PV10		
Intercept	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Exposure												
log(inhabitants)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Socioeconomics												
Population density	0.78	0.14	0.99	0.07	0.47	0.00	1.00	0.85	1.00	0.85	0.91	0.65
Main house rate	0.62	0.01	0.07	0.00	0.21	0.00	0.99	0.33	0.99	0.31	0.79	0.08
Average age of vehicles	0.99	0.59	1.00	0.89	0.48	0.25	1.00	0.78	1.00	0.78	0.93	0.57
Motorization rate	0.90	0.02	1.00	0.99	0.38	0.01	0.72	0.09	1.00	0.50	0.56	0.02
Cars proportion	0.00	0.21	0.89	0.16	0.26	0.00	0.05	0.14	0.96	0.31	0.23	0.00
Motorcycle proportion	1.00	1.00	0.99	0.68	1.00	0.67	1.00	1.00	1.00	0.93	0.99	0.92
Land use												
Secondary proportion	1.00	1.00	1.00	1.00	1.00	1.00	0.83	0.01	1.00	0.74	0.43	0.00
Tertiary proportion	1.00	0.87	1.00	0.94	0.90	0.73	1.00	0.82	0.96	0.22	0.35	0.06
Land use ENT	0.92	0.43	1.00	0.99	0.76	0.30	0.97	0.38	1.00	1.00	1.00	0.37
POIS per km	0.24	0.00	1.00	0.86	0.93	0.00	0.82	0.02	1.00	0.73	0.76	0.01
Infrastructure												
Residential road proportion	0.27	0.00	1.00	0.93	0.76	0.00	1.00	0.46	1.00	1.00	0.96	0.44
Major road proportion	0.45	0.00	0.12	0.16	0.18	0.00	0.18	0.02	0.00	0.43	0.32	0.00
Motorway proportion	1.00	0.62	0.02	0.00	0.51	0.00	1.00	0.51	1.00	0.79	0.79	0.32
Road intersections per km	0.99	0.08	1.00	0.49	0.62	0.02	1.00	0.71	0.58	0.01	0.64	0.00
Traffic lights per km	0.75	0.00	1.00	1.00	0.96	0.00	1.00	0.06	1.00	0.19	0.74	0.01
log(theta)	1.00	1.00	-	-	1.00	1.00	1.00	1.00	-	-	1.00	1.00

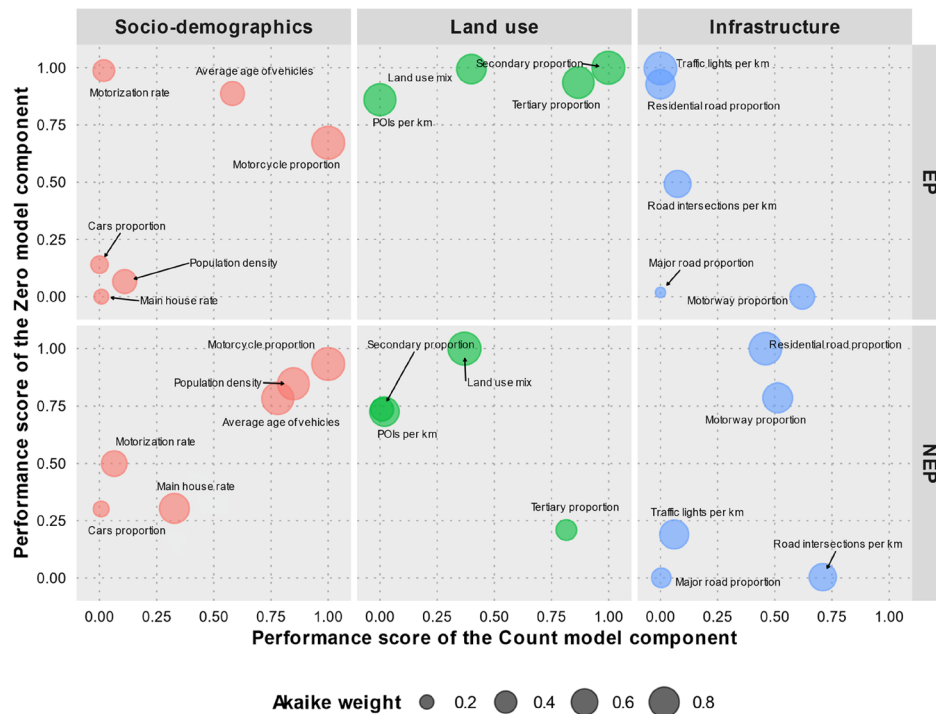


Figure 3. Performance scores of the Count (X-axis) and Zero (Y-axis) model components, and the Akaike weight (size) of each independent variable.

Table 9. HNB models for EP and NEP crashes with the best performance.

Variable	Model EP: Elderly pedestrian traffic accidents							Model NEP: Non-elderly pedestrian traffic accidents							
	Count component			Zero component				Count component			Zero component				
	Estimate	S.E.	p-value	Estimate	S.E.	p-value	Estimate	S.E.	p-value	Estimate	S.E.	p-value	Estimate	S.E.	p-value
Intercept	-9.11	0.27	0.00 ***	-10.45	0.28	0.00 ***	-12.55	0.34	0.00 ***	-13.47	0.45	0.00 ***			
Exposure															
log(inhabitants)	1.25	0.03	0.00 ***	1.36	0.04	0.00 ***	1.42	0.03	0.00 ***	1.38	0.04	0.00 ***			
Socioeconomics															
Population density							0.00	0.00	0.01 **	0.00	0.00	0.04 *			
Motorcycle proportion	4.18	0.96	0.00 ***	3.09	1.35	0.02 *	3.76	0.93	0.00 ***	5.32	1.34	0.00 ***			
Land use															
Secondary proportion	1.33	0.41	0.00 **	2.38	0.56	0.00 ***									
Tertiary proportion	1.04	0.49	0.03 *	2.32	0.84	0.01 **									
Land use ENT							0.64	0.26	0.02 *	1.78	0.34	0.00 ***			
log(theta)	0.74	0.10	0.00 ***				0.44	0.08	0.00 ***						
AIC	8302.65						9566.86								

rate is a fair predictor of NEP accidents, as it is a statistically significant variable in 30% of the models in the CC and ZC. The parameter sign of this variable is stable, and a higher proportion of main houses is related to safer municipalities. Municipalities where most houses are main houses tend to be larger and have more homogeneous vehicle and pedestrian flows during the year, so the infrastructure can be prepared for those flows' fluctuation (traffic variations). Other municipalities, more oriented to second and vacation residences, suffer a significant increase in population and vehicle flow during some months of the year, and the infrastructure might not be prepared for these changes (e.g. lack of traffic lights and visibility). Furthermore,

this variable not being linked to EP road safety indicates that the mobility of elderly people during the year is more stable than the mobility of the rest of the population.

Land use features better performance scores in the EP model than in the NEP model. Secondary and tertiary land use proportions increase the likelihood (ZC) and frequency (CC) of EP accidents, being this effect statistically significant in most of the tested models. The land use entropy mix index increases the likelihood of pedestrian accidents in general (ZC), and municipalities with more homogeneously distributed land uses are riskier for all pedestrians (CC), but this last effect is remarkably less significant at alpha 10% in all tested models than the first one. The number of POIs

per km is not significant in all models for predicting the probability of pedestrian accidents (ZC), and more POIs increase the chance of having zero pedestrian traffic accidents. This effect might seem counterintuitive, but its conclusiveness remains uncertain due to the lack of significance of this variable across all models. However, this fact could potentially associate streets with an increased number of points of interest with better preparation for accommodating high levels of both vehicular and pedestrian traffic, having more road signals, traffic lights or wider sidewalks.

In terms of **infrastructure variables**, the proportion of residential roads is a good predictor for EP and NEP zero components. Having more residential roads in the road network increases the likelihood of having no pedestrian traffic accidents in a municipality, regardless of the pedestrian's age. This result is related to the idea that residential roads are narrower, less transited and the vehicle speed is smaller than in primary ones. Hence, pedestrians may have a low risk of crash regardless of their age. A higher number of traffic lights per km leads to higher probabilities of having at least one EP accident. This fact could be explained these devices are more present in larger municipalities, where it is more common to have at least 1 accident for the studies period. The effect of major streets was not detected for both model components in both EP and NEP models. Higher motorway proportions reduce the number of both EP and NEP accidents as in these locations it is not common to register vehicle-pedestrian collisions. Road intersections per km contribute to a higher number (CC) of NEP accidents and to a higher chance (ZC) of the occurrence of EP traffic accidents. These results show that road intersections are hazardous locations for all pedestrians and their presence may be related to a higher crash risk.

In terms of the best models considering those with all statistically significant variables and the lowest AIC values for the EP and NEP traffic accidents (Table 9), the results are consistent with the overall behaviour of the variables. The exposure variable (elderly and non-elderly inhabitants) and the motorcycle proportion are present in both models, both meaning a rise in the number and probability of crash occurrence. The last variable could be linked to more accidents with slightly injured pedestrians, as accidents were counted regardless of their effects. On the one hand, the secondary and tertiary land uses presences represent that same effect on EP crashes, which could be linked to the elderly being more comfortable walking in residential due to the homogenous road network of these areas. On the other hand, the population density and land use mix index mean more NEP traffic accidents in a municipality. These results could be due to the distinct walking patterns of non-elderly pedestrians, who tend to traverse larger areas within the municipality compared to the elderly population. Elderly individuals often confine their walking activities to the surroundings of their residences. Consequently, non-elderly risk is higher when the municipality present a balanced presence of different land uses, as they are likely to cross different land use categories when walking. Furthermore, higher population density in municipalities often corresponds to an expanded range of leisure activities and a greater need for

commuting, primarily undertaken by the non-elderly pedestrian.

In **summary**, the results suggest that the **road safety** of all pedestrians is related to **socio-demographic** features, being the **elderly pedestrian road safety** linked also to municipality **land use characteristics**, while the rest of the population is more conditioned by **infrastructure**. Despite the differences in the performance scores of the variables, no differences in their effect on road safety of both age groups were found, meaning that actions to improve road safety of the elderly pedestrians should not raise the risk of traffic accidents for the rest of pedestrians. In general, municipalities with older vehicles and less motorcycles, features usually found in smaller municipalities, and unbalanced land uses presence, normally larger cities, appear to be safer for all pedestrians. A reason for unbalanced land use locations being safer for pedestrians could be that drivers and pedestrians mostly circulate in such locations in a homogenous land use. In contrast, areas lacking clear prominent land use lead to heterogeneous road infrastructures, combining different road surfaces, widths, and pedestrian facilities, requiring more attention from road users and increasing the risk of accidents. Particularly for the elderly pedestrian model, less proportion of land covered by manufacturing and service activities was also found to be linked with safer municipalities, which is more common in smaller towns.

Conclusions and future research lines

The objective of this paper was to assess the relationship between elderly pedestrian urban road safety and built environment features at the municipal level. Data were gathered from different sources and processed to produce an ad hoc dataset suitable for performing statistical analysis, resulting in a database with data of 8,028 municipalities for the period 2016-2019. With this information, we developed an exploratory analysis of pedestrian traffic accidents and the characteristics gathered by municipality, with special attention on the population size of the municipalities. Using the hurdle negative binomial regression model, an exhaustive model selection procedure was developed to assess the relative importance of the variables and their main contribution to elderly pedestrian road safety. The non-elderly pedestrian group was also studied to test whether the results of the elderly pedestrian road safety analysis were unique or applicable to the entire population. For each age group, 32,768 HNB regression models were trained, and the results were summarized and interpreted.

The modelling process results show that the elderly pedestrian road safety at the municipal level of detail is linked with land use features, while the rest of the population is more conditioned by infrastructure macro-characteristics. In general, municipalities with older vehicles, less motorcycles and unbalanced land uses presence appear to be safer for all pedestrians. In particular, the elderly pedestrian model indicated that a lower proportion of land devoted to manufacturing and service activities was associated with safer locations for them. In

addition, elderly and non-elderly pedestrian traffic accidents are linked to a set of common built environment features, but also to some specific characteristics of the municipalities. Nevertheless, differences in these relationships were only found regarding the statistical significance of the variables and not the variable effect (sign) on road safety of both age groups, which suggest that special actions to improve elderly pedestrian road safety would not necessarily worsen the road safety situation of the rest of pedestrians.

Based on these modelling results, some possible interventions that should be considered by policymakers to try to reduce the number of elderly pedestrian traffic crashes are outlined here. These interventions focus on infrastructure improvements, as socio-demographic and land use features are assumed to remain fixed in the short term. These actions should include the implementation of traffic signalling in order to advise drivers they are entering an aged urban area. This measure could help drivers understand that they have to control vehicle speed and pay special attention to the road. In addition, in areas with high ageing rates, adapting traffic signalling for the elderly pedestrians by widening the green phase time of traffic lights could help the elderly to cross on time. Furthermore, investing in pedestrian-friendly infrastructure, such as well-maintained sidewalks, pedestrian crosswalks, and pedestrian islands, to prevent older pedestrians from walking on the traffic lanes can improve traffic safety. However, to gain insight into the relationship between infrastructure and crashes, further research should be carried out as exposed in the following paragraphs.

Although this research provides a valuable contribution to elderly pedestrian road safety literature, its main limitation is related to data availability related to built environment data. Some of these limitations are described below:

- The level of detail of the built environmental features, especially socio-demographic variables, limited the use of smaller spatial units of analysis.
- Important variables, such as annual average daily traffic, urban form, and road network area, could not be included in the analysis because of lack of data for all studied municipalities in the vast geographical extension under study.
- Infrastructure macro-variables do not properly reflect the reality of a road network, as the location of the accident inside the unit of analysis is not considered. Hence, infrastructure macro-variables (i.e. traffic lights and road type proportions) may lead to misleading interpretations because accidents may occur in singular hot spots within the unit of analysis.

The limitations of this study show the improvements that can be achieved, and in consequence, new research lines would be focused on the following approaches:

- Smaller units of analysis, such as city districts or neighbourhood, could be used to reduce the heterogeneity inside the studied areas. These smaller units

of analysis have been used in the revised literature, but in studies of a single city, not for an entire country.

- Micro-level analysis, using road segments and intersections as the unit of analysis, should be carried out to study the link between the built environment, especially infrastructure features, and elderly pedestrian road safety, to provide policymakers with recommendations to make urban road design safer for the elderly.

Finally, new research lines will face official data availability, but this fact should not refrain researchers to achieve objectives in order to reduce elderly pedestrian road crashes.

Acknowledgments

The authors gratefully acknowledge the Dirección General de Tráfico (DGT) for providing the traffic accidents data to develop this study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Daniel Gálvez-Pérez is developing his doctoral thesis while he enjoys a grant from the Universidad Politécnica de Madrid through the 'Programa Propio de I+D+I 2020: Ayudas para Contratos Predoctorales'.

ORCID

Begoña Guirao  <http://orcid.org/0000-0003-4138-0073>

References

- Aguero-Valverde, J. (2013). Multivariate spatial models of excess crash frequency at area level: Case of Costa Rica. *Accident; Analysis and Prevention*, 59, 365–373. <https://doi.org/10.1016/j.aap.2013.06.014>
- Anderson, D., & Burnham, K. (2004). Model selection and multi-model inference. In *Second* (vol. 63, pp. 10). Springer-Verlag.
- Apardian, R. E., & Smirnov, O. (2020). An analysis of pedestrian crashes using a spatial count data model. *Papers in Regional Science*, 99(5), 1317–1339. <https://doi.org/10.1111/pirs.12523>
- Asadi, M., Ulak, M. B., Geurs, K. T., Weijermars, W., & Schepers, P. (2022). A comprehensive analysis of the relationships between the built environment and traffic safety in the Dutch urban areas. *Accident; Analysis and Prevention*, 172, 106683. <https://doi.org/10.1016/j.aap.2022.106683>
- Cabrera-Arnau, C., & Bishop, S. R. (2021). Urban population size and road traffic collisions in Europe. *PLoS One*, 16(8), e0256485. <https://doi.org/10.1371/journal.pone.0256485>
- Cabrera-Arnau, C., Prieto Curiel, R., & Bishop, S. R. (2020). Uncovering the behaviour of road accidents in urban areas. *Royal Society Open Science*, 7(4), 191739. <https://doi.org/10.1098/rsos.191739>
- Cai, Q., Abdel-Aty, M., Lee, J., & Eluru, N. (2017). Comparative analysis of zonal systems for macro-level crash modeling. *Journal of Safety Research*, 61, 157–166. <https://doi.org/10.1016/j.jsr.2017.02.018>
- Cai, Q., Lee, J., Eluru, N., & Abdel-Aty, M. (2016). Macro-level pedestrian and bicycle crash analysis: Incorporating spatial spillover effects in dual state count models. *Accident; Analysis and Prevention*, 93, 14–22. <https://doi.org/10.1016/j.aap.2016.04.018>

- Casares Blanco, J., Fernández-Aracil, P., & Ortuño-Padilla, A. (2019). Built environment and tourism as road safety determinants in Benidorm (Spain). *European Planning Studies*, 27(7), 1314–1328. <https://doi.org/10.1080/09654313.2019.1579784>
- Castillo-Manzano, J. I., Castro-Nuño, M., López-Valpuesta, L., & Vassallo, F. V. (2020). An assessment of road traffic accidents in Spain: The role of tourism. *Current Issues in Tourism*, 23(6), 654–658. <https://doi.org/10.1080/13683500.2018.1548581>
- Castro-Nuño, M., & Arévalo-Quijada, M. T. (2018). Assessing urban road safety through multidimensional indexes: Application of multicriteria decision making analysis to rank the Spanish provinces. *Transport Policy*, 68, 118–129. <https://doi.org/10.1016/j.tranpol.2018.04.017>
- Castro-Nuño, M., Castillo Manzano, J. I., & Fageda, X. (2018). The role of road safety in a sustainable urban mobility: An econometric assessment of the Spanish cities. *International Journal of Sustainable Transportation*, 12(3), 205–217. <https://doi.org/10.1080/15568318.2017.1346733>
- Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, 39(5), 829–844. <https://doi.org/10.2307/1909582>
- de España, G. (1985). Ley 7/1985, de 2 de abril, Reguladora de las Bases del Régimen Local. *Boletín Oficial Del Estado*, 80(3), 8945–8964.
- DGT. (2020). Tendencias en la movilidad y siniestralidad en vías urbanas.
- DGT. (2022). DGT in detailed figures. Municipal data 2016, 2017, 2018 and 2019. <https://www.dgt.es/menusecundario/dgt-en-cifras/>
- Esri Inc. (2020). ArcMap (Version 10.8.1). Esri Inc. <https://desktop.arcgis.com/en/arcmap/>
- European Commission. (2021a). Facts and figures seniors. European road safety observatory. Brussels, European commission, directorate general for transport.
- European Commission. (2021b). Road safety thematic report - Seniors. European road safety observatory. Brussels, European commission, directorate general for transport.
- Gálvez-Pérez, D., Guirao, B., Ortuño, A., & Picado-Santos, L. (2022). The influence of built environment factors on elderly pedestrian road safety in cities: The experience of Madrid. *International Journal of Environmental Research and Public Health*, 19(4), 2280. <https://doi.org/10.3390/ijerph19042280>
- Geofabrik. (2022). <https://download.geofabrik.de/africa/canary-islands.html>
- Geofabrik. (2022). <https://download.geofabrik.de/europe/spain.html>
- Hilbe, J. M. (2014). *Modeling count data*. Cambridge University Press.
- IGN. (2022). Geographic reference information. <https://www.ign.es/>
- IGN. (2022). Sistema de Información sobre Ocupación del Suelo de España SIOSE 2014. <https://www.ign.es/>
- INE. (2013). *Population and housing censuses 2011*. Instituto Nacional de Estadística.
- Instituto Nacional de Estadística (INE). (2022). Indicadores de Estructura de la Población. Proporción de personas mayores de cierta edad por provincia. <https://www.ine.es/jaxiT3/Tabla.htm?t=1488>
- International Transport Forum (ITF). (2020). Road Safety annual report. International transport forum.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (vol. 6). Springer.
- Kim, K., Brunner, I. M., & Yamashita, E. Y. (2006). Influence of land use, population, employment, and economic activity on accidents. *Transportation Research Record: Journal of the Transportation Research Board*, 1953(1), 56–64. <https://doi.org/10.1177/0361198106195300107>
- Lee, J., Abdel-Aty, M., Cai, Q., Wang, L., & Huang, H. (2018). Integrated modeling approach for non-motorized mode trips and fatal crashes in the framework of transportation safety planning. *Transportation Research Record: Journal of the Transportation Research Board*, 2672(32), 49–60. <https://doi.org/10.1177/0361198118772704>
- Lee, J., Abdel-Aty, M., de Blasiis, M. R., Wang, X., & Mattei, I. (2019). International transferability of macro-level safety performance functions: A case study of the United States and Italy. *Transportation Safety and Environment*, 1(1), 68–78. <https://doi.org/10.1093/transp/tdz001>
- Lee, J., Abdel-Aty, M., Huang, H., & Cai, Q. (2019). Transportation safety planning approach for pedestrians: An integrated framework of modeling walking duration and pedestrian fatalities. *Transportation Research Record: Journal of the Transportation Research Board*, 2673(4), 898–906. <https://doi.org/10.1177/0361198119837962>
- Lee, S., Yoon, J., & Woo, A. (2020). Does elderly safety matter? Associations between built environments and pedestrian crashes in Seoul, Korea. *Accident; Analysis and Prevention*, 144, 105621. <https://doi.org/10.1016/j.aap.2020.105621>
- Li, Z., Wang, W., Liu, P., Bigham, J. M., & Ragland, D. R. (2013). Using geographically weighted Poisson regression for county-level crash modeling in California. *Safety Science*, 58, 89–97. <https://doi.org/10.1016/j.ssci.2013.04.005>
- Marshall, W. E., & Ferenchak, N. N. (2017). Assessing equity and urban/rural road safety disparities in the US. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, 10(4), 422–441. <https://doi.org/10.1080/17549175.2017.1310748>
- Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, 33(3), 341–365. [https://doi.org/10.1016/0304-4076\(86\)90002-3](https://doi.org/10.1016/0304-4076(86)90002-3)
- Murray, K., & Conner, M. M. (2009). Methods to quantify variable importance: Implications for the analysis of noisy ecological data. *Ecology*, 90(2), 348–355. <https://doi.org/10.1890/07-1929.1>
- Pérez Díaz, J., Abellán García, A., Aceituno Nieto, P., & Ramiro Fariñas, D. (2020). Un perfil de las personas mayores en España, 2020. *Indicadores Estadísticos Básicos. Informes Envejecimiento en Red*, n.º 25, 39p.
- Pew, T., Warr, R. L., Schultz, G. G., & Heaton, M. (2020). Justification for considering zero-inflated models in crash frequency analysis. *Transportation Research Interdisciplinary Perspectives*, 8, 100249. <https://doi.org/10.1016/j.trip.2020.100249>
- Population Reference Bureau (PRB). (2019). Countries with the oldest populations in the world. <https://www.prb.org/resources/countries-with-the-oldest-populations-in-the-world/>
- Quistberg, D. A., Hessel, P., Rodriguez, D. A., Sarmiento, O. L., Bilal, U., Caiaffa, W. T., Miranda, J. J., de Pina, M. D. F., Hernández-Vásquez, A., & Roux, A. V. D. (2022). Urban landscape and street-design factors associated with road-traffic mortality in Latin America between 2010 and 2016 (SALURBAL): An ecological study. *The Lancet. Planetary Health*, 6(2), e122–e131. [https://doi.org/10.1016/S2542-5196\(21\)00323-5](https://doi.org/10.1016/S2542-5196(21)00323-5)
- R Core Team. (2022). R: A language and environment for statistical computing. In *R foundation for statistical computing*. <http://www.R-project.org/>
- Santos, B., & Carvalheira, C. (2019). Pedestrian road accident index for municipalities: The portuguese case. *IOP Conference Series: Materials Science and Engineering*, 603(4), 042084. <https://doi.org/10.1088/1757-899X/603/4/042084>
- Song, Y., Merlin, L., & Rodriguez, D. (2013). Comparing measures of urban land use mix. *Computers, Environment and Urban Systems*, 42, 1–13. <https://doi.org/10.1016/j.compenvurbsys.2013.08.001>
- Stoker, P., Garfinkel-Castro, A., Khayesi, M., Odero, W., Mwangi, M. N., Peden, M., & Ewing, R. (2015). Pedestrian safety and the built environment: A review of the risk factors. *Journal of Planning Literature*, 30(4), 377–392. <https://doi.org/10.1177/0885412215595438>
- Ukkusuri, S., Miranda-Moreno, L. F., Ramadurai, G., & Isa-Tavarez, J. (2012). The role of built environment on pedestrian crash frequency. *Safety Science*, 50(4), 1141–1151. <https://doi.org/10.1016/j.ssci.2011.09.012>
- United Nations, Department of Economic and Social Affairs, Population Division. (2020). World population ageing (ST/ESA/SER.A/444).
- World Health Organization (WHO). (2018). *Global status report on road safety 2018* (No. WHO/NMH/NVI/18.20). World Health Organization.
- Xie, K., Ozbay, K., Kurkcu, A., & Yang, H. (2017). Analysis of traffic crashes involving pedestrians using big data: Investigation of contributing factors and identification of hotspots. *Risk Analysis: An*

- Official Publication of the Society for Risk Analysis*, 37(8), 1459–1476. <https://doi.org/10.1111/risa.12785>
- Yang, Y., Chung, H., & Kim, J. S. (2021). Local or neighborhood? Examining the relationship between traffic accidents and land use using a gradient boosting machine learning method: The case of Suzhou industrial park, China. *Journal of Advanced Transportation*, 2021, 1–30. <https://doi.org/10.1155/2021/8246575>
- Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of Statistical Software*, 27(8), 1–25. <https://doi.org/10.18637/jss.v027.i08>
- Ziakopoulos, A., & Yannis, G. (2020). A review of spatial approaches in road safety. *Accident; Analysis and Prevention*, 135, 105323. <https://doi.org/10.1016/j.aap.2019.105323>

4.4 *Article C: Age-Friendly Urban Design for Older Pedestrian Road Safety: A Street Segment Level Analysis in Madrid*

Article

Age-Friendly Urban Design for Older Pedestrian Road Safety: A Street Segment Level Analysis in Madrid

Daniel Gálvez-Pérez ^{1,*}, Begoña Guirao ¹ and Armando Ortuño ²

¹ Departamento de Ingeniería del Transporte, Territorio y Urbanismo, Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos, Universidad Politécnica de Madrid, Calle del Profesor Aranguren, 3, 28040 Madrid, Spain; begona.guirao@upm.es

² Escuela Politécnica Superior, Universidad de Alicante, 03690 San Vicente del Raspeig, Spain; arorpa@ua.es

* Correspondence: daniel.galvezp@upm.es

Abstract: Walking benefits older pedestrians but exposes them to traffic crashes. With an aging population, designing age-friendly cities is crucial, yet research on older pedestrian safety at a micro-level is limited. This study aims to reduce older pedestrian–vehicle collisions and create more livable environments through infrastructure policies derived from statistical data analysis. Special attention is focused on collecting a holistic set of infrastructure variables to reflect most of the street built environment elements, which helps policymakers implement short-term safety measures. Using Bayesian Poisson regression, this study analyzes factors contributing to the occurrence of crashes involving older and non-older pedestrians on road segments in Madrid, Spain. The results indicate that different factors affect the occurrence of crashes for all pedestrians versus older pedestrians specifically. Traffic crashes involving all pedestrians are affected by leisure points of interest, bus stops, and crosswalk density. Older pedestrian traffic crashes are influenced by population density, the presence of trees and trash containers, and contour complexity. Proposed measures include relocating trees and trash containers, modifying bus stops, and adding crosswalks and traffic lights. This paper also shows that these countermeasures, aimed at creating age-friendly streets for older pedestrians, are not expected to worsen the road safety of other pedestrians.

Keywords: population ageing; age-friendly cities; traffic safety; older adults; older pedestrians



Citation: Gálvez-Pérez, D.; Guirao, B.; Ortuño, A. Age-Friendly Urban Design for Older Pedestrian Road Safety: A Street Segment Level Analysis in Madrid. *Sustainability* **2024**, *16*, 8298. <https://doi.org/10.3390/su16198298>

Academic Editor: Laurie Buys

Received: 6 August 2024

Revised: 11 September 2024

Accepted: 18 September 2024

Published: 24 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Walking offers substantial benefits for older adults (people aged above 65), serving as a cornerstone for maintaining physical health and promoting mental well-being. Regular walking helps to mitigate the risks of chronic diseases, cardiovascular diseases, diabetes, cognitive impairment and dementia, and even improve mental well-being, sleep, and longevity [1]. Moreover, older adults benefit from walking in terms of social inclusion and a reduction in isolation [2,3], which is closely related with their well-being [4–6]. In fact, older adults are particularly vulnerable to loneliness [7] and marginalization because of their age, namely ageism. Ageism can be defined as negative or positive stereotypes, prejudice, and discrimination against older adults [8]. This marginalization was exacerbated during the COVID-19 pandemic due to social isolation and physical limitations [9,10], and ageism was normalized because of the higher level of vulnerability of older adults [11]. Despite the physical activity benefits of walking, this demographic is particularly vulnerable as pedestrians to vehicle collisions.

Older pedestrian (an older adult acting as a pedestrian) road safety is an issue of concern, as the aging population has already impacted crash data. In Europe, the proportion of older adults among those killed in traffic crashes nearly doubled from 17% in 1992 to 29% in 2018 [12]. Specifically in 2020, older pedestrians accounted for 48% of pedestrian fatalities, despite comprising only 21% of the total population. This age group represents

28% of all road-related deaths [13]. Therefore, despite the effects of COVID-19, which significantly reduced traffic volumes, older pedestrians continued to be highly vulnerable on the roads. This vulnerability is particularly pronounced in urban areas. To illustrate, in Spain, 64% of pedestrians killed on urban roads in 2021 were aged 65 or above, while this group represented only 20% of the population [14]. Older pedestrians exhibit characteristics that increase their crash risk, such as chronic medication use, risk-taking behavior, distractions, and diminished self-regulation. In addition, their inherent physical frailty puts them at increased risk of injury [15]. The aim of this study is to identify key factors contributing to traffic crashes within urban scenarios involving older pedestrians and to propose countermeasures to create more age-friendly cities.

The population of older people is expected to grow in the next decades [16], so cities should be prepared to be safer and more inclusive to promote the benefits of walking and avoid social isolation because of poor safety perceptions related to traffic. Although there is an institutional willingness for the inclusion of the older population in public policies [17], the body of research with real traffic crash data is not substantial. Traffic safety studies should be more inclusive to recognize and address the unique needs and risks associated with older pedestrians, as older people are often forgotten among other priorities [18]. There is a need to detect which factors make streets riskier, particularly for older pedestrians, while countermeasures should promote safety but not reduce the operational performance of roads [19], conflict with other goals [20], or reduce the safety of other pedestrians. This is aligned with the definition of an age-friendly city, which is a place that adapts to be more inclusive to the needs of all ages [21] and to promote healthy, active, and successful aging [22], as it is based on minimizing the risk of disability (caused by traffic crashes) and maintaining physical functions (more safety implies more walking activity). This study is particularly in line with the health pillar of active aging [23], especially preventing and reducing disabilities and premature mortality by creating safe environments. Regarding the United Nations' Sustainable Development Goals [24], this approach is aligned particularly with goals 10 ('reduced inequalities') and 11 ('sustainable cities and communities'), and with the European 'Vision Zero' initiative [25], which aims to reduce road traffic fatalities and serious injuries to as close to zero as possible by 2050.

Research on older pedestrian traffic crashes reveals several key factors influencing older pedestrian safety, and methodologically, this research can be divided into four groups: observational studies, surveys, simulations, and crash database analyses. Observational and pedestrian survey analyses [26–29] have highlighted that street crossings are particularly problematic areas, which is shown in the traffic crash data as older adults are over-represented in crashes at intersections, although about 70% of older adult fatalities occur on road segments [30]. For example, Oxley et al. [26] noted increased crossing difficulties in complex environments, like two-way undivided roadways, where older pedestrians struggled to process information from multiple sources, thus being at greater risk compared to younger individuals. This led to the recommendation of installing median strips in such areas to simplify the crossing process, which was supported by findings by other authors [31,32]. Bernhoft et al. [33] showed that older pedestrians prefer routes with signalized intersections and smooth sidewalks, while younger pedestrians focus on the most direct route. Vine et al. [27] found that dangerous pedestrian crossings, cyclist activity, overcrowded or busy road pathways, sidewalks in bad condition, a lack of shading, and public seating undermined older pedestrians' sense of safety. Likewise, Ravi et al. [34] identified older people's preferences for better infrastructure, including better crosswalks with longer green lights, better-maintained sidewalks, and more lighting.

Simulation studies [35,36] underscored that education and training should be complemented with adjustments to infrastructure. Consequently, a focus on real crash data analysis is essential for reducing older pedestrian crashes. For example, early research by Zegeer et al. [31] using data from North Carolina and the U.S. Fatal Pedestrian Crashes database over 11 years highlighted that older adults are more likely to be fatally injured in crashes, particularly at intersections and wide streets due to slower walking speeds and

difficulties in complex traffic situations. A similar study by Martin et al. [37] found that most crashes involving older pedestrians in Ireland occurred during daylight and in good weather conditions. Abou-Raya et al. [32] noted that older pedestrians often did not notice approaching vehicles with falls being a common cause of injuries.

Another group of studies has aimed to identify unique characteristics of older pedestrian crashes using statistical and machine learning techniques. Kim and Ulfarsson [38] utilized random-intercept logistic regression to discover that older pedestrians are more likely to be involved in crashes with older drivers, in scenarios involving vehicle turns, and they are less likely to be involved in crashes during darkness and summer. Das et al. [39] identified that failure to yield while crossing and while a vehicle is turning at intersections were associated with older pedestrians. In particular, females were associated with backing vehicles and males were associated with segment-related crashes and crashes while crossing the expressway at night.

Using traffic crash databases, research into the injury severity of older pedestrian crashes has also been conducted. Wang et al. [40] used an ordered probit model to analyze crash severity in Singapore, finding that more severe crashes tended to occur at night, on roads with high-speed limits, at three-legged intersections, and away from proper crossings, and they were lower at signalized intersections. Laković et al. [41] also highlighted that older men and those over 65 were more likely to suffer severe injuries especially during the day and in traffic conditions that transition from infrequent to normal traffic. These studies, however, did not separate older pedestrian data for specific comparisons with other age groups, which limits the conclusions that can be drawn.

Apart from these previous studies, a detailed built environment analysis using real traffic crash data at the micro-level [42] road segment level is largely missing. This gap in research prevents a comprehensive understanding of the infrastructure features that contribute to the frequency of crashes involving older pedestrians. A built environmental analysis of urban pedestrian crashes can be addressed at a territorial macro-level (e.g., at district or census tract level) or at a territorial micro-level (e.g., street intersection or street segment level). As explained below, micro-level studies are less common in the literature than macro-level ones or those only including exploratory analyses, and their scope is also quite different.

On the one hand, macro-level studies often analyze the number of crashes in larger spatial units like districts or neighborhoods. For example, Dumbaugh et al. [43] found that the number of older adults, arterial roads and big box stores increased crashes involving older pedestrians, while dense networks of lower-speed roads reduced them. Social equity in pedestrian crashes was analyzed by Lee et al. [44], studying the number of older and non-older pedestrians in Seoul (Korea) at the census block group level. This research showed that four-way intersections and crosswalks impacted the number of pedestrian crashes and the severity of older pedestrians only in low-priced-housing areas, implying that pedestrian measures may be unevenly distributed across the city. Research by Gálvez-Pérez et al. [45], developed at the city district level in Madrid (Spain), found that built environment features such as vehicle flow, the presence of traffic lights, and sidewalk surface had a more statistically significant effect on older pedestrian collisions than on younger pedestrian crashes. A more recent study by Gálvez-Pérez et al. [46], using all the Spanish municipalities as the case study, concluded that municipalities above 50,000 inhabitants were safer for older pedestrians, and that more populated provincial capitals had lower rates of older pedestrian traffic crashes. They also stated that countermeasures to improve the road safety of older pedestrians are not likely to worsen the safety of the rest of the pedestrians.

In relation to the scope of these territorial studies, the macro-level approach is suitable for medium- and long-term transportation planning, but short-term measures should be applied to street infrastructure, which is considered in these studies in an aggregated way. The specific location where the crashes take place is not considered with this spatial unit approach; thus, it is not possible to determine which specific road elements should be treated or redesigned.

Conversely, micro-level studies focus on specific road elements like segments or intersections. This approach allows for a more detailed examination of where crashes occur and the factors contributing to them. Unfortunately, due to data limitations, such detailed studies are scarce. Two exceptions are the paper by Kim [47] and Lv et al. [48]. Kim [47] analyzed pedestrian traffic crashes at intersections in the County of Los Angeles (U.S.). Multinomial logistic regression was used to examine factors related to intersections with a high concentration of older and younger pedestrian collisions. The research found that certain features like raised medians and proximity to recreational areas improved safety for older pedestrians, whereas bus stops increased collision risks. However, the results of this study cannot be extended to road segments, where a large number of crashes take place. Lv et al. [48] studied the number of older pedestrian traffic crashes occurring at each road segment in a Shanghai district using Poisson and Geographically Weighted Poisson regressions, revealing that the safety of older pedestrians was significantly influenced by the presence of green spaces, sidewalks, and road junctions. Additionally, this study identified roads near nursing homes, schools, bus stops, metro stations, traditional markets, and supermarkets as particularly dangerous for older pedestrians. However, the research did not compare results with traffic crashes involving the rest of the pedestrians (aged below 65). This omission is crucial for policymakers, as some road features might be riskier for all pedestrians, while others could affect older adults only. If only crashes involving older adults are studied, special measures for this demographic that should be applied to the entire population might be proposed. Furthermore, road segments are defined as 200 m segments and might include road junctions. Hence, this study includes crashes on road segments and road junctions but analyzes only the characteristics of the road segments. Finally, although they use a novel data collection technique through street-view imagery, some physical features (variables) that might interact with pedestrians were not included in the study.

In this literature context, there is an absence of studies on older pedestrian road safety at the street segment level covering a holistic set of infrastructure variables to reflect most of the elements present in the streets, such as the presence of trees, trash containers, fences or terraces, and the intersections at the edges of the segment. There is also a lack of literature comparing older and non-older pedestrian traffic crashes at the segment level to identify factors related to the occurrence of all pedestrian crashes and those that affect specifically older pedestrians. The objective of this study is to identify critical factors influencing traffic crashes involving older pedestrians and propose countermeasures to create more age-friendly cities, especially focusing on infrastructure features. This approach holds significant value for policymakers and urban planners, as it will enable the identification of short-term infrastructure measures to mitigate older pedestrian traffic crashes. At this stage, traffic crashes on road segments were studied, as in Spanish urban roads, nearly two thirds of pedestrians were killed or injured on road segments [49], outside of intersections. The methodology relies on a Poisson model in a Bayesian framework utilizing the Integrated Nested Laplace Approximation (INLA) approach, which is employed to predict separately older and non-older pedestrian traffic crashes per road segment. The analysis utilizes a dataset spanning a 5-year period of pedestrian crashes in Madrid, Spain, in which 872 road segments have been studied.

The structure of this paper is as follows: Section 1 comprises the introduction, Section 2 describes the materials and methods, including the case study, database elaboration and description, and the statistical modelling approach; Section 3 includes the modeling results; Section 4 contains the discussion, and finally, Section 5 provides the conclusions, including future research lines.

2. Materials and Methods

2.1. Madrid Case Study

Spain is among the most aged countries globally, alongside Japan, Finland, Sweden, Greece, Italy, and Germany. In 2022, older adults (those over 65) represented 20% of the

Spanish population with 6% over 80. Spain has a projected increase in the population over 65 from 20% in 2022 to 29% by 2040, and it also has notable life expectancies of 81 years for men and 87 years for women [16].

Regarding traffic crashes, older adults in Spain have a substantially higher fatality rate (number of dead people over number of injured people in traffic crashes) compared to the rest of the population (3.3 vs. 1.1). In 2021, the majority of older pedestrian deaths (82%) occurred in urban areas with 118 of the 144 total deaths [14]. This is higher than the 61% of all pedestrian traffic crashes that occur on urban roads, indicating older adults might walk more on urban roads. As stated by DGT [49], nearly two thirds of pedestrian fatalities in urban areas occurred on road segments (outside intersections) between 2016 and 2018 in Spain. These figures highlight the importance of studying traffic crashes involving older pedestrians on urban street segments.

Madrid, as Spain's most populated city with a diverse population distribution, is a suitable case study. In 2022, the city had about 3.3 million residents, with those over 65 and over 80 comprising 20.3% and 7.1% of the population, respectively, which is above the national average. The city is divided into 21 districts and 131 neighborhoods. To illustrate these differences, Figure 1 shows the proportion of older adults in each neighborhood in 2022.

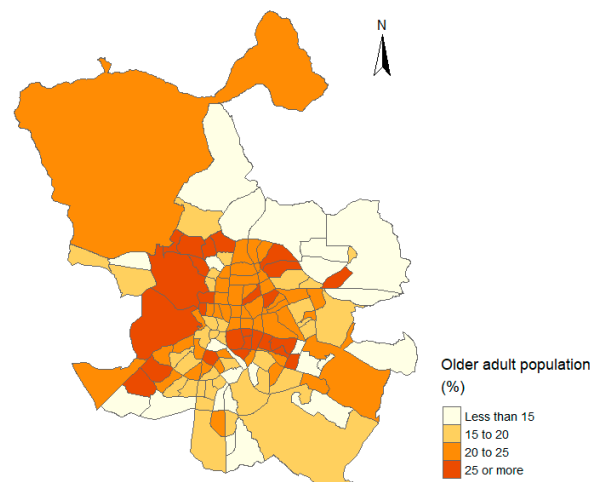


Figure 1. Proportion of older adults per city neighborhood (2022).

This study analyzed comprehensive traffic crash data over five years (2017, 2018, 2019, 2021, and 2022), excluding 2020 due to the COVID-19 pandemic's impact on mobility patterns [50]. The data focused on pedestrian crashes on road segments, excluding intersection-related crashes. Street segments are defined as areas between road junctions, so no road junctions or crashes related to them were included in this analysis. Only single vehicle-pedestrian collisions (one pedestrian and one vehicle) were analyzed to focus on a single type of incident and to simplify assigning the pedestrian's age. Older pedestrians were defined as those aged 65 and above. The final dataset included 3535 traffic crashes with 842 involving older pedestrians and 2692 involving non-older pedestrians. Figure 2 shows the locations of these crashes during the study period.

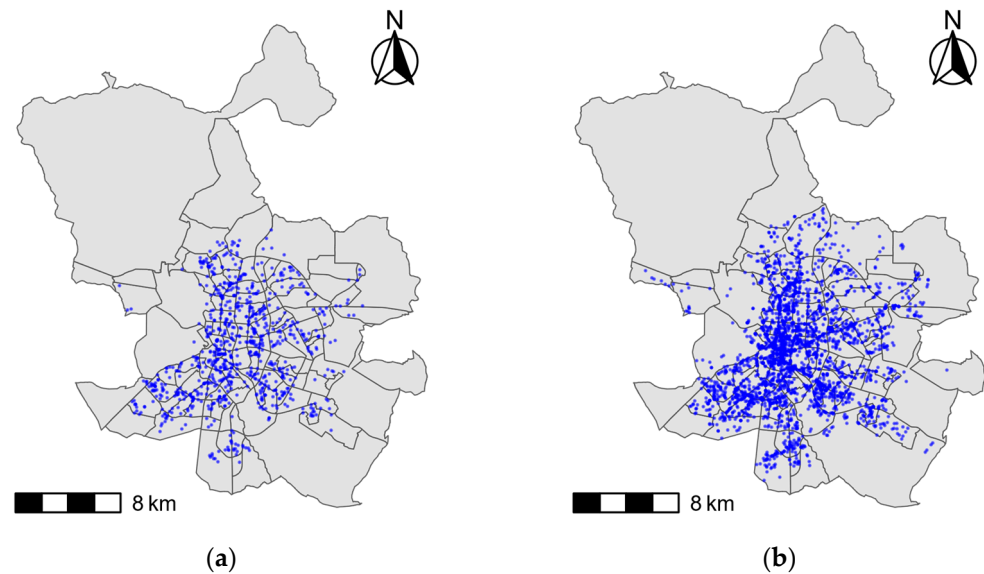


Figure 2. Location of traffic crashes suffered by older (a) and non-older (b) pedestrians in Madrid (2017–2019, and 2021–2022).

2.2. Database Description, Collection and Processing

2.2.1. Overview

Data collection and processing for this study involved intricate procedures. Detailed explanations are provided in the subsequent sections, but an overview is presented here for clarity and reproducibility, as illustrated in Figure 3.

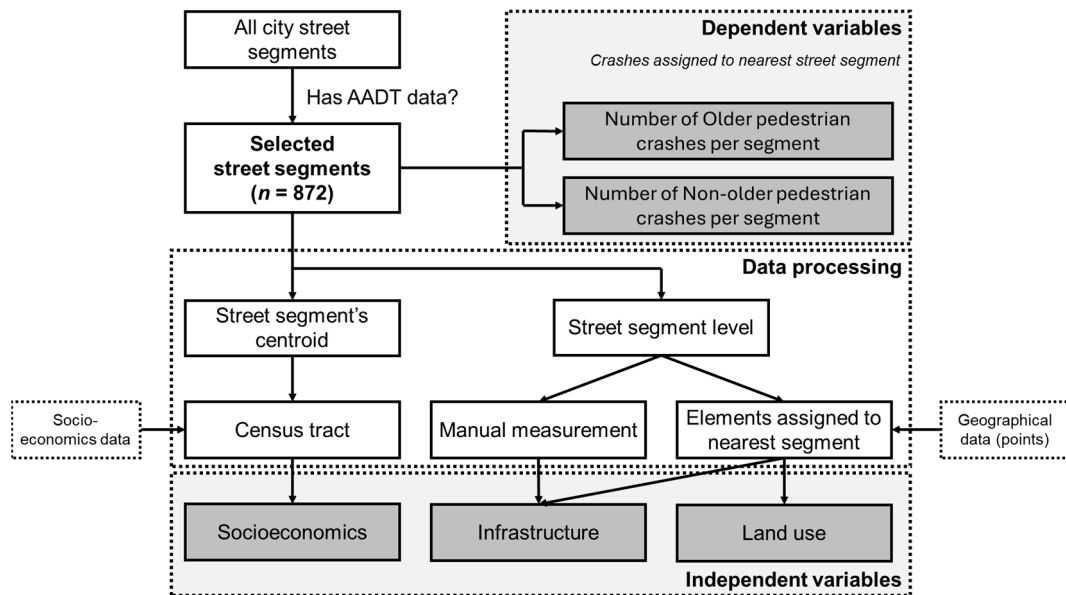


Figure 3. Scheme of the data collection approach.

Initially, all city road segments were considered. For this study, a subset of 872 road segments was selected because they had available Average Annual Daily Traffic (AADT) data, which provided crucial information for our analysis.

The dependent variables were identified by assigning crashes to the nearest road segment, categorizing them into older pedestrian crashes (individuals aged 65 and above) and non-older pedestrian crashes. This categorization helped with understanding the different factors that might affect these two groups differently. The number of crashes for each group was then calculated for each road segment.

Independent variables were collected for each road segment from various sources, systematically when possible, due to the high number of studied segments and to favor reproducibility. Socioeconomic data were derived from the census tract of the road segment's centroid. Land use data, including points of interest (POIs) such as public, health, education, and leisure facilities, were assigned to the nearest road segment. Infrastructure data were obtained similarly whenever possible. When not feasible, such as with sidewalk width, manual measurements were taken to provide detailed information about each road segment's physical characteristics.

2.2.2. Road Segment Selection

A crucial variable for each studied street segment is the AADT. In Madrid, there are more than two thousand AADT measurement points available for 2019. Two-way street segments had two measurement points to distinguish between traffic flows. The geometry of these road segments was carefully reviewed and corrected if necessary. Road segments containing tunnel entrances and bridges were excluded to avoid erroneous data collection. The remaining road segments all had sidewalks on both sides, except for five cases, which were removed from the sample. Therefore, only road segments with two sidewalks were considered. Finally, 872 road segments were included in this study. Despite this selection procedure, these segments are widely spread throughout the city (Figure 4), representing the most common segment types in a city.

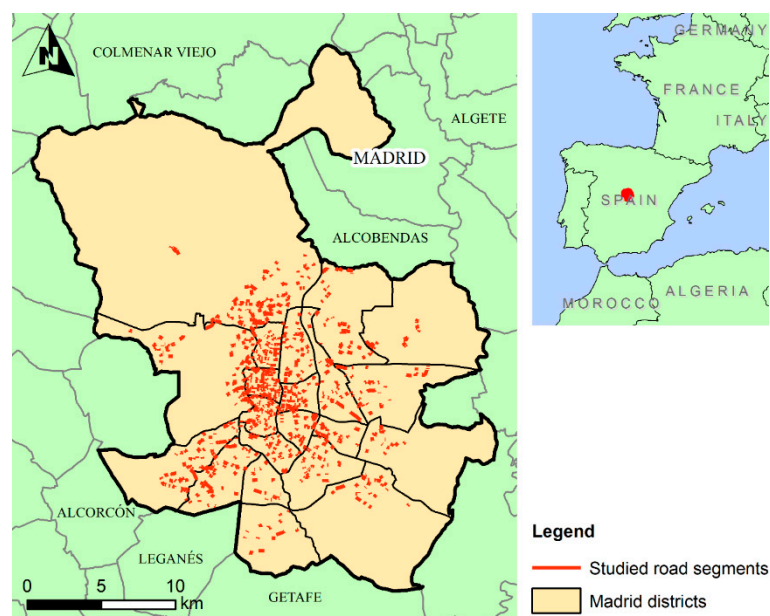


Figure 4. Map of road segments included in the study.

2.2.3. Independent Variables

Independent variables were collected using established variables from the literature and additional factors deemed to influence the road safety of older pedestrians. This section describes the independent variables with the next section outlining data sources and data preparation.

The exposure variables are the length and Annual Average Daily Traffic (AADT) of the road segment, which are commonly used in road safety studies. Although pedestrian flow data would be ideal, they are only available for a few segments and do not account for the age of pedestrians [51]. Consequently, we could not include any pedestrian exposure variable, as is the case in other pedestrian road safety studies [52].

For socioeconomics, the population density and per capita income of the segment's census tract were considered. These variables capture features that are difficult to measure directly, such as economic well-being. Higher-income areas typically have better infras-

structure maintenance, so wealthier neighborhoods are expected to have better-maintained sidewalks and road surfaces.

The studied land use features are the points of interest (POIs) at the segment level. A POI is a specific location of significance within a geographic area. In this study, public (e.g., a police station), education and health (e.g., a school), and leisure (e.g., a restaurant) POIs were considered. Land use independent variables include the type-based number of POIs and the POIs distribution disparity index (PDDI) along the street segment. The PDDI (Equation (1)) is an ad hoc designed index that takes values from 0 (totally balanced distribution) to 1 (totally imbalanced distribution). If there are no POIs on the street segment, the PDDI value is set to 0.

$$PDDI_i^k = \begin{cases} 0, & \text{if } POIs_{i,A}^k = 0 \text{ and } POIs_{i,B}^k = 0 \\ \text{abs} \left(\frac{POIs_{i,A}^k - POIs_{i,B}^k}{POIs_{i,A}^k + POIs_{i,B}^k} \right), & \text{otherwise} \end{cases} \quad (1)$$

where $PDDI_i^k$ is the imbalance index of the POIs of the k th type in the i th road segment, $POIs_{i,A}^k$ is the number of POIs of type k in the side A of the i th road segment and $POIs_{i,B}^k$ is the number of POIs of type k in the side B of the i th road segment.

Infrastructure features were classified based on three primary issues regarding older pedestrians: ease of walking on the sidewalk, ease of crossing the street, and ease of detecting traffic. Certain features may fit into multiple categories, but they are grouped by their most pronounced impact. Figure 5 shows a diagram of the considered infrastructure variables.

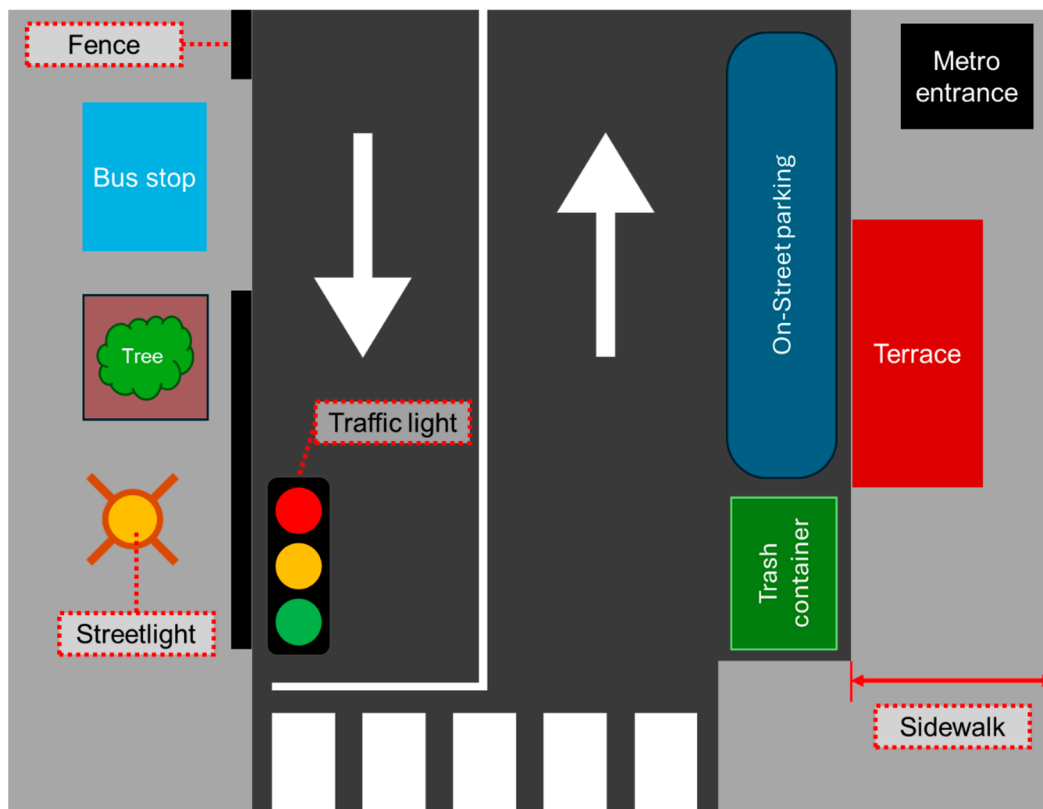


Figure 5. Infrastructure variables considered in the study of street segments.

Factors influencing the ease of walking on the sidewalk include sidewalk width and distribution (Equation (2)), the presence of trees, terraces, bus stops, metro entrances, streetlights, and street mean slope. Factors affecting the ease of crossing the street include

the number of lanes, crosswalk density, traffic light density, and the presence of pedestrian fences. Factors influencing the ease of detecting traffic include the number of ways, the presence of on-street parking, and the presence of trash containers. An interaction term between the number of ways and the presence of a median was proposed based on the conclusions by Oxley et al. [26]. Detecting and processing traffic on a two-way street is easier on divided streets with medians. Hence, a two-way divided street is more similar to a one-way street than to a two-way undivided street.

$$\text{sidewalk distribution} = \text{abs} \left(\frac{\text{sidewalk width A} - \text{sidewalk width B}}{\text{sidewalk width A} + \text{sidewalk width B}} \right) \quad (2)$$

The contour conditions of road segments were considered, focusing on the intersection types at both ends of each segment. At three-legged intersections, pedestrians on the sidewalk opposite the third leg (pedestrian A, Figure 6) have an advantage in detecting oncoming traffic compared to those on the other sidewalk (pedestrian B, Figure 6), who must turn their heads more. Older pedestrians, who often have difficulty with neck rotation [15,53], are particularly affected. This disadvantage is mitigated at intersections with more than three legs. Intersections were categorized into three types: three legs, four legs, and more than four legs or roundabouts. Including these variables in the statistical model is challenging because the order of the intersections must be considered (which intersection is Intersection 1 and which is Intersection 2, Figure 6). To address this, these variables were transformed into a single feature representing the road segment's level of complexity, ranging from A to C, as shown in Table 1.

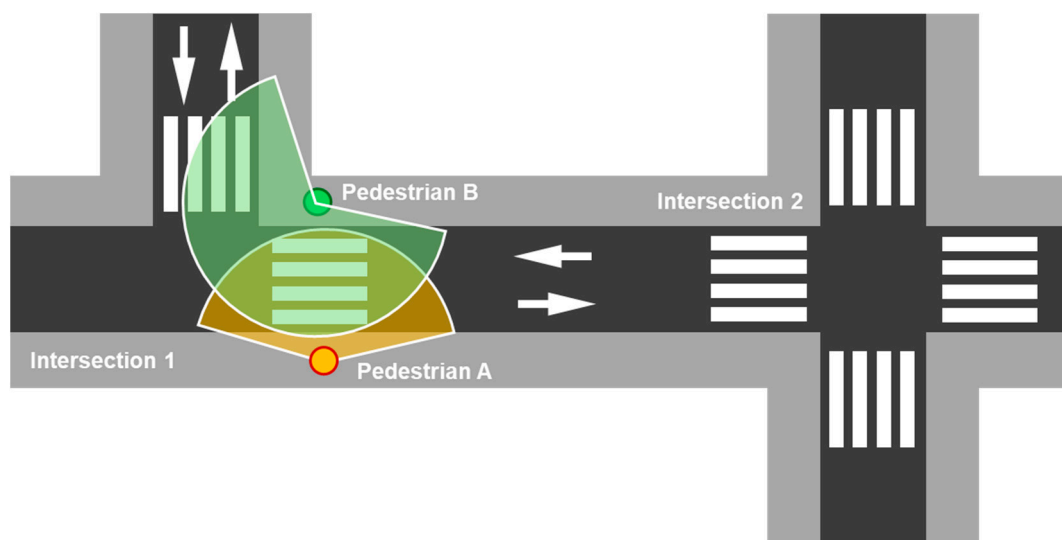


Figure 6. Example of street segment with contour complexity B and pedestrians' field of view.

Table 1. Combinations of possible types of intersections at street edges and definition of level of complexity of the road segment contour.

Intersection 1	Intersection 2	
	3 Legs	+3 Legs or Roundabout
3 Legs	A	B
+3 Legs or Roundabout	B	C

2.2.4. Data Preparation

Data preparation consumed a significant portion of the overall time for this article. Independent variables were gathered from multiple data sources, involving both spatial and tabular data. Spatial data were processed using ArcMap software [54] (version 10.8.1), while

tabular data were managed using R software [55] (version 4.4.1). Wherever possible, the characteristics of the studied street segments were acquired through automatic processes. Table 2 shows the descriptions of the collected variables.

Table 2. Variable description.

Variable	Description	Unit
Dependent variables		
Older pedestrian traffic crashes	Number of older pedestrian crashes in 5-year period	#
Non-older pedestrian traffic crashes	Number of non-older pedestrian crashes in 5-year period	#
Independent variables		
Exposure		
Length	Length of the road segment	km
AADT	Annual average daily traffic in the road segment	veh/day
Socioeconomics		
Population density	Population density of the census tract of the centroid of the road segment	inh/ha
Per capita income	Per capita income of the census tract of the centroid of the road segment	€
Land use		
Public POIs	Public POIs per segment length	#/km
Health and education POIs	Health and education POIs per segment length	#/km
Leisure POIs	Leisure POIs per segment length	#/km
Public PDDI	Public PDDI	
Health and education PDDI	Health and education PDDI	
Leisure PDDI	Leisure PDDI	
Infrastructure		
<i>Ease of walking on the sidewalk</i>		
Sidewalk width	Mean width of both sidewalks	m
Sidewalk disparity	Sidewalk disparity index (Equation (2))	
Trees	If trees in sidewalk = 1, otherwise = 0	
Terrace	If terraces in sidewalk = 1, otherwise = 0	
Bus stop	If bus stops in sidewalk = 1, otherwise = 0	
Metro entrance	If metro entrances in sidewalk = 1, otherwise = 0	
Streetlight	If streetlights in sidewalk = 1, otherwise = 0	
Slope	Mean slope of the road segment	%
<i>Ease of crossing the street</i>		
Number of lanes	Total number of traffic lanes of the road segment	#
Crosswalk density	Number of pedestrian crosswalks per road segment length	#/km
Traffic lights	If traffic lights in segment = 1, otherwise = 0	
Pedestrian fence	If pedestrian fence in sidewalk and/or in median = 1, otherwise = 0	
<i>Ease of detecting traffic</i>		
Number of ways & median	One-way = 0, Two-ways and no median = 1, Two-ways and median = 2	
On-street parking	If on-street parking = 1, otherwise = 0	
Trash containers	If trash containers = 1, otherwise = 0	
Contour complexity	Contour level of complexity A to C (Table 1)	

Dependent variables, including the count per segment of traffic crashes involving both older and non-older pedestrians, were determined by assigning each crash to its closest road segment. The count of crashes for each road segment was then tabulated and analyzed.

For **exposure** variables, the length of each road segment was calculated using geometry from Madrid statistics in GIS format. The AADT was extracted from Madrid open data for 2019, excluding 2020 due to COVID-19 impacts. We meticulously curated this data, ensuring each point was as near as possible to the respective road segment.

Socioeconomic data were linked to each road segment by associating it with the census tract of its central point. The population density and per capita income from the census tract were then assigned to the corresponding road segment.

For **land use**, points of interest (POIs) in Madrid were extracted from OpenStreetMap (OSM [56]). POIs were categorized into public, education and health, and leisure groups. Each POI was assigned to the nearest road segment, capturing the distance and angle from each POI to the segment using ArcMap's 'Generate Near Table' tool. Data processing derived the count of POIs by type on each side of the road segment, depending on the positive or negative sign of the angle (Figure 7). A threshold of 300 m was used to assign a POI to its nearest road segment. The categorized POI counts and the categorized POI distribution disparity index (PDDI) were computed.

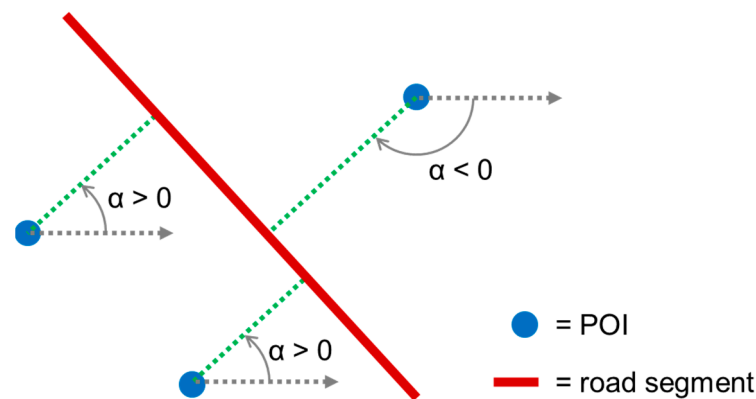


Figure 7. Sign convention for assigning the side of the POIs in a road segment.

Infrastructure features constituted most of the independent variables. Some variables were automatically collected, including bus stops, metro entrances, terraces, streetlights, traffic lights, on-street parking areas, and trash containers. However, the acquisition of the remaining variables could not be automated due to reliability concerns. For instance, the measurement of sidewalk width faced challenges due to the road axis geometry not perfectly aligning with the sidewalks.

Table 3 presents the summary statistics for the sample variables. Linear relationships among the independent variables were assessed using different techniques, depending on the nature of the variables. For pairs of numerical variables, Pearson's correlation coefficient was applied; for pairs of dichotomous categorical variables, the phi coefficient was used; and for pairs of numerical and dichotomous categorical variables, the point-biserial correlation coefficient was employed. These three coefficients share a common interpretation: a value of ± 1 indicates a perfect linear relationship between the variables, while a value of 0 indicates no relationship. Additionally, categorical variables with more than two levels (i.e., road type and contour complexity) were excluded from the analysis. The correlation matrix is presented in Table 4. Strong correlations (above 0.80) were identified for two pairs of variables: (1) the density of health and education POIs and health and education PDDI, and (2) the density of public POIs and public PDDI. These correlations are likely due to the fact that streets typically have either no health and education or public POIs, or only one in most cases. If there are 0 POIs, the PDDI is 0; if there is only 1 POI, the PDDI is 1, as that POI is necessarily located on one side of the street. Consequently, the PDDI for health and education POIs, as well as the PDDI of public POIs, does not add information to the database in most cases, and these variables have been excluded from the modeling process. There is also a notable correlation between AADT and the number of lanes (0.72); however, both variables were included in the modeling process, as AADT is a measure of crash risk exposure.

Table 3. Summary statistics of the sample.

Variable	Mean	SD	Median	Min.	Max.
Older pedestrian crashes	0.10	0.31	0	0	2
Non-older pedestrian crashes	0.27	0.58	0	0	4
Length	0.15	0.10	0.12	0.02	0.99
AADT	11,632.71	10,110.38	8821	160	72,331
Population density	240.85	182.09	219.37	0.07	881.77
Avg. income per person	20,133.15	7282.38	19,261	5801	32,183
POI density leisure	14.32	23.88	4.29	0	176.23
POI density health education	1.86	4.15	0	0	29.56
POI density public	1.16	3.69	0	0	35.23
PIID leisure	0.36	0.44	0	0	1
PIID health education	0.20	0.40	0	0	1
PIID public	0.12	0.32	0	0	1
Sidewalk width	4.26	2.14	3.8	0.9	20
Sidewalk distribution	0.12	0.15	0.07	0.00	0.75
Trees	0.70				
Terraces	0.31				
Bus stop	0.36				
Metro entrance	0.04				
Streetlights	0.99				
Slope	0.02	0.02	0.02	0.00	0.14
Number of lanes	3.20	1.82	3	1	11
Traffic lights	0.69				
Crosswalk density	10.64	6.45	10.32	0.00	40.7
Fence	0.13				
Road type					
One-way	0.31				
Two-way no median	0.46				
Two-way median	0.23				
On-street parking	0.77	0.42	1	0	1
Trash container	0.42	0.49	0	0	1
Contour complexity					
Complexity A	0.22				
Complexity B	0.43				
Complexity C	0.35				

Table 4. Correlation matrix of the independent variables of the sample.

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]
[1] Length	-	-0.01	-0.29	0.05	-0.15	-0.06	-0.07	0.01	0.05	0.03	-0.07	-0.04	-0.14	0.01	0.18	-0.04	-0.09	-0.07	0.02	0.08	-0.27	-0.12	0.04	0.13
[2] AADT	-0.01	-	-0.04	0.15	0.08	0.08	0.02	0.08	0.12	0.02	0.32	0.02	0.18	0.20	0.23	0.12	0.02	-0.10	0.72	0.40	-0.18	0.24	-0.24	-0.14
[3] Population density	-0.29	-0.04	-	-0.27	0.26	0.16	0.12	0.15	0.12	0.09	-0.04	-0.14	0.10	0.22	-0.09	0.09	0.13	0.03	-0.11	0.00	0.14	0.16	0.11	0.07
[4] Avg. income per person	0.05	0.15	-0.27	-	0.06	0.03	-0.03	0.08	0.06	0.02	0.10	-0.05	0.09	0.11	0.03	-0.03	-0.05	-0.03	0.11	0.05	-0.02	0.00	-0.08	-0.21
[5] Leisure POIs	-0.15	0.08	0.26	0.06	-	0.30	0.10	0.24	0.20	0.09	0.00	-0.11	0.09	0.36	-0.03	0.08	0.06	-0.07	0.00	0.08	0.10	0.21	-0.07	-0.07
[6] Health and education POIs	-0.06	0.08	0.16	0.03	0.30	-	0.02	0.10	0.81	0.02	-0.04	-0.12	0.09	0.17	0.02	0.04	0.05	0.00	0.01	0.07	0.03	0.16	0.03	-0.01
[7] Public POIs	-0.07	0.02	0.12	-0.03	0.10	0.02	-	0.08	0.02	0.80	0.04	-0.03	0.06	0.08	0.04	0.05	0.03	0.00	0.00	0.05	0.05	0.06	0.01	0.04
[8] Leisure PDDI	0.01	0.08	0.15	0.08	0.24	0.10	0.08	-	0.12	0.13	0.08	-0.10	0.09	0.23	0.10	0.07	0.04	-0.06	0.05	0.11	0.01	0.12	0.03	0.07
[9] Health and education PDDI	0.05	0.12	0.12	0.06	0.20	0.81	0.02	0.12	-	0.05	-0.03	-0.12	0.06	0.21	0.08	0.06	0.05	0.01	0.05	0.08	0.01	0.06	0.06	0.08
[10] Public PDDI	0.03	0.02	0.09	0.02	0.09	0.02	0.80	0.13	0.05	-	0.06	-0.04	0.05	0.13	0.08	0.01	0.04	0.00	0.02	0.07	0.02	0.06	0.01	0.08
[11] Sidewalk width	-0.07	0.32	-0.04	0.10	0.00	-0.04	0.04	0.08	-0.03	0.06	-	0.21	0.38	0.17	0.17	0.13	-0.06	-0.13	0.45	0.23	-0.07	0.10	-0.05	-0.02
[12] Sidewalk disparity	-0.04	0.02	-0.14	-0.05	-0.11	-0.12	-0.03	-0.10	-0.12	-0.04	0.21	-	-0.08	-0.08	-0.02	-0.02	-0.05	0.05	0.01	-0.01	-0.06	-0.04	0.02	-0.01
[13] Trees	-0.14	0.18	0.10	0.09	0.09	0.09	0.06	0.09	0.06	0.05	0.38	-0.08	-	0.19	0.09	0.07	0.02	-0.08	0.26	0.12	0.00	0.10	0.02	0.00
[14] Terraces	0.01	0.20	0.22	0.11	0.36	0.17	0.08	0.23	0.21	0.13	0.17	-0.08	0.19	-	0.14	0.12	0.07	-0.11	0.16	0.19	-0.01	0.12	0.03	0.14
[15] Bus stop	0.18	0.23	-0.09	0.03	-0.03	0.02	0.04	0.10	0.08	0.08	0.17	-0.02	0.09	0.14	-	0.10	0.01	-0.07	0.27	0.29	-0.06	0.07	-0.10	0.00
[16] Metro entrance	-0.04	0.12	0.09	-0.03	0.08	0.04	0.05	0.07	0.06	0.01	0.13	-0.02	0.07	0.12	0.10	-	0.02	-0.02	0.13	0.12	0.02	0.20	-0.09	-0.02
[17] Streetlights	-0.09	0.02	0.13	-0.05	0.06	0.05	0.03	0.04	0.05	0.04	-0.06	-0.05	0.02	0.07	0.01	0.02	-	0.01	0.04	0.02	0.10	0.04	0.12	0.03
[18] Slope	-0.07	-0.10	0.03	-0.03	-0.07	0.00	0.00	-0.06	0.01	0.00	-0.13	0.05	-0.08	-0.11	-0.07	-0.02	0.01	-	-0.14	-0.09	0.05	-0.06	-0.03	-0.04
[19] Number of lanes	0.02	0.72	-0.11	0.11	0.00	0.01	0.00	0.05	0.05	0.02	0.45	0.01	0.26	0.16	0.27	0.13	0.04	-0.14	-	0.40	-0.18	0.29	-0.22	-0.16
[20] Traffic lights	0.08	0.40	0.00	0.05	0.08	0.07	0.05	0.11	0.08	0.07	0.23	-0.01	0.12	0.19	0.29	0.12	0.02	-0.09	0.40	-	0.03	0.16	-0.15	-0.03
[21] Crosswalk density	-0.27	-0.18	0.14	-0.02	0.10	0.03	0.05	0.01	0.01	0.02	-0.07	-0.06	0.00	-0.01	-0.06	0.02	0.10	0.05	-0.18	0.03	-	0.03	0.05	-0.02
[22] Fence	-0.12	0.24	0.16	0.00	0.21	0.16	0.06	0.12	0.06	0.06	0.10	-0.04	0.10	0.12	0.07	0.20	0.04	-0.06	0.29	0.16	0.03	-	-0.36	-0.19
[23] On-street parking	0.04	-0.24	0.11	-0.08	-0.07	0.03	0.01	0.03	0.06	0.01	-0.05	0.02	0.02	0.03	-0.10	-0.09	0.12	-0.03	-0.22	-0.15	0.05	-0.36	-	0.38
[24] Trash container	0.13	-0.14	0.07	-0.21	-0.07	-0.01	0.04	0.07	0.08	0.08	-0.02	-0.01	0.00	0.14	0.00	-0.02	0.03	-0.04	-0.16	-0.03	-0.02	-0.19	0.38	-

2.3. Statistical Modelling

2.3.1. Model Components

In traffic crash data modeling, current state-of-the-art approaches involve using a model that includes crash exposure, crash risk, a heterogeneity term, and a spatial autocorrelation term. The data generation process commonly follows a Poisson [57], Poisson gamma [58–61], also known as Negative Binomial, or a Poisson log-normal [62–64] distribution.

Crash exposure is typically measured using Annual Average Daily Traffic (AADT) and segment length, or vehicle–kilometers (a product of both). In pedestrian road safety studies, pedestrian flow volume is also an important factor. Researchers have used pedestrian counts [52], available only for segments with pedestrian sensors, and pedestrian density [65], which is measured in average walking meters per surface per day. However, the lack of such data is common, making it difficult to include pedestrian flow in studies [58,60]. Infrastructure factors affecting crash risk include exposure, road type, road surface, road environment, the presence of work zones, road alignment, cross-section, and traffic control measures [66].

The heterogeneity term accounts for unobserved heterogeneity by adding a unique term to each road element. The spatial term is usually calculated using a conditional autoregressive (CAR) model [67]. Other approaches include simultaneous autoregressive (SAR) models, generalized linear mixed models, generalized estimation equations [68], and geographically weighted regression [59]. The CAR model using a binary first-order neighborhood matrix, as defined by Wall [69], shows fair results [70] and is one of the most used models [63,71,72]. However, this approach is not suitable in our case, mostly due to the lack of neighboring segments for most studied segments, which is caused by selecting only those segments with available traffic volume data. Approximately 73% of these segments have no neighbors, and 21% have only one, making the introduction of a spatial term in the model unsuitable.

2.3.2. Model Formulation

The modeling process was conducted within a Bayesian framework using the INLA (Integrated Nested Laplace Approximations) approach, as proposed by Rue et al. [73]. INLA performs inference both faster and more accurately than the traditional Markov Chain Monte Carlo (MCMC) approach [74]. INLA was implemented using the R-INLA package (<http://www.r-inla.org>, accessed on 30 January 2024) in R software [55] (version 4.4.1).

In this study, the data generation process was set to follow a Poisson distribution, which is suitable for count data like traffic crashes. The number of traffic crashes in a road segment (Y_i) depends on the Poisson parameter (λ_i), as shown in Equation (3).

$$Y_i | \lambda_i \sim \text{Poisson}(\lambda_i) \quad (3)$$

The logarithm of the Poisson parameter (λ_i) is modeled as a linear function of the covariates (Equation (4)).

$$\log(\lambda_i) = \alpha + \log(L) + \gamma \log(\text{AADT}) + x_i^T \beta + v_i \quad (4)$$

where α is the intercept, γ is the parameter of the exposure variables AADT , x_i is the vector of covariates of the i th segment, and β is the vector of parameters of the covariates. Note that the exposure variable L is multiplied by 1. Finally, v_i is the heterogeneity term (Equation (5)).

$$v_i \sim N\left(0, \frac{1}{\tau_v}\right) \quad (5)$$

where τ_u is the precision of the normal distribution (inverse of the variance). Each v_i is sampled from the same normal distribution. The hyperprior distribution of the precision is a log-gamma distribution with parameters 1 and 0.001 [$\tau_u \sim \text{loggamma}(1, 0.001)$].

3. Results

This section outlines the results derived from statistical modeling. Unlike deterministic methods, Bayesian statistics express uncertainty directly through inferred posterior distributions of the parameters. The mean of each parameter is presented along with its 95% and 90% Bayesian credible intervals (BCIs), which are represented by the 2.5% and 97.5% quantiles, and 5% and 95% quantiles of the posterior distribution, respectively. The effect of a variable is considered credible at a specific level when the corresponding BCI does not include zero. Unlike confidence intervals in frequentist statistics, which provide a range that would contain the true effect in a certain percentage of repeated samples, BCIs directly reflect the degree of certainty about the parameter's value given the observed data. These results are illustrated in Table 5 (older adult pedestrian crashes model) and Table 6 (non-older adult pedestrian crashes model).

Table 5. Main statistics of the older adult pedestrian model (credible variables at 10% in bold).

Variable	Mean	SD	2.5%	5%	95%	97.5%
Intercept	−5.691	2.093	−9.795	−9.135	−2.247	−1.587
log(AADT)	0.470	0.212	0.055	0.121	0.818	0.885
Population density	0.002	0.001	0.001	0.001	0.003	0.003
Avg. income per person	−0.000	0.000	−0.000	−0.000	0.000	0.000
Leisure POIs	0.011	0.004	0.003	0.004	0.019	0.020
Health and education POIs	0.024	0.025	−0.026	−0.018	0.066	0.074
Public POIs	0.010	0.024	−0.036	−0.029	0.049	0.057
Leisure PDDI	−0.031	0.268	−0.557	−0.472	0.410	0.495
Sidewalk width	−0.052	0.069	−0.187	−0.165	0.061	0.083
Sidewalk disparity	1.023	0.835	−0.614	−0.351	2.396	2.659
Trees	0.501	0.295	−0.078	0.016	0.987	1.080
Terraces	0.009	0.267	−0.514	−0.430	0.447	0.531
Bus stop	0.484	0.239	0.015	0.090	0.877	0.952
Metro entrance	0.342	0.407	−0.457	−0.328	1.011	1.140
Streetlights	−0.212	1.068	−2.307	−1.970	1.546	1.882
Slope	−4.346	6.645	−17.376	−15.281	6.589	8.684
Number of lanes	0.002	0.109	−0.211	−0.177	0.182	0.216
Crosswalk density	0.062	0.021	0.021	0.028	0.096	0.103
Traffic lights	−0.046	0.340	−0.713	−0.606	0.515	0.622
Fence	−0.120	0.367	−0.840	−0.724	0.485	0.601
Road type: Two way no median	−0.153	0.318	−0.777	−0.677	0.370	0.470
Road type: Two way with median	0.063	0.406	−0.733	−0.605	0.731	0.859
On-street parking	−0.271	0.339	−0.935	−0.828	0.286	0.393
Trash container	0.535	0.272	0.001	0.087	0.983	1.069
Contour: complexity B	−0.317	0.296	−0.897	−0.804	0.169	0.263
Contour: complexity C	−0.779	0.327	−1.421	−1.318	−0.241	−0.138
τ_v	104.54					
DIC	574.16					
Effective number of parameters	29.31					

Both models reveal that certain variables influence all pedestrians with credible effects. These include Average Annual Daily Traffic (AADT), the density of leisure POIs, the presence of bus stops, and the density of crosswalks, all of which increase the expected number of crashes for all pedestrians. The precision value of the heterogeneity effect indicates that the covariates explain most of the data dispersion.

Apart from these effects, the number of traffic crashes involving older pedestrians is also affected by population density, the presence of trees on the sidewalk, the presence of trash containers, and the level of contour complexity. Higher population density increases collision risks. Trees and trash containers are also related to riskier streets. Roads ending in three-legged intersections (complexity A) are riskier compared to those without three-

legged intersections (complexity C), while no credible effect is noted for complexity B. These variables have consistent but non-credible effects in the non-older pedestrian model.

Table 6. Main statistics of the rest of pedestrian model (credible variables at 10% in bold).

Variable	Mean	SD	2.5%	5%	95%	97.5%
Intercept	−4.483	1.260	−6.956	−6.558	−2.412	−2.016
log(AADT)	0.511	0.128	0.259	0.300	0.722	0.763
Population density	0.000	0.000	−0.001	−0.001	0.001	0.001
Avg. income per person	−0.000	0.000	−0.000	−0.000	0.000	0.000
Leisure POIs	0.011	0.003	0.006	0.006	0.015	0.016
Health and education POIs	−0.014	0.017	−0.049	−0.043	0.014	0.020
Public POIs	0.013	0.015	−0.017	−0.012	0.038	0.043
Leisure PDDI	0.204	0.162	−0.114	−0.063	0.471	0.522
Sidewalk width	0.061	0.038	−0.013	−0.001	0.124	0.136
Sidewalk disparity	−0.205	0.534	−1.252	−1.083	0.674	0.842
Trees	0.054	0.174	−0.288	−0.233	0.341	0.396
Terraces	0.356	0.163	0.036	0.088	0.625	0.677
Bus stop	0.329	0.143	0.049	0.094	0.565	0.610
Metro entrance	0.729	0.210	0.318	0.384	1.076	1.143
Streetlights	−0.014	0.614	−1.219	−1.025	0.996	1.190
Slope	−0.973	4.026	−8.870	−7.600	5.650	6.919
Number of lanes	−0.037	0.067	−0.168	−0.146	0.073	0.094
Crosswalk density	0.034	0.014	0.007	0.012	0.057	0.061
Traffic lights	−0.130	0.210	−0.543	−0.476	0.216	0.282
Fence	0.143	0.211	−0.270	−0.204	0.489	0.556
Road type: Two way no median	−0.015	0.187	−0.381	−0.322	0.293	0.352
Road type: Two way with median	−0.685	0.261	−1.195	−1.113	−0.256	−0.174
On-street parking	−0.168	0.191	−0.542	−0.481	0.146	0.206
Trash container	0.022	0.168	−0.307	−0.255	0.298	0.351
Contour: complexity B	−0.163	0.194	−0.544	−0.483	0.157	0.218
Contour: complexity C	−0.241	0.200	−0.633	−0.570	0.087	0.150
τ_v	84.28					
DIC	1047.70					
Effective number of parameters	35.45					

Apart from the common patterns between both models stated above, the safety of non-older pedestrians is influenced by terraces, metro stations, and street type. Streets with terraces or metro entrances are riskier. Compared to one-way streets, two-way streets with a median are safer, while two-way streets without a median show no credible difference in safety. Interestingly, the effects of terraces and two-way streets with medians are opposite for older pedestrians, but these effects are not credible.

The remaining covariates did not show credible impacts in either model. Streets with no streetlights, less slope, fewer lanes, no traffic lights, and no on-street parking are riskier for all pedestrians. Additionally, two-way streets without medians and streets with intermediate contour complexity (complexity B) tend to have fewer pedestrian collisions. On the contrary, some variables demonstrated opposite effects between the two models. For instance, streets with more health and education POIs, fewer public POIs, and a more balanced distribution of leisure POIs are riskier for older pedestrians. Features contributing to safer conditions for older pedestrians included wider and better-balanced sidewalks and pedestrian fences. These effects are opposite in the non-older pedestrian crashes model.

4. Discussion

This section discusses the practical policies derived from the modeling results. The exposure variables, length and AADT, were associated with all pedestrian traffic crashes. Segment length was used as an offset variable, fixed at 1, which is in line with other

authors [57,75]. AADT consistently increased pedestrian crashes, affirming its significance across all pedestrian groups.

Regarding the socioeconomic factors, population density showed a positive credible impact on older pedestrian crashes, which is consistent with previous literature [45], although it was not credible in the non-older pedestrians' model. Higher average income per capita reduced all pedestrian collisions, which was expected due to better infrastructure maintenance and safer pedestrian behaviors [76]. The same result was found by other authors [45], but the effect of this variable was not credible.

Among land use factors, the presence of leisure points of interest (POIs) was the only credible variable, increasing crashes for both older and non-older pedestrians. Leisure POIs attract pedestrians who might walk around them without paying special attention to traffic and may feel safer due to crowded and well-maintained areas, leading to increased crashes. Similar results were reported by Lee et al. [44] for older pedestrians and by Zhu et al. [65] for pedestrians in general.

The most notable differences between the age groups were found in the infrastructure variables, which can be modified in the short-term to reduce crashes. Trees on sidewalks increased crashes for older pedestrians due to narrowed sidewalks and navigation hazards, making them more likely to walk on the pavement. This factor did not affect non-older pedestrians.

The most notable and interesting differences between the two age groups were found in the infrastructure variables, which are the most feasible factors to modify in order to reduce pedestrian crashes for older individuals or the general population.

Regarding the ease of walking on the sidewalk, the presence of trees was linked to increased crashes involving older pedestrians due to two main reasons. First, trees narrow the effective width of sidewalks with uncovered tree pits posing navigation hazards. This issue is exacerbated on narrow sidewalks, leading older adults to walk on the pavement, increasing their crash risk. Second, older individuals might prefer walking in areas with vegetation, resulting in higher pedestrian traffic and more crashes. This factor did not affect younger pedestrians, who navigate these sidewalks more easily.

The occurrence of non-older pedestrian crashes is influenced by two factors. First, restaurant terraces on sidewalks might impede pedestrian transit and reduce the ability to detect oncoming vehicles, increasing crashes. Non-older pedestrians may cross streets with terraces in undesignated areas, unlike older pedestrians who show a greater willingness than non-older people to cross at designated areas only [33]. Interestingly, the effect on older pedestrians is opposite, although not credible, which could indicate that terraces have a benefit for them serving as a barrier to vehicles. Second, metro station entrances on sidewalks are linked to more non-older pedestrian crashes, which is likely due to their attraction. Older pedestrians do not show this effect, which is possibly due to reduced metro usage by people over 64, who prefer the use of bus [77]. Similar results were found by Gálvez-Pérez et al. [45]. Streets with bus stops are riskier for all pedestrians, which was something already noted by Lv et al. [48] for the older pedestrian case. Buses might disrupt traffic flow and obstruct the view of oncoming vehicles with visibility further reduced at some bus stops due to opaque billboards.

In examining variables linked to the ease of crossing the street, crosswalk density is the only credible predictor, correlating with increased pedestrian crashes due to higher interactions between pedestrians and vehicles. Li et al. [78] found that older people feel safer in areas with more intersections, typically characterized by increased crosswalks, although these crosswalks would be located at intersections, which is outside the scope of our study. There is a need to protect these areas, especially for older pedestrians. Other non-credible variables show interesting effects. Traffic lights mitigate crashes by controlling traffic and reducing speeds. Pedestrian fences should lower crashes by acting as barriers, restricting crossings to designated areas. Older pedestrians, in particular, perceive a road safety benefit in streets with fences. However, younger pedestrians experience the opposite effect. This finding is similar to streets with terraces, where older individuals may tend

to use designated crossing points, whereas younger ones might choose alternate paths, including potentially jumping fences, which is easier for them.

Regarding the ease of detecting traffic, credible variables are only found in the older pedestrian model. Trash containers, often placed on sidewalks or pavements, obstruct vision and make it more difficult for older individuals to detect traffic, especially those with visual or hearing impairment [15]. Streets with these containers were found to be riskier for older pedestrians. Streets with contour complexity level C (other than three-legged intersections at both ends) are associated with fewer crashes involving older pedestrians compared to streets with level A complexity (three-legged intersections).

Road safety education for older adults, which is already planned in Madrid [79], should be complemented with other measures. Regarding practical applications of this research for policymakers, modeling results suggest short-term road infrastructural measures that could be implemented to decrease the incidence of traffic crashes involving older pedestrians. Trees, while enhancing urban livability and providing benefits, including economic, social, and health benefits [80,81], may pose obstacles for older individuals on narrow or crowded sidewalks by significantly reducing available space. Furthermore, older adults may perceive fall risk if the tree pits are not properly maintained. It is advisable to avoid placing trees on streets with narrow sidewalks. More ideally, relocating trees to areas outside the sidewalks would potentially reduce on-street parking spaces, which would extend additional space for pedestrian use. It is important to select tree species that are appropriate for the given urban environment, particularly those that are suitable for narrower spaces. In wider sidewalks, covering tree pits with tree grates, usually metallic, can enhance safety by preventing potential trip hazards. In addition to improving the traffic safety, these measures would reduce the potential risk of falls, which is something especially beneficial for older people due to their longer recovery processes.

Trash containers, typically located near parking lots, present a more significant visual barrier for older pedestrians than parked vehicles due to their size and shape. Their placement can impair traffic detection, especially when positioned between pedestrians and approaching traffic. To improve traffic detection for older pedestrians, trash containers should be placed away from crosswalks and road junctions, avoiding any reduction in sight distance. Placing the containers between the pedestrians and the oncoming traffic should be avoided. It is recommended to position them close to the sidewalk edge despite the logistical challenges of moving them for emptying. Even minor increases in the gap between the sidewalk and a container can significantly improve sight distance. Marking optimal container placements on the pavement could ensure proper positioning.

Bus stops and crosswalks were found to increase all pedestrian crashes. Like trash containers, bus stops near crosswalks can block vehicle sightlines. To mitigate this, bus stops should avoid obstructing vision and could be replaced with posts. High-traffic bus stops crowd sidewalks and block sightlines, raising pedestrian crash risks and potentially leading to more severe injuries if a pedestrian is hit by a bus [82]. Safety could be improved by marking these areas and installing pedestrian fences where feasible. Regarding crosswalks, the main objective should be to properly mark and condition existing crosswalks, especially on wide streets, and extending green traffic light phases, as older pedestrians walk at slower speeds [28,29,83]. In older neighborhoods, this could enhance safety. Although the effect of traffic lights was not statistically significant, their installation could be beneficial by slowing down vehicle speeds, which is associated with less severe injuries in collisions [84,85].

It should be noted that all credible variables in both models show a consistent effect on the safety of both older and non-older pedestrians. Additionally, there are variables showing opposite effects on those age groups, but these are only credible in one or neither of the models. Consequently, none of the proposed countermeasures aimed at enhancing the safety of older pedestrians should result in a deterioration of safety for other pedestrians.

5. Conclusions

This study employed a Poisson model within a Bayesian framework, utilizing the Integrated Nested Laplace Approximation (INLA) approach, to analyze factors influencing traffic crashes involving both older and non-older pedestrians. A holistic view of the road segments was used through a wide set of infrastructure variables, including less common ones like restaurant terraces, trash containers, and on-street parking. These variables were analyzed to propose short-term countermeasures. Based on the results, we propose measures such as a better placement of trees, trash containers, and bus stops, installing more crosswalks and traffic lights, and extending the green phase of existing traffic lights. These improvements aim to enhance the safety of older pedestrians without compromising the safety of other pedestrians. Furthermore, they should be implemented without compromising the road's operational performance or contributing to traffic jams.

The limitations of this study are mainly related to data availability. Average Annual Daily Traffic (AADT) data, crucial for analysis, was not available for all street segments, only for certain segments. Fortunately, these street segments are spread across various street types and areas, making them representative of the city network. Another limitation is the absence of pedestrian flow data, which is only collected on major streets in Madrid. For a more detailed analysis, pedestrian flow data categorized by age groups, specifically older versus non-older pedestrians, would be beneficial to better understand the dynamics and risks associated with each group.

Future research should address these issues to enhance the understanding of older pedestrian traffic safety. Ad hoc measurements of street and sidewalk widths limit data collection to a few records. While micro-studies are effective in developing infrastructure countermeasures, using small areal units in macro-studies (such as regular grids) could simplify data collection, reduce manual efforts, and minimize measurement errors. Additionally, calculating the 'effective width' of sidewalks and its impact on traffic crashes involving older pedestrians is a promising research direction. This would require precise measurements of actual sidewalk widths and the identification of elements that constrict these widths.

Despite a general willingness to develop inclusive urban spaces, prioritizing the integration of older pedestrians is crucial. This demographic faces a higher risk of fatalities and serious injuries from vehicle collisions and is vulnerable to social exclusion and isolation, which can result from reduced walking activity. Given the expected aging population in developed countries, additional efforts are needed to ensure their safety and inclusion in age-friendly urban planning initiatives. Such inclusion should not only create safer streets for older adults and promote walking activity among this group but also help to reduce ageist stereotypes.

The scientific community, particularly road safety experts, must be sensitive to the needs of the older population and their social integration challenges. This approach is essential for generating practical, applied research with effective policy proposals for road infrastructure interventions, aiding in the design of age-friendly cities.

Author Contributions: Conceptualization, D.G.-P. and B.G.; methodology, D.G.-P.; software, D.G.-P.; validation, B.G. and A.O.; formal analysis, D.G.-P.; investigation, D.G.-P., B.G. and A.O.; resources, D.G.-P.; data curation, D.G.-P.; writing—original draft preparation, D.G.-P., B.G. and A.O.; writing—review and editing, D.G.-P., B.G. and A.O.; visualization, D.G.-P.; supervision, B.G. and A.O.; project administration, B.G. All authors have read and agreed to the published version of the manuscript.

Funding: Daniel Gálvez-Pérez is conducting his doctoral thesis with the support of a grant from the Universidad Politécnica de Madrid through the "Programa Propio de I + D + I 2020: Ayudas para Contratos Predoctorales".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors wish to express their gratitude to the Dirección General de Tráfico (DGT) for providing the traffic crash data used in this study.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Ungvari, Z.; Fazekas-Pongor, V.; Csiszar, A.; Kunutsor, S.K. The multifaceted benefits of walking for healthy aging: From Blue Zones to molecular mechanisms. *GeroScience* **2023**, *45*, 3211–3239. [[CrossRef](#)] [[PubMed](#)]
2. Shellito, N.; Roldan, N.V. Walking away from loneliness: The mediating role of social isolation. *Innov. Aging* **2019**, *3* (Suppl. S1), S836–S837. [[CrossRef](#)]
3. Liu, J.; Zhong, S.; Huang, Y.; Wang, Z. How does the preference heterogeneity affect the elderly’s evaluation of bus accessibility? Evidence from Guangzhou, China. *J. Transp. Health* **2021**, *22*, 101094. [[CrossRef](#)]
4. Gabriel, Z.; Bowling, A. Quality of life from the perspectives of older people. *Ageing Soc.* **2004**, *24*, 675–691. [[CrossRef](#)]
5. Musselwhite, C.; Holland, C.; Walker, I. The role of transport and mobility in the health of older people. *J. Transp. Health* **2015**, *2*, 1–4. [[CrossRef](#)]
6. Nordbakke, S.; Schwanen, T. Well-being and Mobility: A Theoretical Framework and Literature Review Focusing on Older People. *Mobilities* **2014**, *9*, 104–129. [[CrossRef](#)]
7. Mizuta, S.; Uchida, K.; Sawa, R.; Nakamura, J.; Encho, H.; Akisue, T.; Ono, R. Context of walking and loneliness among community-dwelling older adults: A cross-sectional study. *BMC Geriatr.* **2023**, *23*, 326. [[CrossRef](#)]
8. Iversen, T.N.; Larsen, L.; Solem, P.E. A conceptual analysis of Ageism. *Nord. Psychol.* **2012**, *61*, 4–22. [[CrossRef](#)]
9. Lee, K.; Hyun, K.; Mitchell, J.; Saha, T.; Gibson, N.O.; Krejci, C. Exploring Factors Enhancing Resilience Among Marginalized Older Adults During the COVID-19 Pandemic. *J. Appl. Gerontol.* **2022**, *41*, 610–618. [[CrossRef](#)]
10. Irvine, K.N.; Fisher, D.; Marselle, M.R.; Currie, M.; Colley, K.; Warber, S.L. Social Isolation in Older Adults: A Qualitative Study on the Social Dimensions of Group Outdoor Health Walks. *Int. J. Environ. Res. Public Health* **2022**, *19*, 5353. [[CrossRef](#)]
11. D’Cruz, M.; Banerjee, D. ‘An invisible human rights crisis’: The marginalization of older adults during the COVID-19 pandemic—An advocacy review. *Psychiatry Res.* **2020**, *292*, 113369. [[CrossRef](#)] [[PubMed](#)]
12. European Commission. *European Road Safety Observatory. Road Safety Thematic Report—Seniors.*; European Commission: Brussels, Belgium, 2021.
13. European Commission. *European Road Safety Observatory. Facts and Figures—Pedestrians 2023.*; European Commission: Brussels, Belgium, 2023.
14. DGT. Las principales cifras de la Siniestralidad Vial España 2021. 2021. Available online: https://www.dgt.es/export/sites/web-DGT/.galleries/downloads/dgt-en-cifras/24h/Principales_cifras_2021_web_accesible_con_Meta.pdf (accessed on 15 June 2024).
15. European Commission. *ElderSafe-Risks and Countermeasures for Road Traffic of Elderly in Europe* (No. MOVE/C4/2014-244); European Commission–Directorate-General for Mobility and Transport (DG-MOVE): Brussels, Belgium, 2015.
16. European Commission. *2024 Ageing Report: Underlying Assumptions and Projection Methodologies*; European Commission: Brussels, Belgium, 2023.
17. Rémillard-Boilard, S.; Buffel, T.; Phillipson, C. Developing Age-Friendly Cities and Communities: Eleven Case Studies from around the World. *Int. J. Environ. Res. Public Health* **2021**, *18*, 133. [[CrossRef](#)] [[PubMed](#)]
18. World Health Organization. *Hidden Cities: Unmasking and Overcoming Health Inequities in Urban Settings*; World Health Organization, Centre for Health Development: Geneva, Switzerland, 2010.
19. Dumbaugh, E. Designing Communities to Enhance the Safety and Mobility of Older Adults. *J. Plan. Lit.* **2008**, *23*, 17–36. [[CrossRef](#)]
20. Schmeidler, K. Safe and Secure Cities for Pedestrians and Senior Citizens. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *960*, 042099. [[CrossRef](#)]
21. World Health Organization. *Global Age-Friendly Cities: A Guide*; World Health Organization: Geneva, Switzerland, 2007.
22. Estebansari, F.; Dastoorpoor, M.; Khalifehkandi, Z.R.; Nouri, A.; Mostafaei, D.; Hosseini, M.; Esmaeili, R.; Aghababaeian, H. The Concept of Successful Aging: A Review Article. *Curr. Aging Sci.* **2019**, *13*, 4–10. [[CrossRef](#)]
23. World Health Organization. *Active Ageing: A Policy Framework*. 2002. Available online: <https://iris.who.int/handle/10665/67215> (accessed on 15 June 2024).
24. United Nations. *Resolution Adopted by the General Assembly on 25 September 2015*; United Nations: New York, NY, USA, 2015.
25. European Commission. *EU Road Safety: Towards “Vision Zero”*; European Commission: Brussels, Belgium, 2022.
26. Oxley, J.; Fildes, B.; Ihsen, E.; Charlton, J.; Day, R. Differences in traffic judgements between young and old adult pedestrians. *Accid. Anal. Prev.* **1997**, *29*, 839–847. [[CrossRef](#)]

27. Vine, D.; Buys, L.; Aird, R. Experiences of Neighbourhood Walkability Among Older Australians Living in High Density Inner-City Areas. *Plan. Theory Pract.* **2012**, *13*, 421–444. [CrossRef]
28. Ortuño, R.R. La regulación de los semáforos peatonales en España: ¿tienen las personas mayores tiempo suficiente para cruzar? *Rev. Española Geriatria Y Gerontol.* **2010**, *45*, 199–202. [CrossRef]
29. Van Hoof, J.; Dikken, J.; Buttigieg, S.C.; van den Hoven, R.F.; Kroon, E.; Marston, H.R. Age-friendly cities in the Netherlands: An explorative study of facilitators and hindrances in the built environment and ageism in design. *Indoor Built Environ.* **2020**, *29*, 417–437. [CrossRef]
30. European Commission. *European Road Safety Observatory. Facts and Figures Seniors 2021*; European Commission: Brussels, Belgium, 2021.
31. Zegeer, C.V.; Stutts, J.C.; Huang, H.; Zhou, M.; Lodgman, E. Analysis of Elderly Pedestrian Accidents and Recommended Countermeasures. *Transp. Res. Rec.* **1993**, *1405*, 56–63.
32. Abou-Raya, S.; ElMeguid, L.A. Road traffic accidents and the elderly. *Geriatr. Gerontol. Int.* **2009**, *9*, 290–297. [CrossRef]
33. Bernhoft, I.M.; Carstensen, G. Preferences and behaviour of pedestrians and cyclists by age and gender. *Transp. Res. Part F Traffic Psychol. Behav.* **2008**, *11*, 83–95. [CrossRef]
34. Ravi, K.E.; Fields, N.L.; Dabelko-Schoeny, H. Outdoor spaces and buildings, transportation, and environmental justice: A qualitative interpretive meta-synthesis of two age-friendly domains. *J. Transp. Health* **2021**, *20*, 100977. [CrossRef]
35. Dommes, A.; Cavallo, V.; Vienne, F.; Aillerie, I. Age-related differences in street-crossing safety before and after training of older pedestrians. *Accid. Anal. Prev.* **2012**, *44*, 42–47. [CrossRef]
36. Dommes, A.; Cavallo, V.; Oxley, J. Functional declines as predictors of risky street-crossing decisions in older pedestrians. *Accid. Anal. Prev.* **2013**, *59*, 135–143. [CrossRef] [PubMed]
37. Martin, A.; Hand, E.; Trace, F.; O’neill, D. Pedestrian Fatalities and Injuries Involving Irish Older People. *Gerontology* **2010**, *56*, 266–271. [CrossRef] [PubMed]
38. Kim, S.; Ulfarsson, G.F. Traffic safety in an aging society: Analysis of older pedestrian crashes. *J. Transp. Saf. Secur.* **2019**, *11*, 323–332. [CrossRef]
39. Das, S.; Bibeka, A.; Sun, X.; Zhou, H.; Jalayer, M. Elderly Pedestrian Fatal Crash-Related Contributing Factors: Applying Empirical Bayes Geometric Mean Method. *Transp. Res. Rec. J. Transp. Res. Board* **2019**, *2673*, 254–263. [CrossRef]
40. Wang, Y.; Haque, M.; Chin, H.-C. Elderly pedestrian injuries in Singapore. *J. Transp. Saf. Secur.* **2017**, *9*, 273–300. [CrossRef]
41. Laković, S.; Tollazzi, T.; Gruden, C. Elderly Pedestrians and Road Safety: Findings from the Slovenian Accident Database and Measures for Improving Their Safety. *Sustainability* **2023**, *15*, 1631. [CrossRef]
42. Ziakopoulos, A.; Yannis, G. A review of spatial approaches in road safety. *Accid. Anal. Prev.* **2020**, *135*, 105323. [CrossRef] [PubMed]
43. Dumbaugh, E.; Zhang, Y. The Relationship between Community Design and Crashes Involving Older Drivers and Pedestrians. *J. Plan. Educ. Res.* **2013**, *33*, 83–95. [CrossRef]
44. Lee, S.; Yoon, J.; Woo, A. Does elderly safety matter? Associations between built environments and pedestrian crashes in Seoul, Korea. *Accid. Anal. Prev.* **2020**, *144*, 105621. [CrossRef] [PubMed]
45. Gálvez-Pérez, D.; Guirao, B.; Ortuño, A.; Picado-Santos, L. The Influence of Built Environment Factors on Elderly Pedestrian Road Safety in Cities: The Experience of Madrid. *Int. J. Environ. Res. Public Health* **2022**, *19*, 2280. [CrossRef]
46. Gálvez-Pérez, D.; Guirao, B.; Ortuño, A. Analysis of the elderly pedestrian traffic accidents in urban scenarios: The case of the Spanish municipalities. *Int. J. Inj. Control. Saf. Promot.* **2024**, *31*, 376–395. [CrossRef]
47. Kim, D. The transportation safety of elderly pedestrians: Modeling contributing factors to elderly pedestrian collisions. *Accid. Anal. Prev.* **2019**, *131*, 268–274. [CrossRef]
48. Lv, M.; Wang, N.; Yao, S.; Wu, J.; Fang, L. Towards Healthy Aging: Influence of the Built Environment on Elderly Pedestrian Safety at the Micro-Level. *Int. J. Environ. Res. Public Health* **2021**, *18*, 9534. [CrossRef]
49. DGT. Las principales cifras de la siniestralidad de los peatones España 2018. 2018. Available online: <https://www.dgt.es/menusecundario/dgt-en-cifras/dgt-en-cifras-resultados/dgt-en-cifras-detalle/Las-principales-cifras-de-la-Siniestralidad-de-los-Peatones-2018/> (accessed on 15 June 2024).
50. Shaik, M.E.; Ahmed, S. An overview of the impact of COVID-19 on road traffic safety and travel behavior. *Transp. Eng.* **2022**, *9*, 100119. [CrossRef]
51. Ayuntamiento de Madrid. Estado de la Movilidad de la Ciudad de Madrid. Informe Anual 2019. Madrid. 2019. Available online: <https://transparencia.madrid.es/FWPProjects/transparencia/Movilidad/Trafico/InformesMovilidad/InformeMovilidad2019.pdf> (accessed on 15 June 2024).
52. Zhu, M.; Sze, N.; Newnam, S. Effect of urban street trees on pedestrian safety: A micro-level pedestrian casualty model using multivariate Bayesian spatial approach. *Accid. Anal. Prev.* **2022**, *176*, 106818. [CrossRef]
53. Dunbar, G.; Holland, C.A.; Maylor, E.A. *Older Pedestrians: A Critical Review of the Literature*; Department of Transport: London, UK, 2004.
54. *ArcMap, 10.8.1*; Esri Inc.: Redlands, CA, USA, 2020.
55. R Core Team. *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*; R Core Team: Viena, Austria, 2013.
56. OpenStreetMap Contributors. OpenStreetMap. Available online: <https://www.openstreetmap.org/> (accessed on 15 June 2024).

57. Satria, R.; Aguero-Valverde, J.; Castro, M. Spatial analysis of road crash frequency using Bayesian models with Integrated Nested Laplace Approximation (INLA). *J. Transp. Saf. Secur.* **2021**, *13*, 1240–1262. [[CrossRef](#)]
58. Kraidi, R.; Evdorides, H. Pedestrian safety models for urban environments with high roadside activities. *Saf. Sci.* **2020**, *130*, 104847. [[CrossRef](#)]
59. Liu, J.; Khattak, A.J.; Wali, B. Do safety performance functions used for predicting crash frequency vary across space? Applying geographically weighted regressions to account for spatial heterogeneity. *Accid. Anal. Prev.* **2017**, *109*, 132–142. [[CrossRef](#)] [[PubMed](#)]
60. Soltani, N.; Saffarzadeh, M.; Naderan, A. Modeling crash frequencies by transportation mode using micro/macro level variables. *AUT J. Civ. Eng.* **2021**, *5*, 161–170. [[CrossRef](#)]
61. Thakali, L.; Fu, L.; Chen, T. Model-Based Versus Data-Driven Approach for Road Safety Analysis: Do More Data Help? *Transp. Res. Rec. J. Transp. Res. Board* **2016**, *2601*, 33–41. [[CrossRef](#)]
62. Aguero-Valverde, J. Full Bayes Poisson gamma, Poisson lognormal, and zero inflated random effects models: Comparing the precision of crash frequency estimates. *Accid. Anal. Prev.* **2013**, *50*, 289–297. [[CrossRef](#)]
63. Barua, S.; El-Basyouny, K.; Islam, T. A Full Bayesian multivariate count data model of collision severity with spatial correlation. *Anal. Methods Accid. Res.* **2014**, *3-4*, 28–43. [[CrossRef](#)]
64. Ma, J.; Kockelman, K.M.; Damien, P. A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. *Accid. Anal. Prev.* **2008**, *40*, 964–975. [[CrossRef](#)]
65. Zhu, M.; Li, H.; Sze, N.; Ren, G. Exploring the impacts of street layout on the frequency of pedestrian crashes: A micro-level study. *J. Saf. Res.* **2022**, *81*, 91–100. [[CrossRef](#)]
66. Papadimitriou, E.; Filtness, A.; Theofilatos, A.; Ziakopoulos, A.; Quigley, C.; Yannis, G. Review and ranking of crash risk factors related to the road infrastructure. *Accid. Anal. Prev.* **2019**, *125*, 85–97. [[CrossRef](#)]
67. Besag, J. Spatial Interaction and the Statistical Analysis of Lattice Systems. 1974. Available online: <https://www.jstor.org/stable/2984812> (accessed on 15 June 2024).
68. Dormann, C.F.; McPherson, J.M.; Araújo, M.B.; Bivand, R.; Bolliger, J.; Carl, G.; Davies, R.G.; Hirzel, A.; Jetz, W.; Kissling, W.D.; et al. Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography* **2007**, *30*, 609–628. [[CrossRef](#)]
69. Wall, M.M. A close look at the spatial structure implied by the CAR and SAR models. *J. Stat. Plan. Inference* **2004**, *121*, 311–324. [[CrossRef](#)]
70. Aguero-Valverde, J.; Jovanis, P.P. Analysis of Road Crash Frequency with Spatial Models. *Transp. Res. Rec. J. Transp. Res. Board* **2008**, *2061*, 55–63. [[CrossRef](#)]
71. Wang, C.; Quddus, M.A.; Ison, S.G. Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England. *Accid. Anal. Prev.* **2009**, *41*, 798–808. [[CrossRef](#)] [[PubMed](#)]
72. Ahmed, M.; Huang, H.; Abdel-Aty, M.; Guevara, B. Exploring a Bayesian hierarchical approach for developing safety performance functions for a mountainous freeway. *Accid. Anal. Prev.* **2011**, *43*, 1581–1589. [[CrossRef](#)]
73. Rue, H.; Martino, S.; Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2009**, *71*, 319–392. [[CrossRef](#)]
74. Rue, H.; Riebler, A.; Sørbye, S.H.; Illian, J.B.; Simpson, D.P.; Lindgren, F.K. Bayesian Computing with INLA: A Review. *Annu. Rev. Stat. Its Appl.* **2017**, *4*, 395–421. [[CrossRef](#)]
75. Aguero-Valverde, J. Direct Spatial Correlation in Crash Frequency Models: Estimation of the Effective Range. *J. Transp. Saf. Secur.* **2014**, *6*, 21–33. [[CrossRef](#)]
76. Rosenbloom, T.; Sapir-Lavid, Y.; Perlman, A. Risk factors in road crossing among elderly pedestrians and readiness to adopt safe behavior in socio-economic comparison. *Accid. Anal. Prev.* **2016**, *93*, 23–31. [[CrossRef](#)]
77. CRTM. Encuesta domiciliaria de movilidad en día laborable de 2018 en la Comunidad de Madrid. Caracterización socioeconómica de la movilidad. 2019. Available online: https://www.crtm.es/media/jjcn0svt/edm18_doc2_socio_econ.pdf (accessed on 15 June 2024).
78. Li, F.; Fisher, K.J.; Brownson, R.C.; Bosworth, M. Multilevel modelling of built environment characteristics related to neighbourhood walking activity in older adults. *J. Epidemiol. Community Health* **2005**, *59*, 558–564. [[CrossRef](#)]
79. Ayuntamiento de Madrid. Madrid, Ciudad Amigable Con Las Personas Mayores. Plan de Acción 2021–2023. 2021. Available online: https://transparencia.madrid.es/FWProjects/transparencia/PlanesYMemorias/Planes/Mayores/Plan%20Madrid_Amigable%202021_2023.pdf (accessed on 15 June 2024).
80. Roy, S.; Byrne, J.; Pickering, C. A systematic quantitative review of urban tree benefits, costs, and assessment methods across cities in different climatic zones. *Urban For. Urban Green.* **2012**, *11*, 351–363. [[CrossRef](#)]
81. Turner-Skoff, J.B.; Cavender, N. The benefits of trees for livable and sustainable communities. *Plants People Planet* **2019**, *1*, 323–335. [[CrossRef](#)]
82. Moradi, A.; Ameri, P.; Rahmni, K.; Najafi, M.; Jamshidi, E.; Fakhri, Y.; Khazaei, S.; Moeini, B.; Amjadian, M. Factors affecting the severity of pedestrian traffic crashes. *Arch. Trauma Res.* **2019**, *8*, 46–56. [[CrossRef](#)]
83. Bailey, S.S.; Jones, S.A.; Stout, R.J.; Bailey, J.H.; Kass, S.; Morgan, B.B. Issues of Elderly Pedestrians. *Transp. Res. Rec.* **1992**, *1375*, 68–73.

84. Rosén, E.; Stigson, H.; Sander, U. Literature review of pedestrian fatality risk as a function of car impact speed. *Accid. Anal. Prev.* **2011**, *43*, 25–33. [[CrossRef](#)] [[PubMed](#)]
85. Hussain, Q.; Feng, H.; Grzebieta, R.; Brijs, T.; Olivier, J. The relationship between impact speed and the probability of pedestrian fatality during a vehicle-pedestrian crash: A systematic review and meta-analysis. *Accid. Anal. Prev.* **2019**, *129*, 241–249. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

4.5 *Article D: The effect of COVID-19 on older pedestrian road safety: A holistic analysis through pandemic phases*

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Journal of Transport & Health

journal homepage: www.elsevier.com/locate/jth

The effect of COVID-19 on older pedestrian road safety: A holistic analysis through pandemic phases

Daniel Gálvez-Pérez^{*}, Begoña Guirao, David del-Villar-Juez

Departamento de Ingeniería Del Transporte, Territorio y Urbanismo, Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos, Universidad Politécnica de Madrid, Calle del Profesor Aranguren, 3; 28040, Madrid, Spain

ARTICLE INFO

Keywords:

Road safety
Older adults
Population ageing
Pedestrians
Covid-19 pandemic
Age-friendly cities

ABSTRACT

Introduction: The COVID-19 pandemic drastically altered daily life, impacting mobility and road safety. Older pedestrians (above 65) are especially vulnerable to vehicle collisions due to physical and cognitive characteristics. While the immediate effects on traffic are well-documented, the medium- and long-term impacts on older pedestrians' road safety are still unclear. This study aims to provide a holistic analysis of the pandemic's influence on older pedestrian traffic safety in urban areas.

Methods: In the first stage, Poisson regression models were used to analyze urban pedestrian crash frequency in Spain, considering pedestrian age (older and non-older) and injury severity (total, fatal, serious, slight). In the second stage, two binary logistic regression models (for older and non-older pedestrians) classified injured pedestrians into pre- and post-pandemic periods in Madrid, examining changes in crash-related factors.

Results: During the lockdown, crashes for both older and non-older pedestrians declined across all injury levels. In the subsequent phase of the pandemic, 'new normality', crash numbers stayed below pre-pandemic levels, except for fatalities among non-older pedestrians. Post-vaccination, total and slightly injured older and non-older pedestrians stayed reduced, and this reduction was double for older than for non-older pedestrians. The injury severity rates among pedestrians were not linked to pandemic phases. Older pedestrian crashes in the post-pandemic period shifted towards locations with more bike lanes and occurred less frequently during evening and night hours.

Conclusions: The pandemic significantly impacted older pedestrian safety, with crash reductions lasting beyond the lockdown. Slight modifications in crash-related factors suggest potential shifts in older adults' mobility patterns. However, the profile of older pedestrian crashes remains largely unchanged, indicating that current safety measures are still effective. Further research is needed to assess the long-term effects on walking behavior of older adults induced by the pandemic, as older pedestrians remain a vulnerable group in road safety.

1. Introduction

The COVID-19 pandemic and governments' measures have caused unprecedented changes to many aspects of our daily life, including mobility patterns and, consequently, road safety. Policy interventions, such as social distancing or lockdown, proved to

^{*} Corresponding author.

E-mail addresses: daniel.galvezp@upm.es (D. Gálvez-Pérez), begona.guirao@upm.es (B. Guirao), david.delvillar@upm.es (D. del-Villar-Juez).

<https://doi.org/10.1016/j.jth.2025.101986>

Received 2 October 2024; Received in revised form 30 December 2024; Accepted 6 January 2025

Available online 17 January 2025

2214-1405/© 2025 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

reduce the transmission of the virus (M. Zhang et al., 2022). These measures and the pandemic itself modified transportation dynamics worldwide (Shaik and Ahmed, 2022). Scientific research has demonstrated significant reduction in traffic volumes leading to changes in traffic crash rates during lockdown periods. Despite the immediate impact on road safety being well-documented, there remains a substantial gap in understanding medium- and long-term effects of the pandemic on traffic safety. Even more important, the effect of the pandemic on the urban road safety of older pedestrians is practically unknown.

Older adults (people above 65) as pedestrians show physical and cognitive characteristics that makes them especially vulnerable to vehicle collisions, such as slower walking speed, balance, vision and hearing impairments, and higher reaction times (European Commission, 2015). As for the rest of population, walking provides older adults multiple physical benefits (Hanson and Jones, 2015; World Health Organization, 2022), but they are also benefited in social terms, as walking implies major level of social inclusion and well-being (Gabriel and Bowling, 2004; Musselwhite et al., 2015; Nordbakke and Schwanen, 2014). However, older adults are especially vulnerable to a notable walking externality: suffering a vehicle collision, and this fact is backed by official traffic crash databases, as older pedestrians are overrepresented in road crashes, and they are more likely to result killed or seriously injured after a vehicle collision than the rest of pedestrians. In Europe, in fact, almost half of the pedestrians killed in traffic crashes in 2019 were older adults (European Commission, 2021).

At the same time, older adults were the most affected group by the epidemic in medical and social terms. This demographic presents a higher COVID-19 mortality rate compared to the rest of population (Nanda et al., 2020; Shahid et al., 2020). With the pandemic evolution, the number of deaths and the number of days with stringent epidemic control measures were related with increased feelings of sadness and depression among older adults (Atzendorf and Gruber, 2022) and decreased social life and fewer social interactions led to less quality of life and increased depression (Lebrasseur et al., 2021). In addition to the health worries, the exacerbation of ageism in the society has been related with a rise in anxiety (Bergman et al., 2020). Ageism encompasses both negative and positive stereotypes, prejudices, and discriminatory behaviors directed toward older adults (Iversen et al., 2012). Ageism rose during the pandemic (Brooke and Jackson, 2020; Lebrasseur et al., 2021), what was clearly shown in social media. For instance, Jimenez-Sotomayor et al. (2020) analyzed tweets from Twitter on the pandemic and older people and found that almost one quarter of them showed offensive content toward older adults.

The COVID-19 pandemic has significantly affected the mobility of older adults worldwide, with multiple studies documenting a decline in various forms of active travel and an increase in sedentary behaviors. During the pandemic, the duration of physical activity, active travel, life-space mobility (displacements of people in their daily lives), and quality of life among older adults drastically decreased because of social distancing measures (Choe et al., 2022; Esain et al., 2021; Nowossadeck et al., 2023; Perracini et al., 2021; Rantanen et al., 2021; Shaer and Haghshenas, 2021). Particularly, displacement outside of the home environment were reduced (Perracini et al., 2021). The frailty of individuals further intensified the impact of restricted mobility on their quality of life, as frail older adults experienced greater declines in life-space mobility (Saraiva et al., 2021). Factors such as residential density, availability of recreational facilities, and depressive symptoms were identified to influence mobility among older adults (Choe et al., 2022; H. Kim et al., 2023). Moreover, older adults faced social pressures and technological barriers that hindered their mobility, which was shown in Kunming, China (Q. Liu et al., 2021). Regarding walking behavior, P. Liu et al. (2024) found that there was a high decrease in park visitation during COVID-19 for visitors from areas with a high ageing population, and that older adults changed their preference from visiting small parks before the pandemic to large parks during it.

These findings underline the critical need for targeted interventions to address the psychosocial and physical challenges faced by older adults during such unprecedented times, ensuring that mobility and quality of life are preserved. This study is placed in the mixed context of road safety, active mobility and ageing, under the framework of the COVID-19 pandemic.

The following paragraphs shows an overview of the evolution of the existing literature on the effects of the pandemic on road safety, starting from the main published results of the general impacts of the COVID-19 on mobility (as mobility is one of the exposition risk factors of road safety) to the scarce specific studies on older pedestrian-vehicle crashes in urban areas.

There is much more literature focused on the **pandemic impacts on mobility** around the world than on road safety. Travel proposals changes, mode of transportation election, mileage traveled, and number of trips have suffered modifications during the pandemic (Shaik and Ahmed, 2022). Some of the mobility patterns during the pandemic, were common in many countries and the studies clearly distinguish between the lockdown period and the following months after the lockdown, in which a 'new normality' was registered before starting vaccination. For example, regarding active mobility, reductions of walking activity were observed in Israel during policy restrictions, such as lockdown, which recovered after these were removed, although no modifications in the preferred locations to walk were observed (Angel et al., 2023). Similarly, the number of steps showed a widespread decline in the US even after the pandemic restrictions were lifted (Desine et al., 2023), so long-term behaviors could have been affected. Some of studies focused on the mobility behavior of different city districts, attending their socioeconomic features; for example, in Glasgow (Hong et al., 2023) residents in less deprived neighborhoods were more likely to have affirmative attitudes towards walking due to the pandemic and people tended to enjoy less walking as their age increased.

In relation to public transport use, ridership practically disappeared in big metropolitan areas. For example, in Madrid it decreased by 95% during the pandemic peak and recovered slowly to only half of pre-pandemic levels in September 2020, showing lower ridership decreases in low-income districts (Fernández Pozo et al., 2022). In relation with older adults, senior public transport tickets registered the second largest decline, and months after the lockdown the relative decrease was roughly 70% compared to pre-covid levels, being the average drop 50% (Fernández Pozo et al., 2022). A drop in public transportation usage by older adults was also found by Choe et al. (2022) using Hong Kong data. Considering the transportation modal split, a study developed in Detroit (Yao et al., 2021) detected that bus, pedestrian, and car flows decreased during the lockdown, and increased again after it. Conversely, bicycle and motorcycle modes showed a transition in usage which increased after the lockdown.

Public vaccination, other important measure, involved a reduction in public's perceived risk of infection, thereby changing transport modal choice (Bei et al., 2023). Data collected in Brooklyn (US), showed that vaccination led to an increase in driving, transit and walking, but the most important factor on transportation was removing the pandemic policies, which increased travel more than vaccinations (Drummond and Hasnine, 2023).

Due to these changes in mobility patterns **because of the pandemic, traffic safety figures have suffered modifications worldwide**. The effects of the lockdown measures have been widely studied, and there is a consensus that traffic volume decreased during lockdown periods led to a reduction in overall traffic crashes and road deaths, which has been showed worldwide (Yasin et al., 2021), in Europe (ETSC, 2020), at the national level, for instance in USA (Brodeur et al., 2021), Qatar (Alhajyaseen et al., 2022) and Nepal (Sedain and Pant, 2020); state level (Barnes et al., 2020); province level (Saladié et al., 2020); and city level (Aloi et al., 2020; J. Zhang et al., 2021). For instance, at the city level, traffic crashes were reduced by 67% in Santander, Spain (Aloi et al., 2020). Nevertheless, the reduction of road crashes and deaths were generally lower than the decrease of traffic volumes (ETSC, 2020; Sedain and Pant, 2020; Sekadakis et al., 2021; Shaik et al., 2021).

In relation to the traffic crash profile during lockdown measures, single vehicle crashes rates increased during lockdown period (Chand et al., 2021; Doucette et al., 2021b), with a higher presence of individuals aged 25 to 64, males, and non-white drivers (Barnes et al., 2020). In Detroit, angle crashes increased, which is related with more severe and driver behavior issues (Yao et al., 2021). At the state or provincial level, crashes were expected to be more severe in highways (Dong et al., 2022), and the reduction of crashes was higher in urban than in interurban environments (Saladié et al., 2020). Analogously, at the city level, Chand et al. (2021) found a change in the location of crashes in Sydney, where more incidents were recorded in suburban areas than before.

In this scenario, it seems clear that road mobility reduction during the lockdown period led to speeding as drivers found emptier roads, which provoked a decrease in less severe injuries, but not in injuries resulting in serious or fatal outcomes (Brodeur et al., 2021; Qureshi et al., 2020; Shaik and Ahmed, 2022). A study by Sekadakis et al. (2021) using Greece data confirmed that the reduction in traffic crashes was lower than the reduction in traffic volume, so a worse road safety performance was detected. For instance, in Spain the amount of speed violations observed on non-urban roads increased by 39% during lockdown (ETSC, 2020). Other risky driving behaviors such as failing to signal were more present during the lockdown (Dong et al., 2022).

The period after lockdown brought with it, as expected, a general increase in road traffic crashes. For example, in Missouri, an increase in traffic crashes involving slight or no injuries was observed after the expiration of lockdown, but not in more severe crashes (Qureshi et al., 2020). Similarly, crash rates returned to previous values after the stay-at-home order in Connecticut, although fatal crashes were not affected by the stay-at-home order (Doucette et al., 2021). Saladié et al. (2023) found that interurban road crashes in Catalonia (Spain) decreased during lockdown and returned to normal levels in September 2020, about two months after lockdown. The study also found that metropolitan areas remained hotspots throughout the period study, while rural areas presented a more pronounced rebound in traffic crashes during the post-lockdown period.

Research on the effect of the pandemic on older pedestrian road safety is scarce, but some researchers set the spotlight on pedestrians and older adults. Pedestrian deaths were reduced during lockdown periods in New York City and Toronto, but this reduction was not sustained after the lockdown period (Redelmeier and Zipursky, 2021). During the pandemic, the rate of pedestrian crashes and the proportion of injured pedestrians aged 0 to 18 decreased in Budapest, while the severity and location of these crashes did not change compared to the previous situation (Krizsik and Pauer, 2023). In Sydney, the rate of crashes involving pedestrians also decreased (Chand et al., 2021). Regarding drivers' behavior, the willingness to yield the right of way to pedestrians at a pedestrian crossing decreased during the pandemic (Krizsik and Pauer, 2023).

Turning to research on older adults during the pandemic, Rapoport et al. (2021) focused on the oldest old (people above 80) and found a not statistically significant reduction in the odds of an oldest old pedestrian being in an injury or fatal traffic crash during the pandemic compared to the pre-pandemic period. Other analysis with traffic crashes involving older adults in Adelaide (US) found that the number of traffic crashes decreased by 20% comparing before and after the pandemic, but the fundamental factor conditioning the crashes remained the same (Soltani et al., 2023). The proportion of crashes during daylight and the rate of severe injuries increased during the pandemic, and during April to June decreased because of the lockdown measures. Despite these previous studies, no comparison with the rest of population age groups has been accomplished.

There is a need to study the variation in the number and injury severity of vehicle collisions suffered by older pedestrians in urban roads, compared to other age groups of population, not only during the lockdown period, but also in the subsequent periods of the pandemic, such as after public vaccination. In addition, the effect of the pandemic on the location and crash-related factors of older pedestrian crashes should be studied to assess if countermeasures proposed before COVID-19 to mitigate the number of these crashes are still suitable nowadays. From the methodological point of view, in a first phase, Poisson regression models were used to assess the reduction of crashes suffered by older pedestrians at the national level during and after lockdown restrictions, with Spain as the case study. In a second stage of the methodology, a spatial analysis was carried out to evaluate if the hotspots of pedestrian crashes have been altered and a binary logistic regression was used to detect changes in the crash-related features before and after the pandemic, with Madrid as the case study.

In this context, the paper aims to clarify whether the pandemic and post-pandemic measures caused a reduction of older pedestrian traffic crashes, a change of the injury severity of older pedestrian traffic crashes; and whether the epidemic changed the location of crashes and the built environment characteristics of the location of crashes. This article is structured as follows: Section 1 contains the introduction and the state of the art; Section 2 describes the case study, Spain (used in the first stage of the methodology) and Madrid (used in the second stage); Section 3 gives a detailed description of the database and the models used in each phase of the methodology; Section 4 shows and explains the modeling results for each phase of the methodology. Finally, Section 5 presents the discussion of results, conclusions of the paper and future research lines drawn from this research.

2. Case study

For the first stage of the methodology, Spain has been selected for this case study due to its significant ageing population. As of 2021, 19.8% of the Spanish population was over 65 years old, a figure expected to rise to 36.8% by 2050 (United Nations, 2020). This rapid ageing process has directly impacted road safety, with 70% of total pedestrian fatalities in Spanish urban areas in 2019 involving older adults (Dirección General de Tráfico, 2020).

Spain's population of 47 million is distributed across 8131 municipalities, with a majority living in urban areas. Specifically, 79.7% of the population resides in cities with more than 10,000 inhabitants, highlighting the urban focus of this study. There is an inverse relationship between ageing rates and population size, with smaller municipalities having a higher ageing ratio. For example, small towns have an ageing ratio of 28.5%, compared to 18.5% in large cities like Madrid and Barcelona (Pérez Díaz et al., 2020).

Key dates during the COVID-19 pandemic in Spain are crucial for understanding the context of the study, as this research aims to assess the effect of the pandemic mainly on the post-lockdown periods (the lockdown period has been widely studied in the literature) forming representative and long-enough time spans. Important dates regarding the pandemic in Spain are shown in Table 1.

The first period considered, the state of alarm, includes lockdown measures and it is, inevitably, the shortest considered period, lasting roughly three months. The Spanish government declared a state of alarm on March 14, 2020, (Gobierno de España, 2020), which included strict confinement measures that lasted until June 20, 2020. In this study, the first considered day of state of alarm is March 15, 2020. During this period, mobility was severely restricted, impacting the number of trips. The subsequent considered period, following the state of alarm, is the termed 'new normality' starting on June 21, 2020, characterized by gradual easing of restrictions, though measures like mandatory mask-wearing remained in place.

The effect of vaccination should also be assessed as it might lead to a lower perceived risk of infection. The vaccination campaign in Spain began on December 27, 2020, prioritizing older adults and other high-risk groups. By mid-2021, a significant portion of the population, including most older individuals, had received at least one dose of the vaccine. The chosen date to define the last analysis phase is when the full vaccination rate surpassed 70% nationally, which happened on August 31, 2021 (Fig. 1). In the context of this study, this phase ends on December 31, 2022, as national traffic crash data is not available yet for 2023 and next years.

These three phases –the state of alarm, new normality, and 70% vaccination period– provide a framework for analyzing changes in pedestrian crash frequency and severity among older adults. By examining data across these distinct periods, the study aims to analyze the impact of the pandemic and its associated restrictions on urban pedestrian safety for older adults. The weekly number of infected people cases, deaths and full vaccination rate are shown in Fig. 1.

For the second stage of the methodology, Madrid, the capital and the most populated city in Spain, serves as an ideal case study the evolution of the features of the pedestrian traffic crashes due to its high ageing rate and diverse urban landscape. In 2022, Madrid had approximately 3.3 million inhabitants, with individuals over 65 and 80 years old representing 20.3% and 7.1% of the population, respectively, slightly above the national average. The city is administratively divided into 21 districts, further subdivided into 131 neighborhoods.

3. Data and methods

This section describes the databases and methods used in this research in order to analyze the evolution of older pedestrian traffic crashes in urban scenarios through the COVID-19 pandemic. As the study of the evolution of crash frequency and injury severity has been complemented with an analysis of the evolution crash-related factors, two methodologies with their corresponding statistical models have been defined. For better understanding, Fig. 2 shows a scheme of the two stages of the methodology.

3.1. Stage 1. analysis of the evolution of urban older pedestrian traffic crashes frequency and injury severity during COVID-19 in Spain

The first stage of the methodology aims to analyze trends in the frequency and injury severity of pedestrian crashes involving individuals over the age of 65 in urban areas. In order to have sufficient monthly data of the number of injured pedestrians, especially for the killed pedestrians, and avoid an excess of zeros (no observed crashes), Spain was used as the case study (all the Spanish urban pedestrian crashes). This analysis covers the period before the pandemic as well as during different COVID-19 phases, including the state of alarm, new normality, and 70% vaccination periods. By examining these trends, the study aims to understand how the pandemic and its associated phases have impacted the safety of older pedestrians.

Table 1
Important dates of COVID-19 events and measures in Spain.

Event or Measure	Date
First confirmed case of COVID-19 in Spain	January 31, 2020
Declaration of state of alarm (lockdown)	March 14, 2020
End of state of alarm (new normality)	June 21, 2020
Start of public vaccination campaign	December 27, 2020
Full vaccination rate surpassed 70%	August 31, 2021

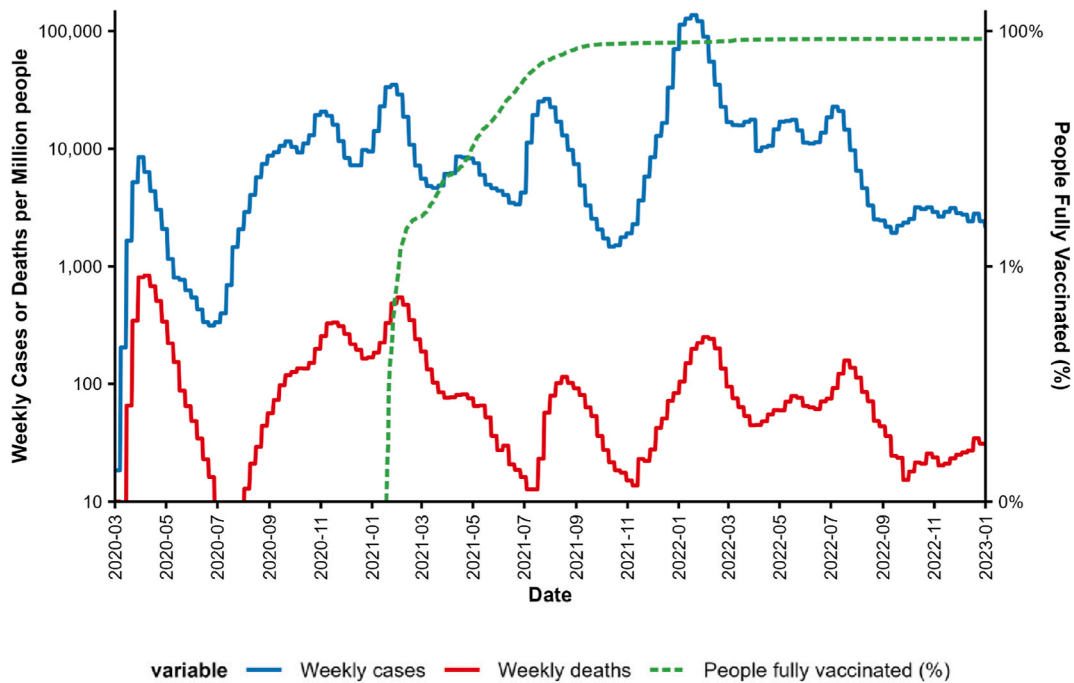


Fig. 1. Weekly infection cases and deaths per million caused by COVID-19 and percentage of people fully vaccinated in Spain.

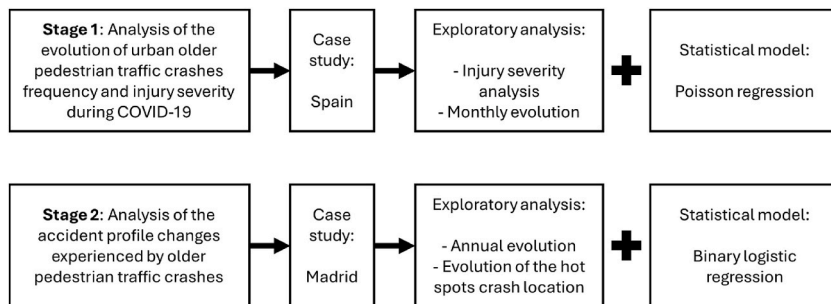


Fig. 2. Scheme of the methodology.

3.1.1.1. Database

The total number of pedestrians involved in vehicle collisions in urban environments was calculated for each month and year, considering the age of the pedestrian (i.e., more than 65 years old). Similarly, the number of pedestrians by their injury level (fatalities, serious and slight injuries) was also computed. The period of study was assigned to each record in terms of the month and year. Four periods were used in the study: (1) pre-covid, (2) state of alarm (3) new normality, and (4) full vaccination rate over 70%. The former one was used as the baseline situation, so the rest periods were compared with the situation before covid-19 lockdown. Although these periods are identified between specific dates, these were defined considering complete months (as shown in Table 2).

An alternative method for defining these periods is the Stringency Index (SI) from the Oxford COVID-19 Government Response Tracker (Hale et al., 2021), which consolidates governmental policy responses to COVID-19. However, defining specific periods was preferred for three key reasons. First, traffic crashes are random events, requiring broad timeframes for meaningful analysis. Using SI

Table 2
Periods defined for the pedestrian traffic crash occurrence study.

Period	Date	Months
Pre-COVID	January 1, 2016–March 14, 2020	January 2016–February 2020
State of alarm	March 15, 2020–June 21, 2020	March 2020–June 2020
New normality	June 22, 2020–August 30, 2021	July 2020–August 2021
Full vaccination >70%	August 31, 2021–December 31, 2022	September 2021–December 2022

as an independent variable would necessitate daily crash counts as the dependent variable, resulting in too many zero observations, especially for fatalities and seriously injured pedestrians. Second, not all SI measures, such as international travel restrictions, potentially influence pedestrian behavior. Third, the main objective of this part of the study is to assess whether pedestrian crash levels have returned to pre-pandemic conditions, a goal more effectively addressed by defining distinct periods, including the final phase (i. e., full vaccination rate above 70%).

3.1.2. Data exploratory analysis

The number of injured pedestrians in vehicle-pedestrian collisions considering the month, period of study, and age of the pedestrian is shown in Fig. 3. At first glance, the data shows a clear decrease in the number of older and non-older pedestrian traffic crashes per month during the pandemic, especially in the state of alarm and new normality periods. The first month of full state of alarm (April 2020) exhibits a remarkable minimum for the series.

Additionally, the database was analyzed using multiple chi-squared tests comparing the pre-pandemic situation with the three considered pandemic phases. Results of this analysis are shown in Table 3. Regarding the crash type, the share of vehicle-pedestrian collisions was reduced compared to the pre-pandemic situation. About injury severity of all traffic crashes, more seriously injured and less slightly injured victims were registered during the pandemic. Notably, seriously injured victims rose from 5.5% to 6.8% during the state of alarm. A statistically significant but less pronounced difference is shown for the rest of periods.

Regarding vehicle-pedestrian collisions, the injury severity of the pedestrians did not statistically change during any of the pandemic phases, but more seriously injured pedestrians were registered during the state of alarm than before (13.2% vs 12.0%). After the full vaccination rate reached 70%, injury severity of pedestrians got back to pre-pandemic levels. Younger pedestrians (0–24) were less present among injured pedestrians than before the pandemic, while middle-aged pedestrians (45–64) were more present during all the pandemic phases. During the state of alarm, the presence of older pedestrians was reduced (28.3%–26.7%). In subsequent phases, their presence was slightly reduced, but near pre-pandemic levels. Finally, the injury severity of older pedestrians was altered during the state of alarm period, killed older pedestrians were reduced (3.4% vs 4.6%) and seriously injured increased (19.8% vs 17.7%). In the next phases, the injury severity of older pedestrians returned to pre-pandemic levels.

3.1.3. Poisson regression

The Poisson regression was used to explain the influence of the different COVID periods on the number of older and non-older pedestrian victims at urban locations by month in Spain, considering the age of the pedestrian and the injury severity. The Poisson

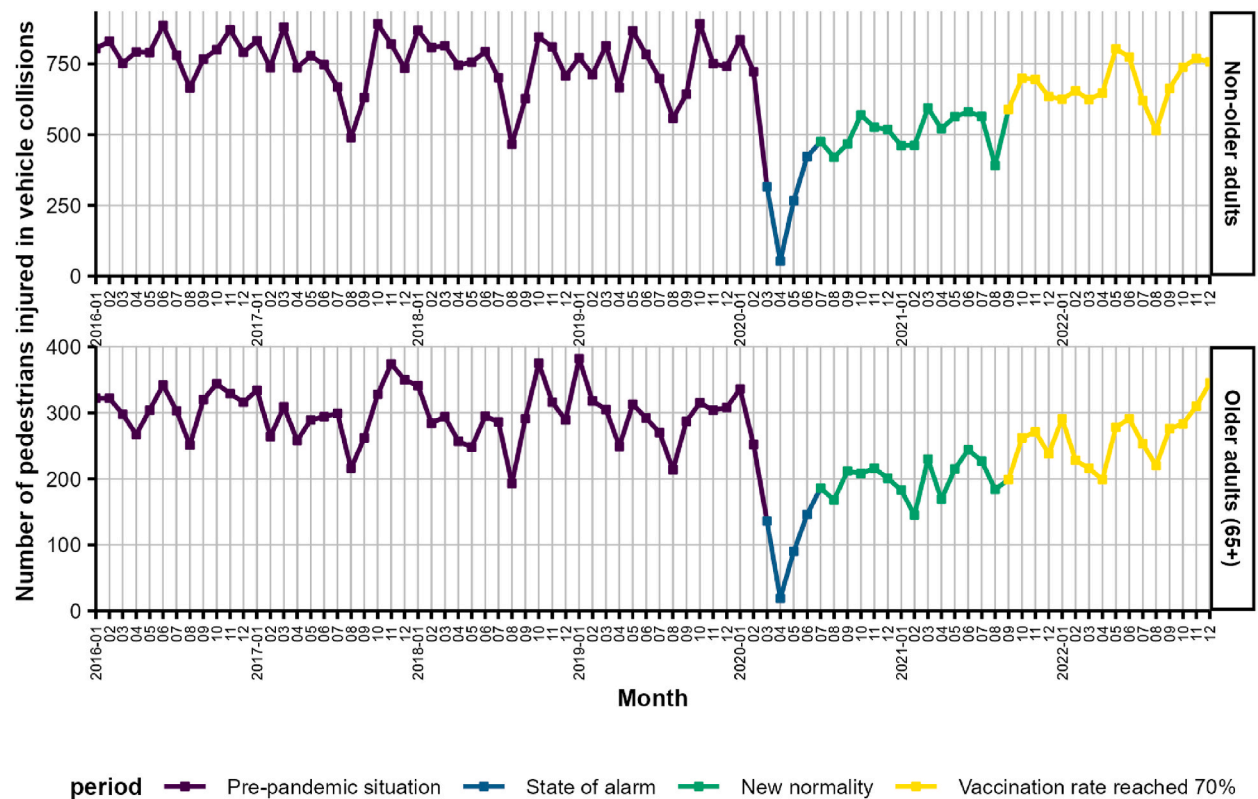


Fig. 3. Monthly evolution of the number of injured pedestrians in Spain (2016–2022) at urban areas, considering the age of the pedestrian (above, non-older adults; below, older adults).

Table 3

Comparison of crash type and injury severity of traffic crashes, and comparison of age distribution and injury severity of injured pedestrians in the different pandemic phases.

	Before COVID-19	State of alarm	χ^2	p-value	New normality	χ^2	p-value	Vaccination above 70%	χ^2	p-value
All urban traffic crashes										
Crash type										
Vehicle-pedestrian collision	53632 (19.5)	907 (14.3)	106.64	0.000	10404 (16.2)	386.6	0.000	15270 (17.9)	119.47	0.000
Other	220786 (80.5)	5415 (85.7)			53931 (83.8)			70230 (82.1)		
Injury severity of victims										
Killed	2134 (0.6)	54 (0.7)	1.32	0.250	509 (0.6)	0.6	0.456	624 (0.6)	0.5	0.471
	19153 (5.5)	511 (6.8)	23.47	0.000	4542 (5.7)	3.7	0.054	6145 (5.8)	17.8	0.000
Seriously injured	327557 (93.9)	6974 (92.5)	24.85	0.000	75160 (93.7)	4.3	0.038	98631 (93.6)	14.3	0.000
Slightly injured										
Vehicle-pedestrian collisions										
Age of injured pedestrians										
0 to 14	7115 (13.4)	93 (10.7)	5.31	0.021	1337 (13.3)	0.06	0.813	1852 (12.4)	10.82	0.000
	6601 (12.4)	94 (10.8)	2.01	0.156	1126 (11.2)	11.71	0.000	1884 (12.6)	0.25	0.619
15 to 24	5297 (10.0)	87 (10.0)	0.00	0.962	965 (9.6)	1.28	0.258	1466 (9.8)	0.44	0.506
25 to 34	6163 (11.6)	106 (12.2)	0.31	0.580	1183 (11.8)	0.24	0.622	1700 (11.3)	0.73	0.394
35 to 44	6435 (12.1)	127 (14.6)	5.06	0.025	1276 (12.7)	2.70	0.100	1884 (12.6)	2.35	0.125
45 to 54	6531 (12.3)	130 (15)	5.67	0.017	1354 (13.5)	10.79	0.001	2034 (13.6)	17.55	0.000
55 to 64	15028 (28.3)	232 (26.7)	1.04	0.309	2816 (28.0)	0.29	0.590	4172 (27.8)	1.10	0.295
Above 65										
Injury severity of pedestrians										
Killed	998 (1.8)	13 (1.5)	0.65	0.420	199 (1.9)	0.51	0.474	274 (1.8)	0.14	0.707
	6599 (12.0)	118 (13.2)	1.17	0.280	1193 (11.5)	2.11	0.146	1860 (12.0)	0.00	0.973
Seriously injured	47391 (86.2)	764 (85.4)	0.50	0.480	8985 (86.6)	1.19	0.276	13352 (86.2)	0.01	0.910
Slightly injured										
Injury severity of older pedestrians										
Killed	698 (4.6)	8 (3.4)	0.74	0.389	124 (4.4)	0.31	0.575	178 (4.3)	1.07	0.300
	2665 (17.7)	46 (19.8)	0.69	0.408	500 (17.8)	0.00	0.978	750 (18.0)	0.13	0.716
Seriously injured	11665 (77.6)	178 (76.7)	0.11	0.745	2192 (77.8)	0.07	0.798	3244 (77.8)	0.03	0.853
Slightly injured										

distribution is a widely used regression model in studies examining road crashes. The negative binomial (NB) distribution was also tested, which is derived from the Poisson distribution, and it is particularly useful for handling data with overdispersion (i.e., the variance exceeds the mean), which is a limitation in the Poisson regression model (Hilbe, 2011). We tested for overdispersion using Negative Binomial models. However, the estimated overdispersion parameter (α) was consistently close to zero across all models (max $\alpha \approx 0.015$), indicating that overdispersion was negligible. Consequently, Poisson regression was deemed the preferable approach.

The dependent variable is the number of injured pedestrians by month. For each age group of pedestrians (older and non-older), four regressions were designed to model the total number of injured pedestrians, along with killed, severely injured and slightly injured pedestrians. The number of injured pedestrians follows a Poisson distribution (Eq. (1)). Independent variables are the national population of each age group, the month of the year, the year, and the phase of the COVID-19 pandemic (pre-covid, lockdown, new normality, and post-vaccination). Eq. (2) shows the natural logarithm of the parameter of the Poisson distribution, which has a linear relationship with the independent variables.

$$Y_{y,m}^{a,s} \sim \text{Poisson}(\lambda_{y,m}^{a,s}) \tag{1}$$

$$\ln(\lambda_{y,m}^{a,s}) = \text{offset}(\log(\text{population}^a)) + \text{Period} + \beta_{\text{year}} * \text{Year} + \text{Month} \tag{2}$$

Where $Y_{y,m}^{a,s}$ is the number of pedestrians injured in vehicle-collisions in urban areas of the a^{th} age group (older or non-older) of s^{th} injury severity (total, killed, seriously or slightly injured), in the m^{th} month of the y^{th} year, $\lambda_{y,m}^{a,s}$ is the parameter of the Poisson regression, population is the number of inhabitants of the specific age group, period is categorical variable that can be pre-covid, state of alarm, new normality, and post-vaccination, year is the specific year, and month is a categorical variable to consider the effect of the month.

The model will provide the significance of the impact of each pandemic phase on the frequency and severity of pedestrian traffic crash. Table 4 shows the monthly mean number of injured pedestrians considering the pedestrian’s age and the injury severity. The results of the modeling are showed and explained in section 4.1 (Table 6).

3.2. Stage 2. comparison between pre- and post-pandemic factors related with the occurrence of older pedestrian traffic crashes in Madrid

The second stage of the methodology aims to identify and compare the variables related to pedestrian crashes before and after the pandemic. By examining the differences in crash-related factors such as time of day, environmental characteristics, and other relevant variables, the study seeks to understand in the pandemic has modified potential crash causes. For this purpose, Madrid, the capital of Spain was used as the case study. Instead of dividing the pandemic in three phases as before, only two period were considered. Pre-pandemic crashes were collected during the period 2018–2019 and post-pandemic during 2021–2022 in order to have a similar number of crashes from each period, thereby the year 2020, was not considered. Before modeling pedestrian crashes in Madrid and as part of the exploratory analysis, an analysis of the pedestrian victims through the two periods has been carried out together with a study on the evolution of the crash hot spot locations.

3.2.1. Database

In this research, each data sample represents an injured or killed pedestrian in a vehicle collision. The urban location of each pedestrian crash from the Spanish traffic crashes database (postal address) has been used to geolocate these crashes, and we were able to accurately locate almost 90% of the crashes. These crashes were then classified considering the age of the pedestrian (above 65, older pedestrian; otherwise, non-older pedestrian) and the period (2018–2019, pre-pandemic; 2021–2022, post-pandemic).

Each injured pedestrian was associated with two groups of variables. The first group of features, crash-specific data, consists of unique variables of each crash, specifically the time of the day, day of the week (weekday or weekend), crash location (inside or outside intersection), lighting and weather conditions of the crash, and the pedestrian gender. These features are present in the Spanish traffic crash database provided by DGT. The second group of features is related to crash location-specific variables, including socioeconomic, land use, and infrastructure, as examined in previous research on older pedestrian road safety (Gálvez-Pérez et al., 2022, 2024; D. Kim, 2019; Lee et al., 2020; Lv et al., 2021).

Turning to location-specific data, socioeconomics variables include population density, ageing rate (proportion of people above 65), rate of Spanish inhabitants, and average income per person. These variables were captured from the census tract where each crash occurred. For simplicity and because socioeconomics features show little modifications in the short term, these were recorded for the middle year of the study (2020) and obtained from the Spanish National Statistics Institute (INE).

The rest of location-specific variables, land use and infrastructure, were recorded from spatial buffers created around each crash with a radius of 400 m (1/4 mile), which has been previously used in road safety studies (Clifton et al., 2009; Park and Bae, 2020). It is also a common measure for the 5-min walk. Although all buffers have equal surface, they might contain lighter or denser road networks, so features were normalized considering the total length of the road network in each buffer. Land use and infrastructure features variables were determined for the last or only available data, regardless of the year each crash occurred.

Land use features included the number of points of interest (POIs), extracted from OpenStreetMap (OSM), and the proportion of the main land uses (i.e., primary, secondary, tertiary, and residential), inside each buffer, obtained from the Spanish land occupation information system (SIOSE). Primary and secondary were excluded from the final dataset as these are negligible in urban scenarios.

Table 4
Summary of injured pedestrians monthly mean used in the Poisson regression of the Stage 1 of the methodology.

Pedestrian’s age group	Period	Number of Months	Monthly mean of number of injured pedestrians			
			Total	Fatalities	Seriously injured	Slightly injured
Older adults	(0) Pre-COVID	50	298.2	13.8	52.9	231.5
	(1) State of alarm	4	97.8	4.0	18.5	75.3
	(2) New normality	14	199.1	8.9	35.6	154.6
	(3) Post-vaccination	16	260.0	11.1	46.4	202.5
Non-older adults	(0) Pre-COVID	50	757.3	5.8	75.8	675.7
	(1) State of alarm	4	264.5	2.0	23.8	238.8
	(2) New normality	14	507.9	4.9	48.4	454.6
	(3) Post-vaccination	16	675.3	5.7	67.1	602.5

The calculation of each land use share was performed in term of the total surface of the buffer (Eq. (3)). Land use mix was computed using the Herfindahl-Hirschman index (HHI) (Eq. (4)) with the land use proportions previously calculated.

$$\%LU_{ij} = \frac{A(LU_{ij})}{\sum_{j=1}^k A(LU_{ij})} \quad (3)$$

Where $\%LU_{ij}$ is the proportion of the j^{th} land use in the buffer around the i^{th} traffic crash, $A(LU_{ij})$ is the surface of the j^{th} land use in the buffer around the i^{th} traffic crash, and k is the number of land uses included in the i^{th} buffer.

$$HHI_i = \sum_{j=1}^k (\%LU_{ij})^2 \quad (4)$$

Where HHI_i is the HHI land use mix index around the i^{th} traffic crash, $\%LU_{ij}$ is the proportion of the j^{th} land use in the buffer around the i^{th} traffic crash, and k is the number of land uses included in the i^{th} buffer.

Finally, the infrastructure features considered in the analysis were the mean width of roads and sidewalks; share of major (i.e., primary, secondary and tertiary) roads over the total street network length, the number of bus stops and metro stations, traffic lights, and crosswalks in each buffer; and total length of bike lanes. Street and sidewalk polygonal geometry, traffic lights devices and bike lanes geometry were retrieved from Madrid Open Data Portal; linear street geometry and crosswalks were obtained from OSM; bus and metro points were extracted from CRTM Open Data Portal.

Table 5

Summary of the variables included in the model of the Stage 2 of the methodology.

Variable	Injured older pedestrians		Injured non-older pedestrian	
	Pre-pandemic	Post-pandemic	Pre-pandemic	Post-pandemic
	(N=697)	(N=454)	(N=2104)	(N=1584)
Weekday				
Weekday	592 (84.9%)	374 (82.4%)	1647 (78.3%)	1267 (80.0%)
Weekend	105 (15.1%)	80 (17.6%)	457 (21.7%)	317 (20.0%)
Time of the day				
Morning	210 (30.1%)	154 (33.9%)	557 (26.5%)	365 (23.0%)
Early morning	2 (0.3%)	0 (0.0%)	116 (5.5%)	95 (6.0%)
Afternoon	284 (40.7%)	199 (43.8%)	704 (33.5%)	549 (34.7%)
Evening	150 (21.5%)	80 (17.6%)	464 (22.1%)	384 (24.2%)
Night	51 (7.3%)	21 (4.6%)	263 (12.5%)	191 (12.1%)
Lighting condition				
Daylight	539 (77.3%)	377 (83.0%)	1340 (63.7%)	997 (62.9%)
Dawn or dusk	16 (2.3%)	12 (2.6%)	85 (4.0%)	60 (3.8%)
Artificial light	138 (19.8%)	64 (14.1%)	615 (29.2%)	515 (32.5%)
Dark no light	4 (0.6%)	1 (0.2%)	64 (3.0%)	12 (0.8%)
Weather condition				
Fair weather	308 (44.2%)	411 (90.5%)	834 (39.6%)	1394 (88.0%)
Inclement weather	26 (3.7%)	38 (8.4%)	111 (5.3%)	173 (10.9%)
Unknown	363 (52.1%)	5 (1.1%)	1159 (55.1%)	17 (1.1%)
Crash location				
Road section	407 (58.4%)	255 (56.2%)	1288 (61.2%)	955 (60.3%)
Road intersection	290 (41.6%)	199 (43.8%)	816 (38.8%)	629 (39.7%)
Pedestrian gender				
Man	278 (39.9%)	181 (39.9%)	942 (44.8%)	753 (47.5%)
Woman	419 (60.1%)	273 (60.1%)	1162 (55.2%)	831 (52.5%)
Population density	30300 (17700)	32000 (18300)	26200 (18100)	27700 (18300)
Ageing rate	22.1 (6.50)	21.7 (7.02)	20.0 (7.29)	19.7 (7.29)
Spanish population	83.3 (8.47)	82.9 (8.94)	83.3 (8.99)	82.6 (9.12)
Average income	22400 (10500)	22000 (10400)	23500 (10900)	22500 (10800)
POIs per km	0.0110 (0.0147)	0.0113 (0.0147)	0.0126 (0.0177)	0.0122 (0.0179)
Land use proportion				
Tertiary	0.278 (0.125)	0.277 (0.131)	0.291 (0.137)	0.284 (0.128)
Residential	0.318 (0.127)	0.323 (0.134)	0.301 (0.138)	0.303 (0.135)
Land use mix HHI	0.335 (0.0498)	0.342 (0.0509)	0.340 (0.0622)	0.335 (0.0544)
Proportion of major roads	0.302 (0.147)	0.300 (0.161)	0.317 (0.163)	0.314 (0.169)
Mean width of the roads	10.9 (2.72)	10.8 (2.81)	11.0 (2.96)	10.9 (2.90)
Mean width of the sidewalks	5.99 (1.99)	5.82 (1.88)	5.93 (2.16)	5.91 (1.98)
Bus stops per km	1.55 (0.675)	1.52 (0.624)	1.56 (0.723)	1.50 (0.655)
Metro stations per km	0.112 (0.132)	0.110 (0.129)	0.126 (0.157)	0.121 (0.154)
Traffic lights per km	20.0 (9.24)	19.4 (9.83)	19.7 (9.98)	19.1 (9.89)
Crosswalks per km	7.07 (3.25)	7.08 (3.26)	6.96 (3.30)	6.90 (3.41)
Total length m bike lanes m	0.0491 (0.070)	0.0554 (0.080)	0.0583 (0.084)	0.0588 (0.081)

The summary of these variables is shown in Table 5. For the modeling process, the weather condition variable was removed as there were unknown for roughly half of the crashes of the pre-pandemic situation, and for only 1% of the post-pandemic crashes, so removing crashes with unknown weather would unbalance the two classes of the dependent variable (pre- and post-pandemic). The final datasets consisted of 1,151 injured older pedestrians and 3,688 injured non-older pedestrians.

3.2.2. Data exploratory analysis

In 2020, like most cities worldwide, Madrid experienced a decline in road traffic crashes, particularly pedestrian traffic crashes. Fig. 4 illustrates the annual evolution of the number of pedestrians killed or injured in vehicle collisions by age, along with the number of drivers involved in crashes of all type. It also includes the mean data for the pre-COVID years (2016–2019), as these trends are expected to continue two years post-pandemic (2021–2022).

Considering the number of drivers involved in traffic crashes, the difference between the expected and actual number of drivers is greater among younger people, and it decreases while age increases. The highest difference is found for drivers between 25 and 34 years old, as the expected number is above 3000 and the actual one is roughly 2500. Conversely, the number of drivers above 65 returned to pre-pandemic levels in 2022, while the number of injured pedestrians show the opposite result. Regarding injured pedestrians, the level of injured pedestrians was restored to the pre-pandemic situation for young ages in 2022. As the pedestrian's age increases, the gap between expected and real number of pedestrians increases. It is interesting to underline that older pedestrians show the highest gap, as the number of involved pedestrians two years after the state of alarm (2022) is 287 and the expected number is about 400. In fact, during the year after the pandemic started (2021) it was registered about the same number of pedestrians as during 2020, when stricter mobility measures were active.

In the context of these figures, it could be argued that the use of private vehicles among young drivers has been reduced after the pandemic, but older drivers are not such affected. In addition, walking activity of most of the population, which is related to higher risk exposure to suffer a vehicle collision, was restored to normal levels two years after the pandemic. Older pedestrians are an exception for this, and they might have reduced their walking activity because they are the most vulnerable demographic for the disease, even after lockdown measures ended.

Apart from the study of the evolution of pedestrian victims in Madrid post-pandemic, changes in the location of pedestrian crashes were analyzed using the Getis-Ord G_i^* statistic. This statistic G_i^* is already a z-score, and it identifies hot and cold spots with significant levels, with high values indicating hot spots and low values indicating cold spots. This technique requires a spatial data layer, here

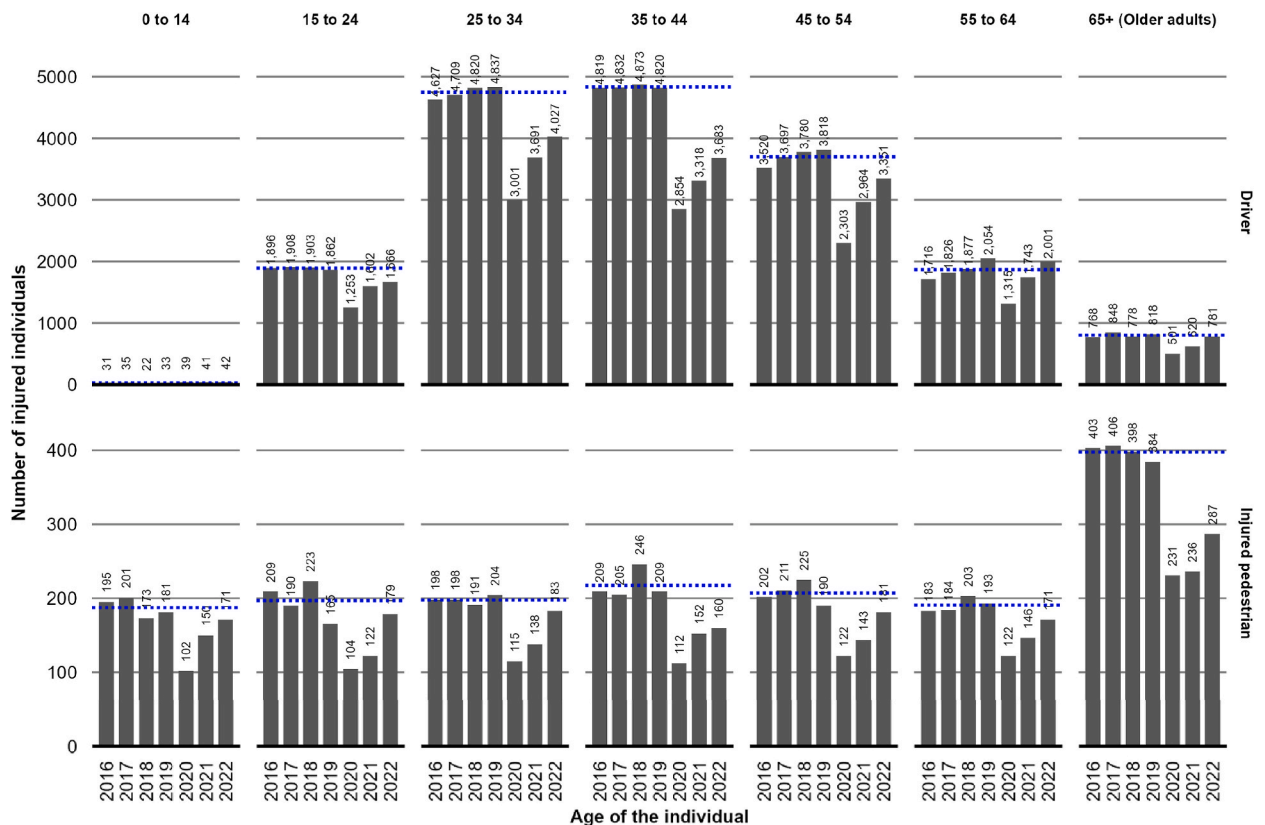


Fig. 4. Annual evolution of the number of injured or killed pedestrians and drivers involved in traffic crashes in Madrid, disaggregated per age group.

represented by a 500×500 m grid, using the number of injured pedestrians as the attribute variable. This grid was preferred over administrative divisions to avoid assigning crashes near major roads to polygons inaccurately. The analysis was conducted separately for older and non-older pedestrians, as well as pre- and post-pandemic periods.

The Getis-Ord G_i^* statistic was calculated for each cell, as shown in Fig. 5. The city center remained a hot spot for both age groups post-pandemic. For older pedestrians, hot spots in the south and east intensified and shifted slightly towards the periphery, while those in the west diminished or disappeared. For non-older pedestrians, hot spots increased on the periphery, particularly in the east, while northern hot spots notably decreased after the pandemic.

3.2.3. Binomial logistic regression

A binomial logistic regression was conducted to examine differences in pedestrian crashes factors before and after the pandemic. Crashes from two years before (2018–2019) were compared to those from two years after (2021–2022). As in the previous analysis, crashes involving older and non-older pedestrians were modeled separately. This approach allows for analyzing both crash-specific and location-specific data. Instead of examining crashes by city divisions, each crash is treated as a data sample, with the dependent variable indicating the crash period: before (2018–2019) or after (2021–2022) the pandemic.

In this research, logistic regression predicts the probability of a crash occurring during the post-pandemic period (class '1'), with the pre-pandemic period serving as the baseline (class '0'). The term $\log(\text{odds})$ (Eq. (5)) represents the natural logarithm of the odds, which is the probability of a crash occurring during the post-pandemic period relative to the pre-pandemic period. The regression parameters were estimated using the maximum likelihood method.

$$\log(\text{odds}) = \log(P/1 - P) = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n = x\beta \quad (5)$$

Where P is the probability that a sample belongs to the group '1' (i.e., the traffic crash has occurred in the post-pandemic period), $\beta_0, \beta_1, \dots, \beta_n$ are the regression parameters, and x is the vector of independent variables.

An automatic model selection procedure was used to enhance explanatory power and simplify interpretation. Exhaustive model selection was impractical and computationally prohibitive due to the over 250 million possible independent variable combinations (2^{28} , with 28 independent variables). Instead, stepwise model selection was employed, which iteratively adds or removes variables until no further improvement is observed, as evaluated by the Akaike Information Criterion (AIC). All stepwise approaches (backward, forward, and both directions) were utilized. The modeling results are presented in Section 4.2 (Table 7).

4. Results

4.1. Results obtained in the stage 1 of the methodology

The methodology used in the Stage 1 was the Poisson regression. The results of this analysis, which examined the monthly count of older and non-older pedestrians injured in vehicle collisions in urban environments, categorized by injury severity, are presented in Table 6.

As expected, total pedestrian crashes and all injury severities decreased significantly during the state of alarm period for both older and non-older pedestrians. Specifically, total crashes dropped by 65.7% ($\text{IRR} = e^{-1.07}$, $p < 0.01$) for both older and non-older pedestrians. Fatalities, serious injuries, and slight injuries also declined notably, reflecting the strong impact of lockdown measures.

In the subsequent 'new normality' period, following the lockdown, total crashes continued to decrease compared to the pre-COVID situation, although this reduction was less pronounced than during the lockdown. For older pedestrians, crashes dropped by 30.2%

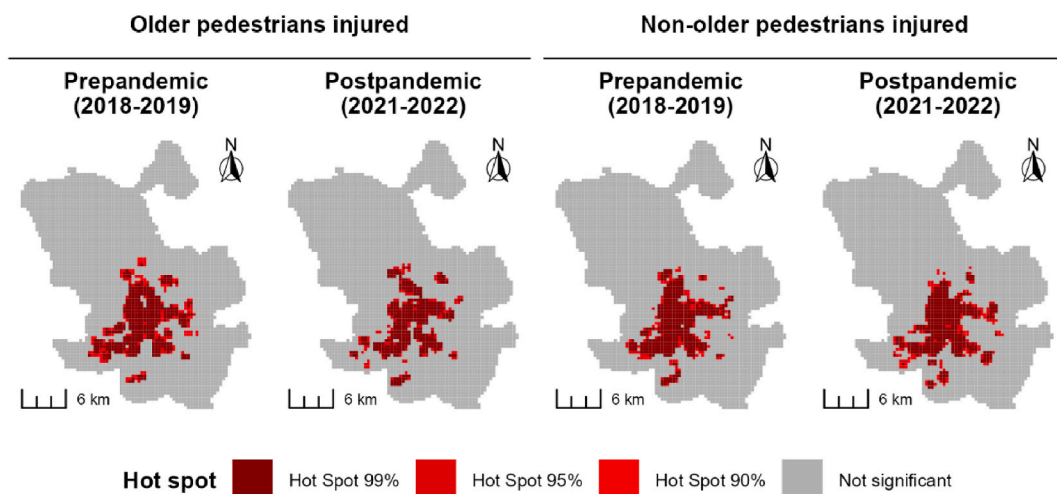


Fig. 5. Hot spots of injured pedestrians in Madrid by age group and period (pre- and post-covid-19) using Getis-Ord G_i^* statistic.

Table 6
Results of the Poisson regressions of the Stage 1 of the methodology for older and non-older pedestrians.

Variable	Older pedestrians (65+)							Non-older pedestrians								
	Total		Fatality		Seriously injured		Slightly injured	Total		Fatality		Seriously injured		Slightly injured		
Intercept	39.29	***	48.95		85.16	***	27.33	*	18.20	**	68.38		68.26	**	11.72	
Period																
State of alarm	-1.07	***	-1.03	***	-0.93	***	-1.11	***	-1.07	***	-0.85	**	-1.12	***	-1.07	***
New normality	-0.36	***	-0.39	***	-0.28	***	-0.37	***	-0.34	***	-0.02		-0.31	***	-0.35	***
Post-vaccination	-0.11	***	-0.17		0.00		-0.14	***	-0.06	***	0.09		0.04		-0.07	***
Year	-0.02	***	-0.03		-0.05	***	-0.02	**	-0.01	***	-0.04		-0.04	***	-0.01	**
Month																
February	-0.19	***	-0.21		-0.23	***	-0.18	***	-0.05	***	0.00		0.02		-0.06	***
March	-0.10	***	-0.11		-0.28	***	-0.06		0.02		-0.06		0.06		0.02	
April	-0.33	***	-0.56	***	-0.39	***	-0.30	***	-0.12	***	-0.06		-0.09		-0.12	***
May	-0.13	***	-0.39	**	-0.26	***	-0.08	**	0.03		0.00		0.09		0.02	
June	-0.04		-0.66	***	-0.09		0.01		0.06	***	-0.16		0.05		0.06	***
July	-0.14	***	-0.28	*	-0.25	***	-0.10	***	-0.10	***	-0.08		-0.16	**	-0.09	***
August	-0.37	***	-0.22		-0.40	***	-0.37	***	-0.35	***	-0.28		-0.28	***	-0.36	***
September	-0.15	***	-0.28	*	-0.30	***	-0.11	***	-0.16	***	0.16		-0.14	**	-0.17	***
October	-0.02		-0.05		-0.17	**	0.02		0.05	***	0.38	*	0.17	***	0.04	*
November	-0.02		-0.15		-0.05		0.00		0.02		0.23		0.03		0.01	
December	-0.05		-0.11		-0.13	*	-0.03		-0.05	***	0.25		-0.02		-0.06	***
n	84		84		84		84		84		84		84		84	
AIC	922.36		454.38		652.82		845.48		1,275.50		408.67		630.20		1231.20	
Log-likelihood at start (ℓ_0)	-1,087.50		-254.74		-425.31		-913.95		-2,153.95		-203.36		-483.94		-1,950.24	
Log-likelihood at convergence (ℓ_1)	-445.18		-211.19		-310.41		-406.74		-621.76		-188.34		-299.10		-599.62	
$\Delta\ell$ ($\ell_1 - \ell_0$)	642.31		43.55		114.90		507.21		1,532.19		15.03		184.84		1,350.62	
Likelihood Ratio Test (p-value)	0.00		0.00		0.00		0.00		0.00		0.01		0.00		0.00	

Statistical significance codes: '*' for 10%, '**' for 5%, and '***' for 1%.

Table 7

Results of the binary logistic regression of the Stage 2 of the methodology for older and non-older pedestrians.

Variable	Injured older pedestrians			Injured non-older pedestrians		
	Parameter	p-value		Parameter	p-value	
Intercept	-1.08E+00	0.02	**	3.08E-01	0.15	
Time of the day (base: morning)						
Early morning	-1.31E+01	0.97		-	-	
Afternoon	-6.18E-02	0.66		-	-	
Evening	-3.07E-01	0.08	*	-	-	
Night	-5.96E-01	0.04	**	-	-	
Lighting condition (base: daylight)						
Dawn dusk	-	-		-6.86E-02	0.69	
Artificial light	-	-		1.38E-01	0.06	*
Dark no light	-	-		-1.40E+00	0.00	***
Weekday: Weekend	-	-		-1.37E-01	0.10	
Pedestrian gender: Woman	-	-		-1.18E-01	0.08	*
Population density	5.42E-06	0.11		5.07E-06	0.01	***
Land use mix (HHI)	2.59E+00	0.03	**	-1.34E+00	0.02	**
Sidewalk mean width	-6.20E-02	0.07	*	-	-	
Bike lanes length per km	1.91E+00	0.03	**	-	-	
Bus stops per km	-	-		-1.33E-01	0.01	***
Number of samples (n)		1,151			3,688	
AIC		1,539.8			5,001.3	
Log-likelihood at start (ℓ_0)		-771.97			-2,519.55	
Log-likelihood at convergence (ℓ_1)		-760.89			-2,491.63	
$\Delta\ell$ ($\ell_1 - \ell_0$)		11.08			27.92	
Likelihood Ratio Test (p-value)		0.005			0.000	

Statistical significance codes: '**' for 10%, '***' for 5%, and '****' for 1%.

(IRR = $e^{-0.36}$, $p < 0.01$), and for non-older pedestrians, by 28.8% (IRR = $e^{-0.34}$, $p < 0.01$). Significant reductions in fatalities, serious, and slight injuries persisted for older pedestrians, while the decrease in non-older pedestrian fatalities was not statistically significant, indicating a return to pre-COVID fatality rates for this group.

When the vaccination rate surpassed 70%, the subsequent analysis period began. During this period, total crashes and slight injuries in older pedestrians decreased by 10.4% (IRR = $e^{-0.11}$, $p < 0.01$) and 13.1% (IRR = $e^{-0.14}$, $p < 0.01$), respectively, compared to the pre-pandemic situation. Similarly, the number of total crashes and slight injured non-older pedestrians were below the normal situation, but this decrease was roughly half than the one for older pedestrians, with a decrease of 5.8% (IRR = $e^{-0.06}$, $p < 0.01$) and 6.8% (IRR = $e^{-0.07}$, $p < 0.01$) for total and slightly injured non-older pedestrians. Fatality and serious injury rates for older and non-older pedestrians showed no significant differences from the pre-pandemic period, suggesting stabilization.

Additional monthly observations revealed varying crash reductions. For instance, in April, total crashes for older pedestrians decreased by 28.1% (IRR = $e^{-0.33}$, $p < 0.01$), with serious injuries dropping by 32.3% (IRR = $e^{-0.39}$, $p < 0.01$). Significant reductions in crashes and slight injuries were noted for non-older pedestrians for this month.

Overall, the data show a significant initial reduction in pedestrian crashes and injuries during the state of alarm, followed by a more moderate decline during the new normality, during which period non-older pedestrian fatalities registered normal pre-pandemic levels. The vaccination period presented less consistent reductions, with significant decreases among total and slightly injured older and non-older pedestrians. These reductions were roughly double for older pedestrians than for non-older pedestrians, indicating a potential higher reduction in the mobility of older adults.

4.2. Results obtained in the stage 2 of the methodology

After stepwise model selection, the same model was identified for older pedestrian samples using the forward, backward and both directions, making it the model chosen for interpretation. For non-older pedestrians, the model with the lowest AIC was selected using the forward direction method. The logistic regression results are presented in Table 7, with odds ratios (OR) calculated by exponentiating the regression coefficients (e^{β}).

The interpretation of the results is based on positive parameters (OR > 1), indicating that crashes are more common during the post-pandemic period when these variables are present (categorical) or higher (numerical). Conversely, negative parameters (OR < 1) indicate that crashes were more common pre-pandemic under those conditions.

Both models indicate that population density and land use mix (HHI) play significant roles in the determining the likelihood of crashes occurring during post-pandemic. Areas with higher population densities exhibit higher probabilities of crashes after the pandemic, the effect is higher for older pedestrians but not statistically significant. The land use mix (HHI) has a contrasting impact: for older pedestrians, crashes are more common in areas with higher land use mix (OR = 13.36, $p = 0.03$), whereas for non-older pedestrians, it significantly reduces the likelihood of post-pandemic crashes (OR = 0.26, $p = 0.02$).

The model for older pedestrian crashes includes also the time of the day, sidewalk mean width, and bike lanes. Evening and night crashes are 27% (OR = 0.73, $p = 0.08$) and 45% (OR = 0.55, $p = 0.04$), respectively, less likely to have occurred post-pandemic

compared to morning crashes. Older pedestrian crashes are more frequent in areas with narrower sidewalks and more bike lanes after the pandemic.

For non-older pedestrians, crashes under artificial lighting are 15% more likely to occur post-pandemic (OR = 1.15, $p = 0.06$), while those in unlit areas are 75% less likely (OR = 0.25, $p < 0.01$). Weekend crashes are 13% less likely post-pandemic, though not statistically significant (OR = 0.87, $p = 0.10$). Female pedestrians' likelihood of post-pandemic crashes decreases by 11% (OR = 0.89, $p = 0.08$). Fewer bus stops per kilometer are inversely related to post-pandemic crash likelihood, indicating a higher crash probability in areas with fewer bus stops.

5. Discussion, conclusions and future research lines

The impact of the COVID-19 pandemic on mobility and traffic safety has extended beyond the lockdown measures, particularly affecting older pedestrians. This demographic is already vulnerable to both traffic crashes and the virus. The evolution of traffic crashes involving older pedestrians has not been thoroughly explored in the existing research. This study adopts a holistic approach by analyzing the effects of the COVID-19 pandemic on older pedestrian safety using a two-stage methodology that incorporates both national-level data from Spain and city-level data from Madrid. Findings indicate that the pandemic has had a more pronounced impact on the traffic safety of older pedestrians than on the rest of the population, especially in terms of the number of recorded crashes.

Firstly, the monthly number of older and non-older pedestrians injured in vehicle collisions, considering three distinct pandemic periods, were statistically modeled. The results showed that the total number of injured pedestrians, including fatalities, serious injuries, and slight injuries, decreased for both groups during the lockdown period. This aligns with previous research on all (Alhajyaseen et al., 2022; Saladié et al., 2020; Yasin et al., 2021), pedestrian (Chand et al., 2021; Krizsik and Pauer, 2023), and older adults (Soltani et al., 2023) traffic crashes. Unlike previous studies suggesting a return to pre-pandemic traffic crash levels after lockdown (Doucette et al., 2021; Redelmeier and Zipursky, 2021; Saladié et al., 2023), our study found that injuries among older pedestrians and non-fatal injuries among non-older pedestrians remained below pre-pandemic levels even during the 'new normality.'

After the vaccination campaign, the total and slight injuries among older and non-older pedestrians remained lower compared to the pre-pandemic situation, although these reductions were roughly double for older pedestrians. Soltani et al. (2023) noticed a similar result regarding a decrease in older adult traffic crashes using April–December 2020 data, but our study covers a wider period. This could be due to older adults walking less compared to non-older pedestrians, given their increased vulnerability to COVID-19 disease (Nanda et al., 2020; Shahid et al., 2020). Nevertheless, vaccination seemed to reduce the perceived risk among all pedestrians, prompting more mobility and walking activity (Bei et al., 2023), although this walking activity is inferred lower for older adults, as older adult crashes are more reduced than the rest of pedestrian crashes. Another factor could be underreporting, a common issue in road safety studies, especially for less severe crashes (Abay, 2015; Watson et al., 2015). Some researchers have suggested that underreporting of less severe crashes increased during the pandemic due to the health system's overload (Sobreira et al., 2022). If the underreporting rates varied during the pandemic, it could significantly impact on the reliability of our findings. Therefore, fatalities and serious injuries may provide more credible data for assessing the evolution of pedestrian crashes if underreporting differed before and after the pandemic, so pedestrian crashes for both age groups would have returned to normal values after vaccination.

Secondly, the injury severity rates among older pedestrians did not show statistically significant changes across the three pandemic phases. However, a slight shift was observed during the lockdown, where the proportion of fatalities decreased, and serious injuries increased, mirroring the results found in other studies (Krizsik and Pauer, 2023). There was also a reduction of younger pedestrians in pedestrian crashes, as noticed previously by Krizsik and Pauer (2023). Notably, the proportion of vehicle-pedestrian collisions relative to total urban traffic crashes decreased even after two years of the pandemic declaration (2022), something already noted by Krizsik and Pauer (2023) for data until 2021. This fact may be linked to a reduction in walking activity, especially during the lockdown (Angel et al., 2023; Choe et al., 2022; Desine et al., 2023; Shaer and Haghshenas, 2021).

Lastly, the profile and location of older pedestrian crashes in Madrid showed slight changes before and after the pandemic. Post-pandemic crashes were more likely to occur in areas with higher population density, suggesting that walking activities may have become more localized. However, this variable was not significant for older pedestrians, indicating that their walking habits may not have shifted in relation to their residence. The land-use mix presented contrasting effects: crashes involving older pedestrians shifted to areas with a higher land-use mix index (indicative of less mixed land uses), while those involving non-older pedestrians occurred more frequently in mixed land-use locations.

Regarding older pedestrian crashes, there was also a significant change in the time of day when crashes occurred. Post-pandemic, crashes during evening and night hours decreased, suggesting that older pedestrians may have adjusted their walking habits, favoring daylight hours. This change, previously linked to nightlife restrictions (Soltani et al., 2023), was observed in our study even when nightlife restrictions were lifted (2021–2022). This suggests that the pandemic may have permanently influenced older adults' walking behavior.

Additionally, the presence of more bike lanes is associated with post-pandemic crashes involving older pedestrians. This could be because these environments are more complex for drivers, although this factor does not appear in the model for non-older pedestrians. It is important to note that bike lane infrastructure is considered as it exists today, so these lanes might not have been present before the pandemic. This association suggests two possibilities: (1) older pedestrians might find these environments more challenging, requiring vigilance for both bicycles and motorized vehicles, and (2) the use of bicycles increased during the pandemic, as documented in the literature (Shaer and Haghshenas, 2021; Yao et al., 2021).

Regarding non-older pedestrian crashes, these occurred more often in areas with fewer bus stops post-pandemic, suggesting a shift

toward peripheral city areas where public transport is less available, as noted by Chand et al. (2021). This shift could also be due to reduced public transport use during the pandemic (Fernández Pozo et al., 2022). Lighting conditions have changed, with more crashes occurring under artificial lighting and fewer in unlit areas. While this might relate to the time of day, night crashes remained consistent before and after the pandemic. This suggests non-older pedestrians may not have changed their walking times but now prefer to walk on safer, better-lit streets. Additionally, the proportion of vehicle collisions involving non-older pedestrians on weekends is lower post-pandemic, although not statistically significant, and the proportion of injured female pedestrians has also decreased.

Based on the modeling results, specific recommendations emerge regarding bicycle infrastructure, micromobility usage, and sidewalk elements. Post-pandemic, older pedestrian crashes near bicycle infrastructure have become more likely. Micromobility vehicles, with their small size and high speeds, present heightened risks to this demographic due to their lower detectability. Despite regulations, these vehicles often use sidewalks, increasing interactions with pedestrians. Stricter enforcement of regulations prohibiting sidewalk use by such vehicles, coupled with improved infra-structure, such as enhanced visibility and clearly marked interaction points between pedestrian and micromobility areas, can facilitate safer coexistence. These measures would enable both pedestrians and micromobility users to detect each other more easily in shared spaces. Furthermore, older pedestrian collisions are more likely to occur in areas with less side-walk density (i.e., sidewalk surface by street length) after the pandemic. Addressing this issue by removing obstacles on narrow sidewalks can prevent older pedestrians from walking on carriageways, especially when sidewalks are crowded or too narrow.

About the limitations of this research, we could not analyze the impact of subsequent pandemic milestones, such as the end of mask mandates in public transport, as traffic crash data was only available until December 2022. The decline in traffic crashes involving older pedestrians after the pandemic might be linked to a reduction in the “more active” older pedestrian population, who were at greater risk of contracting COVID-19 and experiencing severe outcomes, potentially including mortality, compared to less active older individuals. However, this information is unavailable in the traffic crash database. Additionally, the analysis of crash-related factors was limited to static features due to data availability, even though these factors may have evolved over time. Further research is necessary to assess whether the mobility patterns of older pedestrians have changed due to the pandemic even after WHO has declared the end of COVID-19 as a global health emergency (May 5, 2023). Surveys focusing on pre- and post-pandemic mobility habits could provide deeper insights. Additionally, as more traffic crash data becomes available, future studies should examine whether the number, severity, and characteristics of crashes involving older pedestrians have returned to pre-pandemic levels following the end of all pandemic restrictions.

In summary, crashes involving older pedestrians remained below pre-pandemic levels even after the public vaccination campaign. In contrast, non-older pedestrian road safety returned to normal levels across all injury severities. If traffic safety reflects mobility levels, this may suggest that older adults are walking less, shifting to other transport modes, or acting more cautiously due to the pandemic. In any case, older adults are clearly more vulnerable compared to the rest of the population to changes in mobility and social anxiety caused by disruptive events like the pandemic. Additionally, older pedestrians may have permanently altered their mobility patterns, such as walking times, transport modes, or distances traveled. Nevertheless, the profile of older pedestrian crashes has not undergone significant changes, suggesting that current measures to mitigate their road safety risks remain effective. The reduction in crashes involving older pedestrians in 2022 is a positive outcome, but this finding should be reassessed in subsequent years and further analyzed in relation to changes in the mobility patterns of older adults. Future research should focus on how pandemic-induced behavioral changes may persist, as older adults remain a vulnerable group in road safety and the design of age-friendly cities should be prioritized in the sustainable urban design policies.

CRedit authorship contribution statement

Daniel Gálvez-Pérez: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Begoña Guirao:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Investigation, Formal analysis, Conceptualization. **David del-Villar-Juez:** Writing – original draft, Validation, Software, Investigation, Formal analysis.

Declaration of competing interest

None.

Data availability

Data will be made available on request.

References

- Abay, K.A., 2015. Investigating the nature and impact of reporting bias in road crash data. *Transport. Res. Pol. Pract.* 71, 31–45. <https://doi.org/10.1016/j.tpra.2014.11.002>.
- Alhajyaseen, W.K.M., Almkudat, A., Hussain, Q., Almallah, M., Al Malki, M.A., Singaravelu, J., Zammataro, S., 2022. Road safety status during COVID-19 pandemic: exploring public and road safety expert's opinions. *Int. J. Inj. Control Saf. Promot.* 29 (2), 135–151. <https://doi.org/10.1080/17457300.2021.1962915>.

- Aloi, A., Alonso, B., Benavente, J., Cordera, R., Echániz, E., González, F., Ladisa, C., Lezama-Romanelli, R., López-Parra, A., Mazzei, V., Perrucci, L., Prieto-Quintana, D., Rodríguez, A., Sañudo, R., 2020. Effects of the COVID-19 lockdown on urban mobility: empirical evidence from the city of Santander (Spain). *Sustainability* 12 (9). <https://doi.org/10.3390/su12093870>.
- Angel, A., Cohen, A., Dalyot, S., Plaut, P., 2023. Impact of COVID-19 policies on pedestrian traffic and walking patterns. *Environ. Plan. B Urban Anal. City Sci.* 50 (5), 1178–1193. <https://doi.org/10.1177/23998083221113332>.
- Atzendorf, J., Gruber, S., 2022. Depression and loneliness of older adults in Europe and Israel after the first wave of covid-19. *Eur. J. Ageing* 19 (4), 849–861. <https://doi.org/10.1007/s10433-021-00640-8>.
- Barnes, S.R., Beland, L.-P., Huh, J., Kim, D., 2020. The effect of COVID-19 lockdown on mobility and traffic accidents: evidence from Louisiana standard-nutzungsbedingungen: the effect of COVID-19 lockdown on mobility and traffic accidents: evidence from Louisiana. <https://hdl.handle.net/10419/222470>.
- Bei, H., Li, P., Cai, Z., Murcio, R., 2023. The impact of COVID-19 vaccination on human mobility: the London case. *Heliyon* 9 (8), e18769. <https://doi.org/10.1016/j.heliyon.2023.e18769>.
- Bergman, Y.S., Cohen-Fridel, S., Shrira, A., Bodner, E., Palgi, Y., 2020. COVID-19 health worries and anxiety symptoms among older adults: the moderating role of ageism. *Int. Psychogeriatr.* 32 (11), 1371–1375. <https://doi.org/10.1017/S1041610220001258>.
- Brodeur, A., Cook, N., Wright, T., 2021. On the effects of COVID-19 safer-at-home policies on social distancing, car crashes and pollution. *J. Environ. Econ. Manag.* 106. <https://doi.org/10.1016/j.jeem.2021.102427>.
- Brooke, J., Jackson, D., 2020. Older people and COVID-19: isolation, risk and ageism. *J. Clin. Nurs.* 29 (Issues 13–14), 2044–2046. <https://doi.org/10.1111/jocn.15274>.
- Chand, S., Yee, E., Alsultan, A., Dixit, V.V., 2021. A descriptive analysis on the impact of covid-19 lockdowns on road traffic incidents in sydney, Australia. *Int. J. Environ. Res. Publ. Health* 18 (21). <https://doi.org/10.3390/ijerph18211701>.
- Choe, E.Y., Du, Y., Sun, G., 2022. Decline in older adults' daily mobility during the COVID-19 pandemic: the role of individual and built environment factors. *BMC Publ. Health* 22 (1). <https://doi.org/10.1186/s12889-022-14780-8>.
- Clifton, K.J., Burnier, C.V., Akar, G., 2009. Severity of injury resulting from pedestrian-vehicle crashes: what can we learn from examining the built environment? *Transport. Res. Transport Environ.* 14 (6), 425–436.
- Desine, S., Master, H., Annis, J., Hughes, A., Roden, D.M., Harris, P.A., Brittain, E.L., 2023. Daily step counts before and after the COVID-19 pandemic among all of us research participants. *JAMA Netw. Open* 6 (3), e233526. <https://doi.org/10.1001/jamanetworkopen.2023.3526>.
- Dirección General de Tráfico, 2020. In: *Tendencias de la movilidad y siniestralidad en vías urbanas*.
- Dong, X., Xie, K., Yang, H., 2022. How did COVID-19 impact driving behaviors and crash Severity? A multigroup structural equation modeling. *Accid. Anal. Prev.* 172. <https://doi.org/10.1016/j.aap.2022.106687>.
- Doucette, M.L., Tucker, A., Auguste, M.E., Gates, J.D., Shapiro, D., Ehsani, J.P., Borrup, K.T., 2021a. Evaluation of motor vehicle crash rates during and after the COVID-19-associated stay-at-home order in Connecticut. *Accid. Anal. Prev.* 162. <https://doi.org/10.1016/j.aap.2021.106399>.
- Doucette, M.L., Tucker, A., Auguste, M.E., Watkins, A., Green, C., Pereira, F.E., Borrup, K.T., Shapiro, D., Lapidus, G., 2021b. Initial impact of COVID-19's stay-at-home order on motor vehicle traffic and crash patterns in Connecticut: an interrupted time series analysis. *Inj. Prev.* 27 (1), 3–9. <https://doi.org/10.1136/injuryprev-2020-043945>.
- Drummond, J., Hasnine, M.S., 2023. Did the COVID-19 vaccine rollout impact transportation demand? A case study in New York City. *J. Transport Health* 28. <https://doi.org/10.1016/j.jth.2022.101539>.
- Esain, I., Gil, S.M., Dunabeitia, I., Rodríguez-Larrad, A., Bidaurrezaga-Letona, I., 2021. Effects of covid-19 lockdown on physical activity and health-related quality of life in older adults who regularly exercise. *Sustainability* 13 (7). <https://doi.org/10.3390/su13073771>.
- ETSC, 2020. PIN briefing. The Impact of COVID-19 Lockdowns on Road Deaths in April 2020.
- European Commission, 2015. *ElderSafe-Risks and Countermeasures for Road Traffic of Elderly in Europe (No. MOVE/C4/2014-244)*. European Commission-Directorate-General for Mobility and Transport (DG-MOVE).
- European Commission, 2021. *Road safety thematic report – Seniors*. European Road Safety Observatory.
- Fernández Pozo, R., Wilby, M.R., Vinagre Díaz, J.J., Rodríguez González, A.B., 2022. Data-driven analysis of the impact of COVID-19 on Madrid's public transport during each phase of the pandemic. *Cities* 127. <https://doi.org/10.1016/j.cities.2022.103723>.
- Gabriel, Z., Bowling, A.N.N., 2004. Quality of life from the perspectives of older people. *Ageing Soc.* 24 (5), 675–691.
- Gálvez-Pérez, D., Guirao, B., Ortuño, A., 2024. Analysis of the elderly pedestrian traffic accidents in urban scenarios: the case of the Spanish municipalities. *Int. J. Inj. Control Saf. Promot.* 1–20.
- Gálvez-Pérez, D., Guirao, B., Ortuño, A., Picado-Santos, L., 2022. The influence of built environment factors on elderly pedestrian road safety in cities: the experience of Madrid. *Int. J. Environ. Res. Publ. Health* 19 (4). <https://doi.org/10.3390/ijerph19042280>.
- Gobierno de España, 2020. Real Decreto 463/2020, de 14 de marzo, por el que se declara el estado de alarma para la gestión de la situación de crisis sanitaria ocasionada por el COVID-19. *Boletín Oficial del Estado (BOE)* 67, 25390–25400.
- Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., Tatlow, H., 2021. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nat. Human Behav.* 5 (4), 529–538. <https://doi.org/10.1038/s41562-021-01079-8>.
- Hanson, S., Jones, A., 2015. Is there evidence that walking groups have health benefits? A systematic review and meta-analysis. *Br. J. Sports Med.* 49 (11), 710–715. <https://doi.org/10.1136/bjsports-2014-094157>.
- Hilbe, J.M., 2011. *Negative Binomial Regression*. Cambridge University Press.
- Hong, J., McArthur, D.P., Lee, J., 2023. Neighbourhood deprivation, changes in walking attitudes during the COVID-19 pandemic and walking demand in the post-pandemic period. *Cities* 141. <https://doi.org/10.1016/j.cities.2023.104517>.
- Iversen, T.N., Larsen, L., Solem, P.E., 2012. A conceptual analysis of ageism. *Nord. Psychol.* 61 (3), 4–22. <https://doi.org/10.1027/1901-2276.61.3.4>.
- Jimenez-Sotomayor, M.R., Gomez-Moreno, C., Soto-Perez-de-Celis, E., 2020. Coronavirus, ageism, and twitter: an evaluation of tweets about older adults and COVID-19. *J. Am. Geriatr. Soc.* 68 (8), 1661–1665. <https://doi.org/10.1111/jgs.16508>.
- Kim, D., 2019. The transportation safety of elderly pedestrians: modeling contributing factors to elderly pedestrian collisions. *Accid. Anal. Prev.* 131, 268–274. <https://doi.org/10.1016/j.aap.2019.07.009>.
- Kim, H., Kim, J., Ha, J., 2023. Factors influencing mobility in community-dwelling older adults during the early COVID-19 pandemic: a cross-sectional study. *BMC Publ. Health* 23 (1). <https://doi.org/10.1186/s12889-023-16553-3>.
- Krzišnik, N., Pauer, G., 2023. Impact of Covid-19 on pedestrian safety. *Transport. Res. F Traffic Psychol. Behav.* 98, 216–225. <https://doi.org/10.1016/j.trf.2023.09.009>.
- Lebrasseur, A., Fortin-Bédard, N., Lettre, J., Raymond, E., Bussièrès, E.L., Lapièrre, N., Faieta, J., Vincent, C., Duchesne, L., Ouellet, M.C., Gagnon, E., Tourigny, A., Lamontagne, M.É., Routhier, F., 2021. Impact of the COVID-19 pandemic on older adults: rapid review. *JMIR Publications Inc* 4 (Issue 2). <https://doi.org/10.2196/26474>.
- Lee, S., Yoon, J., Woo, A., 2020. Does elderly safety matter? Associations between built environments and pedestrian crashes in Seoul, Korea. *Accid. Anal. Prev.* 144, 105621.
- Liu, P., Song, Y., Zou, Q., Tang, J., Fernandez, J., 2024. Environmental press of urban greenspace pre-post COVID on older adults: a big data study in metropolitan Atlanta. *Cities* 145. <https://doi.org/10.1016/j.cities.2023.104733>.
- Liu, Q., Liu, Y., Zhang, C., An, Z., Zhao, P., 2021. Elderly mobility during the COVID-19 pandemic: a qualitative exploration in Kunming, China. *J. Transport Geogr.* 96. <https://doi.org/10.1016/j.jtrangeo.2021.103176>.
- Lv, M., Wang, N., Yao, S., Wu, J., Fang, L., 2021. Towards healthy aging: influence of the built environment on elderly pedestrian safety at the micro-level. *Int. J. Environ. Res. Publ. Health* 18 (18), 9534.
- Musselwhite, C., Holland, C., Walker, I., 2015. The role of transport and mobility in the health of older people. *J. Transport Health* 2 (1), 1–4.

- Nanda, A., Vura, N.V.R.K., Gravenstein, S., 2020. COVID-19 in older adults. *Aging Clin. Exp. Res.* 32 (7), 1199–1202. <https://doi.org/10.1007/s40520-020-01581-5>.
- Nordbakke, S., Schwane, T., 2014. Well-being and mobility: a theoretical framework and literature review focusing on older people. *Mobilities* 9 (1), 104–129.
- Nowossadeck, S., Wettstein, M., Cengia, A., 2023. Physical activity during the covid-19 pandemic. Changes in the frequency of sports and walking among people in the second half of life. In: *In: Ageing in Times of the COVID-19 Pandemic*. Springer, Fachmedien Wiesbaden, pp. 119–141. https://doi.org/10.1007/978-3-658-40487-1_7.
- Park, S.-H., Bae, M.-K., 2020. Exploring the determinants of the severity of pedestrian injuries by pedestrian age: a case study of Daegu Metropolitan City, South Korea. *Int. J. Environ. Res. Publ. Health* 17 (7), 2358.
- Pérez Díaz, J., Abellán García, A., Aceituno Nieto, P., Ramiro Fariñas, D., 2020. Un perfil de las personas mayores en España, 2020. Indicadores estadísticos básicos.
- Perracini, M.R., de Amorim, J.S.C., Lima, C.A., da Silva, A., Trombini-Souza, F., Pereira, D.S., Pelicioni, P.H.S., Duim, E., Batista, P.P., dos Santos, R.B., de Lima, M. do C.C., Carlos, A.G., Santos, A.D., da Silva Fittipaldi, E.O., Campos, H., Gazzola, J.M., Moreira, M., Ferreira, M.B., de Oliveira Cirino, N.T., et al., 2021. Impact of COVID-19 pandemic on life-space mobility of older adults living in Brazil: REMOBILIZE study. *Front. Public Health* 9. <https://doi.org/10.3389/fpubh.2021.643640>.
- Qureshi, A.I., Huang, W., Khan, S., Lobanova, I., Siddiq, F., Gomez, C.R., Suri, M.F.K., 2020. Mandated societal lockdown and road traffic accidents. *Accid. Anal. Prev.* 146. <https://doi.org/10.1016/j.aap.2020.105747>.
- Rantanen, T., Eronen, J., Kauppinen, M., Kokko, K., Sanaslahti, S., Kajan, N., Portegijs, E., 2021. Life-space mobility and active aging as factors underlying quality of life among older people before and during COVID-19 lockdown in Finland - a longitudinal study. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences* 76 (3), E60–E67. <https://doi.org/10.1093/geron/glaa274>.
- Rapoport, M.J., Chee, J.N., Aljenabi, N., Byrne, P.A., Naglie, G., Ilari, F., Elzohairy, Y., Vingilis, E., Mulsant, B.H., 2021. Impact of COVID-19 on motor vehicle injuries and fatalities in older adults in Ontario, Canada. *Accid. Anal. Prev.* 157. <https://doi.org/10.1016/j.aap.2021.106195>.
- Redelmeier, D.A., Zipursky, J.S., 2021. Pedestrian deaths during the COVID-19 pandemic. *Am. J. Lifestyle Med.* 17 (2), 276–279. <https://doi.org/10.1177/15598276211058378>.
- Saladié, Ò., Bustamante, E., Gutiérrez, A., 2020. COVID-19 lockdown and reduction of traffic accidents in Tarragona province, Spain. *Transp. Res. Interdiscip. Perspect.* 8. <https://doi.org/10.1016/j.trip.2020.100218>.
- Saladié, Ò., Domínguez-Mallafre, M., Domènech, A., Gutiérrez, A., 2023. Back to the past: the decline and rebound in traffic accident numbers in Catalonia in the context of COVID-19 lockdown and post-lockdown. *Nor. Geografisk Tidsskr.* 77 (2), 78–97. <https://doi.org/10.1080/00291951.2023.2217818>.
- Saraiva, M.D., Apolinario, D., Avelino-Silva, T.J., De Assis Moura Tavares, C., Gattás-Vernaglia, I.F., Marques Fernandes, C., Rabelo, L.M., Tavares Fernandes Yamaguti, S., Karnakis, T., Kalil-Filho, R., Jacob-Filho, W., Romero Aliberti, M.J., 2021. The impact of frailty on the relationship between life-space mobility and quality of life in older adults during the COVID-19 pandemic. *J. Nutr. Health Aging* 25 (4), 440–447. <https://doi.org/10.1007/s12603-020-1532-z>.
- Sedain, B., Pant, P.R., 2020. Road traffic injuries in Nepal during COVID-19 lockdown. *F1000Research* 9, 1209. <https://doi.org/10.12688/f1000research.26281.1>.
- Sekadakakis, M., Katrakazas, C., Michelaraki, E., Kehagia, F., Yannis, G., 2021. Analysis of the impact of COVID-19 on collisions, fatalities and injuries using time series forecasting: the case of Greece. *Accid. Anal. Prev.* 162. <https://doi.org/10.1016/j.aap.2021.106391>.
- Shaer, A., Haghshenas, H., 2021. The impacts of COVID-19 on older adults' active transportation mode usage in Isfahan, Iran. *J. Transport Health* 23. <https://doi.org/10.1016/j.jth.2021.101244>.
- Shahid, Z., Kalayanamitra, R., McClafferty, B., Kepko, D., Ramgobin, D., Patel, R., Aggarwal, C.S., Vunnam, R., Sahu, N., Bhatt, D., Jones, K., Golamari, R., Jain, R., 2020. COVID-19 and older adults: what we know. *J. Am. Geriatr. Soc.* 68 (5), 926–929. <https://doi.org/10.1111/jgs.16472>.
- Shaik, M.E., Ahmed, S., 2022. An overview of the impact of COVID-19 on road traffic safety and travel behavior. *Transport Eng.* 9, 100119.
- Shaik, M.E., Hossain, Q.S., Rony, G.M.F.F., 2021. Impact of COVID-19 on public transportation and road safety in Bangladesh. *SN Computer Science* 2 (6). <https://doi.org/10.1007/s42979-021-00849-5>.
- Sobreira, L.T.P., Luna, M., dos, S. de, Cunto, F., Hellenga, B., 2022. Impact of COVID-19 on traffic volume, violations, and crashes in fortaleza, Brazil. *J. Transport. Eng., Part A: Systems* 148 (10). <https://doi.org/10.1061/JTEPBS.0000738>.
- Soltani, A., Azmoodeh, M., Qadikolaei, M.R., 2023. Post COVID-19 transformation in the frequency and location of traffic crashes involving older adults. *Transport. Res. Rec.: J. Transport. Res. Board.* <https://doi.org/10.1177/03611981231163866>.
- United Nations, 2020. *World Population Ageing 2019*. Department of Economic and Social Affairs Population Division, New York, NY, USA.
- Watson, A., Watson, B., Vallmuur, K., 2015. Estimating under-reporting of road crash injuries to police using multiple linked data collections. *Accid. Anal. Prev.* 83, 18–25. <https://doi.org/10.1016/j.aap.2015.06.011>.
- World Health Organization, 2022. Cycling and walking can help reduce physical inactivity and air pollution, save lives and mitigate climate change. <https://www.who.int/europe/news/item/07-06-2022-cycling-and-walking-can-help-reduce-physical-inactivity-and-air-pollution-save-lives-and-mitigate-climate-change>.
- Yao, Y., Geara, T.G., Shi, W., 2021. Impact of COVID-19 on city-scale transportation and safety: an early experience from Detroit. *Smart Health* 22. <https://doi.org/10.1016/j.smhl.2021.100218>.
- Yasin, Y.J., Grivna, M., Abu-Zidan, F.M., 2021. Global impact of COVID-19 pandemic on road traffic collisions. *World J. Emerg. Surg.* 16 (Issue 1). <https://doi.org/10.1186/s13017-021-00395-8>.
- Zhang, J., Feng, B., Wu, Y., Xu, P., Ke, R., Dong, N., 2021. The effect of human mobility and control measures on traffic safety during COVID-19 pandemic. *PLoS One* 16 (3 March 2021). <https://doi.org/10.1371/journal.pone.0243263>.
- Zhang, M., Wang, S., Hu, T., Fu, X., Wang, X., Hu, Y., Halloran, B., Li, Z., Cui, Y., Liu, H., Liu, Z., Bao, S., 2022. Human mobility and COVID-19 transmission: a systematic review and future directions. *Spatial Sci.* 28 (4), 501–514. <https://doi.org/10.1080/19475683.2022.2041725>.

4.6 *Additional publications*

During the development of this doctoral thesis, other works have been presented in **international conferences**.

Title	Road safety of elderly pedestrians in the urban context: an approach based on infrastructure and socioeconomic variables
Authors	Daniel Gálvez-Pérez, Begoña Guirao, Armando Ortuño
Conference	XIV Conference on Transport Engineering (XIV Congreso de Ingeniería del Transporte, CIT 2021)
Dates	6-8 July 2021
Journal	Transportation Research Procedia
Issue, page(s)	58, 254-261
Publication date	8 December 2021
DOI	https://doi.org/10.1016/j.trpro.2021.11.035

Title	The effect of built environment features on vehicle-pedestrian collisions in big cities: a comparison between Lisbon and Madrid
Authors	Daniel Gálvez-Pérez, Begoña Guirao, Luis Picado-Santos, Natalia Casado-Sanz
Conference	2022 Transport Research Arena Conference (TRA)
Dates	14-17 November 2022
Journal	Transportation Research Procedia
Issue, page(s)	72, 588-595
Publication date	13 December 2023
DOI	https://doi.org/10.1016/j.trpro.2023.11.443

Title	Analysis of the Elderly Pedestrian Injury Severity in Urban Traffic Accidents in Spain using Machine Learning Techniques
Authors	Daniel Gálvez-Pérez, Begoña Guirao, Armando Ortuño
Conference	XV Conference on Transport Engineering (XV Congreso de Ingeniería del Transporte, CIT 2023)
Dates	14-16 July 2023
Journal	Transportation Research Procedia
Issue, page(s)	71, 6-13
Publication date	22 November 2023
DOI	https://doi.org/10.1016/j.trpro.2023.11.051

At the time of writing this doctoral thesis (January 2025), one **article is under review** in an international journal:

Title	Enhancing Safe Walking in an Ageing Society: Insights into Injury Severity of Older Pedestrian Traffic Crashes in Urban Environments
Authors	Daniel Gálvez-Pérez, Begoña Guirao, Armando Ortuño
Journal	Journal of Transport & Health (JTH); ISSN: 2214-1413

5

Results

In this fifth chapter of the dissertation, the main Results of the developed works are presented. This section includes the **Results of Article A** (Section 5.1), **Results of Article B** (Section 5.2), **Results of Article C** (Section 5.3), and **Results of Article D** (Section 5.4).

5.1 Results of Article A

The first analysis was developed in Article A: The Influence of Built Environment Factors on Elderly Pedestrian Road Safety in Cities: The Experience of Madrid (Gálvez-Pérez et al., 2022). It consisted of a frequency study of older and non-older pedestrian traffic crashes per Madrid city district. Independent variables included exposure, socioeconomics, land use, and infrastructure characteristics.

The variables considering **exposure to traffic crash risk** included the total street length and the median annual average daily traffic (AADT) measured in each city district. The total street length had a detrimental impact on road safety for all pedestrians, generating a higher number of older and non-older pedestrian traffic crashes. On the other hand, the AADT only showed a significant effect for older pedestrian traffic crashes, making city districts with higher median AADT hazardous for older pedestrians.

When examining **socioeconomic variables**, both models reveal that population density has a positive and significant influence on the number of pedestrian collisions. This effect is more pronounced in the model for non-older pedestrian crashes. Population density serves as a proxy for how congested a district is, considering factors such as the average height and density of buildings. Annual average income, indicative of the economic status of a district, shows an inverse relationship with collision frequency. As the average income level rises, the likelihood of pedestrian collisions decreases. This relationship is consistently observed in both models, and the variable is statistically significant in each, suggesting that more affluent areas might have better

infrastructure, safer pedestrian pathways, or reduced pedestrian activity linked to higher car usage. The number of older adult inhabitants per unit length of street has a positive and significant effect on the number of pedestrian collisions involving older adults. Conversely, in the model for non-older pedestrian casualties, the density of non-older residents per kilometer of street does not reach statistical significance. This pattern aligns with the hypothesis that older adults are more inclined to walk within their immediate neighborhoods or districts, whereas younger individuals are more likely to commute or move towards busier, central areas, where pedestrian traffic may be more dispersed or associated with different risk factors.

In relation to **land use variables**, non-older pedestrian crashes model shows a substantial effect of Points of Interest (POIs) per street length, residential land use proportion, and the proportion of green areas on the frequency of non-older pedestrian collisions. In this model, a higher residential proportion corresponds to a reduced risk for non-older pedestrian incidents, potentially indicating safer environments due to less vehicular traffic or the presence of familiar, localized pedestrian routes.

A higher density of POIs per street length is associated with lower risk areas for non-older pedestrian collisions. Although this result was not initially anticipated, it aligns with the hypothesis that denser concentrations of POIs may lead to more crowded streets where pedestrian movement is slower, and individuals remain more vigilant to their surroundings. This heightened awareness and reduced speed may contribute to fewer pedestrian-related incidents.

Conversely, in the older pedestrian crashes model, POIs per street length, residential proportion, and green area proportion do not demonstrate statistical significance. This suggests that land use factors influencing pedestrian collisions differ by age group, with older pedestrians potentially influenced more by other environmental or behavioral factors.

Finally, the proportion of main streets serves as a key indicator of the significance of roadways within a district. Main streets typically feature greater width, higher traffic volumes, and vehicles traveling at faster speeds compared to smaller roads. Both models confirm that a higher proportion of main streets correlates with an increased level of traffic crash risk exposure to pedestrians, irrespective of age. This underscores the need for targeted safety measures on these critical thoroughfares to mitigate collision risk across different demographic groups.

About **infrastructure variables**, both models reveal a high level of significance for junction density and bus stop density. A higher number of street junctions per kilometer correlates with an increased number of pedestrian collisions in a district for both age

groups. Junction density reflects the complexity of the street network and the corresponding potential hazards that pedestrians and drivers may encounter for a given street length. This complexity adds to the risk profile of an area by increasing the number of conflict points where collisions are more likely to occur.

Regarding public transport infrastructure, bus stop density is a noteworthy risk factor for both older and non-older pedestrians. The positive relationship between bus stop density and traffic crash frequency suggests that districts with more bus stops may experience higher pedestrian flow, increased interactions with large vehicles, and potential confusion or risk for older pedestrians alighting from buses. Metro station density, an indicator of zones with high pedestrian trip generation and attraction, also shows a positive effect on collision risk but is only statistically significant for non-older pedestrians. This result supports the idea that younger pedestrians tend to have higher mobility, often traveling beyond their residential districts and moving towards central or busier areas.

On the other hand, sidewalk density, which reflects the sidewalk mean width of an area, exhibits a negative effect on the number of traffic crashes and reaches statistical significance only in the model for older pedestrians. Sidewalk density serves as an indicator of the average sidewalk width available per unit length of street, typically narrower in older, central districts and wider in more recently developed areas. The importance of this variable is underscored by the increased vulnerability of older pedestrians, who face greater risks of tripping, stumbling, or falling while walking. This age group often struggles more with maintaining postural stability and balance, making narrow sidewalks a significant hazard that may compel them to walk outside designated pedestrian areas, thereby increasing their exposure to vehicular traffic. Narrow or inadequate sidewalks can exacerbate anxiety and negatively impact how older pedestrians perceive their built environment.

Traffic light density emerges as a significant safety feature specifically for older pedestrians, reducing the number of crashes involving this demographic. The presence of traffic lights contributes to lower vehicle speeds, as drivers are required to stop at red lights, enhancing pedestrian safety at intersections. However, this variable is not statistically significant for younger pedestrians, potentially indicating a greater degree of cautiousness, self-regulation or different crossing behaviors among older pedestrians.

In conclusion, the comparison of the two models underscores several shared and distinct risk factors associated with pedestrian collisions across different age groups. Districts with greater total street length, higher population density, a higher proportion of main streets, more street junctions per kilometer, and a higher density of bus stops,

alongside lower average income levels, present an elevated risk for pedestrian collisions overall, regardless of age.

For older pedestrian crashes, certain factors play a unique role. Specifically, higher traffic flow (measured by AADT) and a greater number of older residents per kilometer significantly increase the likelihood of collisions, factors that are not statistically significant for non-older pedestrian incidents. Moreover, for older pedestrians, traffic crash occurrence is mitigated by higher sidewalk density and a greater number of traffic lights per kilometer, both contribute positively to safety by offering safer walking spaces and proportionating proper crossing areas.

In contrast, non-older pedestrian collisions are more likely to occur in districts with a higher density of traffic signals and metro stations per kilometer, reflecting the tendency of younger pedestrians to travel through high-mobility and central areas where these features are prevalent. However, these collisions are less influenced by AADT levels, suggesting that traffic volume may not be a critical risk factor for this group, possibly due to different mobility patterns or crossing behaviors.

5.2 *Results of Article B*

In a subsequent study, Article B: Analysis of the elderly pedestrian traffic accidents in urban scenarios: the case of the Spanish municipalities (Gálvez-Pérez et al., 2024b), the frequency of older and non-older pedestrian traffic crashes per municipality at the macro-level was studied and statistically modeled. The dependent variable was the number of older and non-older pedestrian traffic crashes between 2016 and 2019 inside of the urbanized areas of each municipality. The independent variables were grouped, as in the previous study, into exposure, socioeconomics, land use, and infrastructure variables.

Prior to statistical modeling, a ranking of the provincial capitals of Spain considering the number of older pedestrian traffic crashes by 10,000 older adult population was carried out. Given that Spain is administratively divided into 50 provinces, provincial capitals are typically sizable cities with a sophisticated and cohesive urban plan that houses the province's public administrative offices.

The provincial capital cities' populations range from 34,670 (Teruel) to 3,186,620 (Madrid) inhabitants. Because of the heterogeneity in their population size, these cities were grouped as follows: cities with over 500,000 inhabitants, cities between 500,000 and 250,000 inhabitants, cities between 250,000 and 100,000, towns between 50,000 and 100,000, and towns under 50,000. During the research period, the number of older

pedestrian traffic crashes per 10,000 older adult population per year varies from 0.46 (Ciudad Real) to 13.24 (León).

Apart from Palma (5th) and Barcelona (13th), two towns with high tourist traffic, the cities at the top of the ranking (15 most hazardous cities for older pedestrians) are concentrated in the groups between 50,000 and 250,000 residents.

Despite being Spain's most populous cities, Madrid and Barcelona rank 25th and 13th, respectively. Moreover, except for Barcelona, the largest cities in Spain (i.e., those with a population of over half a million) are among the 50% of provincial capitals that are the safest for older pedestrians. These statistics suggest that larger capital cities are safer for older pedestrians, with a few exceptions where tourism may be crucial (raising normal road traffic flows and the running of the city).

Middle-sized provincial capitals may present greater risks compared to small and large capitals, as small capitals typically experience lower traffic volumes, while large capitals benefit from more developed road infrastructure and denser public transportation systems. Hence, capitals from 50,000 to 250,000 inhabitants may present an unfair balance between the volume of vehicles and the road infrastructure/public transport.

In this article, the interpretation of the statistical modeling of the frequency of older pedestrian traffic crashes is highly conditioned by the structure of the model used, hurdle negative binomial regression, which comprises two components: the zero component (ZC) and the count component (CC). The zero component estimates the likelihood of at least one traffic crash occurring, while the count component predicts the frequency of traffic crashes in municipalities where at least one traffic crash has been recorded during the study period. In addition, to aid interpretation, performance scores (count, zero, and overall) for each dependent variable were computed. These scores provided a basis for understanding the distinct effects of predictors within each model component. Two statistical models were built to predict the frequency of vehicle-pedestrian collisions involving older pedestrians (model OP) and non-older pedestrians (model NOP).

The exposure variable, representing the number of inhabitants, emerged as a robust predictor in both models and across both components (CC and ZC), with a positive and statistically significant parameter.

Regarding **socioeconomic characteristics**, the average age of vehicles and the proportion of motorcycles emerged as significant variables (at the 10% alpha level) in both

components of the OP and NOP models across most tested variable combinations. For both age groups, the presence of older vehicles is associated with safer municipalities, possibly due to their lower speeds and usage patterns. Conversely, a higher proportion of motorcycles indicates riskier areas. Population density plays a notable role in non-older pedestrian (NOP) crash frequency (CC component), with higher densities corresponding to increased crash occurrences. The motorization rate (number of registered vehicles per inhabitant) serves as a strong predictor in the zero component (ZC) of the OP model, with higher rates linked to safer municipalities for older pedestrians. This relationship may reflect the nature of smaller municipalities, where higher motorization rates are often observed but private vehicles are primarily used for inter-municipal travel rather than within the municipality. These interurban trips are outside the scope of this study, which is focused on intra-municipal collisions.

The proportion of main residences is a moderate predictor of NOP collisions, achieving statistical significance in 30% of tested models for both the count and zero components. The parameter consistently suggests that municipalities with a higher proportion of main residences are safer for non-older pedestrians. Larger municipalities, where most houses are main residences, typically experience more consistent and predictable vehicle and pedestrian flows throughout the year, allowing infrastructure to be better suited for steady traffic patterns. By contrast, municipalities with a high proportion of second or vacation residences experience seasonal surges in population and vehicle flow, which can strain infrastructure and contribute to higher collision risks during peak periods due to insufficient traffic controls or visibility features (e.g., inadequate traffic lights or clear signage). This variable does not appear to influence older pedestrian road safety, suggesting that older adult residents maintain relatively stable mobility patterns year-round, less affected by seasonal population fluctuations.

Land use features demonstrate better predictive performance in the OP model compared to the NOP model. Specifically, the proportions of secondary and tertiary land use significantly increase both the likelihood (zero component, ZC) and frequency (count component, CC) of older pedestrian (OP) crashes, with this effect achieving statistical significance in most tested models. These types of land use may expose older pedestrians to environments with mixed traffic patterns and heightened risk levels, particularly in areas not primarily designed for residential or pedestrian-oriented purposes. The land use entropy mix index, which measures the diversity of land uses within a municipality, also increases the likelihood of pedestrian traffic crashes in general (ZC). Municipalities with a more even distribution of different land uses are riskier for all pedestrians in terms of crash frequency (CC), although this effect is less

pronounced, with lower statistical significance at the 10% alpha level across models. This finding suggests that higher land use heterogeneity could lead to increased interaction between pedestrians and vehicles due to the varied purposes of these areas, though the effect varies in strength.

The number of points of interest (POIs) per kilometer does not significantly predict the probability of pedestrian crashes (ZC), and municipalities with a higher density of POIs are associated with a higher chance of recording zero pedestrian traffic crashes. While this might initially seem counterintuitive, the lack of consistent significance across models leaves the interpretation somewhat inconclusive. However, one possible explanation is that areas with many POIs may be better equipped to handle high pedestrian and vehicular volumes, with infrastructure improvements like increased road signage, traffic lights, and wider sidewalks that facilitate safer pedestrian movement.

In terms of **infrastructure variables**, the proportion of residential roads serves as a strong predictor for the zero components of both OP and NOP models. Municipalities with a greater share of residential roads in their network have an increased likelihood of reporting zero pedestrian traffic crashes, regardless of pedestrian age. This result aligns with the notion that residential roads tend to be narrower, less heavily trafficked, and generally support lower vehicle speeds than primary roads, thereby reducing the overall crash risk for pedestrians of all ages.

Interestingly, a higher density of traffic lights per kilometer correlates with a greater probability of at least one OP traffic crash occurring. This effect may be attributed to the fact that traffic lights are more commonly found in larger municipalities, where pedestrian-vehicle interactions are more frequent, making it more likely for at least one pedestrian crash to occur during the study period.

The presence of major streets did not have a significant effect on either model component (ZC or CC) for both OP and NOP models. However, higher proportions of motorways correlate with fewer OP and NOP crashes. This is consistent with the fact that motorways are typically restricted to vehicle traffic, with pedestrian access limited or prohibited, thus reducing the likelihood of vehicle-pedestrian collisions in these areas.

Road intersections per kilometer have a distinct impact, contributing to a higher frequency (CC) of NOP crashes and an increased likelihood (ZC) of at least one OP crash. This finding underscores that road intersections represent hazardous locations for pedestrians of all ages, as intersections inherently involve complex interactions between vehicles and pedestrians, heightening the potential for conflicts and, consequently, crash risk.

The **statistical models with the best fit** (in terms of AIC) were processed for older and non-older pedestrian crashes. The exposure variables, older and non-older inhabitants, alongside the proportion of motorcycles, are significant predictors in both models, each associated with an increase in the likelihood and frequency of crashes. The presence of motorcycles could be linked to a higher frequency of crashes involving slightly injured pedestrians, as crashes were counted irrespective of injury severity.

For older pedestrian crashes, secondary and tertiary land uses contribute significantly, suggesting that these environments may pose unique risks for older pedestrians. This finding may reflect a preference among older pedestrians for residential and homogeneous areas, where they may feel more comfortable and familiar with the road network.

For non-older pedestrian crashes, population density and the land use mix index are strong predictors of increased crash risk. Non-older pedestrians typically have broader walking patterns, traversing larger portions of the municipality than older pedestrians, who often confine their activities to areas near their residences. Additionally, higher population density often correlates with an expanded range of leisure activities and an increased demand for commuting, both predominantly affecting the non-older population.

5.3 Results of Article C

In the third study developed in the framework of this doctoral thesis, Results of Article C (Gálvez-Pérez et al., 2024a), the frequency of older and non-older pedestrian traffic crashes per street segment at the micro-level was studied and statistically modeled. The dependent variables were the number of older and non-older pedestrian traffic crashes between 2017 and 2022 (omitting 2020 because of the pandemic situation) per street segment in Madrid, modeled separately. As in previous studies, independent variables are divided into socioeconomics, land use, and infrastructure groups. A wider weight was set on the infrastructure variables because as it is a micro-level study.

Unlike deterministic approaches, Bayesian statistics incorporate uncertainty directly, expressing it through the inferred posterior distributions of the model parameters and Bayesian credible intervals (BCIs). An effect is deemed credible at a specific level if the corresponding BCI does not include zero, indicating a meaningful association with the outcome. In this article, 90% BCI was used for this end.

Both models indicate that certain variables have a credible influence on pedestrian crashes across all age groups. These include average annual daily traffic (AADT), the density of leisure points of interest (POIs), the presence of bus stops, and the density of crosswalks, each of which increases the expected number of pedestrian crashes. The inclusion of these factors suggests that high-traffic areas with frequent pedestrian-vehicle interactions, such as leisure spots and public transit access points, are consistently associated with elevated crash risk.

Leisure POIs attract pedestrians who may move around these areas with reduced attention to traffic, possibly feeling safer due to the crowded, well-maintained surroundings. This reduced vigilance can contribute to an increase in pedestrian crashes. Additionally, buses can disrupt traffic flow and obstruct the visibility of oncoming vehicles. This visibility issue is often exacerbated at certain bus stops where opaque billboards further limit sightlines.

In addition to the variables affecting all pedestrians, specific factors uniquely impact the number of traffic crashes involving older pedestrians. Population density, the presence of trees along sidewalks, trash containers, and the level of road contour complexity all play significant roles in increasing collision risks for this demographic. Higher population density is associated with an elevated likelihood of crashes, likely due to increased pedestrian and vehicular interactions in densely populated areas. For socioeconomic factors, population density exhibited a positive and credible association with older pedestrian crashes, aligning with findings from previous studies. However, this relationship was not credible in the model for non-older pedestrians.

Similarly, the presence of trees and trash containers along sidewalks contributes to riskier street environments. The presence of trees was associated with a higher incidence of crashes involving older pedestrians for two primary reasons. First, trees reduce the usable sidewalk width, with open tree pits presenting navigation challenges. This problem is particularly acute on narrow sidewalks, often prompting older adults to walk on the pavement, thus heightening their risk of being involved in crashes. Second, older pedestrians may be more inclined to walk in areas with greenery, leading to increased foot traffic and, consequently, a higher likelihood of crashes. This factor did not influence younger pedestrians, who tend to navigate these sidewalks with greater ease. Trash containers, frequently positioned on sidewalks or pavements, obstruct visibility and hinder older individuals' ability to detect oncoming traffic, particularly for those with visual or hearing impairments. Streets with such containers were identified as presenting a greater risk for older pedestrians.

Road contour complexity also influences crash risk for older pedestrians. Streets that terminate in three-legged intersections (classified as complexity A) are associated with a higher risk of collisions than those without such intersections (complexity C). However, no credible effect is observed for streets with intermediate levels of contour complexity (complexity B), suggesting that only certain types of intersections present distinct hazards for older pedestrians.

Analogously, certain factors distinctly affect the safety of non-older pedestrians. The presence of terraces and metro stations is associated with higher crash risk for this demographic, as these areas tend to draw higher pedestrian traffic and may increase exposure to vehicle interactions.

Restaurant terraces on sidewalks may obstruct pedestrian flow and reduce visibility of approaching vehicles, thereby increasing crash risk. Non-older pedestrians often cross streets with terraces at undesignated areas, unlike older pedestrians, who demonstrate a stronger preference for crossing at designated points. While the effect of terraces on crash risk among older pedestrians appears opposite in roughly half of the parameter's distribution, though not statistically credible, it suggests that terraces might serve as a protective barrier, potentially shielding older individuals from vehicle exposure.

Metro station entrances located on sidewalks are associated with a higher frequency of crashes involving non-older pedestrians, likely due to the increased pedestrian traffic they attract. This effect is not observed for older pedestrians, possibly because individuals above 65 are less likely to use the metro and tend to prefer bus transportation instead. Additionally, street type plays a notable role: while two-way streets with a median offer a safer environment compared to one-way streets, two-way streets without a median do not demonstrate a credible difference in safety.

5.4 Results of Article D

Apart from the results related to articles A-C, Article D: The effect of COVID-19 on older pedestrian road safety: A holistic analysis through pandemic phases (Gálvez-Pérez et al., 2025) considered a different problem: whether the frequency, injury severity and crash-related factors of older pedestrian traffic crashes in urban scenarios have been modified because of the COVID-19 pandemic. The main research question of this article is whether measures proposed before COVID -19 pandemic for improving road safety of older adults are still applicable after the pandemic, meaning that crash-related factors remained roughly constant.

In terms of crash frequency at the national level, the total number of injured pedestrians at all levels of injury remained below pre-pandemic levels for older and non-older pedestrians during the state of alarm phase (which includes lockdown). In the subsequent phase, the new normality, the number of injured older pedestrians also remained below pre-pandemic levels, and it was also the case for non-older pedestrians, except for the fatalities. In the final phase studied, which started after the vaccination rate surpassed 70%, the number of fatalities and seriously injured older and non-older pedestrian victims had already returned to pre-pandemic levels, indicating that older pedestrians had already returned to walk. However, the total number of pedestrian injuries and its main contributor, the number of slightly injured pedestrians, remained below pre-pandemic levels even in this phase for older and non-older pedestrians.

Regarding older pedestrian traffic crash injury severity, the proportions of fatalities, seriously injured and slightly injured pedestrians in the pandemic phases are not statistically different from the ones during the pre-pandemic period, as shown by chi-squared tests.

Finally, in terms of the more in-depth analysis on the modification of crash-related factors during post-pandemic (2021-2022) compared to pre-pandemic (2018-2019) period in Madrid showed that older pedestrian traffic crashes occur less during evening and night after the pandemic, and in areas with narrower sidewalks and with more bike lane supply. Because of these results, avoiding obstructing narrow sidewalks and better marking intersections of bike lanes and pedestrian infrastructure was advised.

6

General Discussion

This final section of the PhD dissertation includes the General Discussion extracted from the developed works to gain more insight into the road safety of older pedestrians in urban environments in an ageing society. First, the Integrated solution considering the results of all the published works and the proposed countermeasures to improve the road safety of older pedestrians are explained in **Section 6.1 Integrated solution and Proposed countermeasures**. Second, the main conclusions, along with the achievement of the thesis' objectives in each published article, are presented in the **Section 6.2 Conclusions**. Finally, the further research lines on older pedestrians' urban road safety are presented in **Section 6.3 Further research**.

6.1 Integrated solution and Proposed countermeasures

This research addresses the study of older pedestrian road safety in urban scenarios from different spatial units of analysis. The separate results presented in the four published papers provide valuable insight into the older pedestrian urban road safety topic. However, a general discussion on a combination of them, considering each research aim and configuration, would provide which results are consistently found regardless of the spatial unit of analysis. Apart from road safety educational campaigns, measures applied to the infrastructure should be carried out to effectively convert our cities into age-friendly cities in terms of older pedestrians' road safety.

The effects of the tested independent variables on the number of traffic crashes suffered by older (OP) and non-older (NOP) pedestrians in the three published articles on crash frequency (Article A, B and C) are shown in Table 6.1. Only variables appearing in more than one article are included in this comparison. A **key finding** of this doctoral thesis is that there is no infrastructure variable influencing both older and non-older pedestrian traffic crashes simultaneously with opposite effects at any level studied. In other words, and this finding is quite important for policymakers, any measure or modification on the infrastructure to improve the road safety level of older pedestrians is not expected to deteriorate the respective situation of non-older pedestrians.

Table 6.1. Comparison of results of the published studies on the frequency of older pedestrian traffic crashes.

Variable	Age group	Article B Macro-level in Spain	Article A Macro-level in Madrid	Article C Micro-level in Madrid
Socioeconomics				
Population density	OP	↓	↑	↑
	NOP	↑	↑	↑
Avg. income per person	OP		↓	↓
	NOP		↓	↓
Land use				
Points of interest (POIs)	OP	↑	↑	↑ (leisure)
	NOP	↑	↓	↑ (leisure)
Infrastructure				
Sidewalk width	OP		↓	↓
	NOP		↓	↑
Residential road	OP	↓		
	NOP	↓		
Major road	OP		↑	
	NOP		↑	
Traffic lights	OP	↓	↓	↓
	NOP	↑	↑	↓
Road intersections	OP	↑	↑	
	NOP	↑	↑	
Crosswalks	OP			↑
	NOP			↑
Bus stops	OP		↑	↑
	NOP		↑	↑
Metro stations	OP		↑	↑
	NOP		↑	↑

OP = Older pedestrians; NOP = Non-older pedestrians;

↑ = increases traffic crashes (statistically significant); ↓ = decreases traffic crashes (statistically significant);
 ↑ = increases traffic crashes (not statistically significant); ↓ = decreases traffic crashes (not statistically significant)

Regarding the road safety of all pedestrians, the presence of **major roads** seems to increase the frequency of crashes and, conversely, **residential roads** seem to decrease the frequency of crashes. These patterns may be attributed to the characteristics of major roads, which, despite offering better protection due to higher vehicle speeds, are often more challenging to cross due to their greater width and the potential lack of adequate crossing locations. In contrast, residential roads typically experience lower vehicle flows and reduced vehicle speeds, making them easier for pedestrians to cross. Additionally, the lower speeds on residential roads improve drivers' ability to detect pedestrians crossing or walking on the roadway.

According to the obtained results, both the high density **street junctions** and **crosswalks** seem to increase the number of all pedestrian traffic crashes. Assuming that street junctions are the usual spot where crosswalks are located, this result is applicable to all units of study in Madrid and Spain case studies. Even though crosswalks are the proper location to cross the street, these spots are also the main points of

interaction between vehicles and pedestrians. Crosswalks should be properly conditioned and, if traffic lights are necessary, they should consider enough time for pedestrians, especially older adults, to cross on time. Even more, road junctions should provide enough visibility for vehicles and pedestrians to see each other at the distance necessary to stop the vehicle and to prevent pedestrians to keep on crossing the street.

Bus stops represent a hazard for all pedestrians in Madrid at both macro- and micro-levels studies. Bus stops might represent a higher movement of vehicles in general, which would result in more traffic crashes. However, this variable also has an impact on the micro-level analysis, which might also be caused by pedestrians being hit by buses or the scenario of buses potentially obstructing the visibility of oncoming vehicles. In addition, bus stops might represent a hazard if they obstruct the visibility of vehicles as they present opaque billboards in the perpendicular direction of oncoming traffic.

The higher density of **traffic lights** in a spatial unit of analysis provoked a decrease in the number of older pedestrian traffic crashes at the macro-level study in Madrid. Traffic lights can be a tool for traffic management while potentially reducing the mean speed of vehicles. In addition, they provide broad benefits for pedestrians as they represent an extra in their security: pedestrians should always look at oncoming vehicles, but traffic lights ensure the right time to cross the street. Given the cautious behavior of older pedestrians, mainly because of their self-regulation, more traffic lights located within an area imply less traffic collisions, and this effect was not obtained in the non-older pedestrian crashes model. This output is consistent with the findings in the other two articles; however, this result was not statistically significant or credible. In the case of the Spanish municipalities, this could be due to the studied zones being large, so traffic lights might be heterogeneously located. In the case of the street segment level study, the result might not be credible because the study unit is small, and traffic crashes might not be recorded during the studied period to reflect the effect.

Nevertheless, this result implies that traffic lights should be proposed when necessary, and the green phase should be extended to account for the speed of older pedestrians, which is remarkably lower than the one of younger pedestrians, and should also consider other street crossings actions, such as dealing with the curbs.

The results of the macro-level analysis of Spanish municipalities suggested the implementation of **traffic signals** to inform drivers of their entry into an **aged urban area**. This measure could help drivers understand that they must control vehicle speed and pay special attention to the road. This measure resembles the existing and widely installed signals pertaining to school zones in Spain.

The **complexity of the street segment contour** showed to reduce older pedestrian crashes when no three-leg intersections are there compared to streets with three-leg intersections at both ends.

Higher values of **mean sidewalk width** are associated with fewer traffic crashes involving older pedestrians at the macro-level in Madrid, and this result was consistent in the micro-level study using the same case study, although this last result was not statistically credible. This indicates the necessity of providing wide enough sidewalks, especially for older pedestrians. Older pedestrians are especially vulnerable to stability issues, which might cause them to fear of falling because of non-smooth surfaces and too narrow or crowded sidewalks, and this might lead to even walking on the carriage-way.

There is a need to analyze the state of the sidewalk surface. More importantly, the width of the sidewalks should be properly analyzed and widened when possible and necessary and considering not the nominal width, but the effective width. In other words, the width of the sidewalks as shown in an Urban Development Plan do not properly reflect if they are wide enough for older pedestrians, but the effective width (i.e., barrier-free space for walking), considering elements on the sidewalks such as trees and bus stops, and how crowded they might be because of leisure or catering establishments.

In line with this issue, **streets with trees** showed to be riskier for older pedestrians at the micro-level in Madrid. Despite having innumerable benefits for everyone's health, trees located on the sidewalk might obstruct normal walking if the sidewalk is too narrow or they are located on the middle of the sidewalk. Moreover, if the tree pit does not continue the surface of the sidewalk, it might represent an additional hazard for older pedestrians to fall. If the correct tree species is not properly selected at the time of planting, roots might end up cracking the sidewalk.

Because of this result, trees should be relocated to proper locations when necessary. In narrow sidewalks, trees should be removed or placed as a separation between pedestrians and vehicles, but not in the middle of the sidewalk, although possible sight obstruction should be always checked. More ideally, in those cases the sidewalk should be widened and trees located to provide comfort and shade. Even in wide sidewalks, a bad disposition of trees can cause the pedestrians walking near them to walk "zigzagging". In any case, trees should not pose an obstacle for older pedestrians. The surface of the sidewalk should be continued as evenly as possible in the tree pit areas, by using metallic pieces or other applications.

Finally, the **presence of trash containers** was also related to riskier streets for older pedestrians at the micro-level in Madrid. Trash containers might obstruct the vision of oncoming vehicles, especially for older pedestrians as they normally present visual impairments, and hearing impairments, which make it more difficult to detect traffic in situations of low visibility. As older pedestrians normally use designated areas to cross the street, a revision of the location of trash containers, especially those with a bigger volume, should be carried out in terms of them obstructing the view of oncoming vehicles. In one-way streets, they should never be located between the oncoming traffic and the pedestrian crossing. Additionally, their allocation in the transversal direction to the street should be fixed and, although it represents a logistical challenge during their operation, they should not be placed too far away from the sidewalk, as a difference of centimeters might change the sight distance of pedestrians and vehicles by a longer distance.

Regarding the **injury severity** of older pedestrian traffic crashes, results obtained in a complementary work of this thesis published in the 2023 Procedia collection of Elsevier proceedings (Gálvez-Pérez, Guirao, et al., 2023), showed that more serious injury severity was found in dark light with artificial lighting conditions (compared to daylight conditions) for older pedestrians. Additionally, if the vehicle involved in the crash was a motorcycle (compared to passenger car), the severity of crash was also more serious. Conversely, if an older pedestrian crash occurs, less severe outcomes are expected if it is during the afternoon (compared to the morning hours) and with higher population densities. In addition, all pedestrian crashes are expected to be more severe with the increasing age of the pedestrian, with dawn or dusk lighting conditions, and if the vehicle is a bus or truck; and crashes are expected to be less severe if the vehicle is a bicycle or if the municipality is above 10,000 inhabitants.

About the **effect of the COVID-19 pandemic on the road safety of older pedestrians**, the frequency and injury severity returned to pre-pandemic levels after the vaccination of the population. The hotspots of older pedestrian traffic crashes in Madrid remained similar during the pandemic compared to the pre-pandemic period. Crash-related factors did not suffer remarkable modifications, except for less crashes occurring during evening and night for older pedestrians and these crashes being in areas with narrower sidewalks and more bike lane supply than before the pandemic. In consequence, even though the pandemic might have changed mobility patterns even after the vaccination of the population, crash-related factors are still similar, so measures to improve their road safety are still applicable.

Built environment variables appear to influence the frequency of traffic crashes involving older pedestrians across both macro- and micro-level units of analysis. While

micro-level units are particularly useful for identifying specific infrastructure problems and proposing targeted countermeasures, their application presents significant challenges for researchers. Nonetheless, advancing the study of road safety for older pedestrians requires the development of more comprehensive and accurate databases to encourage further research in this area. These findings provide substantial insight into the topic of road safety for older pedestrians in urban environments. Broader conclusions regarding these results, along with their implications for other topics addressed in this doctoral thesis and further research lines, are discussed in detail in the following section.

6.2 Conclusions

Older adults are over-represented in pedestrian traffic crashes. Even more, this demographic shows a higher mortality rate than the rest of the population. In urban environments, where most of pedestrian crashes take place, older pedestrians represented 70% of pedestrian fatalities in 2019 in Spain. The main objective of this doctoral thesis was to gain insight into the features influencing the frequency and injury severity of older pedestrian traffic crashes in urban environments. Moreover, a separate and simultaneous analysis of the traffic crashes involving the rest of pedestrians (i.e., non-older pedestrians) was developed to assess if those influencing factors were applicable to the general population or only to older pedestrians. With this aim, the city of Madrid and the country of Spain were both employed as case studies.

The problem has been analyzed using **multiple spatial units of study**. The city of Madrid was first analyzed regarding the frequency of traffic crashes in each city district (Article A). In a subsequent study, Spain was analyzed regarding the frequency of traffic crashes in each municipality (Article B). In a third work, the city of Madrid was again analyzed in terms of the number of traffic crashes recorded in each street segment (Article C), which provided the most interesting results in terms of possible infrastructure countermeasures but led to the most time-consuming data acquisition process accomplished during the development of this thesis.

Table 6.2. Objectives of the thesis and their relationship with the published articles.

Objective	Article			
	A	B	C	D
Objective 1: Identify suitable and complete data sources to gain insight into the urban built environment, conditioning road safety				
Objective 2: Develop a systematic and free procedure to geolocate traffic crashes considering the postal address where they were registered				
Objective 3: Develop dedicated tools to systematically process data on socioeconomics, land use, and infrastructure with different spatial units of analysis				
Objective 4: Design ad-hoc indicators of the urban built environment influencing frequency of traffic crashes, including street infrastructure and traffic variables, both at macro- and micro-level spatial analysis				
Objective 5: Develop statistical models suitable to predict the frequency of older pedestrian traffic crashes				
Objective 6: Propose infrastructure countermeasures to improve the road safety scenario of older pedestrians based on the results of the statistical modeling of traffic crashes involving older adults				
Objective 7: Assess the effect of the covid-19 pandemic on the frequency, injury severity and crash-related factors of older pedestrian traffic crashes in urban scenarios				

Partially addressed



Fully addressed



The objectives of the doctoral thesis and in which published article these were accomplished are shown in Table 6.2. First, multiple data sources were identified as suitable to develop road safety studies in urban scenarios (Table 6.2, Objective 1) in the first three published articles. Because of the casa study, Articles A and C covering the city of Madrid employed national and municipal databases. Article B employed national and global (OpenStreetMap) databases. Overall, this objective has been widely accomplished. Municipal databases proved to be more detailed and complete than national databases. Furthermore, geographical data should be preferred than tabular data, as it allows researchers to perform richer data processing if necessary. For Article D, databases were already identified in previous articles, and this knowledge could be used without further work.

Regarding Objective 2, an efficient and systematic procedure to geolocate traffic crashes from the Spanish crash database was successfully designed in Article A. This procedure takes the location of the crashes (i.e., postal addresses) and transforms them into geographical coordinates. Nevertheless, this procedure needs additional information regarding the geographical coordinates of all the postal addresses of the area studied. Furthermore, the result is inexact because of the nature of the data, as the exact location of the crash is unknown. In consequence, it is advisable that Administrations begin to register the (almost) exact coordinates of traffic crashes instead of collecting the nearest postal address of each one of them.

Data processing systematic techniques were also successfully designed to process data on socioeconomics, land use, and infrastructure with different spatial units of analysis (Objective 3). These techniques were designed for each published article, especially in Article A and C. In Article A, data was processed in terms of administrative divisions (macro-level). In Article B, data was processed in ad hoc divisions (macro-level), extrapolating the techniques developed in Article A. Finally, in Article C, processes regarding street segment micro-level were developed. The techniques developed can be extrapolated and used in any other road safety analysis.

Ad hoc indicators regarding socioeconomics, land use and infrastructure were also designed in the four developed publications (Objective 4), as the spatial unit of analysis is different. As for the processes developed, these indicators can be extrapolated to other road safety studies in the macro- or micro-levels.

In terms of data acquisition and processing, although the understanding and experience of possible data sources and data processing techniques have increased during the development of this doctoral thesis, it has been more time-consuming with each work. The study regarding Spanish municipalities required huge data processing times to process geographical data for a whole country. On the other hand, the study

regarding street segments in Madrid, as stated above, required remarkable ad hoc and manual data acquisition, as there was no available data for that level or detail or for some street features.

In line with the higher effort in data acquisition with the third work carried out, this doctoral thesis **shows the importance of micro-level analysis to gain insight into infrastructure features and possible road safety measures on the infrastructure** are shown in this thesis. Macro-level analyses, despite being interesting, provide results more oriented to urban planning than on specific infrastructure measures. In contrast, micro-level analysis provided results applicable to infrastructure variables affecting the frequency of crashes. At this point, the importance of detailed data, especially on infrastructure features, at a level of detail that enables researchers to perform these analyses should be mentioned. In fact, for physical elements and devices, such as trees, pedestrian fences, or benches, it is recommended for Administrations to register them as geographical data, instead of tabular data, as this method provides more accurate location and representation of features regarding their nature. For example, a tree could be represented as a point or a polygon, and a pedestrian fence as a line following the edge of a sidewalk.

Throughout this doctoral research, a significant portion of time and effort was devoted not only to expected tasks such as mastering statistical modeling, handling spatial data, and reviewing existing literature, but also to the arduous task of data preparation, and especially data cleaning. For instance, considerable effort was spent converting traffic crash locations, initially provided as postal addresses, into geographical coordinates and meticulously reviewing approximately 1,000 traffic volume measurement points in Madrid. This review was necessary because many points were erroneously positioned near unrelated street segments or represented data for only one side of a street.

These challenges highlight the **critical need for road safety and urban infrastructure high-quality open datasets**. While open data offers the reliability of using official information, which is often more credible than data from unofficial sources and sometimes the only source for certain data types, it is not always provided in a ready-to-use format. Researchers frequently need to invest time in cleaning and processing such data, which can pose a significant barrier, particularly given the time constraints associated with conducting research studies.

The substantial effort required for data cleaning highlights a broader issue: while open data is valuable, it often lacks the usability needed for efficient and reliable research. Although access to official datasets offers significant benefits, the burden of cleaning and preparing this data typically falls entirely on researchers, with cleaned datasets

rarely being shareable or reusable. As a result, researchers frequently engage in repetitive tasks that could be reduced if data were released in a standardized, ready-to-use format. This inefficiency underscores the urgent need for open datasets to be not only accessible but also rigorously prepared and documented for direct application in scientific research, ultimately promoting more efficient and collaborative advancements.

Depending on the characteristics of the dependent variable in each study, a wide set of statistical modeling approaches have been employed (Objective 5). Negative binomial and hurdle negative binomial regression models were used with the frequentist approach, and Poisson regression models were used with the Bayesian approach. The Bayesian approach proved to be more flexible, and it provides the real distributions of the parameters of the model. In light of the new methods of approximations of Bayesian inference, and specifically INLA, which provides remarkably faster approximations compared to traditional alternatives such as MCMC methods, Bayesian statistics should be further employed in road safety, which is not the general practice yet in the reviewed literature, in which the frequentist approach is more common.

Infrastructure countermeasures to reduce the number of traffic crashes of older pedestrians were proposed based on the results of statistical modeling (Objective 6). This objective was partially covered in the first two publications (Article A and B) and more deeply covered in the third one (Article C) because of the nature of these analyses. In addition, infrastructure measures were proposed based on the change of crash-related factors before and after the covid-19 pandemic (Article D). Micro-level analysis proved to be more useful than macro-level analysis to analyze and interpret the effect of infrastructure variables on road safety. Macro-level analysis treats crashes count and infrastructure variables in an aggregated way, obstructing a detailed analysis which micro-level provides.

The **main finding** of this PhD dissertation is that **infrastructure variables influencing the frequency or injury severity of both older and non-older pedestrian traffic crashes simultaneously showed the same effect, either increasing or decreasing the number or injury severity of all pedestrian traffic crashes**. Other variables were only found to increase/decrease the number or injury severity of older or non-older pedestrian crashes, or no related to any pedestrian crash. In other words, modifications on the infrastructure to seek higher road safety levels of older pedestrians are not expected to deteriorate the road safety situation of the rest of pedestrians.

Based on the results of the developed works, **older pedestrian education programs should be complemented with infrastructure countermeasures** to adapt our streets to the ageing population in an effort to build age-friendly cities. As education

programs are not obligatory, the group of older adults taking these programs might be biased towards more active older pedestrians who are more aware of their limitations might be more interested in these education programs than others who would need more training. Results provide countermeasures that would benefit all pedestrians and not only older adults, regarding bus stops, street junctions and crosswalks. In addition, other countermeasures are especially focused on older pedestrian road safety, which are related to sidewalk width, traffic lights, street trees, and trash containers. These measures might also improve the road safety of pedestrians younger than 65 years old that share the main intrinsic characteristics of older adults, such as people with reduced mobility.

Finally, the **Objective 7** of assessing the effect of the covid-19 pandemic on the frequency, injury severity and crash-related factors of older pedestrian traffic crashes in urban scenarios was successfully covered in Article D.

Collaborating to reduce the overall number of older pedestrian traffic crashes in urban settings is the final goal of this PhD thesis. This problem's perspective is in line with Vision Zero since the number of older pedestrian fatalities and serious injuries is part of the total number of older pedestrian crashes. The overall number of older pedestrian traffic crashes (i.e., including slight injuries) should be also reduced, encouraging healthy ageing and walking, as these factors are linked to a pedestrian's history of traffic crashes and minor injuries should be prevented especially for older adults because of their intrinsic fragility.

6.3 *Further research*

This doctoral thesis provides valuable insight regarding what mechanisms influence the frequency and injury severity of older pedestrian traffic crashes in urbanized areas. A wide range of methodologies, spatial units of study, and factors were used to analyze the phenomena from different points of view. However, **further research** is needed to address the limitations of the current research.

The use of **alternative statistical models** than those used in the developed works could provide additional interpretations of the problem and confirm the results already obtained. Local models obtained with geographically weighted regressions might help to interpret factors influencing the frequency of older pedestrian traffic crashes in different parts of a city or country. Nevertheless, the used spatial units of analysis should be large enough not to have a remarkable number of zeros (i.e., observations with zero pedestrian crashes), and small enough to have enough different units of analysis.

Global models including spatial terms, such as CAR or SAR model components, could benefit the model's performance, impact on the effect of the independent variables on the dependent variable, and, definitively, enhance the interpretation of the model. Nevertheless, these models require that the studied network or area is connected, necessitating variables such as traffic volume in most of the analyzed street segments. Other interesting modeling approaches would include multivariate models, which have multiple outcome variables that are modeled simultaneously.

The lack of exposure measures, especially regarding pedestrians, is also a drawback of traffic crash studies. In urban scenarios, the traffic flow of vehicles is often not available for all street segments or all the legs of intersections. Information on pedestrian flow is almost non-existent, at least in the employed case studies, Spain and Madrid. The lack of this information can lead to wrong model interpretations. This lack of information could be filled with on-site pedestrian measures for each street segment, what would be remarkably costly. Another option would be **to conduct surveys on pedestrian mobility in an area**, although data from surveys would not be detailed enough. Using social media interactions, such as published tweets on X or posted photos on Instagram, could be a proxy pedestrian flows, and it has the advantage that the exact location would be available. Nevertheless, in terms of demographic groups, a smaller proportion of older people employ these social networks than younger people. In consequence, the volume of younger people would be better reflected than the volume of older people.

A more appropriate approach involves developing a campaign to monitor and analyze the **walking routes of older adults** using **GPS tracking techniques**. This

methodology would include designing a mobile application (app) to record the walking routes of older pedestrians. Once recorded over a reasonable period, these routes could help identify not only walking route preferences as well as micro-level behaviors during walking activities, but such also as the number of times a street is crossed, variations in walking speed, and unsafe movements (e.g., walking outside designated sidewalks or crossing the street using non-designated areas). Additionally, older pedestrians would receive a personalized analysis of their walking habits after the activity. This analysis would assess whether their walking activity adequately fulfills their physical activity needs, potentially motivating them to adopt healthier lifestyles. However, this methodology requires financial resources and careful selection of older adults to be studied, ensuring the inclusion of a diverse group, as older adults exhibit a wide range of physical and cognitive characteristics.

During the research works of this doctoral thesis, older pedestrian traffic crashes have been analyzed without considering sub-groups of older pedestrians (e.g., younger-old, 65-75; and older-old, 75+). This decision was meant to provide a clearer interpretation of the models and a more straightforward comparison with non-older pedestrian traffic crashes. However, once these models are developed, **the use of more detailed grouping strategies would be helpful to understand the age-related differences even inside of the older adults group.**

Road safety should be continually studied because mobility patterns might change over time. The road safety of older pedestrians is not an exception. New mobility patterns, some of them accelerated by the pandemic situation of 2020, might represent new threads for older adults while acting as pedestrians. Bicycles and electric scooters might pose new challenges to older pedestrians. Despite the prevalence of mobility problems of older pedestrians involving uneven or crowded surfaces, especially sidewalks, the most common road safety concern for the older adults registered in Madrid during 2022 and 2023 through the survey developed as part of a qualitative study on older pedestrian behavior and perception of urban road safety (see Section 2.5.5) was micromobility vehicles (i.e., bicycles and scooters). Bicycles and scooters were a road safety concern for almost 80% of participants in this pilot survey. With the consideration of new vehicles (i.e., scooter) in the official DGT's traffic crash database (included in the database as a dedicated vehicle type since 2020), it should be studied whether older pedestrians are remarkably more hit than the rest of the pedestrians by micromobility vehicles. These trends could not be analyzed in this PhD dissertation, as official databases take about 1 year to consolidate, but could be a new research line in the following years.

Because of their lower reflexes, visual and hearing impairments, and difficulties handling complex scenarios, older adults might detect these vehicles with difficulty, as these are small and move at a reasonable speed compared to their size. In addition, it is not common practice for drivers of these vehicles to wear reflective clothing. Furthermore, the presence of these vehicles on sidewalks might force older pedestrians to walk on the carriageway. Regarding cycleways, despite separating micromobility vehicles, regular vehicles, and pedestrians, this represents a more complicated scenario for older pedestrians, who should focus their attention on multiple hazards simultaneously. Because of these characteristics, **electric vehicles**, with a remarkable lower sound production, might suppose in the medium term a thread for older pedestrians, as the lack of combustion sound makes these vehicles difficult to detect.

Other threats, but this time not for the older pedestrians themselves, but to Administrations, are two demographic mega-trends: population ageing and urbanization. Most fatalities and seriously injured pedestrians are among older adults nowadays. This situation can be further deteriorated if more older adults increase in number and in share of population, and if more people move towards urban environments. The first fact is driven by ageing population, which is accused in developed countries such as Spain, and the second fact is due to progressive urbanization of societies, for which Spain shows remarkable projected figures. Both megatrends underline, undoubtedly, the need to develop in the future the new research lines derived from this thesis, which aims to be merely a seed of knowledge in the journey to study the needs of older adults.

References

- Abay, K. A. (2015). Investigating the nature and impact of reporting bias in road crash data. *Transportation Research Part A: Policy and Practice*, 71, 31–45. <https://doi.org/10.1016/j.tra.2014.11.002>
- Abou-Raya, S., & Elmeguid, L. A. (2009). Road traffic accidents and the elderly. *Geriatrics and Gerontology International*, 9(3), 290–297. <https://doi.org/10.1111/j.1447-0594.2009.00535.x>
- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52, 317–332.
- Avers, D., Brown, M., Chui, K. K., Wong, R. A., & Lusardi, M. (2011). Use of the term “elderly.” In *Journal of Geriatric Physical Therapy* (Vol. 34, Issue 4, pp. 153–154). LWW.
- Bailey, S. S., Jones, S. A., Stout, R. J., Bailey, J. H., Kass, S., & Morgan, B. B. (1992). *Safety Research: Enforcement, Speed, Older Drivers, and Pedestrians*.
- Bernhoft, I. M., & Carstensen, G. (2008). Preferences and behaviour of pedestrians and cyclists by age and gender. *Transportation Research Part F: Traffic Psychology and Behaviour*, 11(2), 83–95.
- Blyth, C. R. (1972). On Simpson’s paradox and the sure-thing principle. *Journal of the American Statistical Association*, 67(338), 364–366.
- Borst, H. C., de Vries, S. I., Graham, J. M. A., van Dongen, J. E. F., Bakker, I., & Miedema, H. M. E. (2009). Influence of environmental street characteristics on walking route choice of elderly people. *Journal of Environmental Psychology*, 29(4), 477–484. <https://doi.org/10.1016/j.jenvp.2009.08.002>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.
- Clarke, P., & Gallagher, N. A. (2013). Optimizing mobility in later life: The role of the urban built environment for older adults aging in place. *Journal of Urban Health*, 90(6), 997–1009.
- Corazza, M. V., Di Mascio, P., & Moretti, L. (2017). Management of sidewalk maintenance to improve walking comfort for senior citizens. *WIT Trans. Built Environ*, 176, 195–206.
- Cunningham, C., O’ Sullivan, R., Caserotti, P., & Tully, M. A. (2020). Consequences of physical inactivity in older adults: A systematic review of reviews and meta-

- analyses. In *Scandinavian Journal of Medicine and Science in Sports* (Vol. 30, Issue 5, pp. 816–827). Blackwell Munksgaard. <https://doi.org/10.1111/sms.13616>
- Das, S., Bibeka, A., Sun, X., Zhou, H. “Tracy,” & Jalayer, M. (2019). Elderly Pedestrian Fatal Crash-Related Contributing Factors: Applying Empirical Bayes Geometric Mean Method. *Transportation Research Record*, 2673(8), 254–263. <https://doi.org/10.1177/0361198119841570>
- Davis, R. M., & Pless, B. (2001). BMJ bans “accidents”: Accidents are not unpredictable. In *Bmj* (Vol. 322, Issue 7298, pp. 1320–1321). British Medical Journal Publishing Group.
- Dirección General de Tráfico. (2018). *Las principales cifras de la siniestralidad de los peatones España 2018*.
- Dirección General de Tráfico. (2020). *Tendencias de la movilidad y siniestralidad en vías urbanas*.
- Dirección General de Tráfico. (2021a). *Las principales cifras de la Siniestralidad de los Peatones España 2021*.
- Dirección General de Tráfico. (2021b). *Las principales cifras de la Siniestralidad Vial España 2021*.
- Dirección General de Tráfico. (2022a). *Anuario estadístico de accidentes 2022*.
- Dirección General de Tráfico. (2022b). *Estrategia de seguridad vial 2030*.
- Dirección General de Tráfico. (2023). *Las principales cifras de la siniestralidad vial 2023*.
- Dommes, A., Cavallo, V., & Oxley, J. (2013). Functional declines as predictors of risky street-crossing decisions in older pedestrians. *Accident Analysis and Prevention*, 59, 135–143. <https://doi.org/10.1016/j.aap.2013.05.017>
- Dommes, A., Cavallo, V., Vienne, F., & Aillerie, I. (2012). Age-related differences in street-crossing safety before and after training of older pedestrians. *Accident Analysis and Prevention*, 44(1), 42–47. <https://doi.org/10.1016/j.aap.2010.12.012>
- Dumbaugh, E., & Zhang, Y. (2013). The Relationship between Community Design and Crashes Involving Older Drivers and Pedestrians. *Journal of Planning Education and Research*, 33(1), 83–95. <https://doi.org/10.1177/0739456X12468771>
- Esri Inc. (2020). *ArcMap* (10.8.1).

- European Commission. (2015). *ElderSafe-Risks and countermeasures for road traffic of elderly in Europe (No. MOVE/C4/2014-244)*. European Commission–Directorate-General for Mobility and Transport (DG-MOVE).
- European Commission. (2020). *EU road safety policy framework 2021-2030 – Next steps towards ‘Vision Zero.’*
- European Commission. (2021). *Road safety thematic report – Serious injuries*. European Commission, Directorate General for Transport.
- European Commission. (2022). *EU Road Safety: Towards “Vision Zero”*.
- European Commission. (2024a). *Annual statistical report on road safety in the EU, 2024*. European Commission, Directorate General for Transport.
- European Commission. (2024b). *Facts and Figures. Main Figures*. European Commission, Directorate General for Transport.
- European Commission. (2024c). *Facts and Figures. Pedestrians*. European Commission, Directorate General for Transport.
- European Commission. (2024d). *Facts and Figures. Seniors*. European Commission, Directorate General for Transport.
- European Commission. (2024e). *Facts and Figures. Urban Areas*. European Commission, Directorate General for Transport.
- European Commission. (2024f). *Road safety thematic report – Seniors*. European Commission, Directorate General for Transport.
- Fildes, B., Lee, S. J., Kenny, D., & Foddy, W. (1994). Survey of older road users. Behavioural and travel issues. *Monash University, Accident Research Center*.
- Fried, L. P., Tangen, C. M., Walston, J., Newman, A. B., Hirsch, C., Gottdiener, J., Seeman, T., Tracy, R., Kop, W. J., & Burke, G. (2001). Frailty in older adults: evidence for a phenotype. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 56(3), M146–M157.
- Gabriel, Z., & Bowling, A. N. N. (2004). Quality of life from the perspectives of older people. *Ageing & Society*, 24(5), 675–691.
- Galanis, A., Botzoris, G., & Eliou, N. (2017). Pedestrian road safety in relation to urban road type and traffic flow. *Transportation Research Procedia*, 24, 220–227.
- Gálvez-Pérez, D., Guirao, B., & del-Villar-Juez, D. (2025). The effect of COVID-19 on older pedestrian road safety: A holistic analysis through pandemic phases. *Journal of Transport & Health*, 41, 101986.

- Gálvez-Pérez, D., Guirao, B., & Ortuño, A. (2023). Analysis of the Elderly Pedestrian Injury Severity in Urban Traffic Accidents in Spain using Machine Learning Techniques. *Transportation Research Procedia*, 71, 6–13. <https://doi.org/10.1016/j.trpro.2023.11.051>
- Gálvez-Pérez, D., Guirao, B., & Ortuño, A. (2024a). Age-Friendly Urban Design for Older Pedestrian Road Safety: A Street Segment Level Analysis in Madrid. *Sustainability*, 16(19), 8298.
- Gálvez-Pérez, D., Guirao, B., & Ortuño, A. (2024b). Analysis of the elderly pedestrian traffic accidents in urban scenarios: the case of the Spanish municipalities. *International Journal of Injury Control and Safety Promotion*, 1–20.
- Gálvez-Pérez, D., Guirao, B., Ortuño, A., & Picado-Santos, L. (2022). The Influence of Built Environment Factors on Elderly Pedestrian Road Safety in Cities: The Experience of Madrid. *International Journal of Environmental Research and Public Health*, 19(4). <https://doi.org/10.3390/ijerph19042280>
- Gálvez-Pérez, D., Mohíno Sanz, I., Guirao, B., Molina-Sánchez, R., & Vassallo, J. M. (2023). *Aprendiendo de los peatones senior: una experiencia de aprendizaje-servicio en seguridad vial*. Universidad de Zaragoza, Servicio de Publicaciones.
- Guo, M., Yuan, Z., Janson, B., Peng, Y., Yang, Y., & Wang, W. (2021). Older Pedestrian Traffic Crashes Severity Analysis Based on an Emerging Machine Learning XGBoost. *Sustainability*, 13(2), 926. <https://doi.org/10.3390/su13020926>
- Hanson, S., & Jones, A. (2015). Is there evidence that walking groups have health benefits? A systematic review and meta-analysis. In *British Journal of Sports Medicine* (Vol. 49, Issue 11, pp. 710–715). BMJ Publishing Group. <https://doi.org/10.1136/bjsports-2014-094157>
- Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge University Press.
- Hilbe, J. M. (2014). *Modeling Count Data*. Cambridge University Press.
- Hydén, C. (1987). The development of a method for traffic safety evaluation: The Swedish Traffic Conflicts Technique. *Bulletin Lund Institute of Technology, Department*, 70.
- Irvine, K. N., Fisher, D., Marselle, M. R., Currie, M., Colley, K., & Warber, S. L. (2022). Social Isolation in Older Adults: A Qualitative Study on the Social Dimensions of Group Outdoor Health Walks. *International Journal of Environmental Research and Public Health*, 19(9). <https://doi.org/10.3390/ijerph19095353>

- Jansen, C. P., Klenk, J., Nerz, C., Todd, C., Labudek, S., Kramer-Gmeiner, F., Becker, C., & Schwenk, M. (2021). Association between everyday walking activity, objective and perceived risk of falling in older adults. *Age and Ageing*, *50*(5), 1586–1592. <https://doi.org/10.1093/ageing/afab037>
- Kalache, A., & Kickbusch, I. (1997). A global strategy for healthy ageing. *World Health*, *50*(4), 4–5.
- Kim, D. (2019). The transportation safety of elderly pedestrians: Modeling contributing factors to elderly pedestrian collisions. *Accident Analysis and Prevention*, *131*, 268–274. <https://doi.org/10.1016/j.aap.2019.07.009>
- Kim, S., & Ulfarsson, G. F. (2019). Traffic safety in an aging society: Analysis of older pedestrian crashes. *Journal of Transportation Safety and Security*, *11*(3), 323–332. <https://doi.org/10.1080/19439962.2018.1430087>
- Kitali, A. E., Kidando, E., Sando, T., Moses, R., & Ozguven, E. E. (2017). Evaluating Aging Pedestrian Crash Severity with Bayesian Complementary Log–Log Model for Improved Prediction Accuracy. *Transportation Research Record: Journal of the Transportation Research Board*, *2659*(1), 155–163. <https://doi.org/10.3141/2659-17>
- Koepsell, T., McCloskey, L., Wolf, M., Moudon, A. V., Buchner, D., Kraus, J., & Patterson, M. (2002). Crosswalk markings and the risk of pedestrian–motor vehicle collisions in older pedestrians. *Jama*, *288*(17), 2136–2143.
- Kummeneje, A. M., & Rundmo, T. (2019). Risk perception, worry, and pedestrian behaviour in the Norwegian population. *Accident Analysis and Prevention*, *133*. <https://doi.org/10.1016/j.aap.2019.105294>
- Laković, S., Tollazzi, T., & Gruden, C. (2023). Elderly Pedestrians and Road Safety: Findings from the Slovenian Accident Database and Measures for Improving Their Safety. *Sustainability (Switzerland)*, *15*(2). <https://doi.org/10.3390/su15021631>
- Lalika, L., Kitali, A. E., Haule, H. J., Kidando, E., Sando, T., & Alluri, P. (2022). What are the leading causes of fatal and severe injury crashes involving older pedestrian? Evidence from Bayesian network model. *Journal of Safety Research*, *80*, 281–292. <https://doi.org/10.1016/j.jsr.2021.12.011>
- Langford, J., Bohensky, M., Koppel, S., & Newstead, S. (2008). Do older drivers pose a risk to other road users? *Traffic Injury Prevention*, *9*(3), 181–189.
- Langford, J., Methorst, R., & Hakamies-Blomqvist, L. (2006). Older drivers do not have a high crash risk—A replication of low mileage bias. *Accident Analysis & Prevention*, *38*(3), 574–578.

- Lee, S., Yoon, J., & Woo, A. (2020). Does elderly safety matter? Associations between built environments and pedestrian crashes in Seoul, Korea. *Accident Analysis and Prevention*, 144. <https://doi.org/10.1016/j.aap.2020.105621>
- Li, F. (2005). Multilevel modelling of built environment characteristics related to neighbourhood walking activity in older adults. *Journal of Epidemiology & Community Health*, 59(7), 558–564. <https://doi.org/10.1136/jech.2004.028399>
- Liu, J., Zhong, S., Huang, Y., & Wang, Z. (2021). How does the preference heterogeneity affect the elderly's evaluation of bus accessibility? Evidence from Guangzhou, China. *Journal of Transport and Health*, 22. <https://doi.org/10.1016/j.jth.2021.101094>
- López-Lambas, M. E., Sánchez, J. M., & Alonso, A. (2021). The walking health: A route choice model to analyze the street factors enhancing active mobility. *Journal of Transport and Health*, 22. <https://doi.org/10.1016/j.jth.2021.101133>
- Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice*, 44(5), 291–305. <https://doi.org/10.1016/j.tra.2010.02.001>
- Lv, M., Wang, N., Yao, S., Wu, J., & Fang, L. (2021). Towards healthy aging: influence of the built environment on elderly pedestrian safety at the micro-level. *International Journal of Environmental Research and Public Health*, 18(18), 9534.
- Martin, A. J., Hand, E. B., Trace, F., & O'Neill, D. (2010). Pedestrian fatalities and injuries involving Irish older people. *Gerontology*, 56(3), 266–271. <https://doi.org/10.1159/000258052>
- Mattle, M., Meyer, U., Lang, W., Mantegazza, N., Gagesch, M., Mansky, R., Kressig, R. W., Egli, A., Orav, E. J., & Bischoff-Ferrari, H. A. (2022). Prevalence of Physical Activity and Sedentary Behavior Patterns in Generally Healthy European Adults Aged 70 Years and Older—Baseline Results From the DO-HEALTH Clinical Trial. *Frontiers in Public Health*, 10. <https://doi.org/10.3389/fpubh.2022.810725>
- Mizuta, S., Uchida, K., Sawa, R., Nakamura, J., Encho, H., Akisue, T., & Ono, R. (2023). Context of walking and loneliness among community-dwelling older adults: a cross-sectional study. *BMC Geriatrics*, 23(1). <https://doi.org/10.1186/s12877-023-04043-5>
- Musselwhite, C., Holland, C., & Walker, I. (2015). The role of transport and mobility in the health of older people. In *Journal of Transport & Health* (Vol. 2, Issue 1, pp. 1–4). Elsevier.

- Nembrini, S., König, I. R., & Wright, M. N. (2018). The revival of the Gini importance? *Bioinformatics*, *34*(21), 3711–3718. <https://doi.org/10.1093/bioinformatics/bty373>
- Noh, Y., Kim, M., & Yoon, Y. (2018). Elderly pedestrian safety in a rapidly aging society—Commonality and diversity between the younger-old and older-old. *Traffic Injury Prevention*, *19*(8), 874–879. <https://doi.org/10.1080/15389588.2018.1509209>
- Nordbakke, S., & Schwanen, T. (2014). Well-being and mobility: A theoretical framework and literature review focusing on older people. *Mobilities*, *9*(1), 104–129.
- O’Hern, S., Oxley, J., & Logan, D. (2015). Older Adults at Increased Risk as Pedestrians in Victoria, Australia: An Examination of Crash Characteristics and Injury Outcomes. *Traffic Injury Prevention*, *16*(sup2), S161–S167. <https://doi.org/10.1080/15389588.2015.1061662>
- Oxley, J., Corben, B., Fildes, B., O’Hare, M., & Rothengatter, T. (2004). Older vulnerable road users- measures to reduce crash and injury risk. *Monash University Accident Research Centre Reports*, *218*, 162.
- Oxley, J., Fildes, B., Ihlen, E., Charlton, J., & Days, R. (1997). Differences in traffic judgements between young and old adult pedestrians. In *Accid. Anal. and Prev* (Vol. 29, Issue 6).
- Pae, G., & Akar, G. (2020). Effects of walking on self-assessed health status: Links between walking, trip purposes and health. *Journal of Transport and Health*, *18*. <https://doi.org/10.1016/j.jth.2020.100901>
- Phillips, J., Walford, N., Hockey, A., Foreman, N., & Lewis, M. (2013). Older people and outdoor environments: Pedestrian anxieties and barriers in the use of familiar and unfamiliar spaces. *Geoforum*, *47*, 113–124. <https://doi.org/10.1016/j.geoforum.2013.04.002>
- R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Viena.
- Rankavat, S., & Tiwari, G. (2020). Influence of actual and perceived risks in selecting crossing facilities by pedestrians. *Travel Behaviour and Society*, *21*, 1–9. <https://doi.org/10.1016/j.tbs.2020.05.003>
- Remón, Á. L. C., Díaz-Benito, V. J., Beatty, J. E. J., & Lozano, J. A. S. (2020). Levels of physical activity among older adults in the European Union. *Journal of Aging and Physical Activity*, *29*(2), 242–249.

- Rod, J. E., King, M., & Oviedo-Trespalacios, O. (2023). The impact of perceived injury risk and psychosocial factors on walking equity. *Transportation Research Part D: Transport and Environment*, 116. <https://doi.org/10.1016/j.trd.2022.103590>
- Romero Ortuño, R. (2010). La regulación de los semáforos peatonales en España: ¿tienen las personas mayores tiempo suficiente para cruzar? *Revista Española de Geriatria y Gerontologia*, 45(4), 199–202. <https://doi.org/10.1016/j.regg.2010.04.002>
- Rosenbloom, T., Sapir-Lavid, Y., & Perlman, A. (2016). Risk factors in road crossing among elderly pedestrians and readiness to adopt safe behavior in socio-economic comparison. *Accident Analysis and Prevention*, 93, 23–31. <https://doi.org/10.1016/j.aap.2016.04.004>
- Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. In *J. R. Statist. Soc. B* (Vol. 71, Issue 2).
- Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P., & Lindgren, F. K. (2017). Bayesian computing with INLA: A review. In *Annual Review of Statistics and Its Application* (Vol. 4, pp. 395–421). Annual Reviews Inc. <https://doi.org/10.1146/annurev-statistics-060116-054045>
- Santos, K., Dias, J. P., & Amado, C. (2022). A literature review of machine learning algorithms for crash injury severity prediction. *Journal of Safety Research*, 80, 254–269.
- Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. A. (2011). The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis & Prevention*, 43(5), 1666–1676.
- Saxena, A. (2023). Assessing the factors affecting the perceived crossing speed of pedestrians and investigating the direct and indirect effects of crash risk perception on perceived crossing speed. *Journal of Transport and Health*, 28. <https://doi.org/10.1016/j.jth.2022.101541>
- Shellito, N., & Roldan, N. V. (2019). Walking away from loneliness: The mediating role of social isolation. *Innovation in Aging*, 3(Suppl 1), S836.
- Silva, P. B., Andrade, M., & Ferreira, S. (2020). Machine learning applied to road safety modeling: A systematic literature review. *Journal of Traffic and Transportation Engineering (English Edition)*, 7(6), 775–790.
- Silveira, E. A., Mendonça, C. R., Delpino, F. M., Elias Souza, G. V., Pereira de Souza Rosa, L., de Oliveira, C., & Noll, M. (2022). Sedentary behavior, physical inactivity,

- abdominal obesity and obesity in adults and older adults: A systematic review and meta-analysis. *Clinical Nutrition ESPEN*, 50, 63–73. <https://doi.org/10.1016/j.clnesp.2022.06.001>
- Strath, S., Isaacs, R., & Greenwald, M. J. (2007). Operationalizing Environmental Indicators for Physical Activity in Older Adults. In *Journal of Aging and Physical Activity* (Vol. 15).
- Tournier, I., Dommès, A., & Cavallo, V. (2016). Review of safety and mobility issues among older pedestrians. *Accident Analysis and Prevention*, 91, 24–35. <https://doi.org/10.1016/j.aap.2016.02.031>
- Treat, J. R., Tumbas, N. S., McDonald, S. T., Shinar, D., Hume, R. D., Mayer, R. E., Stansifer, R. L., & Castellan, N. J. (1979). *Tri-level study of the causes of traffic accidents: final report. Executive summary*. Indiana University, Bloomington, Institute for Research in Public Safety.
- UNECE, Eurostat, & ITF. (2019). *Glossary for transport statistics* (European Union, United Nations, ITF, & OECD, Eds.; 5th edition). Publications Office of the European Union.
- Ungvari, Z., Fazekas-Pongor, V., Csiszar, A., & Kunutsor, S. K. (2023). The multifaceted benefits of walking for healthy aging: from Blue Zones to molecular mechanisms. In *GeroScience* (Vol. 45, Issue 6, pp. 3211–3239). Springer Science and Business Media Deutschland GmbH. <https://doi.org/10.1007/s11357-023-00873-8>
- United Nations. (2019). *World Urbanization Prospects. The 2018 Revision (ST/ESA/SER.A/420)*.
- United Nations. (2020a). *Improving global road safety. Resolution adopted by the General Assembly on 31 August 2020*.
- United Nations. (2020b). *World Population Ageing 2019. Department of Economic and Social Affairs Population Division: New York, NY, USA*.
- United Nations. (2023). *World Social Report 2023: Leaving no one behind in an ageing world*. United Nations. Department of Economic and Social Affairs.
- van Hoof, J., Dikken, J., Buttigieg, S. C., van den Hoven, R. F. M., Kroon, E., & Mars-ton, H. R. (2020). Age-friendly cities in the Netherlands: An explorative study of facilitators and hindrances in the built environment and ageism in design. *Indoor and Built Environment*, 29(3), 417–437. <https://doi.org/10.1177/1420326X19857216>

- Vine, D., Buys, L., & Aird, R. (2012). Experiences of Neighbourhood Walkability Among Older Australians Living in High Density Inner-City Areas. *Planning Theory and Practice*, 13(3), 421–444. <https://doi.org/10.1080/14649357.2012.696675>
- Wang, Y., Haque, M. M., & Chin, H. C. (2017). Elderly pedestrian injuries in Singapore. *Journal of Transportation Safety and Security*, 9(3), 273–300. <https://doi.org/10.1080/19439962.2016.1194353>
- Watson, A., Watson, B., & Vallmuur, K. (2015). Estimating under-reporting of road crash injuries to police using multiple linked data collections. *Accident Analysis & Prevention*, 83, 18–25. <https://doi.org/10.1016/j.aap.2015.06.011>
- Watson, K. B., Whitfield, G. P., Bricka, S., & Carlson, S. A. (2021). Purpose-based walking trips by duration, distance, and select characteristics, 2017 national household travel survey. In *Journal of Physical Activity and Health* (Vol. 18, Issue 1, pp. S86–S93). Human Kinetics Publishers Inc. <https://doi.org/10.1123/JPAH.2021-0096>
- Whiting, S., Mendes, R., Abu-Omar, K., Gelius, P., Crispo, A., McColl, K., Simmonds, P., Fedkina, N., Andreyan, D., Gahraman, H., Migal, T., Sturua, L., Obreja, G., Abdurakhmanova, Z., Saparkulovna, I. N., Erguder, T., Ekinci, B., Keskinilic, B., Shukurov, S., ... Breda, J. (2021). Physical inactivity in nine European and Central Asian countries: An analysis of national population-based survey results. *European Journal of Public Health*, 31(4), 846–853. <https://doi.org/10.1093/eurpub/ckab028>
- Wilmot, K., & Purcell, C. (2022). Why are older adults more at risk as pedestrians? A systematic review. *Human Factors*, 64(8), 1269–1291.
- World Health Organization. (2002). *Active Ageing: A Policy Framework*. <http://www.who.int/hpr/>
- World Health Organization. (2007). *Global age-friendly cities: A guide: World Health Organization*.
- World Health Organization. (2015). *World Report on Ageing and Health*.
- World Health Organization. (2018a). *Age-friendly environments in Europe: indicators, monitoring and assessments*. World Health Organization. Regional Office for Europe.
- World Health Organization. (2018b). *Global status report on road safety 2018*. World Health Organization.

- World Health Organization. (2018c). *The global network for age-friendly cities and communities: Looking back over the last decade, looking forward to the next*. World Health Organization.
- World Health Organization. (2021a). *Decade of healthy ageing: baseline report*. World Health Organization.
- World Health Organization. (2021b). *Global Plan Decade of Action for Road Safety 2021-2030*.
- World Health Organization. (2022). *Preventing injuries and violence: an overview*.
- World Health Organization. (2023a). *Global status report on road safety 2023*.
- World Health Organization. (2023b). *National programmes for age-friendly cities and communities: A guide*. World Health Organization.
- Zegeer, C. V, Stutts, J. C., Huang, H., Zhou, M., & Lodgman, E. (1993). Analysis of Elderly Pedestrian Accidents and Recommended Countermeasures. *Transportation Research Record, 1405*, 56–63.
- Ziakopoulos, A., & Yannis, G. (2020). A review of spatial approaches in road safety. *Accident Analysis & Prevention, 135*, 105323.

Annex 1

Code for extracting traffic crashes from the database

It is important to note that the R 'pipeline' was used to process the data. The R 'pipeline' (denoted in code as `%>%`) allows for a more intuitive and readable coding structure, resembling the way we naturally read and interpret procedures, as opposed to standard code writing. For example, in traditional code, if we wish to filter data first by variable A to be equal to 8, and then by variable B to be equal to 7, the code would be written as follows:

```
result <- filter(filter(DATA, variable_a == 8), variable_b == 7)
```

This approach is not straightforward, as it requires reading and writing the code from the inside out. Instead, the 'pipeline' simplifies this process by passing intermediate results directly to the next function. The same operation using the pipeline would be written as:

```
result <- filter(DATA, variable_a == 8) %>%  
  filter(variable_b == 7)
```

This method makes the workflow clearer and allows the code to be written in the same logical order that it is read: first, the dataset is filtered by variable A, and then the already filtered dataset is further filtered by variable B.

Vehicle-pedestrian collisions

Initially, vehicle-pedestrian collisions were extracted from the comprehensive traffic crash database. This operation was conducted identically across both the ARENA and ARENA2 databases. The input table for this process was the crashes table, labeled `CRASHES_TABLE`. The extraction was performed using the following code:

```
veh_ped_col <- CRASHES_TABLE %>%  
  filter(TYPE == "vehicle-pedestrian collision")*  
  
* Note that the first argument of the filter function is passed using  
the pipe (%>%)
```

Single vehicle-pedestrian collisions

Multiple vehicles or pedestrians might be involved in a vehicle-pedestrian collision, although it is not the most common case. In this specific doctoral thesis, considering these traffic crashes might be misleading, as if two pedestrians were involved in a crash one can be older and the other non-older. Even more, maybe the older pedestrian was injured and the non-older was not, or vice versa, so maybe we should only consider the injured as being involved in the crash. There are other cases where we find no records for the pedestrian and/or for the driver of the vehicle, or the vehicle itself. In those cases, it is not possible to determine the age or the injury severity of the pedestrian and/or driver.

Considering only single vehicle-pedestrian collisions, which only include 1 pedestrian, and 1 vehicle involved, makes it easier to relate traffic crash characteristics with pedestrian, driver, and vehicle characteristics. In other cases, when relating the crashes table with the others, the output would contain more rows than the original traffic crashes table, as one-to-many joins would be present. In addition, considering only single vehicle-pedestrian collisions, it is easier to process and interpret the data, as the number of crashes would be equal to the number of pedestrians.

The process starts with creating a table containing, for each vehicle-pedestrian collision, the number of pedestrians and vehicles involved. Later, the number of pedestrians and vehicles is included in the crashes table. Then, the number of pedestrians and vehicles involved in each crash is attached for each crash in the crashes table. Finally, crashes with one pedestrian and one vehicle can be identified in that table. This process is performed slightly differently when dealing with ARENA and ARENA2 databases.

In ARENA database, the process requires the individuals table (*INDIVIDUALS_TABLE*) and the vehicles table (*VEHICLES_TABLE*). With this data we need to filter those individuals involved in vehicle-pedestrian collisions, and those who are pedestrians. Later, we would count how many “rows” are there for each crash, as one row correspond to one pedestrian. The same is carried out for the vehicles table.

```

num_pedestrians <- INDIVIDUALS_TABLE %>%
  filter(ID_CRASH %in% veh_ped_col$ID_CRASH) %>%
  filter(POSITION_VEHICLE == "pedestrian") %>%
  group_by(ID_CRASH) %>% *
  summarize(num_pedestrians = n()) *

num_vehicles <- VEHICLES_TABLE %>%
  filter(ID_CRASH %in% veh_ped_col$ID_CRASH) %>%
  group_by(ID_CRASH) %>%
  summarize(num_vehicles = n()) *

veh_ped_col <- veh_ped_col %>%
  left_join(num_pedestrians, by = "CRASH_ID") %>%
  left_join(num_vehicles, by = "CRASH_ID") %>%
  mutate(num_pedestrians = coalesce(num_pedestrians,
                                     0),
         num_vehicles = coalesce(num_vehicles, 0))

veh_ped_col_1ped_1veh <- veh_ped_col %>%
  filter(num_pedestrians == 1,
         num_vehicles == 1)

* The term n() calculates the number of rows.

```

In the ARENA2 case, it is slightly different, as we have the pedestrians already in a separate table, so we need the pedestrians table (*PEDESTRIANS_TABLE*) and the vehicles table (*VEHICLES_TABLE*).

```

num_pedestrians <- PEDESTRIANS_TABLE %>%
  filter(ID_CRASH %in% veh_ped_col$ID_CRASH) %>%
  group_by(ID_CRASH) %>%
  summarize(num_pedestrians = n())

num_vehicles <- VEHICLES_TABLE %>%
  filter(ID_CRASH %in% veh_ped_col$ID_CRASH) %>%
  group_by(ID_CRASH) %>%
  summarize(num_vehicles = n())

veh_ped_col <- veh_ped_col %>%
  left_join(num_pedestrians, by = "CRASH_ID") %>%

```

```

left_join(num_vehicles, by = "CRASH_ID") %>%
mutate(num_pedestrians = coalesce(num_pedestrians,
                                0),
       num_vehicles = coalesce(num_vehicles, 0))

veh_ped_col_1ped_1veh <- veh_ped_col %>%
  filter(num_pedestrians == 1,
         num_vehicles == 1)

```

Vehicle-pedestrian collisions considering pedestrian's age

Once single vehicle-pedestrian collisions are identified, the age of each pedestrian can be directly assigned or related to its corresponding traffic crash without having duplicate records. Later, the crashes table can be divided depending on the pedestrian being older or not. This process is performed slightly differently in ARENA and ARENA2. For the ARENA case, both the table containing single vehicle-pedestrians collision created above (*veh_ped_col_1ped_1veh*) and individuals table are used as follows:

```

# 1. Create a table with two columns: CRASH_ID, pedestrians' age

ped_age <- INDIVIDUALS_TABLE %>%
  filter(POSITION_VEHICLE == "pedestrian") %>%
  select(CRASH_ID, AGE)

# 2. Pedestrians' age is related to the single vehicle-pedestrian collisions table

veh_ped_col_1ped_1veh <- veh_ped_col_1ped_1veh %>%
  left_join(ped_age, by = "CRASH_ID")

# 3. Split the table depending on the age of the pedestrian

veh_ped_col_1ped_1veh_older <- veh_ped_col_1ped_1veh %>%
  filter(AGE >= 65)

veh_ped_col_1ped_1veh_nonolder <- veh_ped_col_1ped_1veh %>%
  filter(AGE < 65)

```

For the ARENA2 case, both the table containing single vehicle-pedestrians collision created above (*veh_ped_col_1ped_1veh*) and pedestrians table are used as follows:

```

# 1. Create a table with two columns: CRASH_ID, pedestrians' age

ped_age <- PEDESTRIANS_TABLE %>%
  select(CRASH_ID, AGE)

```

```
# 2. Pedestrians' age is related to the single vehicle-pedestrian collisions table
```

```
veh_ped_col_1ped_1veh <- veh_ped_col_1ped_1veh %>%  
  left_join(ped_age, by = "CRASH_ID")
```

```
# 3. Split the table depending on the age of the pedestrian
```

```
veh_ped_col_1ped_1veh_older <- veh_ped_col_1ped_1veh %>%  
  filter(AGE >= 65)
```

```
veh_ped_col_1ped_1veh_nonolder <- veh_ped_col_1ped_1veh %>%  
  filter(AGE < 65)
```