



# Identificação de fatores determinantes da severidade das lesões sofridas por pedestres nas rodovias federais brasileiras entre 2017 e 2019: Análise via regressão logística multinomial

Identification of the determinants of severity for pedestrian injuries at Brazilian federal highways between 2017 and 2019: A multinomial logistic regression-based analysis

R. Miranda<sup>1</sup>; W. P. da Silva<sup>1</sup>; S. Dutt-Ross<sup>2</sup>

<sup>1</sup>Universidade Federal Fluminense, 24220-900, Niterói-RJ, Brasil

<sup>2</sup>Universidade Federal do Estado do Rio de Janeiro, 22290-240, Rio de Janeiro-RJ, Brasil

\*rian\_miranda@id.uff.br

(Recebido em 27 de agosto de 2020; aceito em 20 de abril de 2021)

A mobilidade a pé é uma das principais formas de deslocamento da população brasileira, representando 41% das viagens realizadas no país em 2016. Entretanto, o compartilhamento do espaço entre pessoas e veículos motorizados em locais desprovidos de infraestrutura adequada tem provocado situações de conflito, aumentando os riscos de ferimentos e fatalidades entre os transeuntes. Neste estudo, foi realizada uma análise sobre os fatores determinantes da severidade das lesões sofridas por pedestres nas rodovias federais brasileiras no período entre 2017 e 2019. Para tanto, foram utilizados dados oficiais da Polícia Rodoviária Federal para a construção de um modelo de regressão logística multinomial, que descreveu o estado físico dos pedestres a partir das variáveis explicativas presentes na base de dados. Os resultados obtidos apontam que: pedestres atropelados por veículos de grande porte, como caminhões, ônibus e utilitários apresentam maiores chances de morte; fases do dia com menor luminosidade natural, como a plena noite, o amanhecer e o anoitecer, oferecem maiores riscos de lesões graves e fatalidades; perímetros urbanos oferecem menor risco de fatalidades em relação às áreas rurais; pedestres do sexo masculino têm maiores chances de apresentarem ferimentos graves ou virem a óbito; pistas múltiplas e finais de semana oferecem maiores chances de morte ou lesões graves ao pedestre. Por fim, foi identificada uma tendência de aumento anual nas chances de lesão grave (9,9% ao ano) e morte (3,3% ao ano) entre os pedestres brasileiros, bem como maiores riscos à integridade física dos pedestres das regiões norte, nordeste e centro-oeste.

Palavras-chave: atropelamento de pedestres, segurança pública, regressão logística multinomial.

Walkability is a fundamental means of transportation among the Brazilian society, representing 41% of the travels inside the country in 2016. However, the absence of adequate infrastructure has caused conflict in places where people and vehicles share space, leading to injuries and fatalities among pedestrians. This study analyses the determinants of the severity of pedestrian injuries that occurred at Brazilian Federal Highways between 2017 and 2019. We used official data from the Federal Highway Police to build a multinomial logistic regression model, which described the physical condition of pedestrians based on the explanatory variables available in the database. The results indicate that pedestrian-vehicle crashes involving heavy vehicles, such as trucks, buses, and SUVs, caused higher odds of pedestrian-death. Furthermore, dark daytimes such as night, dawn, and dusk, are risky of both severe injuries and pedestrian-deaths; urban zones have a lower risk of pedestrian-death compared to rural areas; male pedestrians are more threatened by death and severe injuries than female pedestrians; and multiple corridor highways and weekends hold higher odds of pedestrian death or severe injuries. Finally, we identified an annual rise in the odds of severe injuries (9,9% per year) and deaths (3,3% per year) among Brazilian pedestrians, as well as a higher risk of physical damage for northern, northeastern, and middle-western pedestrians.

Keywords: pedestrian-vehicle crashes, public safety, multinomial logistic regression.

## 1. INTRODUÇÃO

A mobilidade a pé é uma das principais formas de deslocamento da população brasileira, representando 41% das viagens realizadas no país em 2016 [1]. Estudos revelam que a caminhada promove importantes benefícios à saúde, como a redução do risco de doenças cardiovasculares e cognitivas, bem como a queda dos níveis de colesterol, gordura corporal e depressão [2-6]. Por este motivo, grupos de pesquisa e organizações de saúde pública têm encorajado o desenvolvimento de

políticas de planejamento urbano que ofereçam segurança e qualidade ao deslocamento dos pedestres [7-10]. Entretanto, o crescimento do número de viagens realizadas por meio de veículos motorizados em todo o mundo, aliada a infraestruturas urbanas insuficientes à realização de traslados a pé, tem elevado consideravelmente os riscos de acidentes de trânsito entre veículos e pedestres [11].

No contexto mundial, os acidentes de trânsito registrados diariamente nas estradas e rodovias representam um quebra-cabeças a ser solucionado pelas autoridades de saúde e segurança pública. Segundo dados da Organização Mundial da Saúde (OMS), os acidentes de trânsito fazem aproximadamente 1,35 milhões de vítimas fatais anualmente, o que representa uma taxa de 18,2 mortes para cada 100.000 habitantes [12]. A organização ainda estima que os custos financeiros associados ao atendimento das vítimas girem em torno de 3% do PIB para a maioria dos países. Estudos realizados pela Organização Pan-Americana de Saúde (OPAS) apontam que esta taxa atinge valores inferiores entre os países situados na região das Américas, com uma média de 15,6 mortes por 100.000 habitantes em 2016 [13]. Entretanto, a OPAS realça que 13 dos 30 países pesquisados apresentaram taxas de mortalidade superiores à média da região, apontando que as diferentes realidades observadas em cada país geram discrepâncias em suas taxas individuais.

Entre os países que registraram uma taxa de mortalidade superior à média da região, encontra-se o Brasil, com 19,7 mortes para cada 100.000 habitantes. Segundo a PAHO (2019) [13], os pedestres são o terceiro grupo de usuários mais afetados pela mortalidade no trânsito brasileiro, representando cerca de 18,1% das mortes ocorridas no país. Ladeira et al. (2017) [14] ainda apontam que os acidentes de trânsito envolvendo pedestres são o segundo principal fator responsável tanto pela redução do tempo de vida dos brasileiros, quanto pelo aumento dos anos vividos com sequelas físicas. Em 2015, os acidentes com pedestres no Brasil resultaram em uma taxa de anos de vida perdidos, ajustados por morte ou incapacidade (DALY - *Disability-Adjusted Life Year*), igual a 430,1 anos por 100.000 habitantes [14]. O elevado número de pedestres impactados pelos acidentes de trânsito no Brasil pode ser explicado pela vulnerabilidade destes usuários em ambientes que priorizam os modos de transporte motorizados, criando assim, condições inseguras para os deslocamentos realizados a pé.

É importante destacar que as vias expressas representam um ambiente de maior risco aos pedestres, principalmente devido aos limites de velocidade superiores aos adotados nas vias locais. A OMS estima que a probabilidade de falecimento dos pedestres em decorrência de colisões com veículos conduzidos à velocidade de 50 km/h, é de aproximadamente 80% [15]. Nas rodovias federais brasileiras, onde o limite de velocidade para automóveis é de 110 km/h, os atropelamentos de pedestres têm atingido números expressivos de lesões e óbitos entre 2007 e 2019. Nesse período, foram registradas 54.324 colisões entre veículos e pedestres, segundo as bases de dados oficiais apresentadas pela Polícia Rodoviária Federal [16]. Ao todo, dos 54.407 pedestres que tiveram seu estado físico afetado após os atropelamentos, 71,9% apresentaram lesões graves ou leves, enquanto 26,9% vieram a óbito e apenas 1,2% saíram ilesos. Nota-se ainda que 70,6% dos acidentes foram registrados em trechos urbanos, onde a maior movimentação de veículos e pedestres eleva os riscos de ocorrência de atropelamentos [11].

Com o objetivo de fomentar o desenvolvimento de políticas de prevenção ao atropelamento de pedestres, Milton et al. (2008) [17] destacam que as agências de transporte têm priorizado cada vez mais o entendimento das relações existentes entre as variáveis causadoras dos acidentes, e o estado físico das vítimas. Desta maneira, pesquisadores de todo o mundo tem adotado a técnica de regressão logística para a modelagem dos acidentes de trânsito segundo a sua morbidade [18-21]. Este método tem apresentado resultados satisfatórios quando aplicado aos dados de acidentes rodoviários, possibilitando a construção de modelos matemáticos a partir de variáveis categóricas e da função de ligação *logit*. Tradicionalmente, adota-se o modelo de regressão logística simples quando a variável resposta é do tipo dicotômica, ou seja, apresenta duas categorias possíveis. Já o modelo logístico multinomial deve ser utilizado quando a variável resposta apresenta mais de duas categorias [22].

Neste sentido, diversos estudos têm adotado a regressão logística para a identificação dos fatores estatisticamente significativos associados a este tipo de acidente, bem como para a análise das contribuições destes fatores sobre a severidade das lesões sofridas pelos pedestres. Como exemplo, Tay et al. (2011) [23] empregaram a regressão logística multinomial na análise dos elementos

causadores de acidentes envolvendo pedestres na Coreia do Sul em 2006, considerando três níveis de severidade: ferimentos leves, ferimentos graves e fatalidades. O estudo concluiu que as fatalidades e ferimentos graves estão mais suscetíveis a ocorrência quando os condutores são do sexo masculino, apresentam idade superior a 65 anos, viajam durante a tarde/noite, estão sob efeito de bebidas alcoólicas ou conduzem veículos pesados. Já Chen e Fan (2019) [19] analisaram os atropelamentos de pedestres registrados no estado da Carolina do Norte, EUA, entre 2005 e 2012. Neste estudo, o estado físico dos pedestres foi classificado em 5 diferentes níveis: sem lesão, lesão possível, lesão evidente, invalidez e fatalidade. Os autores identificaram que os casos de invalidez e fatalidade apresentam maiores chances de ocorrência nas colisões com motocicletas ou ônibus, ao amanhecer ou ao anoitecer, com o pavimento molhado pelas chuvas e quando os pedestres possuem 26 anos ou mais. Por fim, Park e Ko (2020) [20] adotaram a regressão logística simples para a modelagem do estado físico dos pedestres envolvidos em atropelamentos na Coreia do Sul entre 2015 e 2017. O estudo concluiu que os acidentes de maior severidade tendem a envolver a presença de veículos de grande porte, próximo às escolas e no período da noite.

Assim, o objetivo deste estudo é analisar os fatores determinantes da severidade das lesões sofridas por pedestres nas rodovias federais brasileiras entre os anos de 2017 e 2019. Esta análise é realizada a partir das seguintes etapas: construção de um modelo de regressão logística multinomial, relacionando o estado físico dos pedestres às variáveis presentes na base de dados, como a fase do dia, o tipo de veículo, o sexo do pedestre entre outras; verificação da significância estatística de cada variável sobre o estado físico dos pedestres; análise da razão de chances de ocorrência de cada estado físico com base nos valores das variáveis explicativas.

Espera-se que este trabalho contribua com a atividade de gerenciamento de segurança rodoviária realizado pelas concessionárias de vias e agentes de segurança pública. Os resultados apresentados neste estudo buscam apoiar os órgãos competentes na adoção de medidas preventivas que reduzam tanto a quantidade quanto a severidade dos atropelamentos de pedestres nas rodovias federais, como campanhas socioeducativas e intervenções físicas sobre a infraestrutura rodoviária. Tais iniciativas são fundamentais para promoção da segurança e do bem-estar de motoristas e transeuntes, bem como para a economia das administradoras com o eventual pagamento de indenizações às vítimas.

## 2. MATERIAL E MÉTODOS

O fluxograma da Figura 1 apresenta os procedimentos adotados neste trabalho, desde a caracterização da área de estudo até a compreensão dos fatores associados ao atropelamento de pedestres nas rodovias federais. Todas as análises foram desenvolvidas na linguagem R [24].

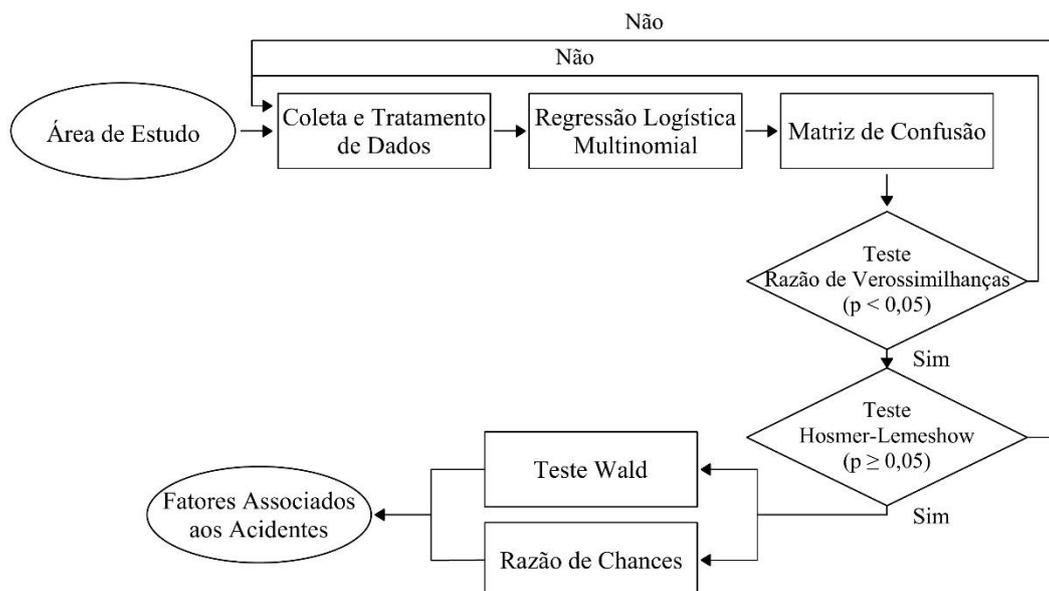


Figura 1: Fluxograma do estudo.

## 2.1 Área de estudo

A área de estudo considerada neste trabalho abrange o território brasileiro e toda a sua malha rodoviária federal. De acordo com pesquisas realizadas pela Confederação Nacional de Transportes (CNT) [25], o Brasil conta com 65.370 km de rodovias federais pavimentadas, dos quais 6.932 km (10,6%) apresentam pista dupla, 57.275 km (87,6%) possuem pista simples, e 1.163 (1,8%) passam por obras de duplicação. Já a frota brasileira atingiu à marca de 100.746.553 veículos em 2018, considerando veículos de todas as categorias [26].

Dados apresentados pelo DENATRAN (2020) [27] revelam que o Brasil apresentava 74.268.555 condutores habilitados até mês de maio de 2020. A distribuição espacial dos condutores revela que a região sudeste concentra o maior número de habilitados, com 37.792.667 motoristas registrados, o equivalente a 50,9% do total apresentado pelo país. São Paulo lidera a estatística por estados, com 23.071.445 condutores (31,1%), seguido por Minas Gerais com 7.388.662 (9,9%) e Rio de Janeiro com 5.809.312 (7,8%).

## 2.2 Coleta e tratamento de dados

Nas rodovias federais brasileiras, a Polícia Rodoviária Federal (PRF) é oficialmente encarregada pela coleta de dados referentes aos acidentes de trânsito. Essa coleta é conduzida a partir de dois meios distintos: 1) Declaração Eletrônica de Acidente de Trânsito (e-DAT): preenchida pelas vítimas envolvidas no acidente, nos casos em que o agente da PRF identifique a inexistência de condutores sob efeito de substâncias psicoativas, ou o acidente não resulte em vítimas feridas ou danos à sociedade. O preenchimento do e-DAT é realizado via internet, através do site oficial da PRF; 2) Boletins de Acidentes de Trânsito (BAT): registrados pelos agentes da PRF nos casos envolvendo vítimas feridas, condutor sob efeito de substâncias psicoativas ou ocorrência de danos à sociedade [28]. Esses documentos registram as variáveis associadas a cada acidente, como a data da ocorrência, condição meteorológica, a fase do dia na qual se deu a ocorrência, o tipo de uso do solo (rural ou urbano), o tipo de pista, o número de pessoas envolvidas, seus sexos e idades, seu estado físico, suas classificações (condutor, passageiro e pedestre), os tipos de veículo envolvidos, entre outros. A classificação do estado físico das vítimas é realizada pelos agentes da PRF, no local de ocorrência do acidente, conforme apresentado pela Tabela 1.

Tabela 1: Classificação do estado físico das vítimas pedestres nas rodovias federais brasileiras.

Estado Físico	Descrição
<b>Ileso</b>	A vítima não apresenta sinais ou sintomas de lesões resultantes do acidente, independentemente de ter sido encaminhada ou não ao atendimento hospitalar.
<b>Lesões Leves</b>	A vítima apresenta uma das seguintes lesões: 1) Dores em área sem órgãos vitais; 2) Contusões, abrasões e lacerações leves; 3) Queimaduras de 1º grau atingindo no máximo 10% da superfície corporal; 4) Fratura dos dentes; 5) Dores na parede torácica, abrasão por cinto de segurança ou por outras partes do veículo; 6) Pequenas hemorragias externas; 7) Pequenas entorses e luxações; 8) Dores de cabeça ou tontura, sem perda de consciência; 9) Dores de pescoço; 10) Contusão e abrasão dos olhos.
<b>Lesões Graves</b>	A vítima não apresenta sinais de morte, porém suas lesões não são classificáveis como leves.
<b>Morto</b>	A vítima apresenta sinais evidentes de morte identificados no local, como decapitação, esmagamento do corpo, carbonização e presença de <i>livor mortis</i> ou <i>vigor mortis</i> . Caso a vítima venha a óbito durante o transporte para o atendimento médico, seu estado físico é classificado como lesões graves.

Fonte: Adaptado de PRF (2015) [28].

Estas informações alimentam as bases de dados eletrônicas da PRF, que são atualizadas anualmente, e então disponibilizadas no site oficial do órgão [16]. Desta maneira, foram obtidas as bases de dados referentes ao período entre 2017 e 2019, nas quais cada registro representa os indivíduos envolvidos nos acidentes relatados. Ao todo, a PRF identificou 104.526 pessoas envolvidas em acidentes de trânsito nesse período, considerando tanto os pedestres quanto os condutores de veículos.

Na sequência foi conduzido o tratamento de dados, com o objetivo de prepará-los para a construção do modelo logístico. Esta etapa foi iniciada com a realização da amostragem, através da qual apenas os registros relativos aos pedestres acidentados foram mantidos na base de dados. Assim, foi identificada a presença de 10.518 pedestres na base de dados, representando 10,1% dos registros. Posteriormente, foi analisada a qualidade de preenchimento dos dados, buscando a identificação de variáveis eventualmente preenchidas com categorias inválidas, como “não informado”, “não identificado”, “ignorado” e “inválido”. Nestes casos, os valores inválidos foram substituídos por “missings” (representados pelo valor NA), disponibilizado pela linguagem R e cuja utilização indica a ocorrência de informação faltante. O mesmo expediente foi adotado para os registros nos quais os pedestres apresentaram idades superiores a 100 anos, seguindo procedimento similar ao utilizado por Barroso et al. (2019) [18].

Em seguida foram selecionadas 14 potenciais variáveis explicativas para a composição do modelo de regressão logística multinomial: dia da semana (domingo a sábado), mês (Janeiro a Dezembro), ano (variável quantitativa discreta, com valores entre 2017 e 2019); fase do dia (Amanhecer, Pleno dia, Anoitecer, Plena Noite), tipo de pista (Dupla, simples ou múltipla), traçado da via, uso do solo (Urbano, Rural), tipo de veículo, estado físico (Ileso, Ferido leve, Ferido Grave, Morto), unidade de federação, condição meteorológica, ano de fabricação do veículo (variável quantitativa discreta), idade (variável quantitativa discreta) e sexo (feminino e masculino).

A construção de modelos de regressão logística demanda que todas as variáveis adotadas estejam completamente preenchidas, sem qualquer ocorrência de informações faltantes. Por este motivo, Milton et al. (2008) [17] recomendam a exclusão de todos os registros que apresentem variáveis com dados faltantes, como uma forma de viabilizar o prosseguimento das análises. Entretanto, dada a elevada ocorrência de variáveis da base de dados com entradas inválidas, fez-se necessária a verificação preliminar do percentual de informações faltantes entre as variáveis selecionadas. Assim, variáveis que apresentaram mais de 2% de valores faltantes (Figura 2), como a “condição meteorológica”, “traçado da via”, “ano de fabricação do veículo” e “idade”, foram excluídas da análise. Esta prática foi adotada com o objetivo de limitar o número de registros excluídos, e assim, garantir uma quantidade de dados robusta para a elaboração do modelo. Após a exclusão desses registros, constata-se que 10.348 casos foram mantidos na base de dados, correspondendo a 98,4% da amostra inicial.

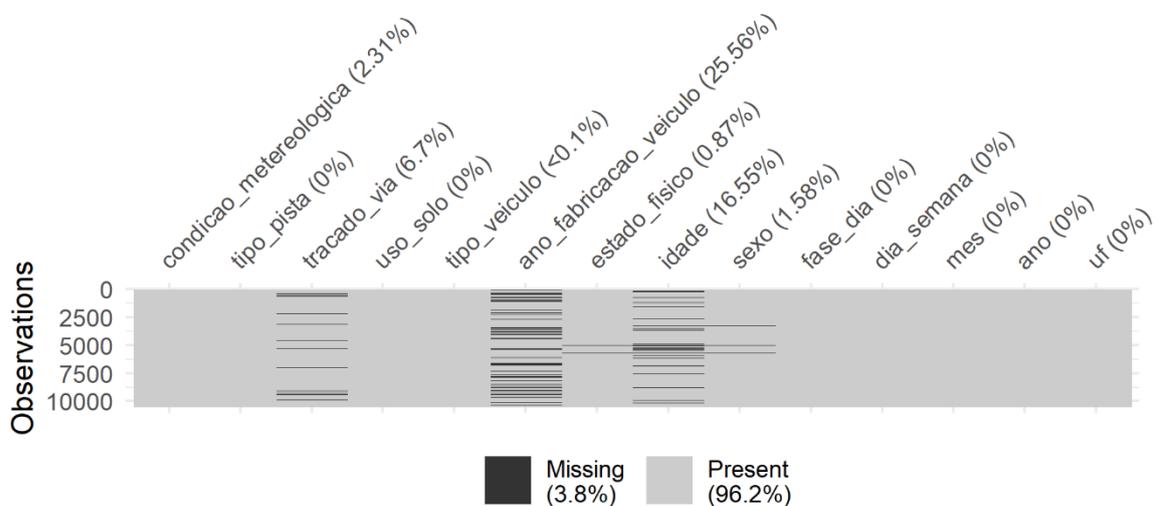


Figura 2: Padrão de dados faltantes nos registros de pedestres acidentados.

Por fim, foram realizadas simplificações sobre as variáveis “unidade de federação”, “mês”, “dia da semana”, “tipo de veículo” e “estado físico”, a fim de reduzir o número de categorias apresentadas por estas. Desta maneira, a variável “unidade de federação”, que apresentava 27 categorias relativas aos estados brasileiros, foi convertida na variável “região”, que possui 5 categorias em referência às macrorregiões brasileiras (norte, nordeste, sul, sudeste e centro-oeste). A variável “mês”, que possuía 12 categorias, deu origem à variável “estação do ano” com 4 categorias (primavera, verão, outono e inverno). A variável “dia da semana”, representada por 7 categorias, foi transformada na variável “tipo de dia”, que apresenta 2 categorias (dia útil e fim de semana). A variável “tipo de veículo”, que inicialmente apresentava 17 categorias, foi reduzida para 6 categorias principais: automóvel; utilitário (referente às categorias utilitário, caminhonete e camioneta); motocicleta (referente às categorias motocicleta e motoneta); caminhão (referente às categorias caminhão, caminhão-trator e semi-reboque); ônibus (referente às categorias ônibus e micro-ônibus); e outros (referente às categorias “outros”, “bicicleta”, “ciclomotor”, “reboque”, “trator de rodas” e “trator misto”). Ademais, não foram realizadas simplificações sobre a variável quantitativa discreta “ano”, e, desta maneira, ela apresentou três valores possíveis (2017, 2018 e 2019).

Já na variável “estado físico”, que inicialmente contava com 4 categorias, nota-se que a frequência relativa da categoria “ileso” (1,4%) foi consideravelmente inferior à observada para as demais categorias (lesões leves: 35,7%; lesões graves: 37,0%; morto: 25,9%). Segundo Elvik e Mysen (1999) [29] e Hakkert e Hauer (1988) [30], os acidentes com menor severidade tendem a apresentar maiores níveis de subnotificação, o que justifica a baixa frequência de pedestres ileso observados na amostra. Este resultado aponta que a categoria “ileso” pode ser classificada como um evento raro, tornando o modelo de regressão logística mais propenso a subestimar a probabilidade de sua ocorrência [31]. Ademais, King e Zeng (2001) [31] apontam que a existência de eventos raros na amostra tende a produzir estimativas enviesadas para os coeficientes do modelo de regressão logística. Como forma de solucionar estes problemas, os registros de pedestres ileso foram excluídos da análise, resultando em uma amostra final com 10.205 casos.

### 2.3 Regressão logística multinomial

A regressão logística multinomial foi utilizada para o ajuste do modelo matemático que descreve o estado físico dos pedestres a partir das variáveis explicativas da base de dados. Desta maneira, a variável resposta apresentou três categorias: 1) Lesões Leves; 2) Lesões Graves; e 3) Morto. Os acidentes podem ser enquadrados em apenas uma dessas categorias, sendo a categoria “Lesões Leves” adotada como referência. O objetivo desta análise foi identificar a influência das variáveis explicativas sobre o estado físico dos pedestres vitimados por atropelamento nas rodovias federais. O modelo de regressão multinomial é expresso pela função *logit*, apresentada na equação 1 [22].

$$g_k(x) = \ln \left[ \frac{\Pr(k|x)}{\Pr(0|x)} \right] = \beta_{k0} + \beta_{k1}x_1 + \dots + \beta_{kp}x_p \quad (1)$$

$$\forall (k, i)/k \in \{0, 1, 2, \dots, K\} \cap i \in \{0, 1, 2, \dots, p\}$$

Onde: k: estado físico do pedestre;

K: número de categorias da variável “estado físico”;

p: número total de graus de liberdade das variáveis explicativas;

x: vetor que contém os valores das variáveis explicativas;

Pr(k|x): Probabilidade condicional do estado físico k dado o vetor x;

Pr(0|x): Probabilidade condicional do estado de físico referência dado o vetor x;

$\beta_{ki}$ : Parâmetro do modelo do estado físico k, referente à variável explicativa i.

Inicialmente, a base de dados foi dividida nas amostras de treinamento e teste, seguindo a proporção de 3:1. Dessa maneira, 75% dos dados foram direcionados para a elaboração do modelo de regressão logística multinomial, enquanto 25% foram reservados para a testagem de sua acurácia. A construção e a avaliação do modelo foram realizadas por meio de seis etapas: 1) Método *Stepwise*; 2) Matriz de Confusão 3) Teste de Razão de Verossimilhança; 4) Teste de Hosmer-

Lemeshow; 5) Teste Wald; 6) Razão de Chances. O nível de significância adotado para os testes estatísticos é igual a 5%.

Durante a construção do modelo, a seleção das variáveis foi realizada por meio do método *stepwise*. Esse método consiste na inclusão de variáveis no modelo por etapas. A cada etapa, o ajuste do modelo foi avaliado por meio do Critério de Informação de Akaike (AIC). Segundo Alvarenga (2015) [32], os modelos de regressão logística apresentam um melhor ajuste quanto menor o seu valor de AIC. Deste modo, o critério de parada do *stepwise* foi satisfeito quando o modelo selecionado apresentou o valor mínimo de AIC em comparação com os modelos estimados em etapas anteriores [33, 34]. Neste contexto, a adição de uma nova variável aumentaria o AIC do modelo resultante, reduzindo assim a qualidade do ajuste. A equação 2 apresenta a fórmula de cálculo do AIC:

$$AIC = -2[LL(\beta_U) - p] \quad (2)$$

Onde:  $LL(\beta_U)$ : Logaritmo de verossimilhança do modelo estimado;  
 p: Número de parâmetros do modelo.

Após a construção do modelo, foram utilizados os dados da amostra de testes para a elaboração de uma matriz de confusão 3x3. Por meio dessa matriz, foi realizada a comparação entre os valores reais de estado físico observados em cada registro da amostra de testes, e os valores de estado físico preditos pelo modelo de regressão. Em seguida, foi aferida a acurácia dos resultados, que consiste na relação entre o número de vezes em que o modelo realizou predições corretas a respeito do estado físico do pedestre, e o número total de predições realizadas. O cálculo da acurácia é realizado por meio da equação 3 [35].

$$Acurácia = \frac{a_{11} + a_{22} + a_{33}}{\sum_{i=1}^{k=3} \sum_{j=1}^{k=3} a_{ij}} \quad (3)$$

Onde:  $a_{11}$ : número de predições corretas da categoria “Ileso/Lesões Leves”;  
 $a_{22}$ : número de predições corretas da categoria “Lesões Graves”;  
 $a_{33}$ : número de predições corretas da categoria “Morto”;  
 $a_{ij}$ : número de predições na linha i e coluna j.

Na sequência, os testes de Razão de Verossimilhanças e Hosmer-Lemeshow foram adotados com o objetivo de aferir a qualidade do modelo. Nesse sentido, o teste da Razão de Verossimilhanças verificou a significância estatística do conjunto de variáveis previamente selecionado, através da comparação entre o logaritmo de verossimilhança do modelo proposto pelo método *stepwise*, e o logaritmo de verossimilhança do modelo nulo. A hipótese nula desse teste preconiza que as variáveis selecionadas não apresentam significância estatística, ou seja, elas somente podem ser aceitas caso o teste apresente um p-valor inferior a 0,05 [22, 36]. Já o teste de Hosmer-Lemeshow foi utilizado para a avaliação da qualidade do ajuste do modelo em relação aos dados da amostra. A hipótese nula desse teste preconiza que o modelo se ajusta bem os dados da amostra [22, 37]. Dessa maneira, caso o teste apresente um p-valor maior ou igual que 0,05, o modelo apresenta um ajuste satisfatório. As estatísticas adotadas nos testes de Razão de Verossimilhanças e de Hosmer-Lemeshow são apresentadas respectivamente pelas equações 4 e 5.

$$G = -2 \times [LL(\beta_R) - LL(\beta_U)] \quad (4)$$

Onde:  $LL(\beta_R)$ : Logaritmo de verossimilhança do modelo nulo;  
 $LL(\beta_U)$ : Logaritmo de verossimilhança do modelo composto pelas variáveis selecionadas.

$$\chi^2 = \sum_{j=1}^J \sum_{k=0}^{K-1} \frac{(O_{jk} - E_{jk})^2}{E_{jk}} \quad (5)$$

Onde: J: Número de grupos no qual a amostra é dividida;  
 $O_{jk}$ : Frequência observada para a severidade k no grupo j;  
 $E_{jk}$ : Frequência esperada para a severidade k no grupo j.

Em seguida, o Teste Wald foi realizado com o objetivo de identificar tanto os parâmetros  $\beta_{ki}$ , quanto a significância estatística das variáveis do modelo. A hipótese nula do teste estabelece que o parâmetro  $\beta_{ki}$  é igual a 0. Dessa maneira, uma variável é considerada estatisticamente significativa caso o p-valor seja menor que 0,05. A estatística do Teste Wald é apresentada pela equação 6.

$$W_{ki} = \frac{\hat{\beta}_{ki}}{SE(\hat{\beta}_{ki})} \quad (6)$$

Onde:  $\hat{\beta}_{ki}$ : Parâmetro estimado referente ao estado físico k e à variável explicativa i;  
 $SE(\hat{\beta}_{ki})$ : Erro padrão de  $\beta_{ki}$ .

Por fim, foi realizada a análise da Razão de Chances, com o objetivo de relacionar as chances do pedestre apresentar um determinado estado físico k, dados os valores apresentados pelas variáveis explicativas do vetor x. Esse procedimento foi adotado individualmente para todas as classes das variáveis explicativas que apresentam significância estatística, a fim de identificar as influências de cada classe sobre a variável resposta [22]. A Razão de Chances e seu intervalo de confiança de 95% foram calculados por meio das equações 7 e 8 respectivamente.

$$OR_k(i, 0) = \frac{\Pr(k | x = i) / \Pr(0 | x = i)}{\Pr(k | x = 0) / \Pr(0 | x = 0)} = \exp(\hat{\beta}_{ki}) \quad (7)$$

$$IC95\% = \exp[\hat{\beta}_{ki} \pm 1,96 SE(\hat{\beta}_{ki})] \quad (8)$$

Onde:  $OR_k(i,0)$ : Razão de chances do estado físico k dada a ocorrência da classe i para a variável explicativa;  
k: estado físico do pedestre;  
x = i: classe analisada para a variável explicativa;  
x = 0: classe de referência para a variável explicativa.

### 3. RESULTADOS

O método *stepwise* resultou na construção de um modelo de regressão logística multinomial composto pelas variáveis “tipo de veículo”, “fase do dia”, “uso do solo”, “sexo”, “tipo de pista”, “tipo de dia”, “região” e “ano”, descartando assim a utilização da variável “estação do ano”. As etapas adotadas durante a execução do método, bem como seus respectivos valores de AIC são apresentados pela tabela 1. Já a matriz de confusão, ilustrada pela tabela 2, apontou uma acurácia de 48% para os resultados preditos pelo modelo, com 52,8% de acertos nas ocorrências com lesões leves, 37,7% de acertos nos casos de lesões graves, e 56,7% de acertos nos casos de morte. Em seguida, o modelo foi validado pelos testes da Razão de Verossimilhanças (p-valor <0,001) e de Hosmer-Lemeshow (p-valor = 0,252). Estes resultados não apenas apontam que o conjunto de variáveis selecionadas é adequado para a descrição do estado físico dos pedestres, como também indica que o modelo resultante do método *stepwise* apresenta qualidade de ajuste satisfatória sobre os dados da amostra.

Tabela 1: Etapas do método stepwise para a construção do modelo a partir da amostra de treinamento.

Etapa	Modelo	AIC
1	estado_fisico ~ 1	16654,92
2	estado_fisico ~ tipo_veiculo	15817,56
3	estado_fisico ~ tipo_veiculo + fase_dia	15492,77
4	estado_fisico ~ tipo_veiculo + fase_dia + uso_solo	15364,12
5	estado_fisico ~ tipo_veiculo + fase_dia + uso_solo + regioao	15310,47
6	estado_fisico ~ tipo_veiculo + fase_dia + uso_solo + regioao + sexo	15275,15
7	estado_fisico ~ tipo_veiculo + fase_dia + uso_solo + regioao + sexo + tipo_pista	15255,65
8	estado_fisico ~ tipo_veiculo + fase_dia + uso_solo + regioao + sexo + tipo_pista + tipo_dia	15249,79
9	estado_fisico ~ tipo_veiculo + fase_dia + uso_solo + regioao + sexo + tipo_pista + tipo_dia + ano	15245,23
10	estado_fisico ~ tipo_veiculo + fase_dia + uso_solo + regioao + sexo + tipo_pista + tipo_dia + ano + estacao	15253,23

Tabela 2: Matriz de confusão para a análise de acurácia do modelo.

Valores Reais (Amostra de Teste)	Valores Preditos (Modelo)		
	Lesões Leves	Lesões Graves	Morto
Lesões Leves	492	283	156
Lesões Graves	387	366	219
Morto	86	195	368

A aplicação do teste de Wald revelou que as lesões graves não estão associadas a fatores como veículos utilitários ( $p$ -valor = 0,608), pista simples ( $p$ -valor = 0,902), região norte ( $p$ -valor = 0,740) e região sul ( $p$ -valor = 0,232). Considerando os atropelamentos ocasionados por caminhões, as chances de os pedestres apresentarem lesões graves são 72,9% superiores em relação aos casos envolvendo automóveis. Já os veículos enquadrados na categoria “outros” apresentaram 37,9% mais chances de provocarem lesões graves aos pedestres. Por outro lado, nos acidentes gerados por ônibus, as chances de lesões graves caem em 6,9%, enquanto as motocicletas apresentaram 13,7% menos chances. Os períodos do amanhecer, anoitecer e plena noite apresentaram maiores chances de ocorrência de lesões graves quando comparados ao período de pleno dia, com 13%, 38,6% e 29,8% mais chances respectivamente. As chances de lesões graves em trechos urbanos foram 10,8% inferiores às obtidas nas zonas rurais. Pedestres do sexo masculino apresentaram 34,1% mais chances de ferimentos graves em relação aos pedestres do sexo feminino. A análise por regiões indicou maiores chances de lesão grave na região nordeste (32% a mais) quando comparada à região sudeste. Já os pedestres da região centro-oeste apresentaram menores chances de lesões graves em relação aos transeuntes da região sudeste (18,9% a menos). Ademais, foi verificado que os pedestres possuem 18,5% mais chances de sofrerem lesões graves nos trechos de pista múltipla em comparação com os trechos de reta, bem como apresentam 15,1% mais chances de lesões graves nos fins de semana em comparação com os dias úteis. Por fim, os dados revelam um crescimento anual de 9,9% nas chances de lesão grave. A tabela 3 apresenta os resultados do Teste Wald e da Razão de Chances para os pedestres diagnosticados com lesões graves.

Tabela 3: Modelo de regressão logística multinomial para a variável resposta igual a “Lesão Grave”.

Variável Explicativa	$\beta$	Razão Chances (OR)	IC95% (OR)	P-valor
(Intercept)	-190,969			<0,001
<b>Tipo de Veículo (ref: Automóvel)</b>				
Caminhão	0,548	1,729	(1,603 ; 1,865)	<0,001
Motocicleta	-0,147	0,863	(0,776 ; 0,96)	0,007
Ônibus	-0,071	0,931	(0,883 ; 0,982)	0,009
Outros	0,321	1,379	(1,252 ; 1,519)	<0,001
Utilitário	0,025	1,025	(0,932 ; 1,129)	0,608
<b>Fase do Dia (ref: Pleno dia)</b>				
Amanhecer	0,122	1,130	(1,112 ; 1,149)	<0,001
Anoitecer	0,326	1,386	(1,353 ; 1,419)	<0,001
Plena Noite	0,261	1,298	(1,171 ; 1,439)	<0,001
<b>Uso do Solo (ref: Rural)</b>				
Urbano	-0,114	0,892	(0,805 ; 0,99)	0,031
<b>Sexo (ref: Feminino)</b>				
Masculino	0,293	1,341	(1,213 ; 1,483)	<0,001
<b>Região (ref: Sudeste)</b>				
Sul	0,060	1,062	(0,962 ; 1,171)	0,232
Nordeste	0,278	1,320	(1,2 ; 1,451)	<0,001
Norte	0,013	1,013	(0,938 ; 1,095)	0,740
Centro-Oeste	-0,209	0,811	(0,733 ; 0,898)	<0,001
<b>Tipo de Pista (ref: Dupla)</b>				
Múltipla	0,170	1,185	(1,05 ; 1,339)	0,006
Simple	0,006	1,006	(0,909 ; 1,115)	0,902
<b>Tipo de Dia (ref: Dia Útil)</b>				
Fim de semana	0,141	1,151	(1,026 ; 1,291)	0,017
<b>Ano</b>	0,094	1,099	(1,099 ; 1,099)	<0,001

Por fim, o Teste Wald revelou que, entre todas as variáveis do modelo, apenas a região sul não está associada à ocorrência de mortes de pedestres por atropelamento ( $p$ -valor = 0,066). A tabela 4 aponta que os pedestres atropelados pelos ônibus apresentaram 52% mais chances de falecerem em relação aos que se envolveram em colisões com automóveis. As chances de falecimento atingem níveis ainda superiores nos casos envolvendo veículos da categoria utilitário, com 54,3% mais chances de falecimento em comparação com os acidentes ocasionados por automóveis. Entretanto, as maiores chances de óbito foram registradas nos atropelamentos causados por caminhões e por veículos enquadrados na categoria “outros”, com respectivamente 264,1% e 163,2% mais chances de falecimento. Já os acidentes causados por motocicletas apresentaram 79,6% menos chances de resultarem em mortes aos pedestres. O período do anoitecer elevou as chances de ocorrência de mortes entre os pedestres em 87,2% quando comparado ao período do pleno dia. Este aumento foi ainda superior nos períodos do amanhecer (240,5%) e da plena noite (227,0%). Os trechos urbanos representaram menor risco aos pedestres, com 53,4% menos chances de mortes em comparação com as áreas rurais. A análise por regiões indicou que as chances de morte de pedestres na região norte são 23,7% superiores em relação a região sudeste, enquanto nas regiões centro-oeste e nordeste, as chances são 26,8% e 89% superiores respectivamente. Pedestres do sexo masculino tiveram 51% mais chances de falecerem em relação aos do sexo feminino. Acidentes ocorridos em pista simples reduziram as chances de falecimento em 24,7% em comparação com os casos registrados em pista dupla, enquanto os acidentes em pista múltipla elevaram as chances de morte

em 14,9%. Já os fins de semana elevaram as chances de falecimento de pedestres em 22,4% em relação aos dias úteis. Os resultados ainda apontaram um aumento anual de 3,3% nas chances de mortes entre pedestres.

Tabela 4: Modelo de regressão logística multinomial para a variável resposta igual a “Morto”.

Variável Explicativa	$\beta$	Razão Chances (OR)	IC95% (OR)	P-valor
(Intercept)	-66,809			<0,001
<b>Tipo de Veículo (ref: Automóvel)</b>				
Caminhão	1,292	3,641	(3,364 ; 3,94)	<0,001
Motocicleta	-1,588	0,204	(0,196 ; 0,213)	<0,001
Ônibus	0,419	1,520	(1,446 ; 1,597)	<0,001
Outros	0,968	2,632	(2,366 ; 2,927)	<0,001
Utilitário	0,434	1,543	(1,413 ; 1,686)	<0,001
<b>Fase do Dia (ref: Pleno dia)</b>				
Amanhecer	1,225	3,405	(3,311 ; 3,501)	<0,001
Anoitecer	0,627	1,872	(1,832 ; 1,914)	<0,001
Plena Noite	1,185	3,270	(2,896 ; 3,692)	<0,001
<b>Uso do Solo (ref: Rural)</b>				
Urbano	-0,765	0,466	(0,418 ; 0,519)	<0,001
<b>Sexo (ref: Feminino)</b>				
Masculino	0,412	1,510	(1,361 ; 1,676)	<0,001
<b>Região (ref: Sudeste)</b>				
Sul	0,095	1,100	(0,994 ; 1,218)	0,066
Nordeste	0,637	1,890	(1,721 ; 2,076)	<0,001
Norte	0,213	1,237	(1,18 ; 1,297)	<0,001
Centro-Oeste	0,237	1,268	(1,174 ; 1,368)	<0,001
<b>Tipo de Pista (ref: Dupla)</b>				
Múltipla	0,139	1,149	(1,059 ; 1,247)	0,001
Simplex	-0,284	0,753	(0,672 ; 0,844)	<0,001
<b>Tipo de Dia (ref: Dia Útil)</b>				
Fim de semana	0,202	1,224	(1,073 ; 1,397)	<0,001
<b>Ano</b>	0,033	1,033	(1,033 ; 1,033)	<0,001

#### 4. DISCUSSÃO E CONCLUSÃO

Este estudo apresentou uma análise sobre os fatores determinantes da severidade das lesões sofridas por pedestres atropelados nas rodovias federais brasileiras entre os anos de 2017 e 2019. Os boletins de acidentes de trânsito da PRF apontaram que 10.518 pessoas foram vítimas de atropelamento durante esse período, das quais 10.205 compuseram a base de dados utilizada para a realização das análises. A partir da amostra de treinamento (75% dos dados), foi construído um modelo de regressão logística multinomial, revelando que a severidade dos danos físicos sofridos pelos pedestres está diretamente relacionada a fatores como: o tipo de veículo que provocou o acidente, a fase do dia, o uso do solo, a região onde a ocorrência foi registrada, o sexo do pedestre, o tipo de pista, o tipo de dia (dia útil ou final de semana) e o ano da ocorrência. De acordo com a matriz de confusão elaborada através da amostra de teste (25% dos dados), o modelo apresentou uma acurácia de 48%.

Analisando os resultados provenientes do modelo, observa-se que os atropelamentos registrados nas rodovias federais brasileiras têm apresentado um aumento anual significativo no risco de lesões

graves e mortes de pedestres, considerando o período entre 2017 e 2019. As chances de ocorrência de mortes entre pedestres das regiões norte, nordeste e centro-oeste foram superiores quando comparadas à região sudeste, enquanto a região nordeste ainda apresentou maior propensão ao registro de pedestres gravemente feridos. Este fato pode ser explicado pelo compartilhamento do espaço rodoviário por um grande número de pedestres e condutores dessas regiões. Como exemplo, a Secretaria de Infraestrutura da Bahia [38] identificou que os atropelamentos de pedestres foram a quarta maior causa de acidentes registrados na região metropolitana de Salvador, que conta com trechos das rodovias BR-324 e BR-420. O estudo ainda apontou que as viagens a pé representam a maior porção dos deslocamentos entre os baianos (35,3%) enquanto os ônibus e automóveis foram respectivamente o segundo e terceiro modos de transporte mais utilizados (45% no total). Esta realidade propicia a ocorrência de situações de risco aos transeuntes nas rodovias federais, devido às altas velocidades de tráfego praticadas pelos condutores, e à escassez de pontos de travessia que atendam as linhas de desejo dos pedestres.

Os resultados da análise ainda revelaram que veículos de grande porte, como caminhões, ônibus e utilitários, tendem a provocar fatalidades entre os pedestres. Além disso, os pedestres atropelados por caminhões estão mais propensos a sofrer lesões graves quando comparados aos que se acidentam com automóveis. Estas constatações estão alinhadas com o estudo conduzido por Rousdari et al. (2004) [39], que aponta os caminhões como fator de risco à integridade física dos pedestres, elevando suas chances de lesões graves em até 210% e de mortes em até 240%, quando atingidos por caminhões. Veículos classificados como “outros” também apresentaram chances elevadas de acidentes graves e fatalidades, o que pode ser explicado pela presença de veículos pesados nesta categoria, como o trator de rodas, o trator misto e o reboque. Já as colisões com motocicletas representaram menor risco aos pedestres, com razão de chances inferior a 1 para a ocorrência de lesões graves e mortes. Este resultado indica que as motocicletas tendem a causar menos lesões graves ou fatalidades aos pedestres do que os automóveis, o que também foi constatado por Tay et al. (2011) [23] e Wang et al. (2013) [40].

A análise da razão de chances para a variável fase do dia indicou que as lesões graves e fatalidades tendem a ocorrer nos períodos do amanhecer, anoitecer e plena noite. Resultados similares foram atingidos por Tay et al. (2011) [23] e Wang et al. (2013) [40], que apontam a variação da luminosidade natural ao longo do dia como um fator agravante dos danos físicos sofridos pelos pedestres. Esta situação é explicada pelo fato de que, na fase de pleno dia, os condutores possuem melhor visibilidade dos pontos de travessia, facilitando tanto a redução da velocidade quanto a realização de manobras de segurança que evitem a colisão com os pedestres. Assim, mesmo nos casos em que o acidente é inevitável, a ação dos condutores permite que eventuais impactos sejam atenuados. Por outro lado, as demais fases do dia apresentam iluminação natural reduzida, dificultando a percepção dos condutores para a existência de pedestres na pista, e desta maneira, aumentando as chances de acidentes graves ou fatais.

Quanto ao uso do solo, nota-se que os acidentes fatais e as lesões graves apresentaram maiores chances de ocorrência nas zonas rurais. Estes resultados são condizentes com as descobertas realizadas por Tay et al. (2011) [23] e Chen e Fan (2019) [19], que também observaram maiores chances de ocorrência de acidentes graves nas zonas rurais. O fato dos trechos urbanos contarem com a maior presença de sinalizações, como faixas de pedestres e placas de travessia de pedestres, bem como dispositivos redutores de velocidades, como faixas elevadas, lombadas e radares, favorece as condições de segurança dos pedestres durante a travessia, reduzindo assim as chances de ocorrência de ferimentos graves e mortes. Por outro lado, as zonas rurais favorecem a adoção de velocidades elevadas pelos condutores, aumentando os riscos de fatalidade para os pedestres, especialmente em localizações desprovidas de passarelas. Conseqüentemente, os perímetros urbanos apresentam 53,4% menos chances de morte aos pedestres quando comparado às zonas rurais.

Pedestres do sexo masculino apresentaram maiores chances de lesões graves e de óbito que as pedestres do sexo feminino. Conforme indicado por Ferenchak (2016) [41], o padrão comportamental observado entre os pedestres do sexo masculino é de maior risco em relação aos do sexo feminino. Em geral, os homens têm uma menor tolerância ao tempo de espera para a realização da travessia, além de utilizarem as faixas de pedestres em menor proporção que as mulheres [41]. Tal comportamento não apenas eleva o risco de ocorrência de acidentes entre os

homens nos trechos urbanos, como pode ocasionar em lesões graves e mortes com maior frequência.

As chances de ocorrência de ferimentos graves e fatalidades ainda foram maiores em rodovias de pista múltipla, enquanto os trechos de pista simples ofereceram menor risco de óbito aos pedestres. A existência de divisões centrais em pistas múltiplas e duplas, e as larguras de seção superiores destes tipos de pista são apontadas como possíveis causas para este fenômeno. Conforme apresentado em estudos anteriores, pistas de maior largura não apenas oferecem maiores distâncias de travessia aos pedestres como aumentam a sensação de segurança por parte dos condutores, tornando-os mais propensos a adoção de altas velocidades [21, 23]. Esta situação ainda se agrava entre os pedestres que naturalmente apresentam velocidades de deslocamento reduzidas, como idosos, crianças e pessoas com deficiência física. Ademais, Tulu et al. (2013) [42] apontam que, nos países em desenvolvimento, os pedestres tendem a pular as barreiras centrais existentes em pistas duplas e múltiplas devido à escassez de passarelas em seus locais de interesse de travessia. Este fenômeno torna os pedestres ainda mais suscetíveis a fatalidades ou lesões graves em decorrência de colisões com veículos em alta velocidade.

Por fim, os resultados indicam que os acidentes ocorridos nos finais de semana acarretaram em maiores riscos de lesões e óbitos aos pedestres. Estes resultados são consistentes com o estudo realizado por Alvarenga (2015) [32], que identificaram o maior risco de acidentes graves aos sábados e domingos. Adanu et al. (2018) [43] sugerem que o risco de lesões graves entre os condutores durante os finais de semana está associado a fatores humanos, como o cansaço durante a direção e a susceptibilidade às distrações. Embora os autores não mencionem o estado físico dos pedestres eventualmente envolvidos nestas ocorrências, a tendência comportamental apresentada pelos condutores durante os finais de semana é uma possível justificativa para o aumento das chances de mortes entre pedestres aos sábados e domingos.

Desta maneira, conclui-se que os fatores que representam o maior risco à integridade física dos pedestres nas rodovias federais brasileiras são: os veículos de grande porte, como caminhões, ônibus e utilitários; os períodos do dia que apresentem menor luminosidade natural, como o amanhecer, o anoitecer e a plena noite; os perímetros rurais; o sexo masculino; os trechos de pista múltipla; os dias de fim de semana; e as regiões norte, nordeste e centro-oeste. Ademais, o estudo ainda revelou que as chances de os pedestres serem diagnosticados com ferimentos graves nas rodovias federais vêm aumentando em 9,9% ao ano. Já as chances de fatalidade apresentaram saltos de 3,3% ao ano, indicando que as elevadas taxas de mortalidade observadas pela OPAS tendem a subir caso medidas de prevenção ao atropelamento de pedestres não sejam adotadas [13].

Assim, estes resultados sugerem a necessidade da implementação de intervenções físicas e operacionais de planejamento viário que protejam os pedestres de possíveis colisões com veículos motorizados. Entre as intervenções com maior potencial para a redução do número de casos envolvendo ferimentos graves ou fatalidades, listam-se: a construção de travessias elevadas que atendam às linhas de desejo dos pedestres nas zonas rurais; a instalação de luminárias em trechos onde há risco de atropelamentos entre o anoitecer e o amanhecer; e a adoção de sinalizações e dispositivos redutores de velocidade em áreas de grande movimentação de pessoas. Nota-se ainda a necessidade de realização de campanhas socioeducativas direcionadas aos condutores de veículos de grande porte, aos pedestres do sexo masculino e aos viajantes de finais de semana. Tais campanhas devem alertar os usuários: para os riscos do uso compartilhado das rodovias federais por pedestres e veículos automotores; para a vulnerabilidade do pedestre em relação aos veículos pesados; para a severidade das lesões sofridas pelos pedestres durante períodos de menor luminosidade natural; para que os pedestres somente cruzem as rodovias em locais de travessia segura; e para que os condutores redobrem a atenção durante os finais de semana e os períodos noturnos. Estas medidas devem ser adotadas prioritariamente nas regiões norte, nordeste e centro-oeste, dadas as maiores chances de mortalidade de pedestres nestas áreas.

Finalmente, sugere-se a realização de pesquisas futuras que priorizem a descoberta de pontos críticos de atropelamento de pedestres nessas regiões, a fim de fomentar a tomada de decisões dos administradores quanto à exata localização onde eventuais intervenções físicas devam ser implementadas. Essas análises podem ser realizadas através de mapas de kernel ou por meio das metodologias apresentadas pelo Programa Pare, do Ministério dos Transportes [44]. Além disso, aconselha-se que os novos estudos testem as demais técnicas de regressão logística, como as

regressões logísticas ordinal, binomial e mista, com o objetivo de compará-las à regressão logística multinomial utilizada no presente estudo. As diferenças apresentadas por técnicas distintas podem se manifestar em termos das acurácias dos modelos resultantes, das suas variáveis significativas, dos valores de seus coeficientes, e das razões de chances. Por fim, recomenda-se que os próximos estudos adotem técnicas de imputação de dados, com o objetivo de preencher as informações inválidas observadas na base de dados da PRF. Esta medida permitiria a utilização de mais variáveis no modelo, como a idade dos pedestres e a condição meteorológica no momento do acidente, possibilitando novas conclusões a respeito da influência desses fatores sobre o estado físico dos pedestres.

## 5. AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior -Brasil (CAPES) - Código de Financiamento 001.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

1. Associação Nacional de Transportes Públicos (ANTP). Sistema de Informações da Mobilidade Urbana da Associação Nacional de Transportes Público - Simob/ANTP: Relatório geral 2016. São Paulo (SP): ANTP; 2016. Disponível em: <http://files.antp.org.br/simob/simob-2016-v6.pdf>
2. Abbott RD, White LR, Ross GW, Masaki KH, Curb JD, Petrovitch H. Walking and dementia in physically capable elderly men. *Jama*. 2004 Sep;292(12):1447-1453, doi: 10.1001/jama.292.12.1447
3. Hamer M, Chida Y. Active commuting and cardiovascular risk: a meta-analytic review. *Preventive Medicine*. 2008 Jan;46(1):9-13, doi: 10.1016/j.ypmed.2007.03.006
4. Hanson S, Jones A, Petrovitch H. Is there evidence that walking groups have health benefits? A systematic review and meta-analysis. *British Journal of Sports Medicine*. 2015 Jan;49(11):710-715, doi: 10.1136/bjsports-2014-094157
5. Tainio M, Nazelle AJ, Götschi T, Kahlmeier S, Rojas-Rueda D, Nieuwenhuijsen MJ, Sá TH, Kelly P, Woodcock J. Can air pollution negate the health benefits of cycling and walking? *Preventive Medicine*. 2016 Jun;87(1):233-236, doi: 10.1016/j.ypmed.2016.02.002
6. Shih YH, Pai MC, Lin HS, Sung PS, Wang JJ. Effects of walking on sundown syndrome in community-dwelling people with Alzheimer's disease. *International Journal of Older People Nursing*. 2020 Jun;87(1):e12292, doi: 10.1111/opn.12292
7. Sallis JF, Cervero RB, Ascher W, Henderson KA, Kraft MK, Kerr J. An ecological approach to creating active living communities. *Annual Review of Public Health*. 2006;27(1):297-322, doi: 10.1146/annurev.publhealth.27.021405.102100
8. World Health Organization (WHO). Global recommendations on physical activity for health. Genebra (CH): WHO; 2011. Disponível em: [https://apps.who.int/iris/bitstream/handle/10665/44399/9789241599979\\_eng.pdf?sequence=1](https://apps.who.int/iris/bitstream/handle/10665/44399/9789241599979_eng.pdf?sequence=1)
9. Panter J, Heinen E, Mackett R, Ogilvie D. Impact of new transport infrastructure on walking, cycling, and physical activity. *American Journal of Preventive Medicine*. 2016 Feb;50(2):e45-e53, doi: 10.1016/j.amepre.2015.09.021
10. Moura F, Cambra P, Gonçalves AB. Measuring walkability for distinct pedestrian groups with a participatory assessment method: A case study in Lisbon. *Landscape and Urban Planning*. 2017 Jan;157(1):282-296, doi: 10.1016/j.landurbplan.2016.07.002
11. Zegeer CV, Bushell M. Pedestrian crash trends and potential countermeasures from around the world. *Accident Analysis & Prevention*. 2012 Jan;44(1):3-11, doi: 10.1016/j.aap.2010.12.007
12. World Health Organization (WHO). Global status report on road safety 2018. Genebra (CH): WHO; 2018. Disponível em: <https://www.who.int/publications/i/item/global-status-report-on-road-safety-2018>
13. Pan American Health Organization (PAHO). Status of Road Safety in the Region of the Americas. Washington (DC): PAHO; 2019. Disponível em: <https://iris.paho.org/handle/10665.2/51088>
14. Ladeira RM, Malta DC, Morais Neto OLD, Montenegro MDMS, Soares Filho AM, Vasconcelos, CH, Mooney M, Naghavi M. Acidentes de transporte terrestre: estudo Carga Global de Doenças, Brasil e unidades federadas, 1990 e 2015. *Revista Brasileira de Epidemiologia*. 2017 Mai;20(1):157-170, doi: 10.1590/1980-5497201700050013
15. World Health Organization (WHO). Control de la velocidad: un manual de seguridad vial. Genebra (CH): WHO; 2017. Disponível em: [https://montevideo.gub.uy/sites/default/files/control\\_de\\_velocidad\\_1.pdf](https://montevideo.gub.uy/sites/default/files/control_de_velocidad_1.pdf)

16. Polícia Rodoviária Federal (PRF). Dados abertos: Acidentes. 2019. Disponível em: <https://portal.prf.gov.br/dados-abertos-acidentes>
17. Milton JC, Shankar VN, Mannering FL. Highway accident severities and the mixed logit model: an exploratory empirical analysis. *Accident Analysis & Prevention*. 2008 Jan;40(1):260-266, doi: 10.1016/j.aap.2007.06.006
18. Barroso Junior GT, Bertho ACS, Veiga ADC. A letalidade dos acidentes de trânsito nas rodovias federais brasileiras em 2016. *Revista Brasileira de Estudos de População*. 2019;36(1):1-22, doi: 10.20947/S0102-3098a0074
19. Chen Z, Fan WD. A multinomial logit model of pedestrian-vehicle crash severity in North Carolina. *International Journal of Transportation Science and Technology*. 2019 Mar;8(1):43-52, doi: 10.1016/j.ijtst.2018.10.001
20. Park S, Ko D. A multilevel model approach for investigating individual accident characteristics and neighborhood environment characteristics affecting pedestrian-vehicle crashes. *International Journal of Environmental Research and Public Health*. 2020 Apr;17(1):3107, doi: 10.3390/ijerph17093107
21. Sanz NC, Guirao B, Attard M. Analysis of the risk factors affecting the severity of traffic accidents on Spanish crosstown roads: The driver's perspective. *Sustainability*. 2020 Mar;12(6):2237, doi: 10.3390/su12062237
22. Hosmer DW, Lemeshow S, Sturdivant RX. *Applied logistic regression*. 3. ed. Hoboken (NJ): Wiley & Sons, Inc.; 2013.
23. Tay R, Choi J, Kattan L, Khan A. A multinomial logit model of pedestrian-vehicle crash severity. *International Journal of Sustainable Transportation*. 2011 Mar;5(4):233-249, doi: 10.1080/15568318.2010.497547
24. R Core Team. *R: A language and environment for statistical computing*. Vienna (AT): R Foundation for Statistical Computing; 2020. Disponível em: <https://www.R-project.org/>
25. Confederação Nacional Do Transporte (CNT). Pesquisa CNT de rodovias 2019. Brasília (DF): CNT: SEST SENAT; 2019. Disponível em: <https://pesquisarodovias.cnt.org.br/downloads/ultimaversao/gerencial.pdf>
26. Instituto Brasileiro de Geografia e Estatística (IBGE). Frota de veículos 2018; 2018. Disponível em: <https://cidades.ibge.gov.br/brasil/pesquisa/22/28120?ano=2018>
27. Departamento Nacional de Trânsito (DENATRAN). Quantidade de Habilitados – DENATRAN: Maio – 2020; 2020. Disponível em: <https://www.gov.br/infraestrutura/pt-br/assuntos/transito/conteudo-denatran/estatisticas-quantidade-de-habilitados-denatran1790>
28. Polícia Rodoviária Federal (PRF). MP-015 Atendimento de acidentes. 2. ed. Brasília (DF): PRF; 2015.
29. Elvik R, Mysen AB. Incomplete accident reporting meta-analysis of studies made in 13 countries. *Transportation Research Record*. 1999 Jan;1665(1):133-140, doi: 10.3141/1665-18
30. Hauer E, Hakkert AS. The Extent and some implications of incomplete accident reporting. *Transportation Research Record*. 1988 Jan; 1185(1): 1-10. Disponível em: [https://www.researchgate.net/publication/243764350\\_The\\_Extent\\_and\\_Some\\_Implications\\_of\\_Incomplete\\_Accident\\_Reporting](https://www.researchgate.net/publication/243764350_The_Extent_and_Some_Implications_of_Incomplete_Accident_Reporting)
31. King G, Zeng L. Logistic regression in rare events data. *Political Analysis*. 2001;9(2):137-163, doi: 10.1093/oxfordjournals.pan.a004868
32. Alvarenga AMT. Modelos lineares generalizados: aplicação a dados de acidentes rodoviários [dissertação]. Lisboa (PT): Universidade de Lisboa; 2015.
33. Burnham KP, Anderson DR. Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*. 2004 Nov;33(2):261-304, doi: 10.1177/0049124104268644
34. Hastie T, Tibshirane R, Friedman J. *The elements of statistical learning*. 2. ed. Stanford (CA): Spring; 2017.
35. Visa S, Ramsay B, Ralescu A, van der Knaap E. Confusion matrix-based feature selection. In: Visa S, Inoue A, Ralescu A, editors. *Proceedings of the Twenty-second Midwest Artificial Intelligence and Cognitive Science Conference*; 2011 Apr 16-17; Cincinnati (OH). Madison (WI): Omnipress; c2011. p. 120-127. Disponível em: <http://ceur-ws.org/Vol-710/MAICS2011Proceedings.pdf>
36. Washington SP, Karlaftis MG, Mannering FL. *Methods for transportation analysis*. 2. ed. London: Chapman and Hall/CRC; 2003.
37. Fagerland MW, Hosmer DW, Bofin AM. Multinomial goodness-of-fit tests for logistic regression models. *Statistics in Medicine*. 2008 Jan;27(21):4238-4253, doi: 10.1002/sim.3202
38. Secretaria de Infraestrutura da Bahia (SEINFRA). Pesquisa de Mobilidade - Região Metropolitana de Salvador. Salvador (BA): SEINFRA; 2012.
39. Roudsari BS, Mock CN, Kaufman R, Grossman D, Henary BY, Crandall J. Pedestrian crashes: higher injury severity and mortality rate for light truck vehicles compared with passenger vehicles. *Injury Prevention*. 2004 Jun;10(3):154-158, doi: 10.1136/ip.2003.003814

40. Wang WW, Haque M, Chin HC, Yun JGJ. Injury severity of pedestrian crashes in Singapore. In: Australasian Transport Research Forum 2013 Proceedings (ATRF). 2013 Oct 1-15; Brisbane (AU). Disponível em: <https://eprints.qut.edu.au/63892/>
41. Ferenchak NN. Pedestrian age and gender in relation to crossing behavior at midblock crossings in India. *Journal of Traffic and Transportation Engineering*. 2016 Aug;3(4):345-351, doi: 10.1016/j.jtte.2015.12.001
42. Tulu GS, Washington S, King MJ, Haque M. Why are pedestrian crashes so different in developing countries? A review of relevant factors in relation to their impact in Ethiopia. *Australasian Transport Research Forum 2013 Proceedings*; 2013 Oct 2-4; Brisbane (AU). p. 1-18. Disponível em: [https://www.australasiantransportresearchforum.org.au/sites/default/files/2013\\_tulu\\_washington\\_king\\_haque.pdf](https://www.australasiantransportresearchforum.org.au/sites/default/files/2013_tulu_washington_king_haque.pdf)
43. Adanu EK, Hainen A, Jones S. Latent class analysis of factors that influence weekday and weekend single vehicle crash severities. *Accident Analysis & Prevention*. 2018 Apr;113(1):187-192, doi: 10.1016/j.aap.2018.01.035
44. Brasil. Ministério dos Transportes. Manual para o tratamento de locais críticos de acidentes de trânsito. Programa Pare. Brasília (DF): Ministério dos Transportes; 2002.