What factors affect the match rates of police and emergency department data in Sweden's road traffic crash database Strada?

Christian Howard, Åsa Forsman
VTI (Swedish National Road and Transport Institute)
Linköping Sweden

Highlights

- The police match rate is 0.345
- The emergency department match rate is 0.197
- The match rate difference is due to police capturing fewer single-bicycle crashes
- Match rates vary greatly between counties
- Bus occupants have low police and emergency department match rates
- Higher injury severity is clearly associated with higher match rates

Introduction

Linking police-based datasets with hospital-based datasets is valuable for a deeper understanding of road traffic crashes. Whereas police crash data often contain information on the location and circumstances of a crash, estimations of injury severity can be inaccurate. Conversely, hospital data can provide detailed injury information but may only provide limited information on crash circumstances. Furthermore, the two sources may capture different parts of the crash and injury population as neither source is complete.

Strada (the Swedish national registry of road traffic crashes) is rather unique in that both the police and the hospital emergency departments (EDs) report to the system. Reports are matched by the system continuously and the resulting dataset is made available to selected authorities through an online portal.

The aim of this study is to investigate what factors affect the match rates of police and emergency department data in Strada.

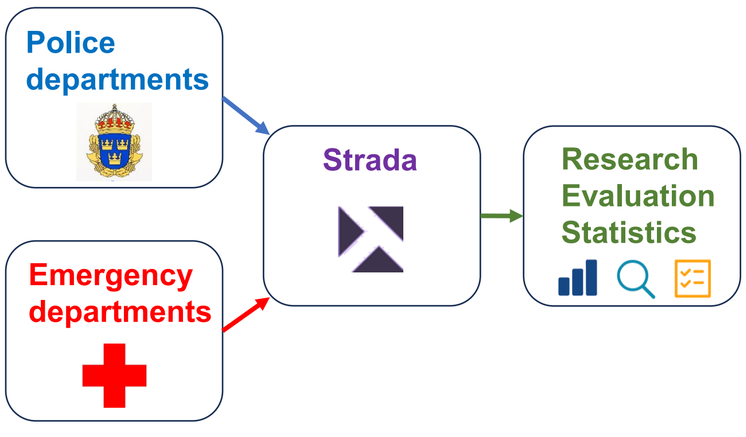

Figure 1. The Strada information system

Method

The linkage of records is based on personal identity numbers, time of the crash, and location of the crash.

The study data are from 2022–2023 and consist of 93,982 persons involved in a road crash. The linking performance was measured as:

$$Police\ match\ rate = \frac{Number\ of\ matched\ records}{Total\ number\ of\ police\ records}$$

$$ED\ match\ rate = \frac{Number\ of\ matched\ records}{Total\ number\ of\ ED\ records}$$

Binary logistic regression is applied to investigate what factors that influence the match rate, considering *injury severity*, *sex*, *age group*, *road user category*, and *county*.

Results

The logistic regression shows significant match rate differences regarding most of the factors. The estimated police match and ED match rates per factor are shown in Figure 2 and Figure 3. Large differences per county were also found.

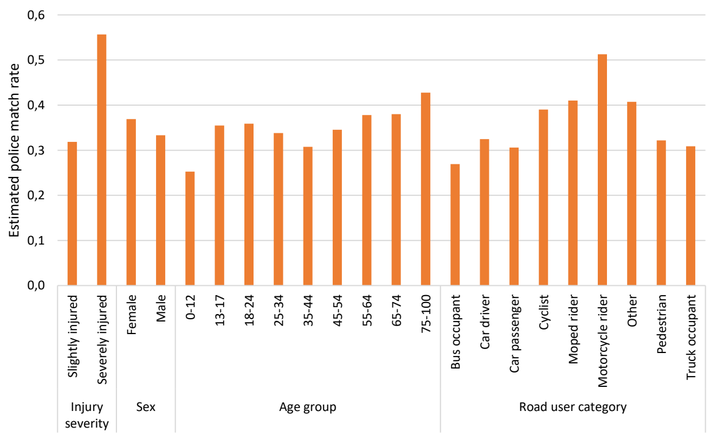

Figure 2. Estimated police match rates

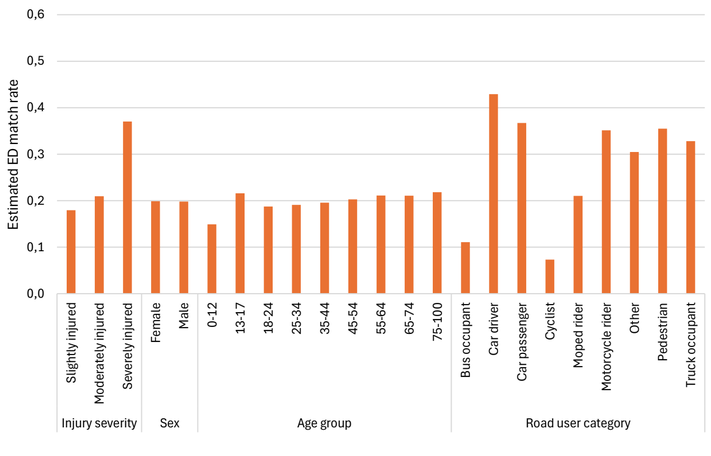

Figure 3. Estimated ED match rates

Conclusions

The police match rate of 0.345 is higher than the ED match rate of 0.197. When excluding single-bicycle crashes, the ED match rate is close to the police match rate.

The match rates also vary significantly between other road user categories, with particularly noteworthy higher match rates for police reported motorcycle riders and lower match rates for both police and ED reported bus occupants.

Consistent with the previous literature, large match rate differences were observed depending on injury severity, where higher injury severities were associated with higher match rates.

A slightly higher police match rate was observed for females compared to males.

The age group 0–12 had significantly lower match rates than all other age groups.

The match rates can vary greatly depending on region. This highlights that results from previous studies focusing on one region or hospital may not be transferable within the same reporting system and country.

How uninjured and unknowns are handled in the analysis impact the estimated police match rates, whereas ED match rates are hardly affected. Therefore, comparisons with other studies are only meaningful if inclusion criteria are considered.

This poster is based on:
Forsman, Å., & Howard, C. (2026). What factors affect the match rates of police and emergency department data in Sweden's road traffic crash database Strada? Manuscript submitted for publication.

Christian Howard
christian.howard@vti.se

Åsa Forsman
asa.forsman@vti.se

vti