

## A Method to Estimate the Number of People Suffering from Long-Term Consequences

Susanne Schönebeck<sup>1</sup>, Albine Chanove<sup>2</sup>, Armira Kontaxi<sup>3</sup>, Ilaria Marino<sup>4</sup>, Marcus Wisch<sup>1</sup>, Jacques Saadé<sup>2</sup>, David Hynd<sup>4</sup>, Lynne Smith<sup>4</sup>, George Yannis<sup>3</sup>

<sup>1</sup>Federal Highway and Transport Research Institute, Bergisch Gladbach, Germany

<sup>2</sup>European Centre for Safety Studies and Risk Analysis (CEESAR), Nanterre, France,

<sup>3</sup>National Technical University of Athens, Department of Transportation Planning and Engineering, Athens, Greece

<sup>4</sup>TRL LIMITED, Wokingham, United Kingdom

### Introduction

Road traffic injuries remain a major public health challenge in Europe, not only because of the number of persons killed each year, but also due to the substantial burden carried by survivors who sustain serious and long-lasting consequences. While fatality figures are well documented, serious non-fatal injuries and their **long-term impacts remain largely undocumented**.

This stems from a structural limitation: **national crash databases rarely include medically validated information** on injury severity or long-term disability outcomes. Bridging this information gap is essential for supporting evidence-based road safety policies and monitoring progress toward EU targets for reducing serious injuries.

### Objectives

This research addresses this gap by **developing and applying a harmonised methodology** to estimate:

- the **number of seriously injured persons** (defined as MAIS3+ according to the Abbreviated Injury Scale, AIS)
- the **number of persons who suffer long-term consequences** (LTC) following road traffic crashes in Europe.

### Methodology

The methodology starts with an **assessment of available in-depth and national data** from Germany, Great Britain, and France. It then applies decision tree models to the in-depth data to identify key factors influencing the likelihood of serious injuries or long-term consequences. These models generate probability estimates, which are subsequently applied to national datasets to determine the share of affected individuals in each country.

Finally, **countries are grouped based on common variables available** in the CARE database European road accidents, and weighting factors are used to link national results to the European level, enabling consistent estimation of serious injuries (MAIS3+) and long-term consequences across Europe.

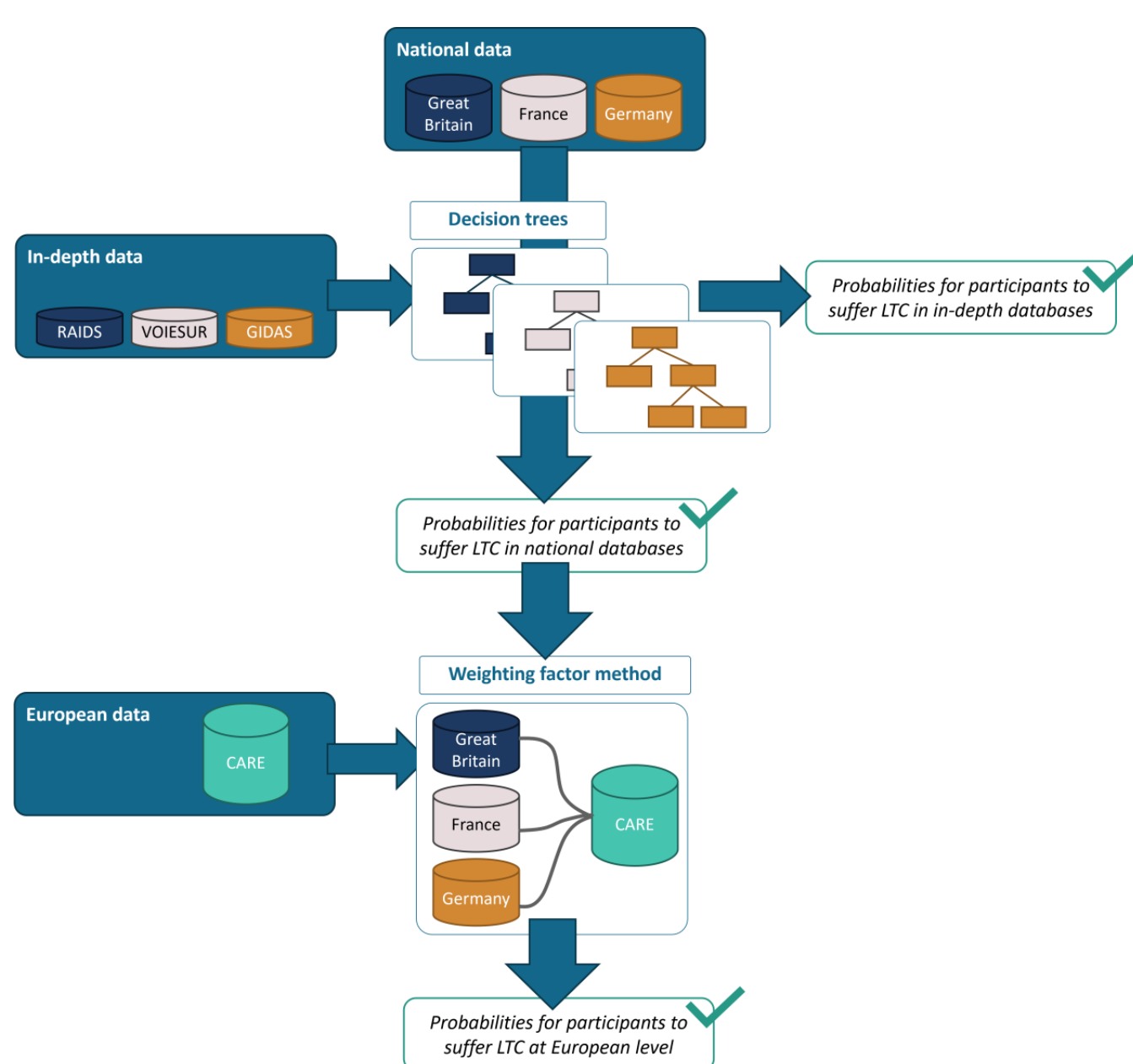


Figure 1: Schema of the approach

### Available datasets

The analysis combines **in-depth and national road crash datasets** from Great Britain, Germany, and France. In-depth databases include RAIDS, GIDAS and VOIESUR providing detailed information on crash mechanisms, vehicles, and injuries, including long-term consequences, though differing in methodology and sample size. National police datasets, STATS19, BAAC and German national statistics, offer large-scale, representative data on crashes and casualties, but with less detail.

At the European level, the **CARE database European road accidents** (2023) is used to upscale results. Using the harmonized CADaS structure, CARE enables cross-country comparisons despite differences in injury definitions, supporting consistent estimation of serious injuries and long-term consequences across Europe.

### Decision trees on the in-depth data

Decision tree models show moderate performance in **predicting MAIS3+ injuries**, with better identification of non-severe cases than severe ones. The **most influential factors** are:

- Road user type
- Impact configuration
- Road environment
- Age

These **patterns are consistent** across France, Germany and Great Britain, supporting their use for EU-level upscaling.

Decision tree models for long-term consequences (**pFCI < 5**) **show similar structure** to MAIS3+ models. Long-term outcomes are influenced by the same key variables, but reflect additional variability linked to recovery and functional impairment.

Applying the models to national datasets enables estimation of long-term consequences, which are not directly recorded in police data.

### Upscaling to National Databases

Assuming that the in-depth databases are representative of their national data, each national data is evaluated using the tree of its **corresponding in-depth database**.

**Only seriously injured participants** are considered (national definition). This ensures that these participants are assigned to a specific node based on the identified factors and receive a ratio estimating the probability of LTC at the national level.

### Upscaling to European Level

The European upscaling uses a **weighting factor method to link national datasets with the CARE**. Harmonised variables, selected based on relevance, expert knowledge, and availability, are used to group national data.

Weighting factors compare the distribution of serious injuries between national datasets and CARE, reflecting how representative national data are and allowing adjustment to EU level.

For each subgroup  $g$ , **weighting factors were computed** as:

$$WF_g = \frac{SI_{g, \text{country}} / SI_{\text{country, total}}}{SI_{g, \text{EU}} / SI_{\text{EU, total}}} \quad (1)$$

The **extrapolated serious injuries** for subgroup  $g$  are then:

$$\widehat{SI}_g = SI_{g, \text{country}} \times WF_g \quad (2)$$

Since CARE does not include MAIS3+ or pFCI data, **severity ratios from national datasets are used**:

$$SR_{\text{country}}^{\text{MAIS3+}} = \frac{\text{MAIS3+}_{\text{country}}}{SI_{\text{country}}} \quad \text{and} \quad SR_{\text{country}}^{\text{pFCI}} = \frac{\text{pFCI}_{\text{country}}}{SI_{\text{country}}} \quad (3)$$

These **ratios are applied to the upscaled counts** to estimate outcomes:

$$\widehat{\text{MAIS3+}}_g = \widehat{SI}_g \times SR_{\text{country}}^{\text{MAIS3+}} \quad \text{and} \quad \widehat{\text{pFCI}}_g = \widehat{SI}_g \times SR_{\text{country}}^{\text{pFCI}} \quad (4)$$

Summing across all subgroups provides the final European estimates for serious injuries (MAIS3+) and long-term consequences (pFCI).

**MAIS3+ cases in the EU-27 plus Great Britain:**

- 83,400 MAIS3+ persons ( $\pm 26\%$ ).

Estimated number of persons suffering from long-term consequences (**LTC in the EU-27 plus Great Britain:**

- 68,883 persons suffering from LTC ( $\pm 10.6\%$ )

### Conclusions

This study demonstrates a **novel and scalable methodology** combining in-depth crash data, national statistics, and European data (CARE) to estimate serious injuries and long-term consequences.

The approach:

- **Bridges the gap** between medical injury data and police records
- Enables **EU-wide comparable estimates**
- Supports **evidence-based** road safety policy

However, **results should be interpreted as first estimates**, due to several limitations:

- Differences in injury definitions across countries
- Limited availability of medically coded data
- Underreporting and data inconsistencies

**Further work** will focus on refining the methodology and improving data harmonisation across Europe.

### Acknowledgments

This project has received funding from the Horizon Europe programme under grant agreement No GAP-101146652.

**IMPROVA**

### Contact Information:

Armira Kontaxi, PhD, Research Associate NTUA  
Department of Transportation Planning and Engineering  
Email: akontaxi@mail.ntua.gr  
Website: <https://www.nrso.ntua.gr/p/akontaxi/>