Closed-Group Injury Risk Functions for Frontal Crashes Integrating Lethality and Long-Term Consequences

Malte Plenter, Mirko Junge

Abstract Conventional field-data-based injury risk functions (IRFs) typically predict whether injury severity exceeds a threshold like MAIS2+ or MAIS3+. While useful for broad safety assessment, such models blur distinctions between different injury severity levels and largely reflect lethality rather than long-term consequences (LTC). This study develops closed-group IRFs for belted front-row occupants in passenger car frontal crashes using data from the German In-Depth Accident Study (GIDAS). Injuries were mapped to RAPEX HARM severity levels, allowing separate modelling of lethality-only and combined lethality/LTC injury outcomes.

Multinomial logistic regression was used to estimate mutually exclusive injury severity probabilities as a function of energy-equivalent speed (EES), age group, sex, frontal impact configuration, and further technical crash descriptors. EES and advanced age showed increasing effect strength with increasing injury severity, while female occupants had elevated risk up to intermediate injury severity levels. Relative to large overlap collisions, non-standard frontal impact configurations increased injury risk for selected severity groups. LTC inclusion produced modest overall changes in frontal crashes but shifted predicted risk towards higher HARM levels in some high-severity small overlap or oblique corner scenarios. Closed-group IRFs were shown to provide a more differentiated view of injury burden than conventional binary models.

Keywords injury risk functions, long-term consequences, GIDAS, frontal crashes, multinomial logistic regression.

I. INTRODUCTION

Advancing traffic safety requires a detailed understanding of the biomechanical and environmental factors that determine occupant injury outcomes in specific crash scenarios. In field-data-based safety research, injury risk functions (IRFs) are the standard epidemiological tool used to correlate technical crash severity and relevant occupant- and crash-related covariates to injury outcome. However, conventional IRFs typically only predict whether a person's injury severity exceeds a predefined severity threshold, most commonly a binary cutoff based on the Maximum Abbreviated Injury Scale (MAIS), such as MAIS2+ or MAIS3+. Although such dichotomous IRFs are valuable for broad safety assessment, they lack the specificity needed to distinguish meaningfully between different types and levels of harm.

- While AIS-coded injuries are a foundational data source, the coding process is prone to miscoding [1]. Systematic post-processing of these initial codes can improve their underlying correlation with true personal injury severity.
- Open-ended classes, such as MAIS2+, combine heterogeneous injury outcomes, effectively blurring distinctions between minor, severe, and fatal injury.
- The standard AIS framework evaluates an injury almost exclusively based on its threat to life, i.e. its acute injury severity (lethality). This paradigm fails to capture non-fatal injuries that result in long-term consequences (LTC) and permanent disability, which carry significant, life-altering impacts on a victim's quality of life but do not influence conventional AIS-based IRFs.

In this paper, we refer to conventional threshold IRFs such as MAIS2+ or MAIS3+ as open-ended models because the positive outcome class contains all injury severities at and above the selected cutoff. By contrast, a closed-group IRF predicts a set of injury severity categories that are mutually exclusive and collectively exhaustive. Thus, each case is assigned to exactly one category, and the category probabilities jointly describe the full injury severity distribution for a given crash scenario. The practical advantage is that the model can show whether a

M Plenter (e-mail: malte.plenter@volkswagen.de) is a PhD candidate in Mechanical Engineering at Volkswagen AG and Technische Universität Berlin. M Junge is Senior Researcher in the Department of Crash Research at Volkswagen AG in Wolfsburg.

change in crash severity, occupant age, or impact configuration mainly shifts probability from no injury to minor injury, from minor to moderate injury, or towards the highest severity levels. Traditional threshold IRFs remain useful for summary assessment, but they cannot directly reveal this redistribution of risk across severity groups.

Recent field-data-based research on injury risk has focused on extending IRFs to other impact configurations than the classic frontal crash and assessing injury risk in low-severity collisions. McMurry *et al.* [2] leveraged US traffic crash databases to develop MAIS3+ and MAIS5+ IRFs for adults involved in all planar motor-vehicle collisions. While this extension to side- and rear-impact collisions is highly valuable, their injury outcome definition remains dichotomous in each model. Östling *et al.* [3] investigated injury occurrence in low- to medium-severity collisions, which are far more prevalent in today's traffic environment than severe ones, and therefore represent a substantial societal injury burden. Remaining within descriptive analyses, no IRFs for low-speed crashes were provided. A multi-level modelling approach could capture how different levels of injury severity occur across the range of crash severity within one model. Krampe & Junge [4], aiming to improve the granularity of the ISO 26262 S-scale, addressed the inaccuracy of AIS coding by abstracting the codes to a system-level classification that can be assigned reliably by trained medical doctors from the provided documentation. Their approach introduced intermediate S-levels while preserving significant distinctions in mortality categories.

To address these limitations, we developed closed-group IRFs for belted front-row occupants of passenger cars in frontal crashes using the German In-Depth Accident Study (GIDAS). In the present context, "closed-group" means that the injury outcome levels are mutually exclusive and collectively exhaustive, in contrast to conventional open-ended threshold models such as MAIS2+ or MAIS3+. Accordingly, the model estimates the probability of belonging to each injury severity group rather than only the probability of exceeding a selected threshold. The IRFs predict an abstracted injury severity description derived from the ISO 26262 S-scale based on the work of Krampe & Junge [4]. Additionally, we applied a novel approach used by [5] to match lethality-based with LTC-based injury severity by mapping both to the European Union RAPEX HARM system, which was designed to detect unsafe consumer products [6]. Using multivariate multinomial logistic regression, we estimated the probability of each mutually exclusive injury severity level across combinations of covariates over the range of technical crash severity, represented by the energy-equivalent speed (EES) [7]. We further evaluated odds ratios (ORs) of the models' predictors across injury severity levels. In this way, the proposed framework aims to provide a more differentiated description of injury burden than conventional dichotomous IRFs and, by doing so, to improve the assessment of occupant injury risk in frontal crashes. In future work, the approach could be applied to other impact configurations and crash types (e.g. VRU collisions) to gain a deeper understanding of injurious trends in the traffic environment.

II. METHODS

The aim of this study was to develop IRFs with two novel characteristics: first, to account for both lethality and LTC; and second, to predict injury risk across as many mutually exclusive severity groups as the data support. Developing such granular IRFs requires a large and detailed dataset together with robust injury coding. Greater reliability and discriminatory information in the injury data enable finer differentiation between distinct injury severity levels.

Dataset and stratifications

The analysis was based on passenger vehicle crashes from the German In-Depth Accident Study (GIDAS), using the database version of December 2025, which sampled cases from 1999 to 2025. GIDAS is a joint accident research project of the German Association for Automotive Research and the Federal Highway and Transport Research Institute. It samples detailed on-scene crash data from the metropolitan regions of Hannover and Dresden and rural Munich using a statistically representative sampling framework. The database contains information on pre-crash circumstances, vehicle characteristics, occupant demographics, injury outcomes, and environmental conditions, documented using standardised technical and medical procedures [8, 9].

Different frontal crash configurations lead to different injury risks [10, 11], which arise from differences in acceleration peaks, compartment intrusion, and occupant angular acceleration. To account for the effect of differing frontal impact configurations, we followed the approach of Brumbelow [11] and assigned frontal impact configurations to all non-trivial frontal crashes in the GIDAS database. Additionally, we coded under- or overridden structures separately and used this information to ensure adequate structural interaction between the crash participants. These crash configurations allowed us to stratify the collisions based on crash type instead of opponent or overlap only, although aggregations were necessary to retain enough observations in each injury severity category. Large overlap collisions were used as the baseline and compared to moderate overlap collisions, a group consisting of near-side small overlap or frontal-oblique corner collisions, and a group of frontal-centre and centre-oblique collisions.

To maximise sample size, both drivers and front-row passengers were included in the analysis. For near-side small overlap crashes, the classification was assigned relative to the seating position of the respective occupant. Within this first attempt at multi-level IRFs, we excluded non-belted occupants, collisions that were involved in rollover events, and severe multi-collision events. The most injurious collision was selected based on the GIDAS variable *kollv*, which relates injuries to collisions and points to the collision that caused the most severe injuries.

Injury code abstraction and aggregation

As AIS-coded injury data have been shown to be prone to miscoding, especially upcoding [1], injury information was abstracted following the approach proposed by Krampe & Junge [4]. Individual injuries were translated to an abstracted level that is more in line with the medical diagnoses used for treatment and therefore can be assigned more reliably by trained medical doctors. This abstraction of AIS-IDs is intended to strengthen the reliability of the injury outcome and thereby support prediction across multiple distinct severity groups.

Following the rationale of the New Injury Severity Score (NISS) [12], the three most severe injuries were identified separately for AIS-CD (AIS-code, the severity rating of the injury corresponding to the AIS-ID) and predicted functional capacity index (pFCI). These injury triples were then used to derive person-level severity ratings on the extended ISO 26262 S-scale or on the four-level RAPEX HARM scale, which assesses the injury severity based on lethality and LTC simultaneously as shown in Table I [5]. In the following, FCI refers to the transformed value derived from AIS2015-pFCI as defined in Table I. Additionally, some AIS9-coded cases could be retained through the assignment of a minimum severity that led to an injury according to the descriptor, which was used when the aggregated severity produced the highest, i.e. unbound, injury severity level. The S-scale classification was used to estimate odds ratios (ORs), because it includes an explicit no-injury reference level. The highest two severity levels, S2.5 and S3, were aggregated to retain enough observations in all injury severity categories. The HARM classification was used for graphical presentation of IRFs because its smaller number of severity levels provides adequate approximation while avoiding overly crowded plots, and because it allows for direct comparisons between a conventional lethality-based assessment and a combined lethality- and LTC-based assessment. To compare the changes in effects that assessing LTC invokes in an IRF, we compared IRFs assessing HARM based on AIS-CD alone with HARM based on AIS-CD and FCI combined. Injury severity levels will be denoted with a preceding S or H, indicating injury severity on the S-scale and the HARM-scale, respectively. Table I shows the mapping of AIS-CD- or FCI-triples to S-levels and H-levels. Note explicitly the improved granularity of MAIS2- or MAIS3-rated injurious outcomes, which can lead to S1, S1.5, S2, S2.5, and S3 ratings depending on the occurrence and severity of other injuries.

TABLE I
INJURY AGGREGATIONS BASED ON AIS-CD OR FCI TRIPLES

RAPEX HARM [5]	ISO 26262 S-scale [4]	Ordered AIS-CD / FCI triples (FCI = 6 – AIS2015-pFCI, x ≤ anything to its left)
H4: Possibly fatal, impact on reproduction, severe loss of function leading to more than ~10% of disability	S3: Life-threatening injuries (survival uncertain)	3/3/x, 4/x/x, 5/x/x, 6/x/x
	S2.5: Life-threatening injuries (survival probable)	3/2/x
H3: Requires hospitalisation, will affect function for more than 6 months or lead to permanent loss of function	S2: Severe injuries	2/2/2, 3/0/0, 3/1/x
H2: May require medical help, but no hospitalisation. Almost complete functional recovery after 6 months at most	S1.5: Moderate injuries	2/2/0, 2/2/1
	S1: Light injuries	2/0/0, 2/1/x
H1: No substantial loss of function or excessive pain after basic treatment	S0.5: Minor injuries	1/x/x
	S0: No injuries	0/0/0

To summarise, we used the injury-level severity rating AIS-CD together with the corresponding long-term severity rating pFCI from the AIS 2015 codebook as the starting point of our assessment. In order to account for coding inaccuracies, we recoded it to a less granular, but more reliable AIS-code as proposed by [4]. Next, we needed to aggregate the injury-level severity assessment to an assessment of a person's overall injury severity. For the sole lethality-based assessment, i.e. the AIS-CD aggregation, we relied on the triple-based AIS-CD aggregation and projection to the ISO 26262 S-scale by Krampe & Junge [4], which the authors have shown to correlate better with lethality than other injury aggregators such as MAIS. Performing this step with the pFCI is rather novel, but Krampe & Junge [5] propose a similar methodology based on FCI-triples that are mapped to the RAPEX HARM system. This system is required because its definition is based on lethality as well as long-term consequences, which the descriptors of the S-scale levels fail to do.

As an example, assume a person suffers a chest contusion (410402.1), a metacarpus fracture (752553.2), a distal ulna fracture (752353.2), an ankle sprain (877110.1), and a malleolus fracture (854361.2). After the AIS abstraction, the three AIS2-rated injuries persist and lead to the AIS-triple "222". While the MAIS-system would rate this as an injurious outcome of MAIS2, the aggregation we used recognises the increased injury severity due to multiple injuries and leads to a rating of S2 or H3, which puts it on a level with AIS-triples such as "300", i.e. MAIS3. In the long-term consequence assessment, the methodology we applied recognises the burden of extremity fractures, which bear the possibility of impaired movement even after the injuries have healed, with an FCI-triple of "421". On the RAPEX HARM scale, this leads to an assessment of H4, an increase against the sole lethality assessment.

Modelling objective and method

The objective of our modelling process was to estimate injury risk functions that describe the probability distribution across mutually exclusive injury severity levels, i.e. closed-group IRFs, as a function of crash, vehicle, and occupant characteristics. The models were not intended to assign a single most likely injury severity level to individual cases, but rather to quantify how injury risk is distributed across severity groups for a given scenario. Accordingly, model quality was assessed based on overall explanatory power and calibration rather than classification accuracy.

To model multiple discrete injury severity levels, ordinal and multinomial logistic regression were considered. Ordinal logistic regression assumes proportional odds across outcome levels, implying that predictors have the same relative effect between all adjacent severity categories. While this assumption was not violated in the present dataset, multinomial logistic regression was selected because it allows estimation of severity-specific effects for each injury outcome category. This was considered particularly valuable for interpreting how risk factors such as age, crash configuration, or vehicle characteristics affect different severity levels differently. All final models were therefore estimated using multinomial logistic regression. We fitted our models using the *nnet* library [13] in RStudio 2024.04.1 [14]. Pointwise 95% confidence bands for predicted class probabilities were obtained via parametric simulation [15].

Model evaluation and quality assurance

Because the aim of the models was probabilistic estimation of injury severity distributions rather than deterministic classification, conventional classification metrics such as confusion matrices were not considered appropriate; due to the nature of multinomial logistic regression, some severity levels might never be assigned because other levels show higher likelihoods over the range of predictor combinations. Instead, model performance was evaluated based on overall explanatory power and calibration.

Model fit was assessed using McFadden's R^2 [16] (further referred to as R^2_{MF}), for which values above 0.2 are commonly interpreted as indicating excellent fit [17]. Models were iteratively extended to include more predictors until this threshold was reached for one of the models. The performance of other models in comparison to this best-performing model was used as an indicator of how well each injury severity assessment is predictable based on the common set of predictors. Model calibration was evaluated using calibration plots and category-wise Hosmer-Lemeshow tests [18], which showed no substantial evidence of lack of fit. Collinearity among predictors was assessed using the variance inflation factor (VIF) [19], with no strong collinearity observed.

Predictor selection

Predictor selection followed a stepwise, hypothesis-driven approach. First, technical crash severity was included as the primary predictor, and several commonly used metrics (energy-equivalent speed, delta-v, and relative collision velocity) were evaluated in separate models. Energy-equivalent speed (EES) was selected for the main analysis because it reflects vehicle damage severity and allows interpretation independent of collision dynamics. ORs based on delta-v and relative collision velocity are supplied in the Appendix (Figs. A1 and A2).

Second, occupant-related variables (age group and sex) were added to account for known human vulnerability effects. Age was grouped into three categories (17–44, 45–64, and 65+) [20], and sex was included to account for sex-based injury risk differences in vehicles of the age range covered by the dataset [21].

Third, crash configuration variables were added to capture differences between frontal impact types. Large overlap collisions were used as the reference category. Finally, vehicle- and opponent-related variables (crash weight of the vehicle, year of market introduction, and collision opponent type) were included to improve model fit and control for known confounding effects. Vehicle crash weight denotes the GIDAS-provided crash weight where available, i.e. the total mass of the vehicle during the crash, otherwise estimated from kerb weight, occupants, and loading information.

At each step, the contribution of additional predictors was assessed using likelihood-ratio tests, and only interpretable predictors that improved model fit were retained in the final models. A table of the final model structure can be found in Table AI.

III. RESULTS

The stratified GIDAS sample enabled the estimation of closed-group IRFs for frontal crashes across multiple mutually exclusive severity levels. The results show that EES and occupant age are strong determinants of increasing injury severity up to the most severe injury severity group, S2.5+, while frontal impact configuration and sex affect selected severity groups. In addition, the closed-group framework revealed how injury risk is redistributed across severity levels rather than reduced to a single binary threshold.

The following sections present the final dataset, model fit, ORs, RAPEX HARM-based injury risk curves, and the injuries most frequently responsible for severity shifts when LTC are included.

Stratified dataset

The stratifications on the GIDAS database led to a set of 2,834 participants of frontal crashes, comprising 2,345 drivers and 489 passengers. Injury distributions according to the HARM classification are shown in Table II, and those according to the extended S-scale in Table III. The model predicting HARM based on AIS alone was the first to cross the R^2_{MF} threshold, with a value of 0.204. Based on otherwise identical model calls, predicting injury severity on the S-scale led to a similar R^2_{MF} of 0.199, while predicting HARM based on AIS and FCI combined led to an R^2_{MF} of 0.182. No VIF values suggested strong multicollinearity between predictors in any model. The most severe collinearity was observed in the model predicting S-levels for the covariate indicative of the collision opponent.

TABLE II
CASE COUNTS PER HARM LEVEL [5]

HARM level	Case count based on AIS	Case count based on AIS and FCI
H1	2,468	2,451
H2	253	247
H3	52	65
H4	61	71
Σ	2,834	2,834

TABLE III
CASE COUNTS PER S-LEVEL [4]

S-level	S0	S0.5	S1	S1.5	S2	S2.5+	Σ
Case count	925	1,543	200	53	52	61	2,834

Odds ratio analysis

Modelling a multinomial logistic regression model to predict the ISO 26262-based injury severity levels as proposed by Krampe & Junge [4], against the baseline of no injuries (S0), led to the ORs shown in Fig. 1. Note that in this model, injuries are assessed based only on their AIS-CD-rating, i.e. their acute injury severity [4]. All predictors we selected are statistically significant risk factors for at least some of the injury severity levels. Most notably, the effect of technical crash severity, represented by EES, increased with injury severity. This was offset by progressively lower intercepts for higher severity levels, resulting in later onsets of high-severity risk curves, as can be seen in the subsequent injury risk curves in Fig. 2 and Fig. 3. A similar phenomenon is observed in the effect of advanced occupant age, which is insignificant at S0.5 and strongest on the higher injury severity levels. In contrast, the additional risk associated with differing impact configurations loses its significance for the highest injury severity level in the case of moderate overlap and small overlap or oblique corner collisions, which may however be caused by low sample sizes. Females show a significantly higher injury risk than males up to the injury severity level S2. HGVs and fixed narrow objects such as poles or trees pose increased injury risk that is not entirely controlled for through higher EES, the exclusion of under- or override-cases, and use of centre impacts as a covariate. Rising vehicle crash weight significantly reduces injury risk across the spectrum of injury severities with rising effect strength across injury severity levels. A more recent year of market introduction tends to be a protective factor, but lacks significance in every category except for S0.5.

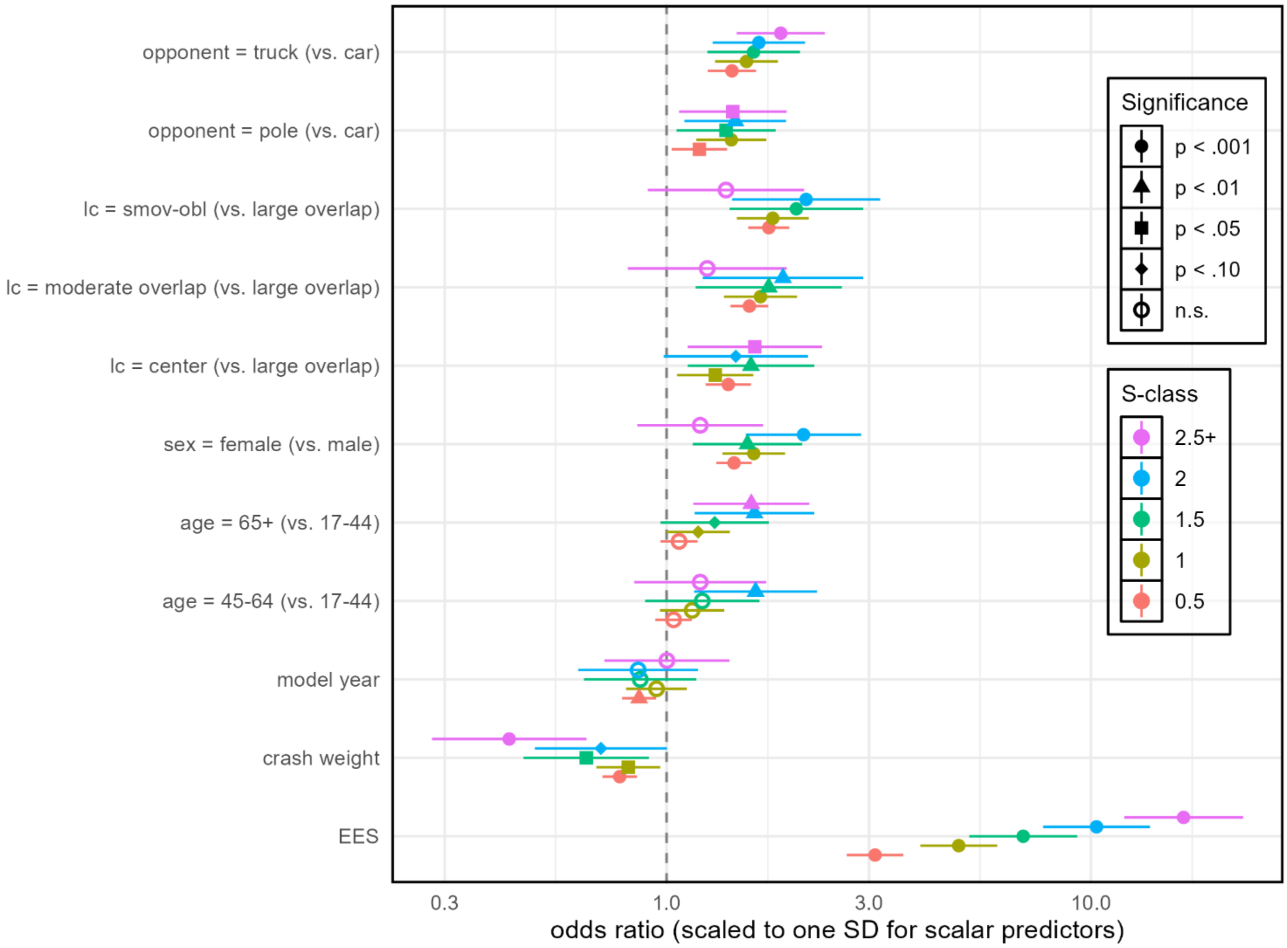

Fig. 1. Odds ratios (ORs) of sustaining an injurious outcome for closed injury severity groups corresponding to the ISO 26262 S-scale derived from a multinomial logistic regression model. ORs are standardised to one standard deviation for continuous predictors, being the EES (SD = 13.6 km/h), the crash weight (SD = 336 kg), and the model year (SD = 4.4 years). Symbols indicate OR significance, while the bands show their 95% confidence interval.

Injury risk curves

Injury risk functions for sustaining an injurious outcome corresponding to one of the four RAPEX HARM levels are shown in Fig. 2 and Fig. 3. We marked the 10% risk threshold in each plot for visual orientation; this threshold is roughly comparable to the 10th percentile used to calculate ISO 26262 S-levels [4], although it is not entirely equivalent as those values are based on open-ended injury severity classes. While Fig. 2 shows the risk of sustaining an injury that is assessed solely on lethality, i.e. the AIS-CD, Fig. 3 shows both the AIS-based model and the model based on the combined AIS-CD and FCI assessment as proposed by Krampe & Junge [5]. Fig. 2 shows variations in impact configuration and occupant age, while Fig. 3 illustrates the effects of varying vehicle crash weight.

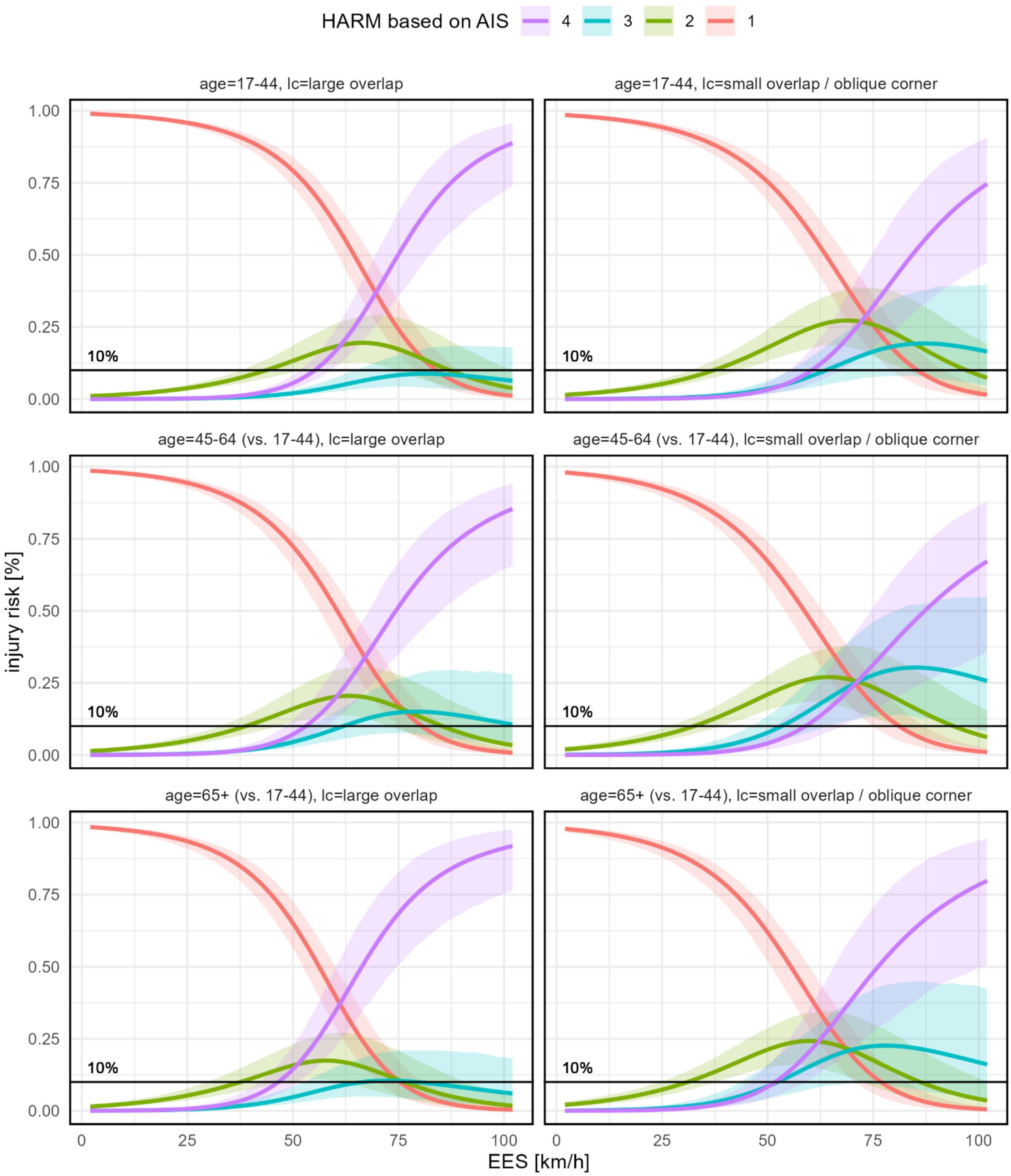

Fig. 2. Injury risk curves on the RAPEX HARM-scale according to Krampe & Junge [5] assessing lethality alone based on the AIS-code. The 10% risk threshold is marked for visual orientation. In each column, the occupant age is varied between the three groups we defined. Each row shows the variation between a large overlap collision and a small overlap or oblique corner collision. Predictors not mentioned are held at reference level. Confidence bands show the 95% confidence interval.

Especially within the injury severity class H2 (comparable with MAIS2), a distinct peak in the risk of sustaining the corresponding injury severity is observed. Those peaks indicate at which collision severity an injurious outcome of the respective severity is most likely to occur. Variations in occupant age and vehicle crash weight show their greatest effect in the onset of the most severe injury severity level. The 45–64 age group generally differs less from the reference group than the oldest age group. Comparing large overlap to small overlap or oblique collisions, an earlier onset of moderate injuries (H2) is observed, but the effect on the most severe injury level is mixed. In some instances, the risk of H4 does not significantly exceed the risk of H3 based on the AIS assessment. Switching from the AIS-based to the AIS-CD- and FCI-based assessment most notably shifts some risk from H3 to H4 in small overlap and oblique corner collisions, which leads to slightly more severe injury risk distributions in those collisions, mostly based on narrower confidence intervals at higher crash severities. In large overlap collisions, barely any differences between the two assessments are noticeable.

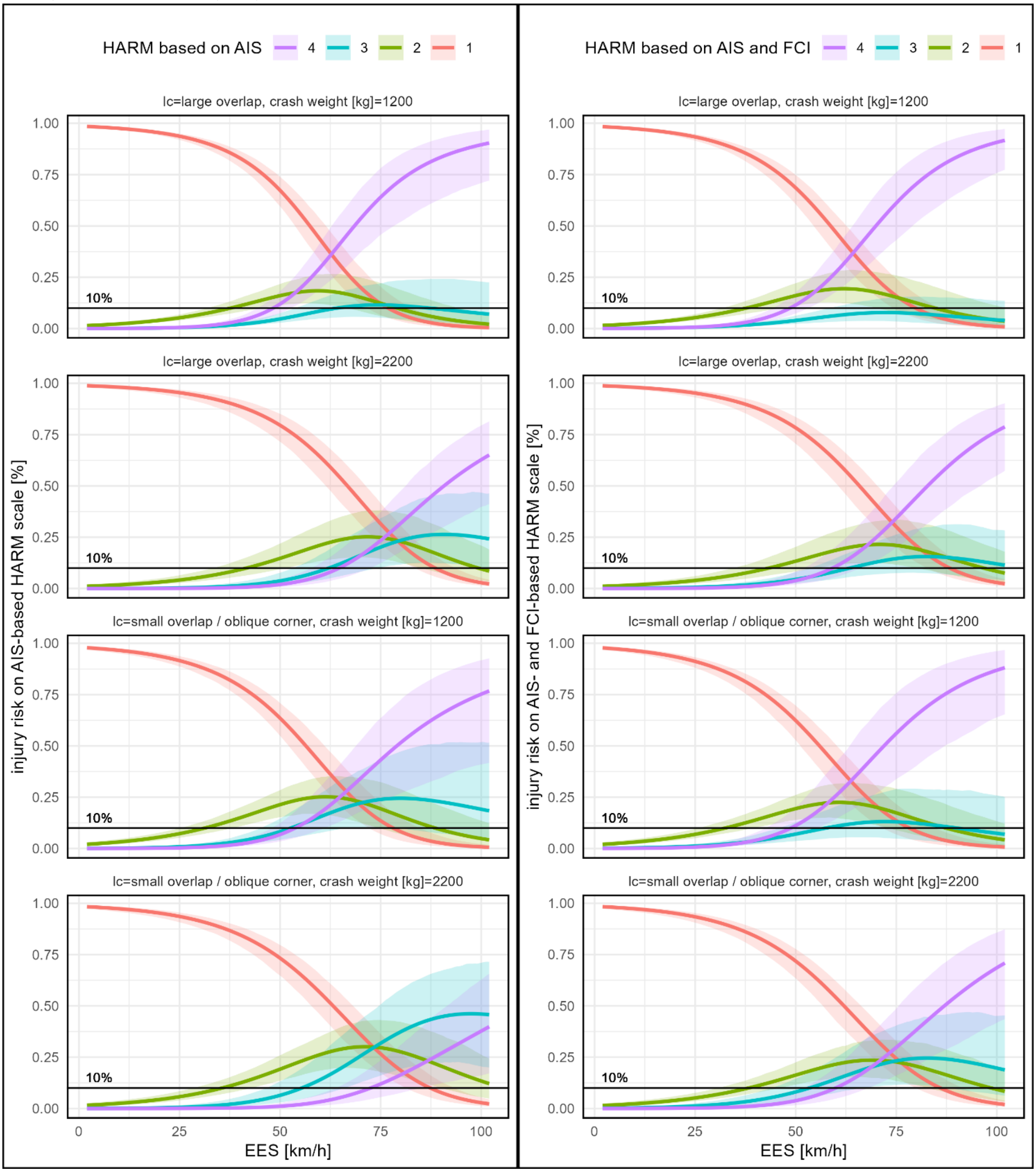

Fig. 3. Injury risk curves on the RAPEX HARM-scale according to Krampe & Junge [5]. The left column shows the injury risk assessment based solely on the AIS-CD, the right column based on the combined assessment of AIS-

CD and FCI. Row-wise, impact configuration and crash weight of the vehicles are varied. The age of the occupants is held at the middle-aged group (45–64), all other predictors are held at their reference level. The 10% risk threshold is marked for visual orientation. Confidence bands show the 95% confidence interval.

Common LTC-intense injuries leading to an increase in HARM in frontal crashes

Several injuries were responsible for increases in the AIS- and FCI-combined injury severity assessment compared to the assessment of AIS alone due to their potential long-term consequences. An overview is given in Table IV, grouping injuries to their specific body regions.

TABLE IV
HARM-SHIFTING INJURED BODY REGIONS IN FRONTAL CRASHES, SORTED BY OCCURRENCE IN THE DATASET OF THIS STUDY

Body region	Corresponding injuries (based on level-1 abstraction [4])	Frequency of corresponding injuries
Head (specifically, ear)	Ear injury n.f.s (including tinnitus and hearing loss)	16
Lower extremities & pelvis	Foot fracture, tibia fracture, ankle joint injury, Chopart joint injury, acetabulum fracture	13
Upper extremities	Carpal joint injury, ulna fracture, forearm fracture n.f.s., traumatic amputation of an upper extremity	5
Spine	Cervical spinal cord injury, cervical nerve root injury	2

IV. DISCUSSION

Model performance and model differences

Our modelling process led to a multinomial AIS-CD-based IRF on the HARM-scale, which is classified as an excellent fit according to its R^2_{MF}, and an IRF on the more granular S-scale that performs only marginally worse. Predicting AIS-CD- and FCI-based injury severity combined, however, led to substantially worse model fit, which indicates that LTC are either more diffuse or more dependent on other crash phenomena than conventional immediate harm-describing AIS-coded injuries. Models based on EES and delta-v were similar in performance, but differed in the injury risk factors they managed to highlight: while the delta-v-based model (Fig. A2) assigns weaker effects to the vehicle crash weight, as that is accounted for in the magnitude of delta-v itself in collisions against other movable objects, the models based on EES (Fig. 1) and on the relative collision velocity (Fig. A1) assign a significant OR to the vehicle crash weight across the spectrum of injury severities. Identifying risk factors in traffic crashes benefits from assessing a crash based on different crash severity indicators and comparing the outcomes.

Risk factors in frontal crashes

Our OR analysis in Fig. 1 shows bundled insights into risk mechanisms of frontal crashes that widely agree with past literature. Increasing occupant age, especially in the oldest age group, led to increases in injury risk, especially for the higher injury severities in our models. The increasing fragility of the elderly is a well-known risk factor [20]; our novel IRFs highlighted that the increase in injury risk is strongest in higher injury severity levels.

Females are at higher risk of injury within the injury severity levels S0.5 through S2 in our dataset, which matches our expectations and recent research [22], given that our dataset consisted of vehicles up to 25 years old. At level S2.5+, the increase in risk diminished and became non-significant, although the dataset contained several data points in this category for females. Recent research that evaluated vehicles of a similar age-spread found similar trends regarding the significance of increased overall injury risk for females [22]. An analysis stratified for very recent passenger cars would have been desirable, as other researchers have also found diminishing injury risk differences between males and females in lower injury severities in those models [21], but we did not have a large enough sample size within our granular injury risk categories based on GIDAS.

Various frontal impact configurations all showed a significant increase in injury risk on at least some injury severity levels compared to the baseline of large overlap frontal collisions. Compared to vehicle mass, age, or EES, the additional risk generally seems less severity-dependent. Past studies have found non-standard frontal impact configurations to be an injury risk-increasing factor on numerous occasions [10, 11]. In contrast to the work of

Brumbelow [11], we did not aggregate large and moderate overlap collisions into one group, therefore we identified moderate overlap collisions to be significantly more injurious on all injury severity levels up to S2. Similar trends were observed for the other impact configurations we examined, although centre impact collisions were the only group that inflicted significantly increased injury risk on level S2.5+ across all crash severity predictors. Based on the OR models, we can robustly infer that small overlap or oblique corner collisions, as well as moderate overlap collisions, increase injury risk up to level S2, and that centre impact collisions are risk-increasing across all injury severity levels.

Injury risk functions for distinct injury severity groups

The main contribution of the closed-group IRFs is that they describe injury risk as a distribution across distinct severity groups rather than as the probability of exceeding a single cutoff. In practical terms, this complements conventional MAIS-based IRFs by showing where the risk is located within the severity spectrum. For example, two crash scenarios may have a similar risk of exceeding MAIS2+, while one primarily increases moderate injuries and the other shifts probability towards severe or life-threatening outcomes. The closed-group framework makes such differences visible and is therefore better suited to examine how occupant, vehicle, and crash characteristics redistribute injury risk. Our lethality-based H1-curves, shown in Fig. 2, which predict the likelihood of sustaining AIS1-injuries at most, can be interpreted as inverse MAIS2+ risk curves. In this regard, they agree well with past configuration-specific IRFs by Brumbelow [11], in which such injurious outcomes were sustained at lower crash severities as well. Through the remaining injury risk curves, our IRFs show that the additional risk sustained in small overlap or oblique corner collisions concentrates in the H2- and H3-groups, which become more likely over the entire spectrum of EES. The most severe injury risk level, H4, indicating injurious outcomes of at least one AIS3 injury and one AIS2 injury and worse, appears less likely even in higher EES ranges, although the change is insignificant. This indicates that when small overlap or oblique corner collisions are the most injurious event, they generally cause less extreme injurious outcomes but more medium- and high-severity ones. However, the risk of multi-collision events caused by small overlap collisions, which often deflect a car and pose the risk of severe secondary collisions, was not assessed in this model because the included collisions are only those that were most injurious to each participant.

In most instances, the H3 injury risk curve crossed the 10% threshold at a higher EES than the H4 curve or not at all. Injurious outcomes of this severity seem to be rather rare and occur at similar technical crash severities as the open class H4, mostly discriminating based on occupant and vehicle characteristics as well as crash constellation.

Distinction between lethality-based and LTC-based injury risk assessment

Aggregating lethality and LTC into a combined injury assessment, as we did in Fig. 3, arguably showed only modest changes in injury risk in our dataset of frontal crashes. A noteworthy factor, although not statistically significant, is that the peak H4 risk of small overlap or oblique corner collisions reached a higher value than in the model that predicted lethality alone due to shifts from H3 to H4. As such collisions are known to carry an increased risk of compartment intrusion, and therefore pose the risk of severe injuries to an occupant's extremities [23], the combined model accounts for severely impairing injuries in a way that the model based on lethality alone could not. The most common severity-shifting injuries, as seen in Table IV, support this finding: aside from tinnitus, which may be inflated by self-reported injuries after a crash, injuries to the lower extremities dominate the list of injuries that cause a shift in the severity rating, followed by upper extremity injuries with slightly lower case counts. Tinnitus should not be disregarded, however, as past research has shown its occurrence in frontal crashes to be realistic depending on airbag timing and overall number of airbags firing [24]. Overall, severe injuries to the extremities are not the greatest risk factor in frontal crashes; a larger shift due to the inclusion of LTC might be observable in IRFs for side-crashes or VRUs.

Limitations and future work

Our study is based on a thoroughly stratified subset of the GIDAS database. These stratifications were necessary to enable a configuration-based injury risk assessment, but led to quite low sample sizes in some categories (e.g. the highest severity class in some impact configurations). This may have caused insignificant effects where a true increase in risk would have been expected, and led to wide confidence bands on the IRFs. Applying the approach to a larger dataset than GIDAS may provide further insights into the differences in risk between different frontal impact configurations.

The methodology we developed to model injury risk in multiple discrete severity levels requires large amounts of crash data. We applied several measures to maintain as large of a dataset from GIDAS as possible without compromising the accuracy of our model, e.g. evaluating drivers and front-row passengers together, limiting the amount of human-descriptive predictors, and assessing collisions against vehicles and objects. Still, severe injuries and crashes are rare events and, therefore, lead to wide confidence intervals at high crash and injury severities. While our approach accounts for this through its confidence intervals, merging information from different datasets may improve the strength of results at higher crash as well as injury severities.

GIDAS is known to be biased towards severe collisions based on its sampling scheme [25]. Additionally, our impact configuration coding strategy prioritised collisions that led to severe injuries. Our dataset being representative of the German traffic crash environment is, therefore, unlikely.

Assessing only the most severe collision of a traffic crash simplifies typical risks of certain impact configurations, such as deflections after small overlap collisions, and therefore may bias the analysis. A detailed comparison of single- versus multi-collision events based on the approach of this study may yield further insights, as secondary collisions may also lead to greater risk of LTC, e.g. through out-of-position occupant engagement.

The inclusion of vehicle model year as a predictor in our models failed to highlight the significant improvements in vehicle crashworthiness in frontal collisions of the past two decades. While this may partially validate our methodology in assessing all vehicles from 2001 onward, modifications to our approach should be done in the future to better highlight safety achievements over the years. For example, assessing body-region-specific injury severity may bear safety improvements where a person-level assessment, such as the one used here, fails to capture.

Research on LTC is an ongoing and much discussed topic. While the consensus is that a deeper understanding of LTC is necessary to further improve road traffic safety, the pathway to achieving this goal is still being devised. The pFCI of the AIS codebook has been shown to struggle with validation and actual correlation with long-term impairment [26]; still, it is currently the best available option. Ongoing research projects such as IMPROVA [27] may help establish a more robust LTC assessment, which could then be used to update the IRFs in this work.

The methodology we used to assess lethality and long-term impairment combined is, just like the FCI itself, a matter of ongoing research. Weighing the risk of lethality against the risk of long-term impairment is a complex, partly philosophical issue. Other means of comparing the two might provide further insights, e.g. assessing the societal cost of injuries. Additionally, assessing the influence of LTC on the overall injury assessment might cover more substantial influences that would be observable in a body-region-specific assessment.

V. CONCLUSION

Closed-group IRFs for frontal crashes were developed from a stratified GIDAS sample. Unlike conventional threshold IRFs based on outcomes such as MAIS2+ or MAIS3+, the proposed models estimate mutually exclusive injury severity probabilities whose sum equals 1, thereby representing injury burden as a full severity distribution rather than as threshold exceedance alone.

Technical crash severity, represented by either EES, delta-v, or relative collision velocity, was the strongest predictor of increasing injury severity that also gained effect strength in the most severe injury risk groups. Occupant age showed a pronounced severity-dependent effect, with older occupants reaching high-severity outcomes at lower crash severities. Vehicle crash weight also showed increasing effect strengths over the injury severity levels, which also shows that occupants of lighter cars are at greater risk, particularly for severe injuries. Sex and frontal impact configuration also influenced injury risk, although their effects were limited to selected severity groups.

Including LTC in the injury assessment led to only modest overall changes in frontal crashes, but slightly shifted risk towards higher severities particularly in small overlap or oblique corner impacts. This indicates that LTC-informed injury assessment can complement lethality-based severity metrics by capturing relevant non-fatal injury burden.

Overall, the proposed framework is feasible in field crash data and offers a more nuanced basis for injury risk assessment in frontal crashes. Future work should extend the approach to other crash modes and improved measures of long-term impairment.

VI. REFERENCES

[1] Twiss, E., Krijnen, P., Schipper, I. (2021) Accuracy and reliability of injury coding in the national Dutch Trauma Registry. *Journal of the International Society for Quality in Health Care*, **33**(1).

[2] McMurry, T. L.; Cormier, J. M.; Daniel, T.; Scanlon, J. M.; Crandall, J. R. (2021) An omni-directional model of injury risk in planar crashes with application for autonomous vehicles. *Traffic Injury Prevention*, **22**(sup1): pp. 122–127.

[3] Östling, M., Jeppsson, H., Forman, J. (2024) Injury Observations in High-Exposure, Low-Severity Frontal Car Crashes - a GIDAS Investigation. *Proceedings of the IRCOBI Conference*, 2024, Stockholm, Sweden.

[4] Krampe, J. & Junge, M. (2025) Injury Severity for Hazard and Risk Analyses: Calculation of ISO 26262 S-parameter Values from Real-World Crash Data. Proceedings of the IRCOBI Conference, 2025, Vilnius, Lithuania.

[5] Krampe, J.; Junge, M. (2026) RAPEX HARM from AIS 2015 Coded Injuries [Unpublished manuscript currently under review, pre-print available under medRxiv, to be published in *Traffic Injury Prevention*].

[6] European Commission (2024) C(2024) 6023 final– Supplementing Regulation (EU) 2023/988. EUR-Lex.

[7] Campbell, K. L. (1974) Energy Basis for Collision Severity. *SAE Technical Paper Series*, SAE-740565.

[8] Otte, D., Krettek, C., Brunner, H., Zwipp, H. (2003) Scientific Approach and Methodology of a New In-Depth Investigation Study in Germany. *Proceedings of the 18th ESV International Conference*, 2003, Japan.

[9] Johannsen, H.; Krettek, C.; Hannawald, L.; Schaser, K. D. (2017) Consideration of accident avoidance technology within GIDAS. *Proceedings of the 25th ESV International Conference*, 2017, Detroit, USA.

[10] Andricevic, N., Junge, M., Krampe, J. (2018) Injury risk functions for frontal oblique collisions. *Traffic Injury Prevention*, **19**(5): pp. 518–522.

[11] Brumbelow, M. L. (2019) Front Crash Injury Risks for Restrained Drivers in Good-rated Vehicles by Age, Impact Configuration, and EDR-based Delta V. *Proceedings of the IRCOBI Conference*, 2019, Florence, Italy.

[12] Osler, T., Baker, S. P., Long, W. (1997) A modification of the injury severity score that both improves accuracy and simplifies scoring. *The Journal of Trauma: Injury, Infection, and Critical Care*, **43**(6): pp. 922–926.

[13] Venables, W. N., Ripley, B. D. (2002) *Modern applied statistics with S. Springer*: New York, NY, USA.

[14] Posit team (2025) RStudio: Integrated Development Environment for R.

[15] King, G., Tomz, M., Wittenberg, J. (2000) Making the Most of Statistical Analyses: Improving Interpretation and Presentation. *American Journal of Political Science*, **44**(2): p. 347.

[16] McFadden, D. (1974) Conditional Logit Analysis of Qualitative Choice Behavior. *Frontiers in Econometrics*, pp. 105–142.

[17] McFadden, D. (1979) Quantitative methods for analyzing travel behaviour of individuals: Some recent developments. *Behavioural Travel Modelling*, pp. 279–318.

[18] Hosmer, D. W. & Lemeshow, S. (1980) Goodness of fit tests for the multiple logistic regression model. *Communications in Statistics - Theory and Methods*, **9**(10): pp. 1043–1069.

[19] Kutner, M. H. (2005) Applied linear statistical models. McGraw-Hill: pp. 408–410.

[20] Liu, C., Utter, D., Chen, C.-L. (2007) *Characteristics of Crash Injuries Among Young, Middle-Aged, and Older Drivers* (Report No. DOT HS 810 857). National Highway Traffic Safety Administration: Washington, DC, USA.

[21] Noh, E. Y., Atwood, J. R. E., Lee, E.; Craig, M. J. (2022) *Female Crash Fatality Risk Relative to Males for Similar Physical Impacts* (Report No. DOT HS 813 358). National Highway Traffic Safety Administration, Washington, DC, USA.

[22] Craig, M. J., Atwood, J. R. E., *et al.* (2026) Sex-Based Differences in Odds of Motor Vehicle Crash Injury Outcomes. National Highway Traffic Safety Administration: Washington, DC, USA.

[23] Sherwood, C. P., Nolan, J. M., Zuby, D. S. (2009) Characteristics of Small Overlap Crashes. *Proceedings of the 21st ESV International Technical Conference*, 2009, Stuttgart, Germany.

[24] Holtz, J., Heidt, R., *et al.* (2022) *Accident injuries in vehicles equipped with airbags*, 94. *Fachverlag NW in der Carl Ed. Schünemann KG*, Bremen, Germany.

[25] Krampe, J. & Junge, M. (2024) Enhancing Data in National Crash Statistics Using an In-depth Study. *Proceedings of the IRCOBI Conference*, 2024, Stockholm, Sweden.

[26] Palmer, C. S., Cameron, P. A., Gabbe, B. J. (2017) A review of the revised Functional Capacity Index as a predictor of 12 month outcomes following injury. *Injury*, **48**(3): pp. 591–598.

[27] IMPROVA Consortium (2026) 'Vision Beyond Zero' road safety approach, https://doi.org/10.3030/101146652.

VII. APPENDIX

Multinomial logistic regression model predictors

Table AI lists the common set of predictors all models in this paper used after the selection process we explained in the methods section.

TABLE AI
MODEL PREDICTORS

Predictor	Type	Unit/Levels (* indicating reference level)
EES	Numeric	km/h
Occupant age	Factor	*17–44 45–64 65+
Occupant sex	Factor	*Male Female
Vehicle crash weight	Numeric	kg
Vehicle year of market introduction	Numeric	years
Loadcase [11]	Factor	*Large overlap Moderate overlap Small overlap & oblique-corner Centre impact & oblique-centre
Collision opponent	Factor	*Car Pole (including trees) HGV

Multinomial logistic regression: underlying equations

The equations below explain how the probability of a specific category in a multinomial logistic regression model is calculated based on the coefficients of the model's predictors.

Let the outcome $Y \in \{0, 1, \dots, K-1\}$ denote mutually exclusive injury severity levels, where Y=0 is the reference category, being S0 or H1 in our models.
For each non-reference category $k = 1, \dots, K-1$, the linear predictor is defined as:

$$\eta_k = \beta_{k0} + \sum_{p=1}^{P} \beta_{kp} x_p$$

The probability of observing injury level *k* given the set of predictors **x** is given by:

$$P(Y = k \mid \boldsymbol{x}) = \frac{\exp(\eta_k)}{1 + \sum_{j=1}^{K-1} \exp(\eta_j)}, \quad k = 1, \dots, K-1$$

The probability of the reference category is:

$$P(Y = 0 \mid x) = \frac{1}{1 + \sum_{j=1}^{K-1} \exp(\eta_j)}$$

Predictor coefficients of the presented IRFs

The tables AII and AIII show the coefficients of the models from which the IRFs as seen in Fig. 2 and Fig. 3 were computed. As in Fig. 1, we exponentiated and standardised the coefficients to one SD for easier interpretation.

TABLE AII
EXPONENTIATED AND SD-STANDARDISED COEFFICIENTS OF THE AIS-BASED HARM IRF (SEE FIG. 2)

	H2	H3	H4
EES	2.123 (1.876–2.403)	3.887 (3.161–4.780)	5.828 (4.582–7.412)
Crash weight	0.918 (0.835–1.011)	0.876 (0.728–1.055)	0.517 (0.409–0.652)
Age: 45-64	1.318 (0.979–1.775)	2.445 (1.629–3.668)	1.414 (0.944–2.117)
Age: 65+	1.446 (1.027–2.036)	2.898 (2.130–3.942)	3.168 (2.134–4.704)
Sex: female (vs. male)	1.621 (1.256–2.093)	2.487 (1.475–4.194)	0.880 (0.481–1.609)
LC: centre impact	0.948 (0.602–1.493)	1.201 (0.850–1.696)	1.464 (1.069–2.004)
LC: moderate overlap	1.318 (0.947–1.834)	1.717 (1.177–2.505)	0.642 (0.398–1.036)
LC: small overlap / oblique corner	1.415 (1.002–1.998)	1.870 (1.244–2.813)	0.606 (0.402–0.915)
Opponent: pole (vs. car)	2.377 (1.590–3.553)	2.298 (1.630–3.239)	2.607 (1.923–3.535)
Opponent: truck (vs. car)	2.085 (1.333–3.263)	2.218 (1.939–2.537)	3.895 (3.236–4.687)

TABLE AIII
EXPONENTIATED AND SD-STANDARDISED COEFFICIENTS OF THE AIS- AND FCI-BASED HARM IRF (SEE FIG. 3)

	H2	H3	H4
EES	2.083 (1.840–2.359)	3.038 (2.531–3.647)	5.107 (4.149–6.286)
Crash weight	0.898 (0.816–0.989)	0.999 (0.851–1.173)	0.664 (0.544–0.809)
Age: 45-64	1.255 (0.928–1.697)	1.702 (1.032–2.808)	1.654 (1.071–2.555)
Age: 65+	1.444 (1.026–2.032)	2.119 (1.241–3.618)	2.651 (1.758–3.996)
Sex: female (vs. male)	1.560 (1.205–2.018)	1.964 (1.228–3.141)	0.945 (0.551–1.622)
LC: centre impact	0.901 (0.569–1.425)	1.894 (1.328–2.701)	2.056 (1.448–2.919)
LC: moderate overlap	1.319 (0.949–1.834)	2.647 (1.795–3.903)	0.963 (0.617–1.503)
LC: small overlap / oblique corner	1.331 (0.939–1.887)	1.966 (1.309–2.953)	1.072 (0.680–1.692)
Opponent: pole (vs. car)	2.345 (1.565–3.514)	1.721 (1.144–2.588)	2.083 (1.489–2.914)
Opponent: truck (vs. car)	2.078 (1.327–3.253)	1.439 (1.132–1.829)	3.111 (2.098–4.613)

Odds ratio sensitivity to change of technical crash severity predictor

To assess the sensitivity of the odds ratios seen in Fig. 1, the model is generated with switched technical crash severity predictors in the following two figures Fig. A1 and Fig. A2. Most notably, the effect of vehicle crash weight is only present in models not based on the momentum-based metric delta-v. Otherwise, the effects are mostly stable.

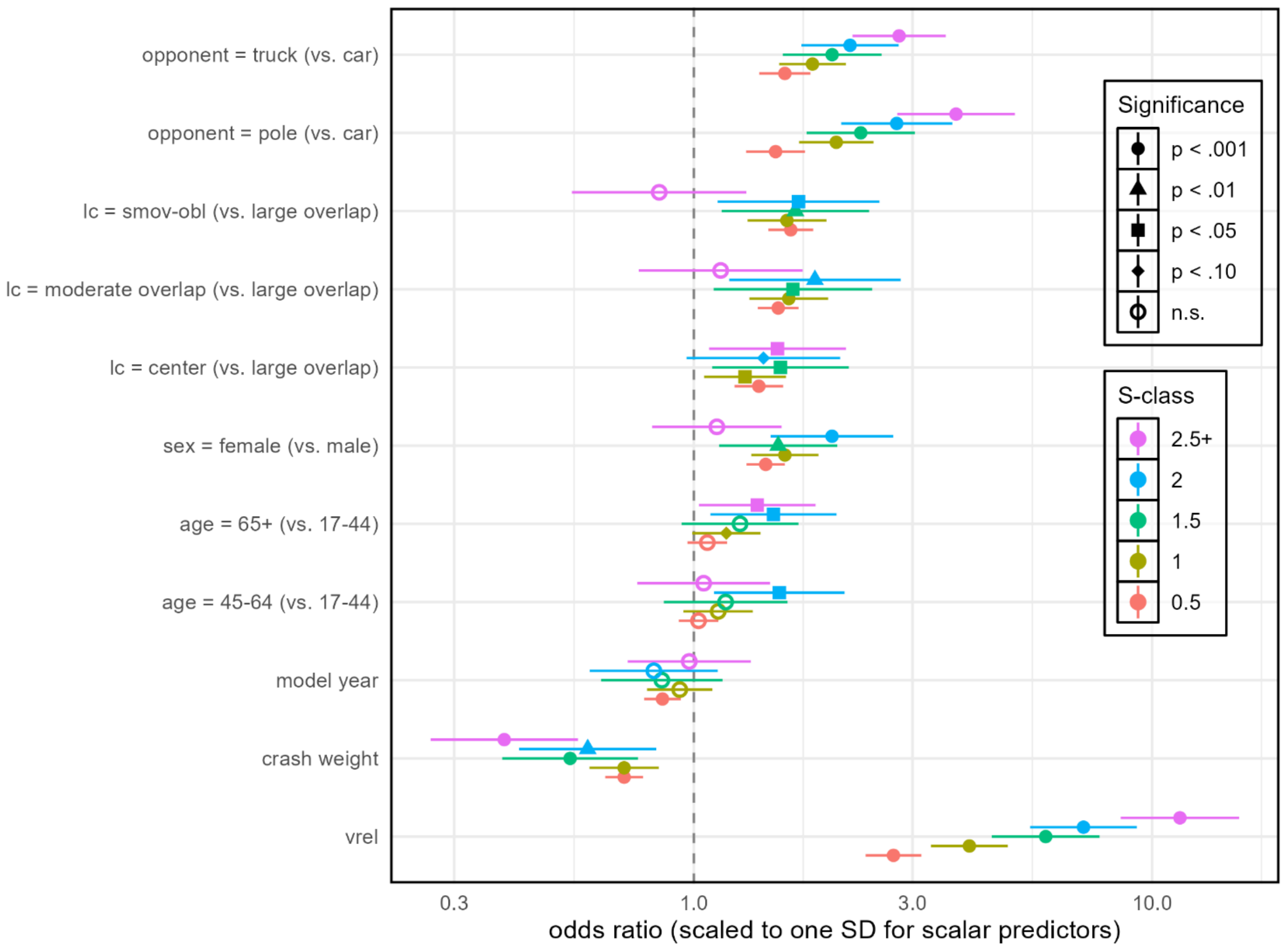

Fig. A1. Odds ratios (ORs) of sustaining an injurious outcome based on the relative collision velocity as the continuous crash severity indicator (SD = 27 km/h). ORs are standardised to one standard deviation for continuous predictors and scaled logarithmically. Symbols indicate OR significance, while the bands show their 95% confidence interval. Notably, the R^2_{MF} for this model dropped to 0.184.

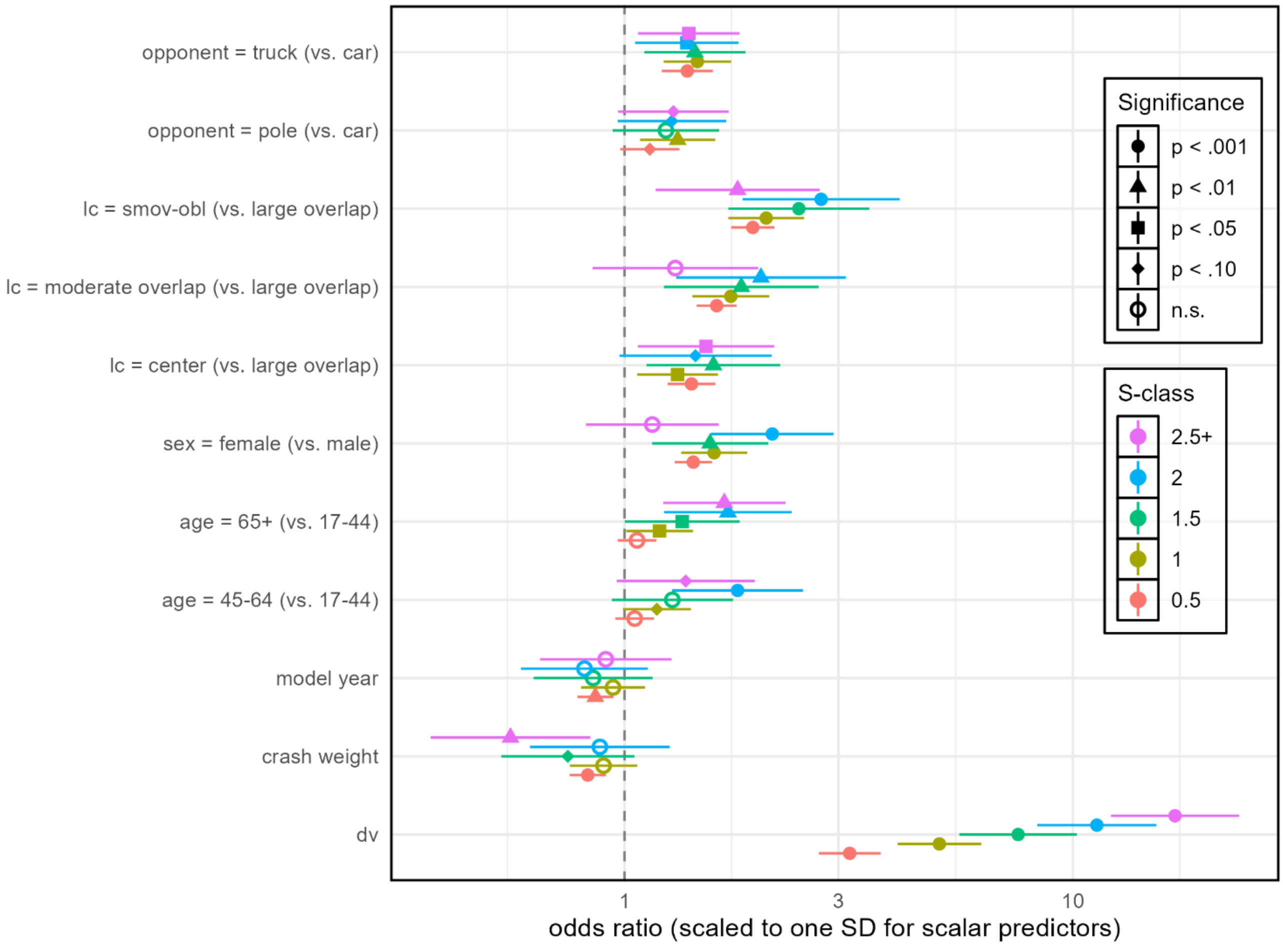

Fig. A2. Odds ratios (ORs) of sustaining an injurious outcome based on delta-v as the continuous crash severity indicator (SD = 15 km/h). ORs are standardised to one standard deviation for continuous predictors and scaled logarithmically. Symbols indicate OR significance, while the bands show their 95% confidence interval. The R^2_{MF} remained at a similar value at 0.197.

VIII. GLOSSARY

AIS
Abbreviated Injury Scale. A standardised anatomical-based coding system used to classify individual injuries by severity, primarily according to their threat to life.

AIS-ID: Unique identifier of AIS injury descriptions
AIS-CD: AIS-Code, rating of the injury severity
AIS Level 1: AIS-ID aggregation and abstraction to account for coding inaccuracies and improve the underlying correlation with mortality, based on [4]
FCI: Functional Capacity Index. A component of the AIS 2015 codebook that quantifies the expected long-term functional impairment (permanent partial disability) associated with a given injury.

Closed-group Injury Risk Function
A multinomial injury risk model that estimates the probability distribution across a set of mutually exclusive and collectively exhaustive injury severity categories (as opposed to conventional open-ended threshold models such as MAIS2+ or MAIS3+). Each case is assigned to exactly one severity group, allowing the full redistribution of risk across severity levels to be observed.

EES
Energy-Equivalent Speed. A crash severity metric that reflects the vehicle's deformation energy and is used as the primary technical predictor of injury risk in this study. As opposed to metrics like the delta-v, it is independent of the collision dynamics, as it only assesses the post-crash static deformation of the vehicle.

GIDAS
German In-Depth Accident Study. A detailed, representative in-depth accident database in Germany that provides high-quality technical crash data and medical injury documentation.

ISO 26262 S-scale
An extended six-level injury severity classification (S0 to S3) adapted from functional safety standards for hazard and risk analysis in the automotive industry.

LTC
Long-Term Consequences. Non-fatal injuries that result in permanent or prolonged functional impairment, reduced quality of life, or disability, which are not adequately captured by traditional AIS-based lethality-focused severity metrics.

MAIS
Maximum Abbreviated Injury Scale, the simplest person-level injury aggregator. The highest AIS severity score of any single injury sustained by an individual. Conventional IRFs often use binary thresholds such as MAIS2+ or MAIS3+.

Multinomial Logistic Regression
The statistical modelling technique used in this study. It estimates separate sets of coefficients for each injury severity outcome category (relative to a reference category, here S0), allowing severity-specific effects of predictors.

Open-ended (threshold) IRF
Conventional binary or dichotomous injury risk function that predicts only whether injury severity exceeds a chosen cutoff (e.g. probability of MAIS2+). All severities at or above the threshold are grouped together.

RAPEX HARM
A four-level injury severity scale developed within the European RAPEX framework that integrates both lethality (threat to life) and long-term consequences (LTC) into a single severity metric.