



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation UVEK
Département fédéral de l'environnement, des transports, de l'énergie et de la communication DETEC
Dipartimento federale dell'ambiente, dei trasporti, dell'energia e delle comunicazioni DATEC

Bundesamt für Strassen
Office fédéral des routes
Ufficio federale delle Strade

MEDUSA: Methodik zur differenzierten Quantifizierung der Dunkelziffer der Strassenverkehrsunfälle in der Schweiz

MEDUSA: Méthode de quantification différenciée du sous-enregistrement des accidents de la route en Suisse

MEDUSA: Methodology for the differentiated quantification of the number of unreported road traffic accidents in Switzerland

BFU, Beratungsstelle für Unfallverhütung

Patrizia Hertach
Steffen Niemann
Markus Deublein

ECOPLAN AG, Forschung und Beratung in Wirtschaft und Politik

Matthias Amacher
Martin Wagenbach
Christoph Lieb

ETH Zürich

Markus Kalisch
Lukas Meier

Forschungsprojekt MFZ_20_06B_03 auf Antrag des ASTRA

Juli 2025

1796

Der Inhalt dieses Berichtes verpflichtet nur den (die) vom Bundesamt für Strassen unterstützten Autor(en). Dies gilt nicht für das Formular 3 "Projektabschluss", welches die Meinung der Begleitkommission darstellt und deshalb nur diese verpflichtet.

Bezug: Schweizerischer Verband der Strassen- und Verkehrsfachleute (VSS)

Le contenu de ce rapport n'engage que les auteurs ayant obtenu l'appui de l'Office fédéral des routes. Cela ne s'applique pas au formulaire 3 « Clôture du projet », qui représente l'avis de la commission de suivi et qui n'engage que cette dernière.

Diffusion : Association suisse des professionnels de la route et des transports (VSS)

La responsabilità per il contenuto di questo rapporto spetta unicamente agli autori sostenuti dall'Ufficio federale delle strade. Tale indicazione non si applica al modulo 3 "conclusione del progetto", che esprime l'opinione della commissione d'accompagnamento e di cui risponde solo quest'ultima.

Ordinazione: Associazione svizzera dei professionisti della strada e dei trasporti (VSS)

The content of this report engages only the author(s) supported by the Federal Roads Office. This does not apply to Form 3 'Project Conclusion' which presents the view of the monitoring committee.

Distribution: Swiss Association of Road and Transportation Experts (VSS)



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation UVEK
Département fédéral de l'environnement, des transports, de l'énergie et de la communication DETEC
Dipartimento federale dell'ambiente, dei trasporti, dell'energia e delle comunicazioni DATEC

Bundesamt für Strassen
Office fédéral des routes
Ufficio federale delle Strade

MEDUSA: Methodik zur differenzierten Quantifizierung der Dunkelziffer der Strassenverkehrsunfälle in der Schweiz

MEDUSA: Méthode de quantification différenciée du sous-enregistrement des accidents de la route en Suisse

MEDUSA: Methodology for the differentiated quantification of the number of unreported road traffic accidents in Switzerland

BFU, Beratungsstelle für Unfallverhütung

Patrizia Hertach
Steffen Niemann
Markus Deublein

ECOPLAN AG, Forschung und Beratung in Wirtschaft und Politik

Matthias Amacher
Martin Wagenbach
Christoph Lieb

ETH Zürich

Markus Kalisch
Lukas Meier

Forschungsprojekt MFZ_20_06B_03 auf Antrag des ASTRA

Juli 2025

1796

Impressum

Forschungsstelle und Projektteam

Projektleitung

Patrizia Hertach

Mitglieder

Matthias Amacher

Markus Kalisch

Steffen Niemann

Lukas Meier

Martin Wagenbach

Christoph Lieb

Markus Deublein

Begleitkommission

Präsident

Christian Kamenik, Bundesamt für Strassen ASTRA

Mitglieder

Joséphine Leuba, Bundesamt für Raumentwicklung ARE

Wernher Brucks, Stadt Zürich, Dienstabteilung Verkehr

Stefan Scholz, Schweizerische Unfallversicherungsanstalt Suva, Sammelstelle für die Statistik der Unfallversicherung SSUV

Alexandra Quandt, Bundesamt für Statistik BFS

Nicole Felder, Bundesamt für Strassen ASTRA

Martin Urwyler, Bundesamt für Strassen ASTRA

Antragsteller

Schweizerisches Bundesamt für Strassen, ASTRA

Arbeitsgruppe MFZ – Mensch und Fahrzeug

Bezugsquelle

Das Dokument kann kostenlos von <http://www.mobilityplatform.ch> heruntergeladen werden.

Inhaltsverzeichnis

Impressum	4
Zusammenfassung	7
Résumé	14
Summary	21
1 Einleitung	27
1.1 Ausgangslage und Problemstellung	27
1.2 Projektziel und Abgrenzung	28
1.3 Verwendete Begriffe	28
2 Methodik.....	31
2.1 (Inter-)Nationaler Stand der Forschung	31
2.2 Bedürfnisse der Stakeholder	31
2.3 Auswahl statistischer Modelle	32
2.4 Potenziell nutzbare Datengrundlagen	32
2.5 Erarbeitung der Methodenvorschläge	32
3 (Inter-)Nationaler Stand der Forschung	34
3.1 Nationale Literatur	34
3.2 Internationale Literatur	34
3.2.1 Methoden zur Schätzung der Dunkelziffer (IRTAD).....	34
3.2.2 Differenzierungsmerkmale mit Einfluss auf die Erfassungswahrscheinlichkeit	35
4 Bedürfnisse der Stakeholder	39
4.1 Anwendungsfelder der Dunkelziffer	39
4.2 Wissenslücken	39
4.3 Frequenz der Erhebung	39
4.4 Zu verwendende «Abgrenzungen» (Begriffe siehe Kapitel 1.3)	39
4.5 Differenzierungswünsche	40
4.6 Priorisierung und Einordnung.....	40
5 Auswahl statistischer Modelle	42
5.1 Grundlegende Überlegungen	42
5.2 Stichprobe bzw. Befragungen	42
5.3 Datenspuren	43
5.3.1 Zu beachtende Aspekte: Filter, Linkage, Erweiterung	44
5.3.2 Realistischeres Beispiel	46
5.3.3 Mögliche Probleme	47
5.4 Maschinelles Lernen	48
6 Potenziell nutzbare Datengrundlagen.....	51
6.1 Strassenverkehrsunfallstatistik des ASTRA.....	51
6.2 Stichproben- bzw. Befragungsdaten	52
6.2.1 Übersicht erschlossene Daten	52
6.2.2 Weitere bisher nicht erschlossene Möglichkeiten / Varianten	54
6.3 Daten für Datenspuren / Linkage	54
6.3.1 Übersicht erschlossene Daten	55
6.3.2 Beurteilung der verfügbaren Datenquellen	57
6.3.3 Nicht erschlossene Quellen	61
6.4 Weitere Datenquellen und Daten für maschinelles Lernen	64
6.5 Zusammenfassung Datenquellen	67

7	Methodenvorschläge	69
7.1	Dunkelziffer für Getötete	69
7.2	Dunkelziffer für Verletzte	71
7.2.1	Methode 1: Stichprobe (Befragung) inkl. Hochrechnung	71
7.2.2	Methode 2: Linkage bestehender Datenquellen inkl. Capture-Recapture	76
7.2.3	Methode 3: Versicherungsliste (SSUV und KVG)	86
7.2.4	Übersicht / Beurteilung der Methoden zur Abschätzung der Dunkelziffer für Verletzte	86
8	Gesamtbeurteilung und Empfehlungen	90
8.1	Gesamtbeurteilung der Methoden zur Berechnung der Dunkelziffer für Verletzte	90
8.1.1	Bewertungshinweise der Stakeholder	90
8.1.2	Methode 1: Befragung	90
8.1.3	Methode 2: Linkage ASTRA-SSUV + CR	91
8.1.4	Methode 3: Versicherungsliste	91
8.1.5	Weitere Datenquellen / Methoden für Verletzte	92
8.2	Empfehlungen	92
8.2.1	Dunkelziffer für Getötete	92
8.2.2	Dunkelziffer für Verletzte	93
9	Forschungsbedarf	96
9.1	Methode 1: Befragung	96
9.2	Methode 2: Linkage ASTRA-SSUV + CR	96
9.3	Methode 3: Versicherungsliste	97
9.4	Weitere Datenquellen / Methoden	97
	Anhänge	98
	Abkürzungsverzeichnis	143
	Literaturverzeichnis	144
	Projektabschluss	150

Zusammenfassung

Einleitung

Die vollständige Erfassung von Strassenverkehrsunfällen ist essenziell für eine effektive Verkehrssicherheitsarbeit. Die offizielle Statistik – die Strassenverkehrsunfallstatistik des Bundesamts für Strassen ASTRA – basiert auf polizeilichen Meldungen. Weil nicht alle Unfälle erfasst werden, wird das tatsächliche Unfallgeschehen unterschätzt. Darüber hinaus ist das Bild des tatsächlichen Unfallgeschehens verzerrt, da die Erfassungswahrscheinlichkeit von verschiedenen Faktoren abhängt, zum Beispiel von der Verletzungsschwere oder vom Unfallhergang: Alleinunfälle weisen beispielsweise eine geringere Erfassungswahrscheinlichkeit auf als Kollisionen.

Eine realistische und ausreichend differenzierte Dunkelzifferbestimmung ist entscheidend für die Schaffung einer verlässlichen Grundlage für Präventionsmassnahmen. Ebenfalls sind diese Daten wesentlich für volkswirtschaftliche Kostenanalysen, Unfallratenberechnungen und Normensetzungen. Das MEDUSA-Projekt hatte zum Ziel, eine Methodik zur differenzierten Bestimmung der Dunkelziffer zu entwickeln. Die Methodik sollte so konzipiert sein, dass sie eine Differenzierung von verschiedenen Verkehrsteilnehmergruppen, Unfallhergängen und der Verletzungsschwere ermöglicht.

Methodik

Zu Beginn des Projekts wurde eine Literaturrecherche durchgeführt, um nationale und internationale Ansätze zur Dunkelzifferermittlung zu analysieren. Insbesondere umfasste die Literaturrecherche auch die Analyse von infrage kommenden statistischen Modellen, auch aus anderen Themenbereichen (z.B. Biologie). Parallel dazu wurde eine breit gefasste Auslegeordnung und Beurteilung bereits existierender oder noch zu erschliessender Datenquellen vorgenommen, die Angaben zu Strassenverkehrsunfällen enthalten oder sich für die Erhebung von Unfalldaten und einen Abgleich mit den offiziellen Strassenverkehrsunfalldaten des ASTRA eignen.

Ein zentraler Bestandteil der Methodik war auch die Einbindung relevanter Stakeholder. In einem ersten Workshop wurden deren Bedürfnisse und Erwartungen beispielsweise bezüglich der Differenzierung oder der Periodizität der Dunkelzifferberechnung erhoben.

Die Ergebnisse und Erkenntnisse der verschiedenen Arbeitsschritte wurden miteinander verknüpft, um konkrete Vorschläge für priorisierte Varianten der Methodik zu erarbeiten. Im Rahmen der Methodenentwicklung wurden Simulationen für drei verschiedene methodische Varianten sowie eine Pilotberechnung für eine dieser Varianten durchgeführt. In einem zweiten Workshop wurden die ausgearbeiteten Methodenvarianten den Stakeholdern vorgestellt, bewertet und gemäss den drei Zielen Differenzierungen, Aufwand und Genauigkeit priorisiert.

Ergebnisse

(Inter-)Nationaler Stand der Forschung zur Dunkelziffer von Strassenverkehrsunfällen

Bereits 1990 wurde die Dunkelziffer von Strassenverkehrsunfällen in der Schweiz anhand von Unfallversicherungsdaten geschätzt. Dabei ergab sich, dass die tatsächliche Zahl der Unfälle etwa viermal höher war als polizeilich registriert. Seit 2015 gibt es für die Schweiz eine jährlich aktualisierte Schätzung, basierend auf einer Haushaltsbefragung von 2011. Die aktuellen Schätzungen für 2021 gehen von etwa 77 000 verletzten Personen aus, während nur rund 20 500 von der Polizei erfasst wurden. Besonders auffällig ist die hohe Dunkelziffer im Veloverkehr, bei dem etwa neunmal mehr Verletzte geschätzt werden als in der amtlichen Statistik verzeichnet sind.

Eine Umfrage unter den Mitgliedsländern der International Traffic Safety Data and Analysis Group (IRTAD) im Jahr 2021 zeigte, dass neun von 15 Ländern Dunkelziffern entweder einmalig oder wiederholt ermittelt haben. Von den Ländern, die Berechnungen zur Unterschätzung von Verletzten im Strassenverkehr in den Polizeidaten vornahmen, wurden in den meisten Fällen Spitaldaten beigezogen. Aber auch andere Datenquellen wie (Haus-)Befragungen oder Versicherungsdaten dienten dem Vergleich mit den polizeilich registrierten Unfällen. Die Differenzierung beschränkt sich meist auf die Verletzungsschwere und Verkehrsteilnahme, vereinzelt wird auch nach Unfallhergang oder Tageszeit unterschieden. Methodisch basieren die Schätzungen meist auf der Verknüpfung und Zusammenführung von (Register-)Daten. Zum Teil werden Gesamtunfallzahlen mittels eines Capture-Recapture-Verfahrens, das in der Biologie zur Schätzung unbekannter Populationsgrössen genutzt wird, hochgerechnet. Eine systematische Übersichtsarbeit aus dem Jahr 2021 ergab, dass in Ländern mit hohem Einkommen 35–97 % der Verkehrstoten und 16–82 % der Schwerverletzten in Polizeistatistiken erfasst sind.

Internationale Studien untersuchten Einflussgrössen auf die Erfassungswahrscheinlichkeit eines Unfalls. Als massgebend für die Erfassung von Unterschieden in der Erfassungswahrscheinlichkeit zeigen sich die Verletzungsschwere, der Unfallhergang (Anzahl Beteiligte), die Verkehrsteilnahme (Fahrzeugkategorie) und das Alter, wobei andere mögliche Faktoren (z. B. Raumtyp, Geschlecht) mit diesen Einflüssen interagieren. Wird nicht nach den ursächlichen Faktoren (Verletzungsschwere, Hergang, Verkehrsteilnahme, Alter) differenziert, – beispielsweise, weil die Daten nicht verfügbar sind –, kann dies zu Verzerrungen in den Schätzungen führen.

Bedürfnisse der Stakeholder

Im Rahmen eines Workshops mit Vertretern aus Bundesverwaltung, Polizei, Stadtverwaltung, der Sammelstelle für die Statistik der Unfallversicherung SSUV sowie der BFU, Beratungsstelle für Unfallverhütung, wurden zentrale Anforderungen an die Erfassung der Dunkelziffer von Strassenverkehrsunfällen gesammelt. Wichtigste Erkenntnisse: Die Daten sind essenziell für Unfallprävention, Verkehrsplanung, politische Entscheidungen und Erfolgskontrollen von Massnahmen in der Verkehrssicherheit. Besonders bei Mikromobilität, Elektromobilität und dem Veloverkehr bestehen erhebliche Wissenslücken zur Dunkelziffer. Eine umfassende Dunkelziffer-Analyse sollte alle fünf bis sieben Jahre vorliegen, während jährliche Fortschreibungen kurzfristige Mobilitätstrends berücksichtigen. Die Ergebnisse sollten spätestens 12 bis 18 Monate nach Verfügbarkeit der ASTRA-Unfalldaten vorliegen. Unfälle sollten optimalerweise nach dem Territorialitätsprinzip erfasst werden. Bezüglich einer Differenzierung werden die Fahrzeugart, die Verletzungsschwere, das Alter und das Geschlecht als zentrale Merkmale erachtet. Als weitere gewünschte Differenzierungsmöglichkeiten wurden der Unfallhergang, die Unfallursache, der Unfallzeitpunkt und der Raumtyp genannt.

Auswahl statistischer Modelle

Da eine vollständige Erhebung aller Unfälle in der Praxis nicht umsetzbar ist, müssen andere Ansätze zur Anwendung kommen, die eine nach den relevanten Merkmalen differenzierte Schätzung des Gesamtunfallgeschehens bzw. der Dunkelziffer erlauben. Dabei kommen drei grundlegende Ansätze in Frage: Stichproben, Datenspuren und maschinelles Lernen.

- **Stichproben bzw. Befragungen:** Mit diesem Ansatz können repräsentative Daten gesammelt werden. Die Methode ist flexibel und bewährt, jedoch mit hohem Aufwand verbunden und anfällig für Verzerrungen durch Erinnerungsfehler oder soziale Erwünschtheit.
- **Datenspuren / Linkage:** Bei diesem Ansatz geht es darum, zusätzliche Datenquellen zu nutzen, die differenzierte Informationen zu Strassenverkehrsunfällen enthalten. Hier kommen beispielsweise Daten von Versicherungen oder Spitälern in Frage. Die Anforderungen an die Datenquellen sind hoch: die Daten müssen es ermöglichen, Stras-

senverkehrsunfälle von anderen Ereignissen abzugrenzen, und nach den gewünschten Differenzierungsmerkmalen zu filtern. Zudem muss eine zuverlässige Verknüpfung (Linkage) mit den ASTRA-Daten möglich sein, um Verunfallte zu identifizieren, die in beiden Datensätzen vorkommen. Da kein Datensatz alle Unfälle enthält, wird in einem weiteren Schritt mit der Capture-Recapture-Methode auf die Grundgesamtheit der Unfälle geschlossen.

- **Maschinelles Lernen:** Mit diesem Ansatz können Unfälle anhand leicht messbarer Variablen vorhergesagt werden. Zum Aufbau der Vorhersagemodelle braucht es umfangreiche, nach den relevanten Merkmalen differenzierte und qualitativ hochwertige Trainingsdaten zu polizeilich sowohl erfassten als auch nicht erfassten Unfällen. Solche Daten sind allerdings in absehbarer Zeit nicht vorhanden oder nicht zugänglich.

Potenziell nutzbare Datengrundlagen

Im Folgenden werden für den Zweck der Dunkelzifferberechnung potenziell nutzbare Datenquellen aufgeführt und deren Eignung beurteilt. Die Einteilung erfolgte nach den oben identifizierten statistischen Ansätzen. Eine wichtige Voraussetzung dafür, dass eine Datenquelle in Frage kommt, ist, dass sich Strassenverkehrsunfälle im Einklang mit der entsprechenden Definition des ASTRA identifizieren lassen.

Stichproben bzw. Befragungen

Folgende Befragungen kommen für eine Dunkelzifferberechnung in Frage:

- **Haushaltsbefragung der BFU:** Die Stichprobe für diese Befragung wurde aus dem Stichprobenrahmen für Personen- und Haushaltserhebungen (SRPH) des Bundesamts für Statistik (BFS) gezogen und erlaubt eine Identifizierung von Strassenverkehrsunfällen sowie eine Differenzierung dieser Unfälle nach Verkehrsteilnahme und Unfallhergang. Die letzte Befragung fand 2024/2025 statt und umfasste ungefähr 25 000 teilnehmende Haushalte. Falls diese Befragung weitergeführt wird, eignet sie sich auch für zukünftige Dunkelzifferabschätzungen.
- **Mikrozensus Mobilität und Verkehr (MZMV):** Die Stichprobe für diese ungefähr alle fünf Jahre vom BFS sowie vom Bundesamt für Raumentwicklung (ARE) durchgeführte Haushaltsbefragung wird ebenfalls aus dem SRPH gezogen. Sie enthält aktuell noch keine Fragen zum Unfallgeschehen, würde sich aber durch eine entsprechende Erweiterung des Fragekatalogs für eine Dunkelzifferberechnung eignen. Am letzten MZMV aus dem Jahr 2021 nahmen rund 55 000 Personen teil.

Nicht geeignet ist hingegen in ihrer aktuellen Form die Schweizerische Gesundheitsbefragung (SGB) des BFS, die zwar bereits Unfälle erfasst und regelmässig wiederholt wird, aber keine präzise Identifikation von Strassenverkehrsunfällen erlaubt sowie kaum Differenzierungsmöglichkeiten bietet.

Ein nicht weiter verfolgter Ansatz ist die gezielte Befragung von Verletzten in Krankenhäusern nach dem Vorbild der European Injury Database (EU-IDB). Dieser könnte zwar wertvolle Daten liefern, ist jedoch personal- und kostenintensiv, erfasst überwiegend mittelschwere bis schwere Unfälle und ist nur repräsentativ, wenn alle Spitäler teilnehmen.

Datenspuren / Linkage

Bereits erschlossene Datenquellen, die für einen Linkage-Ansatz geeignet sind:

- **Statistik der Unfallversicherung UVG (Unfallversicherungsgesetz; SSUV):** Liefert differenzierte Informationen zu Unfällen und lässt sich mit den ASTRA-Daten verknüpfen. Der Datensatz ist jedoch auf die UVG-versicherte Bevölkerung beschränkt, also alle in der Schweiz beschäftigten Arbeitnehmenden, und enthält somit keine Daten zu Kindern, Hausfrauen und -männern, Rentnerinnen und Rentnern sowie Studierenden.
- **Todesursachenstatistik des BFS:** Ermöglicht einen zuverlässigen Abgleich mit den ASTRA-Daten für tödliche Unfälle.

Die Medizinische Statistik der Krankenhäuser (MSK) und Patientendaten Spital ambulant (PSA) enthalten zwar Daten zu Verletzungen, erlauben aber keine präzise Abgrenzung

von Strassenverkehrsunfällen. Ausserdem erfassen sie vor allem schwere Verletzungen. Daten von Arztpraxen könnten auch leichtere Verletzungen aufzeigen, sind jedoch nicht erschlossen. Eine Erschliessung erscheint aktuell unverhältnismässig.

Weitere Datenquellen, die für den Linkage-Ansatz eingesetzt werden könnten, sind die Unfalldaten aus der gesetzlichen Krankenversicherung (Bundesgesetz über die Krankenversicherung, KVG) oder jene der Motorfahrzeugversicherer. Diese beiden Datenquellen sind allerdings aktuell noch nicht erschlossen und eine Erschliessung dürfte sehr aufwendig sein. Die Daten zu KVG-Unfallversicherten (beinhaltet alle nicht UVG versicherten Personen mit Wohnsitz in der Schweiz) könnten zusammen mit den UVG-Daten eine «Versicherungsliste» oder eine repräsentative Stichprobe aller ärztlich behandelten Unfallopfer bilden. Daten von Motorfahrzeugversicherungen könnten Informationen insbesondere zu Sachschäden, verursacht durch motorisierte Fahrzeuge, liefern. Eine Erschliessung dieser Daten erscheint ohne entsprechende gesetzliche Grundlage aktuell jedoch nicht realistisch.

Weitere Datenquellen und Daten für maschinelles Lernen

Folgende weitere Datenquellen wurden auf ihre Eignung zur Erfassung von Strassenverkehrsunfällen in Betracht gezogen: Anamnesetexte in Spitälern, Smartphone-/Smart-Device-Funktionen, Mobile-Apps, Meldesysteme, Videoaufnahmen und Fahrzeugsensoren. Diese Datenquellen sind derzeit aus verschiedenen Gründen für eine Schätzung einer Dunkelziffer von Strassenverkehrsunfällen noch nicht geeignet, sie bergen aber ein grosses Potenzial für eine zukünftige Nutzung. Herausforderungen bestehen unter Anderem im Datenschutz sowie in der Erschliessung der Datenquellen. Die blosser Feststellung und Identifikation eines Strassenverkehrsunfalls genügt dabei nicht, um die Daten für eine differenzierte Dunkelzifferberechnung nutzen zu können. Für eine erfolgreiche Nutzung solcher Daten in Modellen des maschinellen Lernens müssen möglichst vollständige und qualitativ hochwertige Daten zu Strassenverkehrsunfällen und deren Folgen vorhanden sein. Dies ist momentan noch nicht der Fall.

Methodenvorschläge

Aktuell sind die beiden Ansätze «Stichproben» und «Datenspuren» am erfolgversprechendsten. Maschinelles Lernen ist derzeit aufgrund fehlender Trainingsdaten nicht anwendbar, birgt aber ein grosses Potenzial und sollte deshalb zukünftig regelmässig neu evaluiert werden. Im Folgenden wird aufgrund unterschiedlicher Datengrundlagen zwischen Getöteten und Verletzten unterschieden.

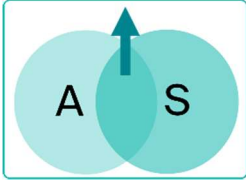
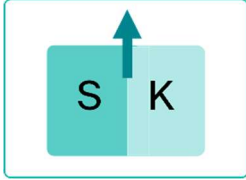
Dunkelziffer für Getötete

Für die Dunkelzifferberechnung für Getötete könnten die Daten der Todesursachenstatistik des BFS hinzugezogen werden. Diese zeigen aber gegenüber den ASTRA-Daten geringe Abweichungen. Diese Differenzen ergeben sich u. a. durch unterschiedliche Zählweisen: das ASTRA richtet sich nach dem Unfalljahr, das BFS nach dem Sterbejahr. Zudem verwendet ASTRA das Territorialprinzip (alle Getöteten in der Schweiz), während das BFS dem Inländerprinzip folgt (Todesfälle der Wohnbevölkerung im In- und Ausland). Auch die 30-Tage-Regel in der Definition für Verkehrstote beim ASTRA sowie die Erfassung über ICD-10-Kodes in der Todesursachenstatistik (V-Codes) spielen eine Rolle. Einige Abweichungen könnten durch eine verbesserte Datenerfassung in der Todesursachenstatistik – z. B. Unfalldatum – ausgeglichen werden, jedoch fehlen diese Informationen häufig oder sind fehlerhaft. Man könnte die statistischen Unschärfen auch durch einen Korrekturfaktor ausgleichen. Die WHO setzt solche Korrekturfaktoren für ihre Abschätzungen ein. Für die Schweiz liegt dieser derzeit bei 1,03.

Dunkelziffer für Verletzte

Für Verletzte sind drei Methoden denkbar. Jede Methode hat Vor- und Nachteile.

- **Methode 1: Befragung mit Hochrechnung** (symbolisiert mit Pfeil): Befragungen ermöglichen eine direkte Erhebung von Strassenverkehrsunfällen. Sie können als eigenständige Erhebung konzipiert oder in bestehende Umfragen integriert werden. Für das Jahr 2024 kann die Haushaltsbefragung der BFU als Grundlage für eine Dunkelzifferberechnung dienen. Ebenfalls wäre eine Integration von Fragen zu Strassenverkehrsunfällen in den MZMV denkbar. Vorteile der Methode sind die Flexibilität und Differenzierbarkeit der Daten. Ausserdem handelt es sich um einen bewährten und kurzfristig umsetzbaren Ansatz. Nachteile liegen darin, dass Befragungen kostenintensiv und anfällig für Recall Bias und sozial erwünschte Antworten sind. Zudem ist eine grosse Stichprobe erforderlich, um ausreichend belastbare Ergebnisse zu erzielen.

- **Methode 2: Linkage der Daten von ASTRA (A) und SSUV (S) mit Capture-Recapture** (Methode zur Hochrechnung, symbolisiert mit Pfeil): Vorteil dieser Methode ist die kostengünstigere Umsetzung. Ein grosser Nachteil ist, dass mit ihr keine Dunkelzifferberechnungen für Kinder und Senioren/-innen vorgenommen werden können. Die Verletzungsschweren wird zwar in beiden Datensätzen differenziert erfasst, jedoch in einer anderen Kategorisierung, was eine Dunkelzifferberechnung nach Verletzungsschwere verunmöglicht. Mit der Capture-Recapture-Methode kann dennoch eine Hochrechnung auf das Gesamtunfallgeschehen vorgenommen und damit die Dunkelziffer geschätzt werden. Eine Verknüpfung zwischen ASTRA- und SSUV-Daten kann mittels probabilistischer Methoden erfolgen, stellt aber eine Herausforderung dar und bedarf deshalb weiterer Forschung. Eine Harmonisierung der beiden Datenquellen, insbesondere bezüglich der Verletzungsschwere, wäre zukünftig gewinnbringend. Auch hierzu braucht es zusätzliche Forschung.

- **Methode 3: Versicherungsliste aus den Daten von SSUV (S) und KVG (K) mit Hochrechnung** (symbolisiert mit Pfeil): Langfristig wäre dieser Ansatz vielversprechend und könnte – sobald aufgegleist – relativ kostengünstig umgesetzt werden. Dafür müssten allerdings erst die Daten der KVG-Versicherer erschlossen werden. Ähnlich wie bei Methode 2 ist eine Harmonisierung der drei Datenquellen erforderlich.


Die Herausforderung bei allen drei Methoden ist die Berücksichtigung von Personen, die in der Schweiz verunfallen, aber nicht in der Schweiz wohnhaft oder erwerbstätig sind. In zukünftigen Forschungsvorhaben sollte geklärt werden, wie diese Personen dennoch in der Ermittlung der Dunkelziffer berücksichtigt werden können. Grundsätzlich sollte auch eine weitere Harmonisierung der verschiedenen Datenspuren (Unfallversicherungen, Gesundheitsversorger, ASTRA) im Zusammenhang mit Unfällen angestrebt werden. Davon könnten auch andere Forschungsbereiche profitieren – beispielsweise die Forschung zu Sport-, Haus- und Freizeitunfällen sowie die Berufsunfallforschung.

Empfehlungen

Für zukünftige Methoden der Dunkelzifferberechnung bei Strassenverkehrsunfällen sollten bereits kurzfristig Massnahmen eingeleitet werden. Dafür wird die Koordination durch eine zentrale Organisation (z. B. ASTRA) empfohlen, um eine einheitliche Strategie zu formulieren.

Dunkelziffer für Getötete

Aufgrund geringer Abweichungen zwischen ASTRA- und BFS-Daten sowie einer kleinen erwarteten Dunkelziffer wird empfohlen, für die Schweiz vereinfachend von einem Faktor 1,0 auszugehen. Das bedeutet, dass die offiziellen ASTRA-Daten direkt als Grundlage für die Anzahl tödlicher Unfälle verwendet werden. Dies gilt auch für alle Differenzierungen.

Dunkelziffer für Verletzte

Zur Berechnung der Dunkelziffer von Verletzten wird eine Kombination aus Befragungen (Methode 1) und Daten-Linkage (Methode 2) empfohlen (s. Abb. I).

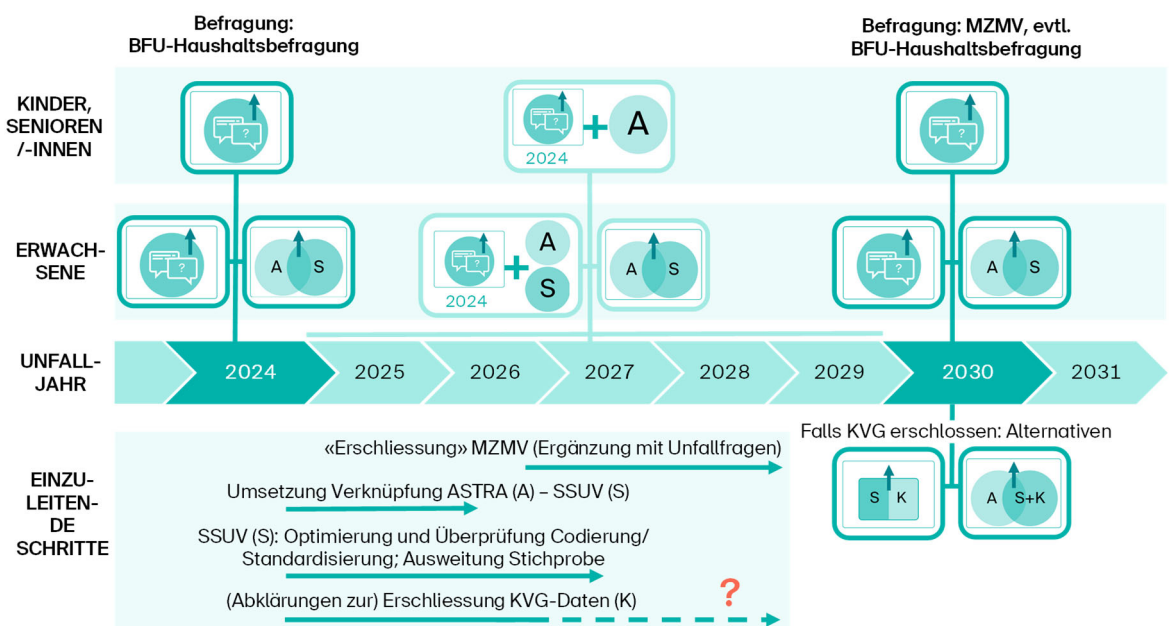


Abb. I: Schematische Darstellung des vorgeschlagenen mittelfristigen Vorgehens zur Berechnung der Dunkelziffer für verletzte Personen bei Strassenverkehrsunfällen (ohne Todesfälle). Eine Erläuterung der verwendeten Piktogramme findet sich in Kapitel 7.2 (ab S. 71). (Hinweis: A = ASTRA-Daten, S = SSUV-Daten, K = KVG-Daten).

1. Befragung als flexibel und gezielt einsetzbares Instrument (Methode 1)

- Für die Berechnung der differenzierten Dunkelziffern für das Unfalljahr 2024 kann die BFU-Haushaltsbefragung genutzt werden.
- Danach sollte alle fünf bis sieben Jahre eine neue Befragung durchgeführt werden.
- Für eine Fortschreibung zwischen den Befragungen können die Daten des ASTRA (und für Erwachsene allenfalls der SSUV) verwendet werden.
- Der MZMV 2030 sollte mit Unfallfragen erweitert werden.
- Wichtig sind ein geeignetes Stichprobendesign und Massnahmen zur Bias-Kontrolle (z. B. für Non-Responder).
- Die Anzahl von Unfällen von Personen, die keinen Wohnsitz in der Schweiz haben, sollte über ihren Anteil in den ASTRA-Daten geschätzt werden. Dabei müssen gewisse Unschärfen akzeptiert werden.

2. Linkage ASTRA–SSUV + Capture-Recapture als Ergänzung (Methode 2)

- Diese Methode sollte jährlich für Erwachsene angewendet werden, um die Befragungsergebnisse zu validieren.
- Dazu müssen Schritte eingeleitet werden, um die Verknüpfung der Daten vornehmen zu können (probabilistische Linkage).
- Zudem ist eine Harmonisierung der beiden Datenquellen bezüglich der Verletzungsschwere essenziell. Auch die Kodierung der beteiligten Objekte (Fahrzeugkategorien) sollte harmonisiert werden.
- Darüber hinaus wäre eine Verbesserung der Qualität und Optimierung der UVG-Statistik sowie eine Erweiterung der Stichprobe (von 5 % auf bis zu 100 %) anzustreben.

3. Erschliessung der KVG-Daten für eine langfristige Lösung (Methoden 2 und 3)

- Um Methode 2 auf alle Altersgruppen ausweiten zu können bzw. um Methode 3 anwenden zu können, müssen KVG-Unfalldaten systematisch erfasst werden. Die Erfassung und Nutzung erfordern eine Anpassung der rechtlichen Rahmenbedingungen.
- Sollte dies gelingen, wäre solch eine vollständige Versicherungsliste (Methode 3) eine kostengünstige Alternative zu Befragungen.

Die vorgeschlagenen Methoden erlauben eine Differenzierung nach Verkehrsteilnahme, Unfallhergang, (teilweise) Verletzungsschwere, Alter, Geschlecht, Raumtyp und Zeitfenster. Dies entspricht den Bedürfnissen der Stakeholder und ermöglicht Vergleiche mit anderen Ländern. Voraussetzungen für diese und weitere Differenzierungen sind eine ausreichende Fallzahl und hohe Datenqualität.

Résumé

Introduction

Le recensement complet des accidents de la route est essentiel à l'efficacité du travail de sécurité routière. La statistique officielle, à savoir la statistique des accidents de la route établie par l'Office fédéral des routes (OFROU), se fonde sur les déclarations d'accident remplies par la police. Vu que les accidents ne sont pas tous enregistrés, l'accidentalité réelle est sous-estimée. L'image de l'accidentalité réelle est par ailleurs faussée, car la probabilité d'enregistrement des accidents dépend de différents facteurs comme la gravité des blessures ou le déroulement de l'accident: la probabilité est par exemple plus faible pour les accidents individuels que pour les collisions.

Une détermination réaliste et suffisamment différenciée du sous-enregistrement est déterminante pour fournir une base fiable aux mesures de prévention. Les données correspondantes sont également essentielles pour l'analyse des coûts socioéconomiques des accidents, le calcul des taux d'accident et le travail de normalisation. Le projet MEDUSA visait à élaborer une méthode de détermination différenciée du sous-enregistrement des accidents de la route en Suisse, devant permettre la différenciation des groupes d'usagers de la route, des déroulements des accidents et de la gravité des blessures.

Méthode

Au début du projet, une revue de la littérature scientifique a été menée pour analyser les approches nationales et internationales en matière de détermination du facteur de sous-enregistrement. Elle a notamment porté sur l'analyse de modèles statistiques envisageables, y compris de ceux utilisés dans d'autres domaines (p. ex. biologie). Parallèlement, on a procédé à l'inventaire à large échelle et à l'évaluation des sources de données existantes ou à mettre en place, contenant des informations sur les accidents de la route ou susceptibles d'être utilisées pour obtenir des données accidentologiques, et qui pourraient permettre une comparaison avec les données officielles de l'OFROU sur les accidents de la route.

Un autre élément clé de la méthode était l'implication des différentes parties concernées. Un premier atelier a permis de déterminer leurs besoins et leurs attentes, par exemple en matière de différenciation ou de périodicité du calcul du facteur de sous-enregistrement.

Les résultats et les enseignements tirés des différentes étapes de travail ont été mis en relation afin d'élaborer des propositions concrètes quant aux variantes méthodologiques privilégiées. Dans le cadre de l'élaboration de la méthode, on a procédé à des simulations pour trois variantes méthodologiques ainsi qu'au calcul pilote pour l'une de ces variantes. Dans le cadre d'un deuxième atelier, les variantes méthodologiques élaborées ont été présentées aux parties concernées, évaluées et priorisées en fonction de trois objectifs: différenciation, ressources nécessaires et précision.

Résultats

État, au niveau (inter)national, de la recherche en matière de sous-enregistrement des accidents de la route

Une estimation du facteur de sous-enregistrement des accidents de la route en Suisse a été réalisée dès 1990 sur la base des données de l'assurance-accidents. Il en est ressorti que le nombre effectif d'accidents était environ quatre fois plus élevé que celui enregistré par la police. Depuis 2015, il existe pour la Suisse une estimation actualisée chaque année, basée sur un sondage mené auprès des ménages en 2011. Les estimations actuelles pour 2021 s'élèvent à environ 77 000 blessés, alors que la police en a enregistrés seulement quelque 20 500. Le sous-enregistrement est particulièrement grand pour le trafic cycliste:

on estime que le nombre de blessés est environ neuf fois plus élevé que celui fourni par la statistique officielle.

En 2021, une enquête menée parmi les pays membres de l'International Traffic Safety Data and Analysis Group (IRTAD) a montré que 9 pays sur 15 ont déterminé le sous-enregistrement une fois ou de manière répétée. La plupart des pays ayant effectué des calculs relatifs à la sous-estimation, dans les données de la police, du nombre de blessés sur les routes ont fait appel à des données hospitalières. D'autres sources de données, comme des sondages (auprès des ménages) ou des registres d'assurance, ont également utilisées pour la comparaison avec les accidents enregistrés par la police. La différenciation concerne le plus souvent la gravité des blessures et le mode de déplacement, parfois aussi le déroulement ou l'heure de l'accident. D'un point de vue méthodologique, les estimations reposent généralement sur le couplage et la liaison de données (provenant de registres). Le nombre total d'accidents est parfois extrapolé à l'aide d'une méthode de capture-recapture utilisée en biologie pour estimer des tailles de population inconnues. Une revue systématique datée de 2021 montre que, dans les pays à hauts revenus, les statistiques de la police recensent 35 à 97 % des tués et 16 à 82 % des blessés graves sur les routes.

Des études internationales se sont penchées sur les facteurs ayant une influence sur la probabilité d'enregistrement d'un accident. Il s'avère que la gravité des blessures, le déroulement des accidents (nombre d'utilisateurs impliqués), le mode de déplacement (catégorie de véhicules) et l'âge ont une influence déterminante sur les différences au niveau de la probabilité d'enregistrement. D'autres facteurs (p. ex. type d'espace routier, sexe) sont susceptibles d'interagir avec ceux précités. L'absence de différenciation en fonction des facteurs causaux (gravité des blessures, déroulement des accidents, mode de déplacement, âge), par exemple car les données sont indisponibles, peut donner lieu à des estimations biaisées.

Besoins des parties concernées

Un atelier réunissant des représentants de l'administration fédérale, de la police, d'administrations communales, du Service de centralisation des statistiques de l'assurance-accidents (SSAA) et du BPA, Bureau de prévention des accidents, a permis de rassembler les principales exigences posées à la détermination du facteur de sous-enregistrement des accidents de la route. Les conclusions les plus importantes sont les suivantes. Les données sont essentielles à la prévention des accidents, à la gestion du trafic, aux décisions politiques et au contrôle du succès des mesures de sécurité routière. Il existe d'importantes lacunes de connaissances sur le facteur de sous-enregistrement, en particulier en ce qui concerne la micromobilité, l'électromobilité et le trafic cycliste. Une analyse exhaustive relative au sous-enregistrement devrait être disponible tous les 5 à 7 ans, tandis que les actualisations annuelles devraient tenir compte des tendances à court terme en matière de mobilité. Les résultats devraient être disponibles au plus tard 12 à 18 mois après la mise à disposition par l'OFROU des données sur les accidents. Dans l'idéal, les accidents devraient être enregistrés selon le principe de la territorialité. Le type de véhicule, la gravité des blessures, l'âge et le sexe sont considérés comme des caractéristiques centrales pour la différenciation. Les autres possibilités de différenciation souhaitées sont le déroulement, les causes et le jour/l'heure des accidents ainsi que le type d'espace routier.

Choix des modèles statistiques

Étant donné qu'un recensement complet des accidents n'est pas réalisable en pratique, il faut recourir à d'autres approches permettant une estimation de l'accidentalité globale ou du sous-enregistrement différenciés en fonction des caractéristiques souhaitées. Trois approches fondamentales sont ainsi envisageables: l'échantillonnage, les traces de données et l'apprentissage automatique (machine learning).

- **Échantillonnage ou enquêtes:** cette approche permet de collecter des données représentatives. La méthode est flexible et éprouvée, mais elle est dispendieuse et sujette au biais de mémoire ou à la désirabilité sociale.

- **Traces / liaison de données:** cette approche vise à exploiter des sources de données supplémentaires contenant des informations différenciées sur les accidents de la route. Les données des assureurs ou des hôpitaux sont par exemple envisageables. Les exigences posées aux sources de données sont élevées: les données doivent permettre de distinguer les accidents de la route des autres événements et de les filtrer en fonction des caractéristiques de différenciation souhaités. Un couplage fiable (liaison) avec les données de l'OFROU doit par ailleurs être possible afin d'identifier les accidentés figurant dans les deux bases de données. Étant donné qu'aucun ensemble de données ne répertorie tous les accidents, la méthode de capture-recapture permet ensuite de déduire la population d'accidents.
- **Apprentissage automatique (machine learning):** cette approche permet de prévoir les accidents à l'aide de variables aisément mesurables. L'élaboration des modèles prédictifs nécessite de vastes données d'entraînement, de haute qualité et différenciées en fonction des caractéristiques souhaitées, relatives aux accidents enregistrés et non enregistrés par la police. De telles données ne seront toutefois pas disponibles ou pas accessibles dans un avenir proche.

Bases de données potentiellement exploitables

Les sources de données potentiellement exploitables en vue du calcul du facteur de sous-enregistrement des accidents de la route sont indiquées ici et leur adéquation est évaluée. Elles ont été classées en fonction des approches statistiques décrites plus haut. La possibilité d'identifier les accidents de la route en adéquation avec la définition correspondante de l'OFROU constitue une condition essentielle à la prise en considération d'une source de données.

Échantillonnage ou enquêtes

Les enquêtes suivantes sont envisageables pour déterminer le sous-enregistrement:

- **Sondage du BPA auprès des ménages:** l'échantillon utilisé pour ce sondage découle du cadre d'échantillonnage pour les sondages auprès des personnes et des ménages (SRPH) de l'Office fédéral de la statistique (OFS). Il permet d'identifier les accidents de la route et de les différencier en fonction du mode de déplacement et de leur déroulement. Le dernier sondage a été mené en 2024/2025 auprès de quelque 25 000 ménages suisses. S'il est reconduit à l'avenir, il se prêterait également aux futures estimations du sous-enregistrement.
- **Micrecensement mobilité et transports (MRMT):** l'échantillon utilisé pour ce sondage auprès des ménages mené environ tous les cinq ans par l'OFS et l'Office fédéral du développement territorial (ARE) découle également du SRPH. À l'heure actuelle, il ne contient encore aucune question sur l'accidentalité, mais en complétant le questionnaire, il se prêterait à un calcul du facteur de sous-enregistrement. Quelque 55 000 personnes ont participé au dernier MRMT en 2021.

En revanche, dans sa forme actuelle, l'enquête suisse sur la santé (ESS) de l'OFS n'est pas appropriée car, bien qu'elle soit menée régulièrement et qu'elle porte aussi sur les accidents, elle ne permet pas d'identifier avec précision les accidents de la route et n'offre guère de possibilités de différenciation.

Une autre approche qui ne sera pas poursuivie est l'enquête ciblée menée auprès des blessés dans les hôpitaux, sur le modèle de la European Injury Database (EU-IDB). Elle pourrait certes fournir de précieuses données, mais seulement à grand renfort de personnel et de moyens financiers. Elle recense par ailleurs surtout les accidents moyennement graves à graves et n'est représentative que si tous les hôpitaux de Suisse y prennent part.

Traces / liaison de données:

Sources de données précédemment exploitées aux fins de détermination du sous-enregistrement qui se prêtent à la présente approche:

- **Statistique de l'assurance-accidents LAA (loi fédérale sur l'assurance-accidents; SSAA):** elle fournit des informations différenciées sur les accidents et permet

un rapprochement avec les données de l'OFROU. Les données se limitent toutefois à la population assurée selon la LAA, c'est-à-dire à l'ensemble des personnes employées en Suisse, et ne renseignent donc pas sur les enfants, les hommes et femmes au foyer, les retraités et les étudiants.

- **Statistique des causes de décès de l'OFS:** elle permet un rapprochement fiable avec les données de l'OFROU sur les accidents mortels.

La Statistique médicale des hôpitaux (MS) et les Données des patients ambulatoires des hôpitaux (PSA) contiennent certes des informations sur les blessures, mais ne permettent pas d'isoler avec précision les accidents de la route. Ces deux sources de données recensent par ailleurs surtout les blessures graves. Les données des cabinets médicaux pourraient renseigner également sur les blessures plus légères, mais elles ne sont pas accessibles à l'heure actuelle, et il paraît aujourd'hui disproportionné de changer cette situation.

D'autres sources qui pourraient se prêter à l'approche de liaison de données sont les données accidentologiques provenant de l'assurance-maladie obligatoire (LAMal) ou celles des assureurs automobiles. Toutefois, ces deux sources de données ne sont pas encore exploitables à l'heure actuelle et il serait complexe de les rendre accessibles. Les données des assurés-accidents selon la LAMal (soit toutes les personnes domiciliées en Suisse et non assurées selon la LAA) pourraient fournir, avec les données LAA, une «liste des données d'assurance» ou un échantillon représentatif de tous les accidentés ayant bénéficié d'une prise en charge médicale. Les données des assureurs automobiles pourraient, quant à elles, fournir des informations en particulier sur les dommages matériels causés par les véhicules motorisés. Faute de base légale, l'exploitation de ces données ne paraît cependant pas réaliste à l'heure actuelle.

Autres sources de données et données pour l'apprentissage automatique (machine learning)

Les sources de données suivantes ont également été prises en considération afin de déterminer leur adéquation au recensement des accidents de la route: données anamnestiques des hôpitaux, fonctions des smartphones/autres appareils intelligents, applications mobiles, systèmes d'alerte, enregistrements vidéo et capteurs automobiles. À l'heure actuelle, ces sources de données ne se prêtent pas encore à une estimation du sous-enregistrement des accidents de la route, ce pour différentes raisons, mais elles présentent un grand potentiel pour l'avenir. Les défis résident notamment dans la protection des données et l'accès à ces sources de données. La simple constatation et identification d'un accident de la route ne suffisent pas à utiliser les données correspondantes dans le cadre d'un calcul différencié du facteur de sous-enregistrement. Pour pouvoir exploiter avec succès de telles données dans des modèles d'apprentissage automatique, il faut disposer de données de haute qualité aussi complètes que possible sur les accidents de la route et leurs conséquences, ce qui n'est pas encore le cas à l'heure actuelle.

Variantes méthodologiques

Actuellement, les deux approches «échantillonnage» et «traces de données» sont les plus prometteuses. L'apprentissage automatique n'est pas envisageable pour le moment en raison du manque de données d'entraînement des modèles. Il recèle cependant un important potentiel et la situation en la matière devrait donc être réexaminée régulièrement à l'avenir. En raison de bases de données différentes, une distinction est faite ici entre les tués et les blessés.

Sous-enregistrement des tués sur les routes

La statistique des causes de décès de l'OFS peut intervenir dans le calcul du facteur de sous-enregistrement des tués sur les routes suisses. Ces données présentent cependant de faibles divergences par rapport à celles de l'OFROU. Ces différences résultent notamment du mode de comptage des cas: l'OFROU se fonde sur l'année de l'accident, tandis que l'OFS se base sur l'année du décès. Par ailleurs, l'OFROU applique le principe territorial (tous les tués en Suisse), alors que l'OFS suit le principe domestique (tués en Suisse ou à l'étranger parmi la population résidente suisse). De même, la règle des 30 jours consécutifs à l'accident dans la définition de la notion de tué sur les routes selon l'OFROU

ainsi que la saisie sur la base des codes CIM 10 dans la statistique des causes de décès (codes V) jouent également un rôle. Certaines divergences pourraient être lissées grâce à une meilleure saisie des données dans la statistique des causes de décès (p. ex. date de l'accident), mais ces informations font souvent défaut ou sont erronées. Les imprécisions statistiques pourraient aussi être atténuées par un facteur de correction. L'OMS fait appel à de tels facteurs de correction dans le cadre de ses estimations. Celui pour la Suisse est actuellement de 1,03.

Sous-enregistrement des blessés sur les routes

Pour les blessés, trois méthodes sont envisageables. Chacune d'elles a ses avantages et ses inconvénients.

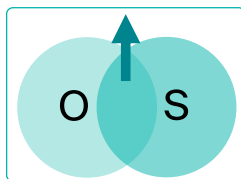
- **Méthode 1: enquête avec extrapolation** (symbolisée par la flèche). Les enquêtes



permettent un relevé direct des accidents de la route. Elles peuvent former des éléments autonomes ou être intégrées dans des enquêtes existantes. Pour 2024, le sondage du BPA auprès des ménages peut servir de base au calcul du facteur de sous-enregistrement. L'intégration de questions relatives aux accidents de la route dans le MRMT est également envisageable.

Les avantages de la méthode résident dans la flexibilité et la possibilité de différenciation des données. Il s'agit en outre d'une approche éprouvée et pouvant être mise en œuvre à court terme. S'agissant des inconvénients, les enquêtes sont dispendieuses et sujettes au biais de mémoire et à la désirabilité sociale des réponses. Un vaste échantillon est par ailleurs nécessaire pour obtenir des résultats suffisamment robustes.

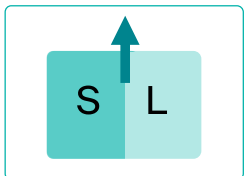
- **Méthode 2: liaison des données de l'OFROU (O) et du SSAA (S) avec capture-recapture** (méthode d'extrapolation, symbolisée par la flèche).



L'avantage de cette méthode est sa mise en œuvre moins coûteuse. Son grand inconvénient réside dans l'impossibilité de procéder au calcul du facteur de sous-enregistrement des accidents de la route pour les enfants et les seniors. Les deux bases de données différencient certes la gravité des blessures, mais les catégorisations sont différentes, ce qui rend impossible tout calcul

du facteur de sous-enregistrement en fonction de la gravité des blessures. La méthode de capture-recapture permet toutefois de procéder à une extrapolation permettant de déterminer l'accidentalité globale et donc d'estimer le sous-enregistrement. Le rapprochement des données de l'OFROU et du SSAA peut être effectué au moyen de méthodes probabilistes, ce qui représente toutefois un défi et nécessite donc un travail de recherche complémentaire. Une harmonisation des deux sources de données, surtout quant à la gravité des blessures, serait bénéfique à l'avenir. Là encore, un travail de recherche additionnel s'impose.

- **Méthode 3: liste des données d'assurance sur la base des données du SSAA (S) et de la LAMal (L) avec extrapolation** (symbolisée par la flèche).



Cette approche est prometteuse à long terme et, une fois lancée, elle pourrait être mise en œuvre avec relativement peu de frais. Les données des assureurs LAMal devraient toutefois être accessibles. Comme pour la méthode 2, une harmonisation des trois sources de données est nécessaire.

Pour les trois méthodes, le défi réside dans la prise en compte des personnes accidentées en Suisse mais qui n'y résident pas ou n'y exercent pas d'activité lucrative. Dans le cadre de travaux de recherche futurs, il convient d'examiner comment il est néanmoins possible de prendre en compte ces personnes dans le calcul du facteur de sous-enregistrement. D'une manière générale, il convient également de s'employer à obtenir une plus grande harmonisation des différentes sources de données (assureurs-accidents, prestataires de santé, OFROU). ce qui pourrait aussi être bénéfique à d'autres domaines de recherche, comme la recherche sur les accidents de sport, les accidents domestiques ou les accidents professionnels.

Recommandations

Pour pouvoir faire appel à l'avenir à de nouvelles variantes méthodologiques en vue du calcul du facteur de sous-enregistrement des accidents de la route, des mesures devraient être engagées à court terme. Un travail de coordination par un organisme central (p. ex. l'OFROU) est recommandé dans l'optique d'une stratégie uniforme.

Sous-enregistrement des tués sur les routes

En raison des faibles divergences entre les données de l'OFROU et de l'OFS ainsi que du faible sous-enregistrement escompté, il est recommandé d'admettre le facteur 1,0 par souci de simplicité. Autrement dit, les données officielles de l'OFROU peuvent servir telles quelles de base au nombre d'accidents de la route mortels. Il en va de même pour l'ensemble des différenciations.

Sous-enregistrement des blessés sur les routes

Pour calculer le facteur de sous-enregistrement des blessés sur les routes, il est recommandé d'associer des enquêtes (méthode 1) à la liaison de données (méthode 2) (cf. Illustration I).

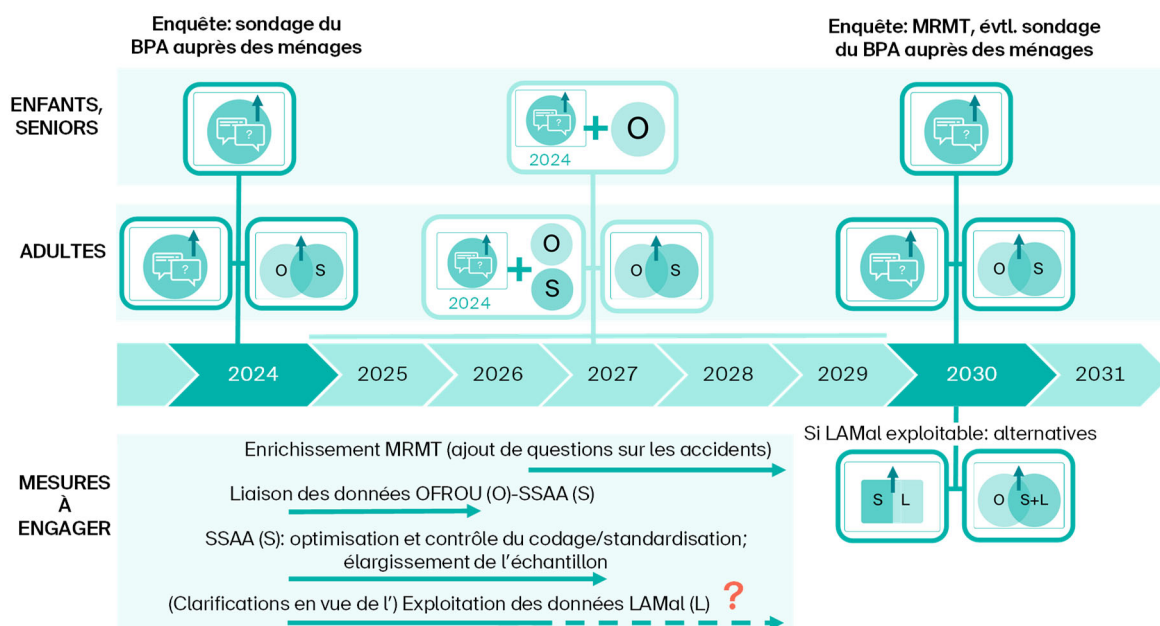


Illustration I: Représentation schématique de la démarche proposée à moyen terme pour calculer le facteur de sous-enregistrement des blessés sur les routes suisses (tués non considérés ici). Les pictogrammes utilisés sont expliqués au chap. 7.2 (dès la p. 71). Remarque: O = données de l'OFROU, S = données du SSAA, L = données LAMal.

1. Les enquêtes, un instrument flexible et ciblé (méthode 1)

- Le sondage du BPA auprès des ménages peut être utilisé en vue du calcul différencié des facteurs de sous-enregistrement des accidents de la route pour l'année d'accident 2024.
- Par la suite, un nouveau sondage devrait être réalisé tous les 5 à 7 ans.
- Les données de l'OFROU (et, pour les adultes, éventuellement celles du SSAA) peuvent servir à l'actualisation des données entre les sondages.

- Il convient d'intégrer dans le MRMT 2030 des questions relatives aux accidents.
 - Un échantillon bien conçu et des mesures de contrôle des biais statistiques (p. ex. non-répondants) sont essentiels.
 - Il convient d'estimer le nombre d'accidents de la route subis par des personnes non domiciliées en Suisse, sur la base de la part de ces dernières dans les données de l'OFROU, si bien qu'il faut accepter certaines imprécisions.
- 2. Liaison des données OFROU-SSAA + capture-recapture en complément (méthode 2)**
- Cette méthode devrait être utilisée annuellement pour les adultes afin de valider les résultats des enquêtes.
 - Pour ce faire, il s'agit d'engager des mesures permettant la liaison des données (liaison probabiliste).
 - Il est par ailleurs essentiel d'harmoniser les deux sources de données pour ce qui est de la gravité des blessures. Il convient également d'harmoniser le codage des objets impliqués (catégories de véhicules).
 - Il convient par ailleurs d'améliorer la qualité et de procéder à une optimisation de la statistique LAA, de même que d'élargir l'échantillon (de 5 % à 100 % dans l'idéal).
- 3. Exploitation des données LAMal pour une solution à long terme (méthodes 2 et 3)**
- Afin d'étendre la méthode 2 à tous les groupes d'âge ou d'appliquer la méthode 3, la saisie des données accidentologiques LAMal doit être soumise à une systématique. Celle-ci ainsi que l'utilisation de ces données nécessitent une adaptation du cadre légal.
 - Si on parvient à l'obtenir, une telle liste complète des données d'assurance (méthode 3) constituerait une alternative peu coûteuse aux enquêtes.

Les méthodes proposées permettent une différenciation en fonction du moyen de locomotion, du déroulement des accidents, de la gravité des blessures (en partie), de l'âge, du sexe, du type d'espace routier et du jour/de l'heure des accidents, ce qui répond aux besoins des parties concernées et permet des comparaisons avec d'autres pays. Ces différenciations ainsi que les autres susceptibles d'être réalisées nécessitent un nombre suffisant de cas et des données de haute qualité.

Summary

Introduction

A complete record of road traffic accidents is essential for effective road safety work. The official statistics – the road traffic accident statistics of the Federal Roads Office FEDRO – are based on police reports. But not all accidents are recorded, which leads to an underestimation of the actual accident occurrence. Moreover, this results in a distorted representation of the actual accident occurrence, as the likelihood of an incident being recorded depends on several factors, such as injury severity or the circumstances of the accident: single-vehicle accidents, for instance, are less likely to be recorded than collisions.

A realistic and sufficiently differentiated estimate of the number of unreported cases is crucial to creating a reliable basis for prevention measures. This data is also essential for economic cost analyses, accident rate calculations and standard setting. The aim of the MEDUSA project was to develop a methodology for a differentiated estimation of the number of unreported cases. The methodology was to be structured in such a way that enables differentiation between different road user groups, accident circumstances and degrees of injury severity.

Methodology

A literature review was carried out at the start of the project to analyse national and international approaches to estimating the number of unreported cases. Specifically, the literature review also included an analysis of relevant statistical models, including those from other subject areas (e.g. biology). Alongside this, a broad-based analysis and assessment was conducted of existing data sources, or those yet to be developed, that contain data on road traffic accidents or could be used to collect accident data, and which could be aligned with the official FEDRO road traffic accident data.

Involving the relevant stakeholders was also a core component of the methodology. An initial workshop was held to ascertain their needs and expectations – for instance with respect to differentiation or the periodicity of estimating unreported cases.

The outcomes and findings of the various steps were interlinked in order to develop concrete proposals for prioritised variants of the methodology. As part of the method development process, simulations were conducted for three different methodological variants as well as a pilot calculation for one of these variants. In a second workshop, the method variants were put to the stakeholders, evaluated and prioritised according to the three objectives «differentiation», «time and cost expenditure», and «accuracy».

Findings

(Inter-)national state of research on the number of unreported road traffic accidents

Switzerland carried out estimations of the number of unreported road traffic accidents on the basis of accident insurance data as early as 1990. The actual accident figures were found to be around four times higher than those recorded by the police. An annually updated estimate for Switzerland has existed since 2015, based on a household survey conducted in 2011. The current estimates for 2021 put the number of injured persons at around 77,000, while only around 20,500 were recorded by the police. The high number of unreported bicycle accidents is particularly striking, with around nine times more casualties estimated than recorded in the official statistics.

A 2021 survey among the International Traffic Safety Data and Analysis Group (IRTAD) members showed that nine out of 15 countries have estimated the number of unreported cases – either once or multiple times. Among the countries that estimated underreported

road traffic injuries in police data, hospital data was used in most cases. Also, other sources of data, such as (household) surveys or insurance data, were used to compare to the number of police-registered accidents. Differentiation is usually limited to injury severity and travel mode; in some cases, a distinction is also made according to accident circumstances or the time of day. Methodologically, the estimates are usually made by linking and merging (register) data. In some cases, overall accident figures are extrapolated using a capture-recapture method, which is used in biology to estimate unknown population sizes. A systematic review from 2021 found that in high-income countries, 35–97 % of road fatalities and 16–82 % of the seriously injured are recorded in police statistics.

International studies analysed factors that affect the probability of an accident being recorded. Injury severity, accident circumstances (number of people involved), travel mode (vehicle category) and age were shown to be decisive in determining differences in recording probability, with other possible factors (e.g. type of area, gender) interacting with these influences. Failure to differentiate according to causal factors (injury severity, circumstances, travel mode, age) – for instance due to unavailability of data – can result in distorted estimates.

Stakeholder needs

Key requirements for recording the number of unreported road traffic accidents were established during a workshop with representatives from the federal administration, the police, the city administration, the Central Office for Statistics under the Federal Act on Accident Insurance (SSUV), and the BFU, Swiss Council for Accident Prevention. Main findings: The data is essential for accident prevention, traffic planning, policy decisions and monitoring the success of road safety measures. There are significant knowledge gaps regarding the number of unreported cases, notably in the areas of micromobility, electromobility and bicycle traffic. A comprehensive analysis of the number of unreported cases should be available every five to seven years, while annual updates should factor in short-term mobility trends. The findings should be ready no later than 12 to 18 months after the FEDRO accident data becomes available. Accidents should preferably be recorded according to the territoriality principle. In terms of differentiation, travel mode, injury severity, age and gender are considered key characteristics. Other requested differentiation characteristics were accident circumstances, accident cause, time of accident and area type.

Selecting statistical models

Since collecting a full record of all accidents is not feasible in practice, other approaches must be applied which allow a differentiated estimation of the overall accident occurrence or the number of unreported cases according to the relevant characteristics. Three basic approaches are possible: Sampling, data tracks and machine learning.

- **Sampling or surveys:** This approach can be used to collect representative data. While this method is flexible and proven, it also involves a high time and cost expenditure and is susceptible to distortion due to memory errors or social desirability.
- **Data tracks / linkage:** The aim of this approach is to exploit additional data sources that provide differentiated information on road traffic accidents. They could include, for example, data from insurers or hospitals. There are high demands on the data sources: the data must allow road traffic accidents to be distinguished from other incidents and be filtered according to the desired differentiation characteristics. Moreover, this requires a reliable linkage with the FEDRO data in order to identify accident casualties which occur in both datasets. Since no dataset contains all accidents, the capture-recapture method is used in a further step to infer the population of accidents.
- **Machine learning:** This approach makes it possible to predict accidents on the basis of easily measurable variables. Developing the prediction models requires extensive, high-quality training data on both officially recorded and unrecorded accidents, differentiated according to the relevant characteristics. However, such data will not be available or accessible in the foreseeable future.

Potentially exploitable data bases

The following describes potentially exploitable data sources and assesses their suitability for estimating the number of unreported cases. The categorisation is based on the statistical approaches identified above. An important criteria for a data source to be considered is that road traffic accidents can be identified in line with the corresponding FEDRO definition.

Sampling or surveys

The following surveys can be considered for estimating the number of unreported cases:

- **BFU household survey:** The sample for this survey was drawn from the sampling frame for personal and household surveys (SRPH) of the Federal Statistical Office (FSO) and enables the identification of road traffic accidents and a differentiation of these accidents according to travel mode and accident history. The last survey took place in 2024/2025 and covered around 25,000 participating households. If this survey is continued, it will also be suitable for future estimates of the number unreported cases.
- **Mobility and Transport Microcensus (MTMC):** The sample for this household survey, which is carried out approximately every five years by the FSO and the Federal Office for Spatial Development (ARE), is also drawn from the SRPH. It does not currently contain any questions on accident occurrence, but would be suitable for estimating the number of unreported cases if the questionnaire were to be expanded accordingly. Around 55,000 people participated in the last MTMC in 2021.

However, the Swiss Health Survey (SHS) by the FSO in its current form is unsuitable. Although it already records accidents and is repeated regularly, it does not allow for a precise identification of road traffic accidents and offers barely any differentiation options.

One approach that has not been pursued further is the targeted surveying of injured persons in hospitals, modelled on the European Injury Database (EU-IDB). Although it could provide valuable data, it is labour-intensive and costly, covers mainly moderate to serious accidents and is only representative if all hospitals participate.

Data tracks / linkage

Existing data sources that are suitable for a linkage approach:

- **Accident Insurance Act UVG Statistics; SSUV:** provides differentiated information on accidents and can be linked to the FEDRO data. However, the dataset is limited to the population insured by the Accident Insurance Act (UVG), i.e. all persons employed in Switzerland, and therefore does not include data on children, housewives and househusbands, pensioners or students.
- **FSO cause-of-death statistics:** enables a reliable comparison with the FEDRO data for fatal accidents.

Although hospital medical statistics (MSK) and outpatient data (PSA) contain details on injuries, they do not enable a clear separation of road traffic accidents. They also primarily record serious injuries. Although data from medical practices could identify minor injuries, it has yet to be developed. Indexing this data is not considered appropriate at present.

Other data sources that could be used for the linkage approach are accident data from the statutory health insurance (HIA) or from motor vehicle insurers. However, both these data sources have yet to be accessed, and developing them is likely to be highly complex. The data on persons insured under the HIA (this includes everyone resident in Switzerland and not insured under the UVG) could, combined with the UVG data, form an "insurance list" or a representative sample of all casualties who have received medical treatment. Data from motor vehicle insurers could specifically provide information on property damage caused by motorised vehicles. However, developing this data without a corresponding legal basis does not currently appear viable.

Other data sources and data for machine learning

The following additional data sources were considered for their suitability for recording road traffic accidents: inpatient medical records, smartphone/smart device functions, mobile apps, notification systems, video recordings and vehicle sensors. Although these data

sources are for various reasons not currently suitable for estimating the number of unreported road traffic accidents, they have considerable potential for future use. Challenges include data privacy issues and developing the data sources. Simply detecting and identifying a road traffic accident is not sufficient to be able to use the data for a differentiated estimation of the number of unreported cases. In order to successfully exploit such data in machine learning models, data on road traffic accidents and their consequences must be as complete and high quality as possible. This is not yet the case.

Method variants

The two approaches «sampling» and «data tracks» are currently the most promising. Machine learning is not currently suitable due to the lack of training data, but holds great potential and should therefore be re-evaluated regularly in the future. The following distinguishes between fatalities and injured persons resulting from different data bases.

Number of unreported fatalities

Data from the FSO cause-of-death statistics could be used to estimate the number of unreported fatalities. However, they show minor deviations from the FEDRO data. These deviations are attributable, among other factors, to different counting methods: the FEDRO bases its calculations on the year of the accident; the FSO on the year of death. Furthermore, the FEDRO applies the territorial principle (all fatalities in Switzerland), while the FSO follows the domestic principle (fatalities among the resident population in Switzerland and abroad). The 30-day rule used to define road fatalities by the FEDRO and the recording via ICD-10 codes in the cause-of-death statistics (V-codes) are also relevant. Some deviations could be offset by improved data collection in cause-of-death statistics, such as the accident date, but this information is often either lacking or inaccurate. The statistical inaccuracies could also be offset by a correction factor. The World Health Organization WHO applies such correction factors in its estimates – for Switzerland, it is currently 1.03.

Number of unreported injured persons

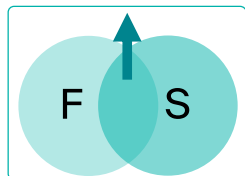
There are three conceivable methods for injured persons. Each method has advantages and disadvantages.

- **Method 1: Survey with extrapolation** (symbolised with arrow): Surveys permit a direct record of road traffic accidents. They can be designed as a standalone survey or integrated into existing surveys. For 2024, the BFU household survey can serve as a basis for estimating the number of unreported cases. Integrating questions on road traffic accidents into the MTMC is also conceivable. The advantages of this method are flexibility and the differentiability of the data. It is also a proven approach that can be implemented



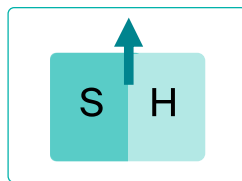
at short notice. The disadvantages of this approach are that surveys are expensive and prone to recall bias and socially desirable responses. They also require a large sample in order to obtain sufficiently robust results.

- **Method 2: Linking data from the FEDRO (F) and SSUV (S) with Capture-Recapture** (extrapolation method, symbolised with arrow): The advantage of this method is that it is less expensive to apply. A major disadvantage is that it cannot be used to estimate the number of unreported cases which affect children and senior citizens. Injury severity is recorded in a differentiated manner in both datasets, albeit using a different categorisation, which makes it impossible to estimate the number of unreported cases



by injury severity. Nonetheless, the capture-recapture method can be used to extrapolate the overall accident occurrence and thus estimate the number of unreported cases. Linking FEDRO and SSUV data can be achieved using probabilistic methods, but this is challenging and requires further research. Harmonising the two data sources, in particular with regard to injury severity, would be beneficial in the future. This approach also needs to be researched further.

- Method 3: Insurance list compiled from the SSUV (S) and HIA (H) data with extrapolation** (symbolised with arrow): This approach would be promising in the long term and, once set up, relatively inexpensive to implement. It would, however, necessitate access to the HIA insurer data. As with Method 2, this requires a harmonisation of the three data sources.



The challenge posed by all three methods is how to account for people who suffer an accident in Switzerland but are neither resident nor employed in Switzerland. Future research projects should clarify how these people can still be included in the estimation of the number of unreported cases. The general aim should also be to continue harmonising the various data tracks (accident insurers, healthcare providers, FEDRO) in relation to accidents. Other research areas, such as research into sport, home and leisure accidents, as well as occupational accident research, could also benefit from this.

Recommendations

Short-term measures should already be introduced for future method variants to estimate the number of unreported road traffic accidents. To this end, coordination by a central organisation (e.g. FEDRO) is recommended in order to define a unified strategy.

Number of unreported fatalities

Due to minor deviations between FEDRO and FSO data and a small expected number of unreported cases, it is recommended that a factor of 1.0 be assumed for simplicity's sake. This means that the official FEDRO data is used as a direct basis for the number of fatal accidents. This also applies to all differentiations.

Number of unreported injured persons

A combination of surveys (Method 1) and data linkage (Method 2) is recommended to estimate the number of unreported injured persons (see Fig. I).

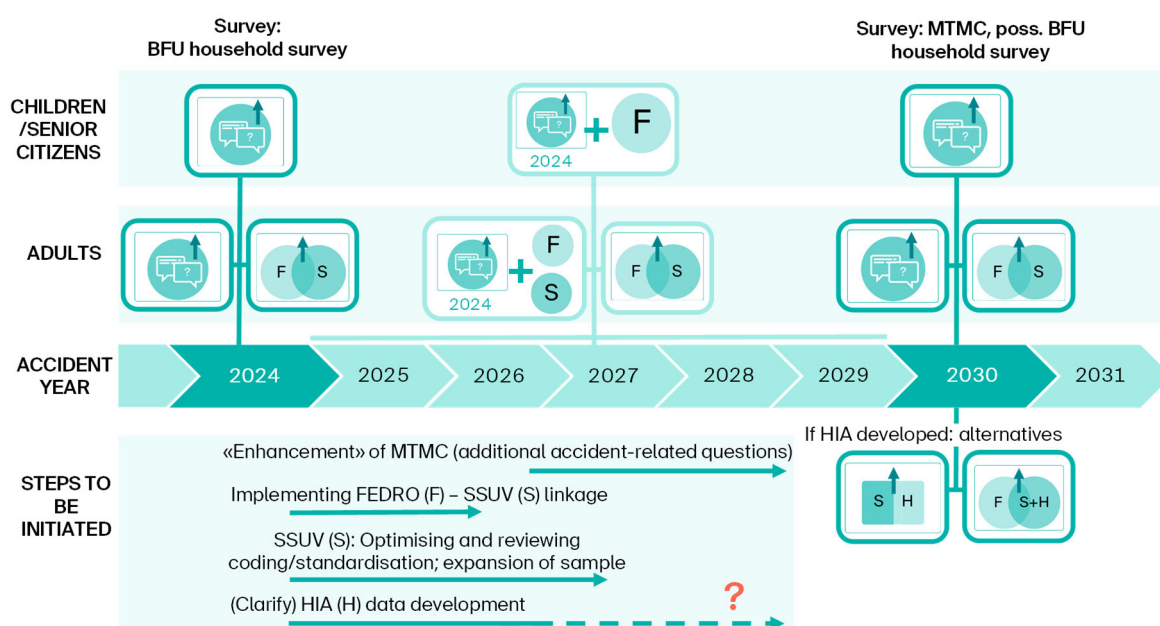


Fig. I: Schematic representation of the proposed medium-term approach to estimating the number of unreported injured persons in road traffic accidents (excluding fatalities). Information on the pictograms can be found in Chapter 7.2 (from p. 71).
 Note: F = FEDRO data, S = SSUV data, H = HIA data.

1. Survey as a flexible and targeted tool (Method 1)

- The BFU household survey can be used to estimate the differentiated number of unreported cases for the 2024 accident year.
- A new survey should be carried out every five to seven years thereafter.
- The FEDRO data (and, for adults, possibly the SSUV data) can be used for updates between surveys.
- The 2030 MTMC should be expanded to include accident-related questions.
- A suitable sample design and bias control measures (e.g. for non-responders) are important.
- The number of accidents involving non-Swiss residents should be estimated via their share in the FEDRO data. This means that a degree of uncertainty must be accepted.

2. Linking FEDRO–SSUV + Capture-Recapture as a supplement (Method 2)

- This method should be applied annually for adults in order to validate the survey results.
- This requires taking steps to enable data linkage (probabilistic linkage).
- Harmonising both data sources in terms of injury severity is essential. The coding of the participating objects (vehicle categories) should also be harmonised.
- Efforts should also be made to optimise and improve the quality of the UVG statistics and to expand the sample (from 5 % to 100 %).

3. Developing the HIA data for a long-term solution (Methods 2 and 3)

- HIA accident data must be systematically collected in order to expand Method 2 to all age groups or to apply Method 3. Collecting and exploiting this data requires an adaptation of the legal framework.
- Provided this is successful, such a complete insurance list (Method 3) would be a cost-effective alternative to surveys.

The proposed methods enable differentiation according to travel mode, accident occurrence, injury severity (partially), age, gender, area type and time window. This corresponds to stakeholder needs and allows for comparisons with other countries. The requirements for these and further differentiations are sufficient case numbers and high-quality data.

1 Einleitung

1.1 Ausgangslage und Problemstellung

Das Gesamtausmass der Strassenverkehrsunfälle (das Gesamtunfallgeschehen) bildet eine zentrale Grundlage für die Verkehrssicherheitsarbeit. Um wirksame Massnahmen zur Unfallprävention zu entwickeln, ist eine möglichst vollständige und verlässliche Erhebung der Strassenverkehrsunfälle unerlässlich. Das tatsächliche Unfallgeschehen wird durch verschiedene Faktoren bestimmt, darunter das Verhalten der Verkehrsteilnehmenden, die Qualität und der Zustand der Infrastruktur, die eingesetzten Fahrzeuge sowie das Verkehrsaufkommen auf unterschiedlichen zeitlichen und räumlichen Ebenen.

In der Schweiz basiert die offizielle Unfallstatistik auf den Meldungen der kantonalen Polizeikorps sowie des Schadenszentrums VBS. Ein Strassenverkehrsunfall muss gemeldet werden, wenn Personen verletzt wurden oder eine unmittelbare Gefahr nicht beseitigt werden kann. Bei reinen Sachschäden erfolgt eine polizeiliche Meldung in der Regel nur dann, wenn ein Unfallbeteiligter dies wünscht oder der Geschädigte nicht über den Unfall informiert werden kann (Art. 51 SVG¹ und Art. 54-56 VRV²). Die von der Polizei registrierten Unfälle werden mit dem schweizweit gültigen Unfallaufnahmeprotokoll (UAP) erfasst [1] und an das Bundesamt für Strassen (ASTRA) übermittelt.

Allerdings werden nicht alle Unfälle der Polizei gemeldet, wodurch eine Dunkelziffer entsteht [2]. Betrachtet man lediglich die polizeilich registrierten Unfälle, wird das tatsächliche Ausmass des Unfallgeschehens – insbesondere im Vergleich zu anderen Bereichen der öffentlichen Gesundheit – unterschätzt. Die Höhe der Dunkelziffer wird unter anderem durch das Meldeverhalten der Unfallbeteiligten beeinflusst, beispielsweise durch deren Pflichtbewusstsein, die Angst vor Sanktionen oder mögliche Konsequenzen für die Versicherungsentschädigung.

Die Erfassungswahrscheinlichkeit – also die Wahrscheinlichkeit, dass ein Unfall gemeldet und registriert wird – hängt von verschiedenen Faktoren ab. Beispiele sind die Schwere des Unfalls, die Art der Verkehrsteilnahme, der Unfallhergang und der Unfallort [3]. Schwere Unfälle werden eher gemeldet, während leichtere Unfälle seltener in die Statistik eingehen (Abb. 1) [4]. Zudem weisen Alleinunfälle oder Unfälle mit nur einer beteiligten Person eine geringere Erfassungswahrscheinlichkeit auf. Dies gilt insbesondere für Velounfälle, die ohne Drittbeteiligung häufig nicht gemeldet werden, obwohl die betroffene Person verletzt wurde [5,6]. Viele verzichten auf eine Meldung, da kein hoher Sachschaden an Fahrzeugen oder Infrastruktur entstanden ist und somit keine Ansprüche Dritter geltend gemacht werden müssen. Darüber hinaus spielt der administrative Aufwand eine Rolle, da die formelle Abwicklung mit der Polizei und Versicherung als unverhältnismässig empfunden werden kann. Nicht zuletzt bestehen trotz der standardisierten Aufnahme der Unfälle mit dem UAP und der rechtlichen Rahmenbedingungen kantonale Unterschiede in der polizeilichen Erfassung von Unfällen.

All die genannten Gründe führen dazu, dass die offizielle Unfallstatistik ein unvollständiges und verzerrtes Bild des tatsächlichen Unfallgeschehens auf den öffentlichen Strassen vermittelt. Die Folgen dieser unvollständigen Erfassung sind zweifach. Zum einen führt sie zu einer Unterschätzung des Unfallgeschehens, was Auswirkungen auf die Legitimation der Verkehrssicherheitsarbeit, die Kosten-Nutzen-Abwägung von Massnahmen und die Berechnung volkswirtschaftlicher Schäden hat. Zum anderen können auf polizeilich registrierten Daten basierende Risikobewertungen und Unfallstatistiken von der Realität abweichen, was in der Analyse und Interpretation der Daten berücksichtigt werden muss.

Erkenntnisse zur Dunkelziffer von Strassenverkehrsunfällen sind daher nicht nur eine wichtige Ergänzung, sondern essenziell für eine fundierte Interpretation und für wirksame Massnahmen zur Verbesserung der Verkehrssicherheit.

¹ Strassenverkehrsgesetz vom 19.12.1958 (SR 741.01)

² Verkehrsregelnverordnung vom 13.11.1962 (SR 741.11)

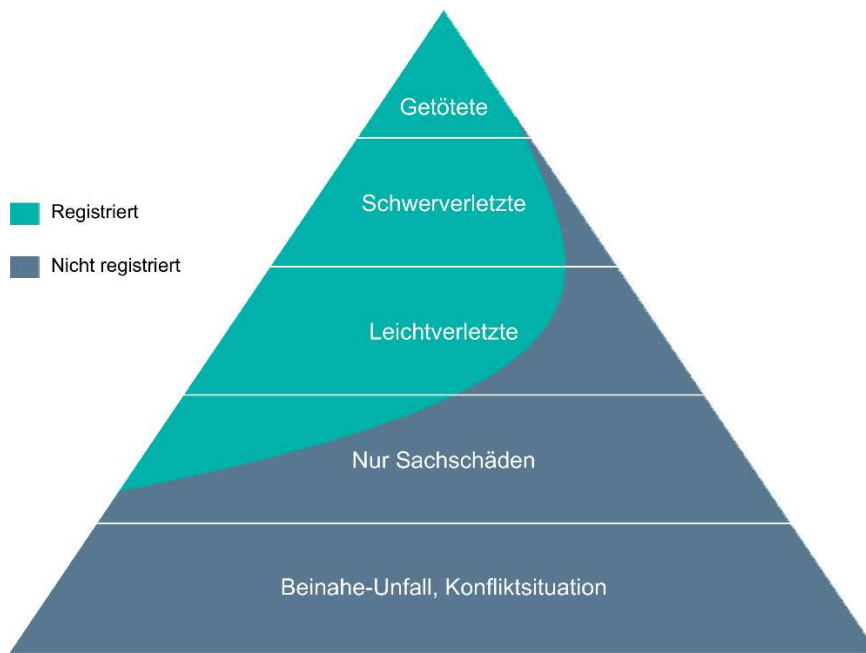


Abb. 1: Vermutete Vollständigkeit der Unfallerkennung nach Verletzungsschwere.

1.2 Projektziel und Abgrenzung

Das Ziel dieses Projekts war die Entwicklung einer Methodik zur differenzierten Bestimmung der Dunkelziffer von Strassenverkehrsunfällen in der Schweiz. Das ASTRA gab vor, dass die Methodik eine Differenzierung von Verkehrsteilnehmergruppen, Unfallhergängen und Verletzungsschwere ermöglichen muss. Darüber hinaus verfolgte das Projekt – gemäss der Ausschreibung des ASTRA – die folgenden Teilziele:

- Auflistung und Bewertung potenzieller methodischer Ansätze zur Bestimmung der Dunkelziffer, basierend auf existierenden Datenbanken sowie weiterer Möglichkeiten (neue Technologien, Befragungen etc.).
- Nachweis der praktischen Umsetzbarkeit der entwickelten Methodik.
- Empfehlung einer konkreten Implementierungsstrategie sowie einer sinnvollen Periodizität der Dunkelzifferabschätzung.

Die Grundlage dieses Projekts bildet die offizielle Definition von Strassenverkehrsunfällen gemäss ASTRA (s. Kapitel 1.3). Diese setzt zwar keinen Personenschaden voraus. Dennoch lag der Fokus dieser Forschungsarbeit auf Unfällen mit Personenschäden, da diese eine höhere Relevanz für die Verkehrssicherheit aufweisen. Alleinunfälle von Fussgängerinnen und Fussgängern wurden nicht berücksichtigt. Darüber hinaus wurde in diesem Forschungsprojekt keine Untersuchung oder Bewertung der Qualität der erfassten Strassenverkehrsunfälle durchgeführt. Insbesondere kantonale oder individuelle Unterschiede in der polizeilichen Erfassung wurden nicht analysiert.

1.3 Verwendete Begriffe

Zur Analyse der Dunkelziffer wird der Begriff **Strassenverkehrsunfall** entsprechend der Definition des ASTRA in den Anweisungen zum UAP 2018 verwendet [7]:

«Ein unvorhergesehenes Ereignis auf einer öffentlichen Verkehrsfläche, das in ursächlichem Zusammenhang mit dem Strassenverkehr und seinen Gefahren steht,

einen Sachschaden und/oder Personenschaden zur Folge hat und an dem mindestens ein Fahrzeug oder ein fahrzeugähnliches Gerät beteiligt ist. Planmässiges Handeln (z. B. Suizid- oder Tötungsabsicht) aller Beteiligten ist dabei ausgeschlossen.»

Die Unterkategorie **Alleinunfall** umfasst Unfälle mit nur einem beteiligten Fahrzeug (z. B. Velo). Dabei können sich mehrere Personen in oder auf diesem Fahrzeug befinden.

Insbesondere in der Kriminalitätsstatistik (z. B. Birkel et al. 2020 [8], Eifler, Pollich 2014 [9]) hat sich das Begriffspaar Hellfeld und Dunkelfeld etabliert. Diese Begriffe sind im vorliegenden Projekt wie folgt definiert:

- **Hellfeld** = Alle in der Statistik der polizeilich registrierten Strassenverkehrsunfälle des ASTRA (SVU-Statistik) polizeilich erfassten Strassenverkehrsunfälle, einschliesslich der beteiligten Objekte, Personen sowie alle dokumentierten Personen- und Sachschäden.
- **Dunkelfeld** = Alle Strassenverkehrsunfälle sowie daraus resultierende Personen- und Sachschäden, die nicht in der SVU-Statistik des ASTRA erfasst wurden.

Diese Unterscheidung ist theoretisch klar abgegrenzt (veranschaulicht in Abb. 2), in der Praxis jedoch nicht immer trennscharf, insbesondere, weil auch die Abgrenzung von Strassenverkehrsunfällen nach obiger Definition nicht immer eindeutig ist. Zudem kann es auch Fälle geben, die einmal dem Hellfeld und einmal dem Dunkelfeld zugeordnet werden, weil eine eindeutige Zusammenführung der Erfassungen nicht möglich war.

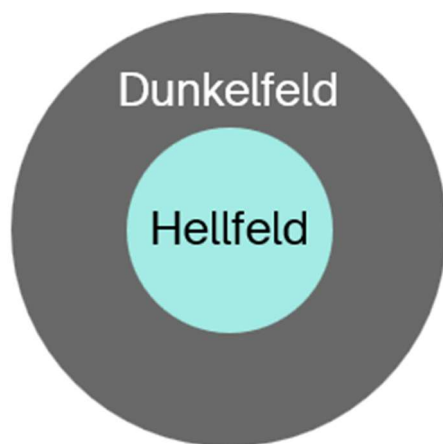


Abb. 2: Gesamtunfallgeschehen mit Differenzierung zwischen Hell- und Dunkelfeld.

Das **Gesamtunfallgeschehen** ergibt sich aus der Summe von Hell- und Dunkelfeld und umfasst somit alle Strassenverkehrsunfälle sowie die daraus resultierenden Personen- und Sachschäden.

Die **Dunkelziffer** wird definiert als das Verhältnis der Anzahl Unfälle oder Verletzten im Gesamtunfallgeschehen zur Anzahl der im Hellfeld registrierten Fälle. Aus der Menge im Hellfeld lässt sich durch Multiplikation mit der Dunkelziffer das Gesamtunfallgeschehen ermitteln.³ Die methodische Herausforderung bei der Quantifizierung der Dunkelziffer besteht in der Bestimmung des Gesamtunfallgeschehens bzw. der Fälle im Dunkelfeld.

Tab. 1 verdeutlicht die im Bericht verwendeten Begriffe anhand einer Beispielberechnung mit fiktiven Zahlen.

³ In der Literatur sind auch andere Konzepte zur Beschreibung der Grösse des Dunkelfelds gebräuchlich. Beispielsweise die «Erfassungsrate», welche den Anteil der im Hellfeld erfassten Fälle am Gesamtunfallgeschehen misst ($= 1 / \text{Dunkelziffer}$). Teilweise kommt auch ein «Untererfassungs-Korrekturfaktor» (engl. Under-reporting Correction Factor UCF) zum Einsatz, der sich aus dem Verhältnis zwischen nicht gemeldeten und gemeldeten Fällen ergibt ($= \text{Dunkelziffer} - 1$)

Tab. 1: Veranschaulichung der Dunkelzifferberechnung

Ausprägung Differenzierungs- merkmal	Hellfeld	Dunkelfeld	Gesamtunfall- geschehen	Dunkelziffer
Schwerverletzt	10	2	$10 + 2 = 12$	$12 / 10 = 1.2$
Leichtverletzt	20	10	$20 + 10 = 30$	$30 / 20 = 1.5$

Je nach Verwendungszweck der Dunkelziffer sind verschiedene räumliche Abgrenzungen für die Feststellung des Gesamtunfallgeschehens relevant:

- **Territorialitätsprinzip** = berücksichtigt alle Unfälle, die sich in der Schweiz bzw. auf Schweizer Staatsgebiet ereignet haben, unabhängig vom Wohnort, Arbeitsort und Nationalität der beteiligten Person(en).
- **Inländerprinzip** = berücksichtigt alle Unfälle von Personen mit Wohnsitz in der Schweiz, unabhängig vom Ort des Unfallereignisses.

In der Schweiz basiert die amtliche Statistik der Strassenverkehrsunfälle auf dem Territorialitätsprinzip, sodass alle Unfälle auf Schweizer Hoheitsgebiet erfasst werden, unabhängig von Wohnort und Nationalität der Beteiligten.

Für die Differenzierung des Gesamtunfallgeschehens (und der Dunkelziffer) lässt sich je nach Verwendungszweck der Dunkelziffern (Prävention, Schutzbedarf, verursachergerechte Finanzierung) eines der folgenden Prinzipien anwenden:

- **Monitoringprinzip** = ein Unfallopfer wird der Ausprägung eines Differenzierungsmerkmals (z. B. Fahrzeugkategorie) zugeordnet, zu der es selbst gehört. Das Monitoringprinzip eignet sich für Aussagen, welche Personengruppe(n) (z. B. Velofahrende) am meisten Schäden erleidet und allenfalls besser geschützt werden sollte(n).
- **Verursacherprinzip** = ein Unfallopfer wird der Ausprägung eines Differenzierungsmerkmals (z. B. Fahrzeugkategorie) zugeordnet, das den Unfall (mutmasslich) verursacht hat. Es spielt dabei keine Rolle, welche Merkmale das Unfallopfer aufweist. Das Verursacherprinzip eignet sich für Aussagen dazu, welche Personengruppe (z. B. Autofahrende) am meisten Schäden verursacht.

Die Wahrscheinlichkeit, dass ein Unfall polizeilich registriert wird (**Erfassungswahrscheinlichkeit**), hängt von verschiedenen Faktoren ab. Wie in Kapitel 1.1 dargestellt, steigt die Wahrscheinlichkeit einer polizeilichen Erfassung mit der Schwere der Verletzungen der Beteiligten (s. auch Abb. 1).

2 Methodik

Für die Bearbeitung des Projekts wurden verschiedene logisch aufeinander aufbauende Arbeitsschritte durchgeführt. Zu Beginn wurde eine Literaturrecherche zum nationalen und internationalen Forschungsstand und insbesondere auch zu den verwendeten statistischen Methoden zum Thema Dunkelziffer durchgeführt. In einem Workshop wurden die Bedürfnisse relevanter Akteure erhoben, beispielsweise bezüglich der Differenzierung oder der Periodizität der Dunkelzifferberechnung. Auf Basis dieser Erkenntnisse wurde ein Methodenvorschlag erarbeitet, der anschliessend mit den Stakeholdern konsolidiert wurde.

Im Folgenden werden die Arbeitsschritte detailliert beschrieben.

2.1 (Inter-)Nationaler Stand der Forschung

In diesem Schritt wurde der nationale und internationale Forschungsstand zur Dunkelzifferberechnung analysiert. Für die Schweiz wurde untersucht, welche Datenquellen und Methoden bisher zur Berechnung der Dunkelziffer verwendet wurden.

Eine wesentliche Grundlage für die Analyse ausländischer Ansätze bildeten die Ergebnisse einer Befragung, die das ASTRA im August 2021 bei den Mitgliedsländern der International Traffic Safety Data and Analysis Group (IRTAD) durchgeführt hatte. Das ASTRA stellte dem Forschungsteam die Antworten dieser Befragung zur Verfügung.

Zusätzlich wurden weitere öffentlich zugängliche, wissenschaftliche Artikel (Fokus Übersichtsartikel) und Buchkapitel berücksichtigt. Dabei wurde auch geprüft, ob die Resultate anderer Forschungsprojekte, die von den ASTRA-Arbeitsgruppen oder vom Bundesamt für Raumentwicklung (ARE) ausgeschrieben wurden, genutzt werden könnten, insbesondere die ASTRA-Projekte MFZ_20_04A «Situation en matière de sécurité du trafic cycliste sur les routes et dans les carrefours» [10] und VPT_20_05B «Velo-Infrastruktur-Sicherheitsinstrumente VISSI» [11] sowie das ARE-Projekt «Externe Kosten des Verkehrs 2021» [12].

Im Fokus der Analyse stand insbesondere, welche Datenquellen und statistischen Methoden zur Schätzung der Dunkelziffern verwendet wurden. Weiter war von Interesse, welche Definitionen verwendet, welche Differenzierungen vorgenommen wurden und welche Differenzierungen als relevant und notwendig erachtet wurden.

2.2 Bedürfnisse der Stakeholder

Da die entwickelte Methodik einen praxisnahen Nutzen für Fachleute und Entscheidungsträger im Bereich der Strassenverkehrssicherheit haben sollte, wurden zu Projektbeginn relevante Stakeholder zu einem halbtägigen Workshop eingeladen. Dieser wurde von den Forschenden moderiert und diente der Erhebung von Erwartungen und Bedürfnissen.

Die Auswahl der Teilnehmenden erfolgte in Abstimmung mit der Begleitkommission (BK) des Forschungsprojekts auf Basis einer Einschätzung zum professionell-wissenschaftlichen Bezug zum Forschungsthema und zum Bedarf an differenzierten Grundlagen zur Dunkelziffer im Strassenverkehrsunfallgeschehen (z. B. Sicherheitsverantwortliche, Entscheidungsträger im Bereich Strassenverkehrssicherheit, Mitglieder der BK).

Ziel des Workshops war eine Übersicht über die von den Stakeholdern erwartete Differenzierung der Dunkelzifferabschätzung, deren Periodizität und die relevanten Anwendungsfälle.

2.3 Auswahl statistischer Modelle

In diesem Arbeitsschritt wurde in einer Literaturrecherche eine Übersicht von infrage kommenden statistischen Methoden erstellt. Auch hier wurden die Resultate der IRTAD-Befragung einbezogen (s. Kapitel 2.1). Dabei wurde auch nach Lösungsansätzen in anderen Branchen mit vergleichbaren Problemstellungen gesucht (z. B. Ökologie, Epidemiologie, Soziologie). Die Suche beinhaltete auch die Recherche nach Software-Paketen zur Implementierung der Methode (z. B. R-Packages). Es wurde berücksichtigt, dass je nach Merkmal (z. B. Alter der verletzten Personen) die methodischen Lösungsansätze unterschiedlich sein können. Ebenfalls berücksichtigt wurden potenziell zusätzlich nutzbare Datenquellen bezüglich Unfallereignissen (Kapitel 2.4) sowie die im Stakeholderworkshop identifizierten Bedürfnisse (Kapitel 2.2). Schliesslich wurden diejenigen statistischen Modelle ausgewählt, die bezüglich der Forschungsfragen dieses Projekts, der Bedürfnislage der Stakeholder und aufgrund ihrer praktischen Anwendbarkeit am geeignetsten erschienen.

2.4 Potenziell nutzbare Datengrundlagen

Dieser Arbeitsschritt beinhaltete eine breit gefasste Auslegeordnung bereits existierender oder noch zu erschliessender Datenquellen, die Angaben zu Strassenverkehrsunfällen enthalten oder sich zur Erhebung von Unfalldaten eignen und sich für einen Abgleich mit den offiziellen Strassenverkehrsunfalldaten des ASTRA anbieten. Ihre Charakteristiken wurden dokumentiert und hinsichtlich ihrer Nutzbarkeit im vorliegenden Kontext beurteilt (z. B. allgemeine Datenqualität, Repräsentativität, Periodizität). Der Arbeitsschritt lief mehrheitlich parallel zum Arbeitsschritt «Auswahl statistischer Methoden» (Kapitel 2.3).

2.5 Erarbeitung der Methodenvorschläge

Die Ergebnisse und Erkenntnisse der vorangegangenen Arbeitsschritte wurden miteinander verknüpft, um konkrete Vorschläge für priorisierte Varianten der statistischen Methodik zu erarbeiten. Ziel war es, einen allgemein gültigen und standardisierbaren methodischen Ansatz zur Berechnung der Dunkelziffern über alle Merkmalskombinationen hinweg zu entwickeln. Jedoch war davon auszugehen, dass bei bestimmten Merkmalskombinationen möglicherweise andere Methodiken erforderlich sind, basierend auf unterschiedlichen Datenquellen und Berechnungsmodellen. Dies wird in Abb. 3 vereinfacht und schematisch dargestellt.

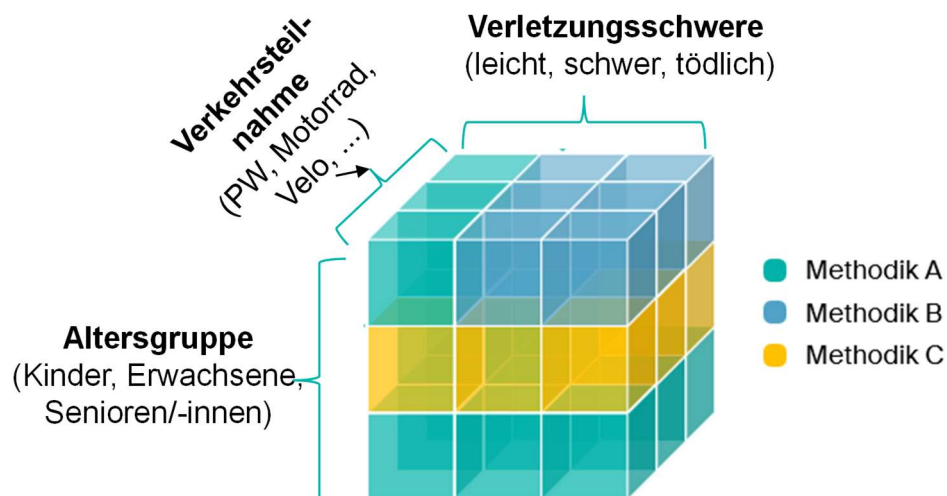


Abb. 3: Vereinfachte, beispielhafte Darstellung möglicher Merkmalskombinationen und möglicherweise erforderlicher Methodenvarianten.

Im Rahmen der Methodenentwicklung wurden Simulationen für verschiedene Methoden sowie eine Pilotberechnung für eine vorgeschlagene Variante durchgeführt. In einem zweiten Stakeholder-Workshop wurden die ausgearbeiteten Methodenvarianten den gleichen Teilnehmenden des ersten Workshops vorgestellt, gemeinsam diskutiert und konsolidiert.

3 (Inter-)Nationaler Stand der Forschung

3.1 Nationale Literatur

Für die Schweiz wurde bereits 1990 die Dunkelziffer der Strassenverkehrsunfälle abgeschätzt: Anhand der Statistik der Unfallversicherung UVG (Unfallversicherungsgesetz) und der darin enthaltenen Information zum Vorliegen eines Polizeiprotokolls bei einem Versicherungsfall wurde die Dunkelziffer für differenzierte Verkehrsteilnehmergruppen und Verletzungsschweren im Unfalljahr 1984 ermittelt [13]. Das Gesamtunfallgeschehen im Strassenverkehr wurde in dieser Studie auf rund 122 000 Personen geschätzt und entsprach damit rund dem 4-Fachen der polizeilich registrierten Unfälle.

Eine erneute Berechnung wurde anhand der Daten einer im Jahr 2011 durchgeführten Haushaltsbefragung im Jahr 2015 publiziert und wird bis heute jährlich aktualisiert, indem die Veränderungen der Anzahl der leicht und schwer verletzten Personen in der Statistik der polizeilich registrierten Strassenverkehrsunfälle zum jeweiligen Vorjahr – differenziert nach fünf Altersklassen, Geschlecht und fünf verschiedenen Verkehrsteilnehmergruppen – auf die Basis-Hochrechnung übertragen werden [2, 14].

Aktuell wird die Anzahl der bei Strassenverkehrsunfällen verletzten Personen auf rund 77 000 geschätzt, 20 500 wurden von der Polizei registriert (Jahr 2021). Für den Veloverkehr werden ungefähr 9-mal mehr Verletzte geschätzt, als in der amtlichen Statistik registriert werden (Polizeirapporte) [14]. Das wahre Ausmass der Strassenverkehrsunfälle in der Schweiz spielt insbesondere bei der Berechnung der volkswirtschaftlichen Kosten des Verkehrs [2, 12, 15–18], von Kennzahlen (Unfallraten) und Kostensätzen in den VSS-Normen (VSS 41 713 und SN 641 724) [19] eine zentrale Rolle.

Im Jahr 2018 wurde in der Stadt Zürich eine Befragung bei Velofahrenden durchgeführt [20]. Für die Rekrutierung der Studienteilnehmenden wurde die Befragung auf verschiedenen Kanälen (z. B. durch Online- und Printmedien oder Social Media) beworben. Die Dunkelziffer in dieser Studie betrug 8,6 für Unfälle nur mit Sachschäden, 6,2 für leichte Verletzungen und 1,7 für schwere Verletzungen.

3.2 Internationale Literatur

3.2.1 Methoden zur Schätzung der Dunkelziffer (IRTAD)

Auch international findet das Problem der Dunkelziffer von Strassenverkehrsunfällen Beachtung. In der OECD-Arbeitsgruppe IRTAD (International Traffic Safety Data and Analysis Group) wurden bereits Übersichten zur Handhabung der Problematik in den IRTAD-Mitgliedsländern publiziert [21–23]. Für die Lösung des Problems wurde die Datenverknüpfung mit Spitaldaten und Todesursachenstatistiken, die Durchführung von Befragungen sowie der Einbezug von Versicherungsdaten vorgeschlagen. Die Weltgesundheitsorganisation (WHO) schlägt für die Erfassung von tödlichen Strassenverkehrsunfällen die Verknüpfung von verschiedenen Datenquellen vor, kombiniert mit der Capture-Recapture-Methode, um das Gesamtausmass der Todesfälle abzuschätzen (s. Kapitel 5) [24]. In einer im Jahr 2021 veröffentlichten systematischen Übersichtsarbeit wurde eine umfassende Literaturrecherche zum Thema Dunkelziffer durchgeführt [25]. Gemäss dieser Studie sind in Ländern mit hohem Einkommen zwischen 35 % und 97 % der Todesfälle und zwischen 16 % und 82 % der Schwerverletzten in den Polizeistatistiken zu finden.

Mitte August 2021 führte das ASTRA eine Befragung der IRTAD-Mitgliedsländer zum Thema Dunkelziffer durch. Neun der 15 an der Befragung teilnehmenden Länder hatten gemäss eigenen Angaben mindestens einmal eine Abschätzung der Dunkelziffer für ihr Land durchgeführt. Von den sechs Ländern, die angegeben haben, keine (aktuellen) Zahlen zu der Unterschätzung der Zahlen in der Polizeistatistik zu haben (Dänemark, Deutschland, Japan, Mexiko, Polen und Slowenien) gab Dänemark als Begründung an, dass ein Vergleich von Polizei- und Spitaldaten an unterschiedlichen Definitionen in den beiden Datensätzen scheiterte, und in Japan lag die Begründung darin, dass bei den Verletzten im

Strassenverkehr keine Dunkelziffer existiere, da die Polizei über alle Unfälle mit Verletzungsfolge informiert sei. Deutschland gab an, über keine kürzlich durchgeführte Abschätzung zu verfügen. In Mexiko wurde zwar das Gesamtausmass von Strassenverkehrsunfällen für tödliche Unfälle mit Todesursachenstatistiken und für nicht tödliche Unfälle mit Angaben von Haushaltsbefragungen geschätzt, jedoch nicht mit der offiziellen Statistik verglichen, wodurch im engeren Sinne keine «Dunkelziffer» berechnet wurde [26].

In den Ländern, die Berechnungen zur Unterschätzung von Verletzten im Strassenverkehr in den Polizeidaten vornahmen, wurden in den meisten Fällen Spitaldaten beigezogen (s. Tab. 15, Anhang I). Dies war auch in der oben genannten Übersichtsarbeit der Fall (ein Teil der eingeschlossenen Studien überschneidet sich mit denjenigen, die in Tab. 15 beschrieben sind). Andere verwendete Datenquellen für einen Vergleich mit den polizeilich registrierten Unfällen (IRTAD-Befragung und Übersichtsarbeit) umfassen (Haushalts-)Befragungen, Todesursachenstatistiken, Strafregister, Versicherungsdaten, ein in-Depth-Unfalldatenset, Daten von Feuerwehr und Brandschutzversicherungen, Notfallregistern, Bestattungsinstituten, Krankenversicherern sowie Zeitungsartikel. Die Differenzierung beschränkte sich meist auf die Verletzungsschwere und Verkehrsteilnahme, vereinzelt wurde auch nach Unfallhergang oder Tageszeit unterschieden. Die «feinste» Differenzierung erfolgte in einer Studie aus Indien, in der ein in-Depth-Unfalldatenset die Datengrundlage bildete (s. Tab. 15, Anhang I).

In Studien zur Berechnung von Dunkelziffern bzw. anderer Kenngrössen zur Unterschätzung des Gesamtunfallgeschehens in den Polizeidaten wurden verschiedene statistische Methoden genutzt (s. dazu auch die Ausführungen in Kapitel 5). Die WHO schätzt das internationale Ausmass tödlicher Unfälle mittels Regressionsmodellen [27]. Dies ist vor allem für Länder nötig, in denen keine Mortalitätsstatistiken verfügbar sind oder diese als qualitativ schlecht eingeschätzt werden. Die Schätzungen der WHO werden auch für Länder durchgeführt, die über gute Registrierungssysteme verfügen. Vergleiche zwischen den WHO-Schätzungen der Todesfälle und den nationalen Registern weisen aber teils starke Differenzen auf [28]. In anderen Studien wurden oft (Register-)Daten verknüpft bzw. zusammengeführt und zum Teil mittels eines Capture-Recapture-Verfahrens, das in der Biologie zur Schätzung unbekannter Populationsgrössen genutzt wird, hochgerechnet. Ausgewiesen wurde dann oft der Prozentsatz der Spitaldaten, die nicht in den Polizeidaten zu finden waren, oder ein Faktor mit dem Verhältnis der errechneten Anzahl aus der Hochrechnung verglichen mit der Anzahl im Helffeld (analog der hier verwendeten Definition der Dunkelziffer).

3.2.2 Differenzierungsmerkmale mit Einfluss auf die Erfassungswahrscheinlichkeit

Angesichts der untersuchten Methodik zur Differenzierung der Dunkelzifferschätzung stellt sich die Frage, welche Differenzierungsmerkmale überhaupt einen relevanten Einfluss auf die Erfassungswahrscheinlichkeit (und damit auf die Dunkelziffer) haben. Die Erfassungswahrscheinlichkeit oder -rate ist dabei nicht mit der Inzidenz – also der Unfallhäufigkeit – zu verwechseln: Risikoreiches Fahrverhalten kann zu einer höheren Unfallhäufigkeit führen, muss jedoch im Vergleich zum normalen Verhalten nicht zwingend zu einer höheren oder tieferen Erfassungsrate führen.

Tab. 2 zeigt eine Auswahl an internationalen Studien zur Dunkelziffer im Strassenverkehrsunfallgeschehen, welche insbesondere der Frage nach den relevanten Einflussfaktoren auf die Erfassungswahrscheinlichkeit nachgehen. Spalten 2 und 3 zeigen die jeweils ausgewiesenen Ausprägungen der in diesen Studien am häufigsten verwendeten Differenzierungsmerkmale Verletzungsschwere und Verkehrsteilnahme, wobei beide Kriterien als statistisch signifikant eingestuft werden. Einzelne Studien haben lediglich den univariaten Einfluss dieser beiden Variablen auf die Erfassungswahrscheinlichkeit untersucht, andere wiederum kombinieren die Ausprägungen (Spalte 4 von Tab. 2). Fünf Studien haben zusätzlich den Einfluss von weiteren Faktoren auf die Erfassungswahrscheinlichkeit betrachtet (Spalte 5). Untersucht wurden u. a. das Geschlecht und Alter, aber auch die Infrastruktur (Anzahl Fahrspuren, signalisierte Geschwindigkeit) und der Raumtyp. Insgesamt haben sich vor allem die Variablen Verletzungsschwere, Verkehrsteilnahme, Unfallhergang (Alleinunfall vs. Kollision), das Alter und der Raumtyp als signifikante Einflussfaktoren erwiesen.

Tab. 2: Studien zum Einfluss ausgewählter Differenzierungsmerkmale auf die Erfassungswahrscheinlichkeit verschiedener Strassenverkehrsunfälle

Studie	Verletzungsschweren	Verkehrsteilnahmen	Kombinationen / Interaktionen getestet bzw. ausgewiesen	Signifikanz weiterer Einflussfaktoren
A Harris (1990) [29] Niederlande	Behandlung: Ambulant, stationär, extra muros	Velo	Nein	Signifikant: Alter (Kinder), Unfallhergang (Alleinunfall)
Alsop, Langley (1999) [30] Neuseeland	AIS (1–5; unbekannt), Spitalaufenthaltsdauer	Auto (Lenker/-in, Passagier/-in), Motorrad	Ja	Signifikant: Alter (< 14), Raumtyp, Monat, Anzahl beteiligter Fahrzeuge Nicht signifikant: Geschlecht, Wochentag, Ethnie
Elvik, Mysen (1999) [31] 13 Länder	Tödlich, schwer, leicht, sehr leicht	Auto (Lenker/-in, Passagier/-in), Motorrad, Velo, Fussgänger/-in (mit Kollision vs. Alleinunfall)	Nein	
Ward et al (2006) [32] Vereinigtes Königreich	Tödlich, schwer, leicht	Alle, Velo, Fussgänger/-in, Auto (Lenker/-in, Passagier/-in), Motorrad	Nein	
Loo, Tsui (2007) [33] Hong Kong	Injury Severity Factors (ISS), Spitalaufenthaltsdauer	Auto (Lenker/-in, Passagier/-in hinten/vorne), öffentlicher Verkehr, Motorrad, Velo, Fussgänger/-in, Unbekannt	Nein	Signifikant: Alter Nicht signifikant: Geschlecht
Yannis et al (2014) [34] 7 Europäische Länder	Leicht, schwer; MAIS 1–2, MAIS > 2	Auto, Motorrad, Velo, Fussgänger/-in, Alle	Ja	
Janstrup et al. (2014) [35] Funen, DK	Tödlich, schwer, leicht	Velo, Fussgänger/-in, Mofa, Motorrad, Auto, Bus, Andere		Signifikant: Alter, Geschlecht, Tragen von Schutzausrüstung (Helm, Gurt), Anzahl Beteiligte, signalisierte Geschwindigkeit, Art der Verletzung
Watson et al. (2015) [36] Queensland, AUS	Leicht, schwer	Motorrad, Velo, Fussgänger/-in, Auto (Lenker/-in, Passagier/-in)	Nein	Signifikant: Alter (< 15), Raumtyp Nicht signifikant: Geschlecht

Insgesamt ist die wissenschaftliche Grundlage zur Einschätzung der Signifikanz einzelner Einflussfaktoren und deren Interaktion klein, zeitlich weit zurückliegend, und deren Repräsentativität für die Schweiz kann nicht geprüft werden. Aus den Erkenntnissen lassen sich aber Hinweise für eine notwendige Differenzierung der Dunkelziffern bei einer Schätzung für die Schweiz gewinnen. Die Ausführungen zeigen dabei die hohe Komplexität und die Vielfalt der möglichen Abhängigkeiten und Interaktionen zwischen den Einflussfaktoren auf, die bei der Auswahl der Differenzierungsmerkmale zu berücksichtigen ist:

- **Verletzungsschwere**

Wissenschaftlich gut belegt ist der direkte positive Zusammenhang zwischen Verletzungsschwere als «Ergebnis» eines Unfalls und der Erfassungswahrscheinlichkeit (alle Studien in Tab. 2): Je schwerer die Verletzung, desto höher ist die Erfassungswahrscheinlichkeit (+) im Helfeld und desto tiefer die Dunkelziffer. Die Verletzungsschwere hat dabei gleichzeitig aus ökonomischer Sicht den grössten Einfluss auf die Kosten eines Strassenverkehrsunfalls, d. h., höhere Kosten stellen einen Anreiz dar, den Unfall zu melden.

- **Verkehrsteilnahme (Fahrzeugkategorie)**

Der Einfluss der Verkehrsteilnahme auf die Erfassungswahrscheinlichkeit ist nachgewiesen (alle Studien in Tab. 2). Inwiefern der Einfluss direkt (kausal) oder indirekt (z. B. über unterschiedliche Anteile an Alleinunfällen oder Verletzungsschweren) ist, kann

jedoch nicht abschliessend beurteilt werden. Mögliche erklärende Faktoren können Unterschiede in der Fahrgeschwindigkeit, Unterschiede in der Versicherungsdeckung oder vorhandener Schutzmechanismen sein. Eine Rolle dürfte auch die Zahl der Plätze im Fahrzeug (Motorrad vs. Kleinbus) und der Besetzungsgrad der beteiligten Fahrzeuge spielen (s. Einflussfaktor «Unfallhergang / Anzahl Beteiligte»). Je auffälliger und schwerer ein Fahrzeug (z. B. LKW, Reisedar), desto eher ist bei einem Unfall auch die Infrastruktur oder der übrige Verkehr betroffen, was tendenziell zu einer höheren Erfassungswahrscheinlichkeit (+) führen dürfte. Mitfahrende in Personenwagen haben gemäss einigen Studien eine geringere Erfassungswahrscheinlichkeit (-) als die Fahrzeugführenden (s. Studien B, E, H in Tab. 2).

- **Unfallhergang / Anzahl Beteiligte** (Alleinunfall (mit oder ohne Mitfahrende) oder Kollision mit anderen Verkehrsteilnehmenden)

Je mehr Fahrzeuge und Personen an einem Unfall beteiligt sind, desto höher ist gemäss einzelnen der obigen Studien die Erfassungswahrscheinlichkeit (+) (s. Studien A, B, G in Tab. 2). Bei Alleinunfällen ist von einer tieferen Erfassungswahrscheinlichkeit (-) im Hellfeld auszugehen, vor allem, wenn nur eine Person am Unfall beteiligt ist (-). Bei Kollisionen ist die Erfassungswahrscheinlichkeit höher (+), weil mehr Objekte und Personen beteiligt sind. Für eine höhere Erfassung bei Kollisionen sprechen voraussichtlich auch versicherungsrechtliche Haftungsfragen (Polizeirapport als Grundlage der Verschuldensbeurteilung). Innerhalb der Kollisionen dürfte die Erfassung auch massgeblich von der aus der Kollision resultierenden Verletzungsschwere abhängen (s. «Verletzungsschwere»). Diese wiederum hängt auch von den Eigenschaften der beteiligten Fahrzeuge ab (s. «Verkehrsteilnahme»). Je grösser die Unterschiede bei Gewicht und Geschwindigkeit der beteiligten Objekte, desto höher dürfte bei Kollisionen die Erfassungswahrscheinlichkeit (+) sein.
- **Alter der Beteiligten bzw. der Opfer**

In den Studien zeigt sich, dass das Alter einen signifikanten Einfluss auf die Erfassung hat (s. Studien A, B, E, G, H in Tab. 2). Der Zusammenhang geht jedoch nicht in allen Studien in die gleiche Richtung. Die meisten Studien finden eine tiefere Erfassungswahrscheinlichkeit bei Kindern und Jugendlichen (-). Jugendliche und junge Erwachsene melden unter Umständen Unfälle weniger häufig, weil sie die Meldepflichten nicht kennen oder Sanktionen oder Einschränkungen erwarten.
- **Geschlecht der Beteiligten bzw. der Opfer**

Die Verkehrsmittelwahl, die Verteilung der Verkehrszwecke (Arbeit, Freizeit, Einkauf etc.) und auch das Risikoverhalten ist teilweise geschlechtsspezifisch. Diese geschlechterspezifische Mobilität ist teilweise bereits in den obigen Kriterien (z. B. Verkehrsteilnahme, Unfallhergang) abgebildet. Es erscheint nachvollziehbar, dass ein geschlechtsspezifisches, risikoreicheres Verkehrsverhalten zu insgesamt mehr Unfällen und Verletzten führt (Inzidenz). Der Einfluss auf die Erfassungswahrscheinlichkeit ist jedoch nicht klar. So kann das Fahren mit überhöhter Geschwindigkeit einerseits zu einem höheren Anteil an Alleinunfällen (tiefere Erfassungswahrscheinlichkeit), andererseits aber auch zu schwereren Verletzungen führen (höhere Erfassungswahrscheinlichkeit). Es gibt keine schlüssigen Hinweise darauf, dass je nach Geschlecht der Verletzten die Strassenverkehrsunfälle systematisch unterschiedlich häufig gemeldet werden. In der Literatur ist die Notwendigkeit einer solchen Differenzierung deshalb eher umstritten (s. Studien B, E, H in Tab. 2). Dennoch kann eine solche Differenzierung beim Ausweis des Gesamtunfallgeschehens erwünscht sein, beispielsweise, um später unterschiedliche ökonomische Folgen eines Unfalls je nach Geschlecht abbilden zu können.
- **Raumtyp (Ort des Unfalls)**

Der Raumtyp erweist sich gemäss zwei Studien als signifikant (s. Studien B, H in Tab. 2). Eine Differenzierung nach Raumtyp stützt sich insbesondere auf die Annahme, dass unterschiedliche Verkehrssysteme, Unterschiede in der Verkehrsmittelwahl und unterschiedliche Kontrollmechanismen zu einer unterschiedlichen Erfassungswahrscheinlichkeit führen. In stark frequentierten städtischen Räumen ist dabei auch die Wahrscheinlichkeit höher, dass der Unfall von Dritten oder Überwachungsinstrumen-

ten beobachtet und gemeldet wird oder dass der Verkehr durch den Unfall beeinträchtigt wird, was die Erfassungswahrscheinlichkeit (+) erhöht. Zusätzlich könnte eine Rolle spielen, dass im urbanen Raum leichtere Fahrzeuge und Fussgängerinnen (s. Verkehrsteilnahme) präsenter sind und der Anteil an Innerortsstrecken mit tieferen Geschwindigkeiten höher ist als im ländlichen Raum oder auf Autobahnen (s. Infrastrukturtyp).

- **Infrastrukturtyp (Ort des Unfalls)**

Verschiedene Infrastrukturtypen wie Innerortsstrassen, Ausserortsstrassen oder Autobahnen werden in unterschiedlicher Intensität überwacht. Auf Autobahnen mit hoher Geschwindigkeit (s. Studie G in Tab. 2) ist nach Unfällen oft die Infrastruktur beschädigt oder der Verkehrsfluss beeinträchtigt, was die Erfassungswahrscheinlichkeit (+) erhöht. Bei hohen Geschwindigkeiten treten zudem öfter schwere Verletzungen auf, was sich ebenfalls positiv auf die Erfassungswahrscheinlichkeit (+) auswirkt. Autobahnen sind zudem durch Fussgänger/-innen und Velos nicht nutzbar (s. Verkehrsteilnahme). Auch innerorts können Unfälle rasch verkehrliche Auswirkungen haben (+). Auf Ausserortsstrassen im ländlichen Gebiet dürfte dieser Effekt schwächer sein (-).

- **Zeitpunkt des Unfalls (Tageszeit / Jahreszeit)**

In der Nacht sind weniger Verkehrsteilnehmende bzw. Passantinnen und Passanten und an gewissen Orten weniger Polizeipatrouillen unterwegs, was für eine tiefere Erfassungswahrscheinlichkeit (-) spricht. Im Winter oder bei schlechter Witterung steigt u. U. der Anteil von Alleinunfällen an, was zu einer tieferen Erfassungswahrscheinlichkeit (-) beitragen kann (s. Unfallhergang).

Als signifikant und prioritär werden in der Regel die Verletzungsschwere, die Verkehrsteilnahme, der Unfallhergang (Alleinunfall vs. Kollision) und das Alter eingestuft. Durch diese Merkmale lassen sich in der Regel bereits grosse Unterschiede in der Dunkelziffer erklären. Diese Merkmale interagieren oft mit weiteren Merkmalen wie beispielsweise dem Geschlecht und dem Unfallort (Infrastrukturtyp, Raumtyp). Gerade aufgrund dieser Interaktionen ist es eine Herausforderung, die darüber hinaus gehenden, für die Schweiz zusätzlich notwendigen Differenzierungen zu identifizieren, d. h. jene, die weitere Teile der Unterschiede in der Erfassungswahrscheinlichkeit erklären können.

Müssen beispielsweise aufgrund von fehlenden Daten einzelne signifikante Differenzierungen weggelassen werden, kann es zu einer Verzerrung kommen, wenn als einzige Alternative die Verteilung aus dem Helfeld zur weiteren Differenzierung des Gesamtunfallgeschehens verwendet werden müsste. Ein solcher Analogieschluss (Anteile einer Gruppe am Gesamtunfallgeschehen sind gleich wie im Helfeld) kann für eine zusätzliche Differenzierung des Gesamtunfallgeschehens jedoch sinnvoll sein, wenn keine systematischen Unterschiede in der Erfassungswahrscheinlichkeit erwartet werden.

In der Literatur kommen auch Teildifferenzierungen vor, z. B. eine zusätzliche Aufschlüsselung der Dunkelziffer für Leichtverletzte nach Altersklasse, aber ein Verzicht auf diese Differenzierung bei den Schwerverletzten, weil hier die Erfassungswahrscheinlichkeit unabhängig vom Alter bereits hoch ist und kleine Unterschiede sich durch kleine Subgruppen nicht gut abstützen lassen würden.

4 Bedürfnisse der Stakeholder

Im Rahmen eines Workshops wurden die Bedürfnisse hinsichtlich der Dunkelziffer im Strassenverkehrsunfallgeschehen aufgenommen. Die Teilnehmenden stammten aus verschiedenen Teilen der Bundesverwaltung (ASTRA, ARE, Bundesamt für Statistik (BFS)), einer kantonalen Verkehrspolizei (Basel-Land), einer Stadtverwaltung (Sicherheitsbeauftragter Stadt Zürich), der Sammelstelle für die Statistik der Unfallversicherung UVG (SSUV) sowie der Beratungsstelle für Unfallverhütung (BFU).⁴ Details zum Workshop können der Aktennotiz im Anhang II entnommen werden. Die wichtigsten Resultate des Workshops sind im Folgenden zusammengefasst.

4.1 Anwendungsfelder der Dunkelziffer

Vertiefte Kenntnisse über die Höhe der Dunkelziffer und deren Differenzierung ist in verschiedenen Bereichen relevant: in der Unfallprävention (Zielgruppenansprache, teilweise Repression), der Verkehrsplanung (Gestaltung und Bewertung von Projekten), als Grundlage für politische Entscheidungen (amtliche Statistik, externe Kosten) sowie für die Erfolgskontrolle von Massnahmen in der Verkehrssicherheit (beispielsweise ISSI).

4.2 Wissenslücken

Derzeit bestehen Wissenslücken und Unsicherheiten insbesondere im Bereich der Mikromobilität, bei neuen Fahrzeugtypen, der Elektromobilität und beim Veloverkehr (einschliesslich kleinerer Verletzungen, Gefährdungen und Sachschäden). Für die Prävention sowie eine sicherheitsorientierte Verkehrs- und Infrastrukturplanung sind umfassendere Informationen zu den Unfallursachen erforderlich. Besonders relevant sind hierbei Daten zum Unfallort («wo?»), zur dortigen Qualität der Infrastruktur («warum?») sowie allgemein zur Unfallursache («wer?»).

4.3 Frequenz der Erhebung

Die vorgeschlagene Methodik sollte eine Abschätzung des Gesamtunfallgeschehens und – im Verhältnis zum Hellfeld – der Dunkelziffer alle fünf bis sieben Jahre ermöglichen. Dazwischen ist eine laufende Fortschreibung notwendig, falls die Methode nicht jährlich angewendet werden kann, damit jährliche Gesamtzahlen ermittelt und kurzfristige Mobilitätstrends aufgenommen werden können. Die Dunkelzifferberechnung sowie Angaben zur jährlichen Fortschreibung sollten rund 12–18 Monate nach Verfügbarkeit der ASTRA-Unfalldaten vorliegen.

4.4 Zu verwendende «Abgrenzungen» (Begriffe siehe Kapitel 1.3)

Die Unfälle sollten im Einklang mit dem Hellfeld im Territorialitätsprinzip vorliegen (alle Unfälle in der Schweiz, unabhängig von Wohnort und Nationalität der Beteiligten), wenn möglich differenziert nach Kanton oder Raumtyp. Es besteht ein Bedürfnis nach Dunkelziffern sowohl im Monitoring- («welche Teilnehmenden sind von Verletzungen betroffen?») als auch im Verursacherprinzip («welche Teilnehmenden haben Verletzungen verursacht?»).

⁴ Eingeladen war auch eine Vertretung aus der privaten Fahrzeugversicherungsbranche.

4.5 Differenzierungswünsche

In Tab. 3 sind die Bedürfnisse hinsichtlich der Differenzierung des **Gesamtunfallgeschehens** und der **Dunkelziffer** aufgeführt. Die Tabelle zeigt für jedes Differenzierungsmerkmal, ob es sich auf die Person, das Objekt (z. B. Fahrzeug) oder das Ereignis (Unfall) bezieht, ob es bereits in der bisher von der BFU erstellten Hochrechnung des Gesamtunfallgeschehens enthalten ist [2] und ob es im Hellfeld (ASTRA Strassenverkehrsunfälle) vorhanden ist.

Tab. 3: Bedürfnisse der Stakeholder zur Differenzierung der Dunkelziffer

Differenzierungsmerkmal	Bezug	Bisher («BFU-Hochrechnung») bzw. Neu	Differenzierung im Hellfeld vorhanden?
Verletzungsschwere	Person	Bisher, aber Abstimmung mit ASTRA-Verletzungsschweren gewünscht	Ja
Verkehrsteilnahme bzw. Fahrzeugart (Monitoring)	Objekt <> Person	Bisher, mehr Kategorien (siehe Text oben) gewünscht	Ja
Fahrzeugart (Verursacher)	Ereignis (Beteiligte) <> Objekt(e) <> Person	Neu	Ja (Ableitung aus Ereignis, Objekten und Person)
Alter	Person	Bisher	Ja
Geschlecht	Person	Bisher	Ja
Unfallhergang	Objekt (Anzahl)	Neu	Ja (Ableitung aus Ereignis und Objekt)
Ursache (persönlicher Zustand / Verhalten der Person)	Person	Neu	Ja (teilweise)
Ursache (Einfluss der Infrastruktur-Gestaltung)	Ereignis (Ort)	Neu	Ja (teilweise)
Zeitpunkt des Unfalls	Ereignis (Zeit)	Neu	Ja
Raumtyp oder Kanton	Ereignis (Ort)	Neu, auch als Teilmenge (pro Kanton) möglich	Ja (Ableitung aus Koordinate)

4.6 Priorisierung und Einordnung

Die durch die Stakeholder im Rahmen des Workshops vorgenommene Priorisierung ergab Folgendes:

- Besondere Bedeutung für die Dunkelziffer kommt der Verletzungsschwere und der Verkehrsteilnahme (Fahrzeugart) zu.
- Eine Differenzierung (des Gesamtunfallgeschehens) nach Alter und Geschlecht (wie bisher) wurde als selbstverständlich erachtet, insbesondere als Grundlage zur Ermittlung der Unfallkosten mit differenzierten Kostensätzen sowie zur gezielten Ansprache der jeweiligen Personengruppen im Rahmen der Präventionsarbeit.
- Die Differenzierung nach dem Zeitpunkt des Unfalls (neu) könnte insbesondere für die Polizeiarbeit nützliche Hinweise liefern.
- Ebenfalls gewünscht wird eine Differenzierung des Gesamtunfallgeschehens nach Raumtyp und/oder nach Kanton (unabhängig von den anderen Differenzierungen für das Total der Unfälle).

Die von den Stakeholdern priorisierten Differenzierungen stimmen grösstenteils mit den bereits in der internationalen Literatur verwendeten und als relevant eingestuften Merkmalen überein (s. Kapitel 3.2). Bei einigen Merkmalen bestand aber der Wunsch, diese noch differenzierter zu betrachten als diese im Hellfeld (UAP) kodiert sind (s. Tab. 3 letzte Spalte). Besonders bei tieferliegenden Unfallursachen, welche die Polizei vor Ort nicht voll-

ständig erkennen kann – wie die Infrastruktur-Gestaltung (nicht nur die objektive Gestaltung der Infrastruktur), der persönliche Zustand oder das Verhalten vor einem Unfall (zum Beispiel Unkonzentriertheit / Überforderung) oder auch, ob Gefährdungen oder Fehlverhalten durch unbekannte Dritte vorlagen – fehlen objektive Angaben auch in den Daten des ASTRA. Diese Differenzierungen sind auch international wenig berücksichtigt.

5 Auswahl statistischer Modelle

5.1 Grundlegende Überlegungen

Um das methodische Vorgehen zur Erfassung der Dunkelziffer von Strassenverkehrsunfällen auszuwählen, lohnt sich ein Gedankenexperiment:

Ein vollständiges Bild des Unfallgeschehens in einem gewissen Zeitraum (z. B. ein Kalenderjahr) könnte durch eine «Vollerhebung» gewonnen werden. Dazu könnte das gesamte (relevante) Strassennetz der Schweiz in überschaubare Segmente (z. B. 100 m Länge) unterteilt werden. Zusätzlich könnte der Beobachtungszeitraum (Kalenderjahr) in Abschnitte von jeweils einer Stunde unterteilt werden. Dadurch entstünde ein Raum-Zeit-Raster, in dem sich alle relevanten Vorfälle – Unfälle, aber evtl. auch Beinaheunfälle – prinzipiell erfassen liessen. In einer solchen Vollerhebung könnte man sich nun vorstellen, dass in jedem Raum-Zeit-Segment eine geschulte Person als Beobachter eingesetzt wird, die alle Vorkommnisse vollständig erfasst und dokumentiert.

Eine derartige Vollerhebung ist in der Praxis nicht umsetzbar. Sie dient jedoch als Ideal, dem man sich auf verschiedene Weise annähern kann. Dabei kommen primär drei Ansätze infrage:

- **Stichprobe:** Anstatt alle Segmente zu untersuchen, wird nur eine zufällige Auswahl untersucht. Die Stichprobentheorie ist ein gut etabliertes Feld der Statistik und bietet eine Vielzahl effizienter Methoden [37,38].
- **Datenspuren:** Falls ein Vorfall nicht direkt beobachtet werden kann, lassen sich möglicherweise nachgelagerte Ereignisse erfassen. Beispielsweise konnte gezeigt werden, dass die Sequenzierung von SARS-CoV-2 in kommunalen Abwasserproben dabei hilft, die Verbreitung verschiedener Linien zu erkennen und zu überwachen [39]. Auf ähnliche Weise werden verschiedene Institutionen bei einem Unfall ab einer gewissen Schwere in der Regel involviert (Polizei, Krankenhaus, Versicherungen etc.). Das heisst, auch wenn der eigentliche Vorfall nicht von der Polizei erfasst wurde, können über diese Institutionen dennoch Informationen zu den Vorfällen eingeholt werden.
- **Maschinelles Lernen:** Hierbei wird versucht, ein Ereignis (z. B. einen Unfall) anhand leichter messbarer Variablen vorherzusagen. Beim maschinellen Lernen wird eine Zielgrösse (z. B., ob ein Unfall in einem bestimmten Raum-Zeit-Segment stattgefunden hat) durch einfach zu erhebende erklärende Variablen modelliert. Voraussetzung ist ein ausreichend grosser Trainingsdatensatz, um ein Modell zu trainieren. Anschliessend lassen sich mithilfe neuer Beobachtungen für die erklärenden Variablen Vorhersagen zu relevanten Ereignissen machen [40].

Im Folgenden werden diese drei Ansätze genauer erläutert.

5.2 Stichprobe bzw. Befragungen

Durch eine zufällige Stichprobe kann die Gesamtzahl der Vorfälle mit prinzipiell beliebiger Genauigkeit erfasst werden. Zudem ermöglicht das zufällige Element der Stichprobe eine Genauigkeitsangabe der Schätzung.

Unfälle sind sehr seltene Ereignisse, d. h., in dem Raum-Zeit-Raster werden die meisten Segmente ohne nennenswerte Vorfälle sein. Eine naive Herangehensweise (z. B. Simple Random Sampling) würde daher eine sehr grosse Stichprobengrösse erfordern, um eine befriedigende Schätzgenauigkeit zu erreichen [41].

In den meisten Raum-Zeit-Segmenten werden sich keine Personen aufhalten (z. B. ein Teilstück einer Ausserortsstrasse um 3 Uhr nachts). Das bedeutet, dass dort keine Unfälle auftreten können. Eine naheliegende Verfeinerung ist daher die Verwendung von Clustern: Personen werden zufällig ausgewählt und zu ihren Unfällen im interessierenden Zeitraum befragt.

Der Vorteil dieser Methode liegt in ihrer Flexibilität: Es können beliebige Fragen gestellt werden, die von Interesse sind. Dabei ist allerdings zu beachten, dass die Teilnahmequote mit steigender Anzahl Fragen sinken dürfte [42]. Zudem ist die Stichprobentheorie ausge-reift, und es gibt zahlreiche Anwendungsbeispiele. Daher scheint es realistisch, dass ein effizientes Vorgehen gefunden werden kann, das eine befriedigende Genauigkeit erzielt.

Nachteile einer gross angelegten Befragung sind zum einen der beträchtliche Aufwand. Zudem handelt es sich lediglich um eine Momentaufnahme. Des Weiteren ist bekannt, dass Antworten bei Befragungen nicht immer zuverlässig sind – teils aufgrund von fehler-haftem Erinnerungsvermögen («Recall Bias» [43], teils mit Absicht, wenn z. B. eine Straf-verfolgung befürchtet wird (z. B. Drogenkonsum, «Social Desirability Bias», siehe z. B. Tourangeau et al., 2007 [44]). Ein weiterer Nachteil besteht darin, dass man bei der Anwen-dung des Territorialprinzips Personen mit Wohnsitz Ausland nicht erreichen kann.

Im Prinzip ist die Befragung von Personen nicht die einzige Möglichkeit, um die Raum-Zeit-Segmente effizient zu untersuchen. Auch das Aufstellen von Kameras, der Einsatz von Sensoren oder Patrouillen wäre eine Möglichkeit. Dabei muss jedoch immer ein Zufallsele-ment berücksichtigt werden, um die Genauigkeit der Schätzung beurteilen zu können. Zu-dem sind Aspekte des Datenschutzes (z. B. bei Kameras) und des Informationsgehalts zu beachten. So könnte eine hochauflösende Kamera aus datenschutzrechtlichen Gründen nicht infrage kommen und gleichzeitig nicht genügend Informationen über die beteiligten Personen liefern (z. B. Alter, Verletzungsschwere etc.).

Einige IRTAD-Mitgliedsstaaten haben die Befragungsmethodik gewählt, um die Zahl der nicht tödlichen Verkehrsunfälle zu ermitteln. So hat Mexiko beispielsweise mehrere Haus-haltsbefragungen durchgeführt [26]. In Ringel et al. (2023) werden die Vorteile einer Befra-gung insbesondere für den Veloverkehr hervorgehoben [20].

5.3 Datenspuren

Wenn Unfälle passieren, ziehen sie Abläufe nach sich, die eine Datenspur in verschiede-nen Datenquellen hinterlassen. Informationen zu einem Unfall werden beispielsweise von Polizei, Krankenhäusern oder Versicherungen erfasst. Das bedeutet, dass selbst wenn ein Unfall im zugehörigen Raum-Zeit-Segment nicht direkt beobachtet wurde, die Informatio-nen dennoch wie durch einen Trichter in gewisse Datenquellen fliessen können und man diese dort zu einem späteren Zeitpunkt prinzipiell abrufen kann. Häufig können jedoch auch in diesen Datenquellen nur Stichproben entnommen werden, sodass auf die Grund-gesamtheit geschlossen werden muss. Diese Herangehensweise, auch bekannt als Cap-ture-Recapture Estimation (oder Multiple Systems Estimation, s. Exkurs auf der nächsten Seite) wird in vielen Bereichen wie Ökologie, Epidemiologie oder Soziologie eingesetzt [26,45]. Während die Methode ursprünglich für die Erfassung von Tierpopulationen ent-wickelt wurde, sind die möglichen Vor-, aber auch Nachteile in der epidemiologischen An-wendung erkannt und diskutiert worden [46]. Die WHO empfiehlt diesen Ansatz zur Erfas-sung tödlicher Strassenverkehrsunfälle [24,47,48].

Diese Methode wird auch in mehreren IRTAD-Mitgliederstaaten angewandt. So hat bei-spielsweise die Bosch-Accident-Research-Gruppe in Indien Daten aus Polizeiberichten, Krankenhäusern und eigenen Patrouillen verknüpft [49]. In Schweden wurden Polizeibe-richte und Krankenhausdaten verglichen und verknüpft [50]. Ein europäisches Projekt führte in mehreren Ländern Polizei- und Krankenhausdaten zusammen [34,51]. In Irland wurden zudem Versicherungsdaten einbezogen [52].

Obwohl diese Daten in der Regel ohne weiteres Zutun «automatisch» gesammelt werden, sind der Zugang, die Auswertung und die Zusammenführung der Daten nicht trivial. Zudem müssen datenschutzrechtliche Aspekte berücksichtigt werden. Unter Umständen wäre z. B. eine Anfrage an eine Datenquelle bzw. eine Verknüpfung von Datenquellen technisch möglich, darf aber aus rechtlichen Gründen nicht durchgeführt werden (z. B. Einsatz von e-Call-Daten oder AHV-Nummer für den Abgleich von Datenquellen [53].

Exkurs zu Capture-Recapture:

Bei der Capture-Recapture Estimation geht es darum, mithilfe von mehreren Stichproben auf die Grundgesamtheit zu schliessen. Ein klassisches Beispiel ist die Schätzung der Anzahl von Fischen in einem See. Die Capture-Recapture-Methode geht dabei folgendermassen vor:

- 1) Capture: Fange zufällig einige Fische, markiere sie und lasse sie wieder frei (konkret könnten das die Unfälle sein, die polizeilich erfasst wurden)
- 2) Recapture: Fange nach einer gewissen Wartezeit für die Durchmischung der Fische erneut einige Fische zufällig. Zähle dabei die markierten sowie unmarkierten Fische (konkret könnten das die Unfälle sein, die polizeilich erfasst wurden (d. h. «markiert» wurden) und auch in der SSUV erscheinen (d. h. im «Recapture»-Schritt)).
- 3) Intuition: Ist die Population sehr gross, werden wir beim zweiten Fang kaum markierte Fische fangen. Ist hingegen die Population sehr klein ist, werden relativ viele markierte Fische erneut gefangen. Aus der Anzahl markierter Fische beim Recapture-Schritt lassen sich also Rückschlüsse auf die Gesamtpopulation ziehen.

5.3.1 Zu beachtende Aspekte: Filter, Linkage, Erweiterung

Um Datenquellen effektiv nutzen und sowohl mit anderen Datenquellen als auch mit dem Hellfeld verknüpfen zu können, müssen drei Aspekte bedacht werden: Filter, Linkage und eventuell Erweiterungen. Im Folgenden werden diese drei Aspekte genauer erklärt und anhand des Beispiels einer fiktiven Verknüpfung von Daten aus Polizeiberichten (mit dem Kürzel P bezeichnet, Hellfeld) und der Krankenhaus-/Spitalstatistik (Kürzel KH, Datenspu- ren) aus einem gegebenen Kalenderjahr illustriert.

Filter: Die Grundgesamtheit von Interesse («Würfelement», s. Kapitel 2.5) muss festgelegt werden (z. B. verunfallte Personenwageninsassen/-innen, die eine schwere Verletzung erlitten haben, oder E-Bike-Fahrende, die eine leichte Verletzung erlitten haben). In allen beteiligten Datenquellen müssen diese Kriterien identisch definiert sein und es muss möglich sein, nach ihnen zu «filtern». Wenn das nicht möglich ist, werden unterschiedliche Ereignisse miteinander verglichen und es ist keine klare Aussage möglich. Möglicherweise ist nur eine grobe Kategorisierung verwendbar.

Beispiel: In unserer fiktiven Verknüpfung werden folgende Filter verwendet: «Schwerverletzt», «Wohnsitz in der Schweiz» und «Unfallland Schweiz». Für dieses Beispiel wurden fiktive Daten mit einem wahren (unbekannten) Total für obigen Filter von 2000 Fällen simuliert. Angenommen, es landen nach Anwendung des Filters 1200 Fälle in P und 1400 Fälle in KH. Ohne weitere Informationen kann nicht festgestellt werden, ob diese Fälle überlappen (Doppelzählungen). Das heisst, in diesen beiden Listen gibt es total mindestens 1400 Fälle und höchstens 2600 Fälle. Die Anzahl der Unfälle, die in keiner der Listen vorhanden ist, kann aber nicht ermittelt werden. Somit kann auch die Dunkelziffer nicht berechnet werden.

Linkage: Verschiedene Datenquellen erfassen in der Regel ein und denselben Unfall. Um die Gesamtzahl der Unfälle (von einer bestimmten Art) erfassen zu können, ist es daher zwingend nötig, die Datensätze der infrage kommenden Datenquellen verknüpfen zu können. Dadurch können Fälle identifiziert werden, die in mehreren Datenquellen vorkommen. Somit werden Mehrfachzählungen von einem Unfall vermieden und Datensätze können komplettiert werden. Rein technisch gesehen kann eine Verknüpfung von Unfällen mit einer eindeutigen Identifizierungsvariable (ID) gemacht werden. Eine solche Variable ist allerdings häufig nicht verfügbar (z. B. aus technischen oder datenschutzrechtlichen Gründen), und daher muss eine Verknüpfung anhand der verfügbaren Variablen (Datum des Unfalls, Geburtsdatum, Wohnort etc.) durchgeführt werden. In Karimi et al. (2024) ist ein Überblick über solche Methoden zu finden, insbesondere für die Linkage von Polizei- und Krankenhausdaten [54]. Je nach Datenlage kann diese Verknüpfung sehr gut oder auch schlecht funktionieren. Damit kann die Anzahl der Unfälle in den vorhandenen Datenquellen ermittelt werden, jedoch nicht die Dunkelziffer.

Beispiel: Mit den Filtern «Schwerverletzt», «Wohnsitz in der Schweiz» und «Unfallland Schweiz» wurden 1200 Fälle in P und 1400 Fälle in KH gefunden. Mit den Variablen «Tag

des Unfalls», «Geburtstag», «Postleitzahl des Wohnorts» wurde eine probabilistische Verknüpfung durchgeführt. Dabei wurden 850 Fälle gefunden, die sowohl in P als auch KH vorkommen. Insgesamt kommen also 850 Fälle sowohl in P und KH vor, 350 Fälle nur in P und 550 Fälle nur in KH vor. Total umfassen die beiden Listen 1750 verschiedene Fälle. Fälle ausserhalb der beiden Listen können nicht ermittelt werden.

Erweiterung (Augmentation; Capture-Recapture): Keine Datenquelle wird die Unfälle mit einer Zuverlässigkeit von 100 % erfassen. Die Fälle in unseren Datenquellen sind daher nur ein Teil der realen Anzahl Fälle. Diese Problemstellung ist weit verbreitet [45] und entsprechend wurde eine umfassende statistische Methode zur Behandlung des Problems entwickelt [46,55] (s. auch Exkurs unter Kapitel 5.3).

Wie bei jeder statistischen Methode werden hier Annahmen gemacht, die im Anwendungsfall plausibilisiert werden sollten. Es gibt zahlreiche Varianten dieser Methode, die sich in ihren Annahmen unterscheiden. Grundlegende Annahmen, die mit fortgeschrittenen Methoden teilweise fallengelassen werden können, sind [47]:

1. Die Grundgesamtheit hat sich während der Untersuchung nicht verändert (die Grundgesamtheit ist geschlossen; «closed»).
2. Strassenverkehrsunfälle können zwischen den verfügbaren Listen zugeordnet werden.
3. Die erfassten Strassenverkehrsunfälle sind unabhängig voneinander.
4. Bei jeder Stichprobe/Liste hat jeder Strassenverkehrsunfall die gleiche Wahrscheinlichkeit, in die Stichprobe/Liste aufgenommen zu werden (falls nicht gegeben: Problem der Heterogenität («heterogeneity») oder Inhomogenität; diese Annahme kann prinzipiell abgeschwächt werden, s. u.)
5. Die Stichproben/Listen sind unabhängig (falls nicht gegeben: Problem der Abhängigkeit («dependence»)).

Die Annahmen (4) und (5) sind vor allem bei der Schätzung von menschlichen Populationen oft nicht realistisch. Daher wurden erweiterte Methoden entwickelt, wie beispielsweise der Einsatz von log-linearen Modellen M_{*} , die diese Annahmen nicht mehr zwingend nötig machen) [47,48,56,57]. Zudem gibt es Methoden, die Heterogenität durch erklärende Variablen (wie z. B. Alter, Geschlecht, Versicherungsstatus, Verletzungsschwere etc.) modellieren können (siehe z. B. [58,59]). Bei zwei Listen reicht es, wenn Annahme (4) bei nur einer der beiden Listen gewährleistet ist (s. Kapitel 3 in [56]).

Um die Annahme (5) fallen zu lassen (Abhängigkeit), braucht es in der Regel mindestens drei Listen. Falls diese Annahme fälschlicherweise gemacht wird, kann es zu einer Über- oder Unterschätzung des wahren Totals kommen [52,60].

Das grundlegende Vorgehen ist wie folgt: Ausgehend von mehreren Listen (z. B. P, KH, Versicherungen) wird unter unterschiedlichen Annahmen abgeschätzt, wie gross die zugrunde liegende Grundgesamtheit ist. Dabei ist auch eine Quantifizierung des Schätzfehlers möglich. Um einfache Varianten dieser Methode anwenden zu können, reicht es in der Regel, wenn die Mengengrössen für jede mögliche «capture-history» bekannt sind (z. B., wie viele Unfälle wurden sowohl in P als auch KH erfasst, wie viele nur in P, aber nicht in KH und wie viele wurden in KH erfasst, aber nicht in P). Bei Bedenken hinsichtlich des Datenschutzes wäre es denkbar, dass nur eine vertrauenswürdige Institution (z. B. das BFS) vollen Einblick in die Datensätze erhält und anschliessend lediglich die Anzahlen der capture-histories weiterleitet.

Beispiel: Es wird folgende Annahme getroffen: Die Unfälle, die dem Filter «Stationäre Behandlung war nötig», «Wohnsitz in der Schweiz» und «Unfallland Schweiz» entsprechen, haben eine gewisse (unbekannte) Erfassungswahrscheinlichkeit bei P und eine möglicherweise andere (ebenfalls unbekannte) Erfassungswahrscheinlichkeit bei KH. Dabei wird insbesondere angenommen, dass eine Erfassung bei P unabhängig von einer Erfassung bei KH ist. Unter diesen Annahmen lässt sich mit der Capture-Recapture-Methode abschätzen, dass gut 200 Unfälle in keiner der Datenquellen erfasst wurden (Modell M_{*t} , siehe [55]). Konkret wird das Total der Unfälle auf 1975 geschätzt (95%-Vertrauensintervall: [1933; 2023]). Diesem Beispiel liegen simulierte Daten zugrunde, sodass in diesem fiktiven Fall bekannt ist, dass das wahre Total der Unfälle 2000 ist. Diese Zahl wird in der Tat mit obigem Vertrauensintervall erfasst. Falls die Annahmen erfüllt sind, kann die totale Anzahl der Unfälle (und somit auch die Dunkelziffer) ermittelt werden, einschliesslich einer Angabe

zur Genauigkeit der Schätzung. Die Details zur Berechnung können dem Anhang III entnommen werden.

5.3.2 Realistischeres Beispiel

Zum besseren Verständnis werden im Folgenden die drei Schritte (Filter, Linkage, Erweiterungen) an einem realistischeren und komplexeren Beispiel erklärt. Dazu werden die beiden Datensätze Polizeiberichte (P) und Spezialstatistik der Sammelstelle für die Statistik der Unfallversicherung UVG (Kürzel SU) verknüpft. Folgende Schwierigkeiten treten hierbei auf: Die Polizei erfasst grundsätzlich alle Unfälle auf Schweizer Boden, insbesondere werden auch Personen mit Wohnsitz im Ausland oder Personen erfasst, die gesetzlich über die Krankenversicherung (Bundesgesetz über die Krankenversicherung, KVG) unfallversichert sind. Die SU erfasst hingegen nur eine Zufallsauswahl von 5 % aller UVG-versicherten Personen. Zudem werden Rentenfälle voll erhoben, was wir aber zur Vereinfachung in diesem Beispiel ignorieren.

Das folgende Beispiel basiert auf einem simulierten, also fiktiven Datensatz mit einem Total von 50 000 Unfällen auf Schweizer Boden. Weil die Daten simuliert wurden, ist die wahre Anzahl Unfälle bekannt, was eine Überprüfung der vorgeschlagenen Schätzmethoden ermöglicht. Für die Simulation wird angenommen, dass 20 % dieser Unfälle von der Polizei erfasst werden. Weiter sind 80 % der Verunfallten UVG-versichert, 10 % KVG-versichert und 10 % nicht in der Schweiz versichert (z. B. Touristinnen oder Fernfahrer). Für die Anwendung in der Praxis müssen diese Zahlen hingegen *nicht* bekannt sein. Zusätzlich wird angenommen, dass die polizeiliche Erfassung unabhängig von Versicherungsstatus erfolgt.

Filter: Die Daten werden hinsichtlich Attributen, die in beiden Datenquellen verfügbar sind, gefiltert, z. B. «Unfälle auf Schweizer Boden», «Unfalljahr 20xx», «Altersklasse: alle», «Fahrzeugart: alle», «Wohnland: alle», «Unfallland: CH», «Verletzungsschwere: alle». Falls hier Fehler passieren, wirkt sich das negativ auf alle folgenden Auswertungsschritte aus. Eine konkrete Schwierigkeit ist z. B. die Verletzungsschwere, die in P und SU nicht einheitlich deklariert wird.

In diesem simulierten Beispiel gibt es 10 000 polizeilich registrierte Unfälle. Zudem sind 2000 Fälle in der UVG-Statistik erfasst. Da die SU eine 5%-Stichprobe darstellt, lässt sich die Gesamtzahl der Fälle in der SU durch Hochrechnung auf 40 000 schätzen (20×2000).

Ausgehend von diesen Zahlen können wir für das Verknüpfen der beiden Datensätze folgende Schlüsse ziehen: Es gab mindestens 40 000 Fälle (bei vollständiger Überlappung der beiden Datensätze) und höchstens 50 000 Fälle (bei keiner Überlappung der Fälle). Der wahre Wert von 50 000 liegt an der oberen Grenze der Abschätzung, das bedeutet, das Total wird tendenziell unterschätzt. Das liegt daran, dass zwar die UVG-Fälle erfasst und hochgerechnet werden können, nicht aber die KVG-versicherten Personen oder die Personen mit Wohnsitz im Ausland.

Linkage: In diesem Schritt werden die beiden Datenquellen verknüpft. Idealerweise geschieht dies mit einer zuverlässigen, eindeutigen ID (z. B. AHV-Nummer etc.). Hierbei gibt es zwei praktische Schwierigkeiten: Eine solche ID ist manchmal nicht oder nur in einer Datenquelle verfügbar, aber selbst wenn eine solche ID verfügbar wäre, darf sie aus datenschutzrechtlichen Gründen nicht immer verwendet werden. In diesen Fällen gibt es verschiedene Methoden, wie die Linkage anhand von verfügbaren Daten (z. B. Geburtsdatum, Unfalldatum, Wohnort etc.) gemacht werden kann [54]. Konkret könnten in der SU alle Fälle zählen, die auch von der Polizei erfasst wurden. Dazu könnte man bei der SU alle Fälle mit den Attributen «Polizeirapport = wahrscheinlich ja» und «Polizeirapport = eindeutig ja» als polizeilich erfasste Fälle gewertet werden. Fälle mit dem Attribut «Polizeirapport = wahrscheinlich nein» würden hingegen als nicht polizeilich erfasst gelten. Eine Ungenauigkeit in diesem Schritt führt unweigerlich zu Ungenauigkeiten in allen folgenden Auswertungen.

In unserem Datenbeispiel wird davon ausgegangen, dass Linkage fehlerfrei möglich ist. Damit ergeben sich folgende Informationen (dies sind die Anzahlen der «capture-histories»): 9597 Fälle erscheinen in P, aber nicht in SU, 1597 Fälle erscheinen in SU, aber nicht in P und 403 Fälle erscheinen sowohl in P als auch in SU. Für die weitere Auswertung

würde es für einfache Varianten der Methode reichen, wenn diese Zahlen von einer vertrauenswürdigen Institution ermittelt und weitergeleitet werden. Weitere Informationen auf Ebene der Individuen sind für die aktuelle Auswertung nicht nötig, werden allerdings für fortgeschrittene Varianten der Methode benötigt.

Ausgehend von diesen Zahlen können folgende Schlüsse gezogen werden: Das Total beträgt $9597 + 20 \times 1597 = 41\,537$. Dabei ist die Hochrechnung mit dem Faktor 20 für die SU-Fälle eine sehr grobe Abschätzung.

Erweiterung: Falls Annahmen über die Erfassungswahrscheinlichkeit bei P und SU getroffen werden, können wir die obigen Informationen mit der Capture-Recapture-Methode erweitern. Konkret wird in diesem Beispiel Folgendes angenommen: Die Erfassungen bei P und SU sind unabhängig voneinander (Annahme 5); P und SU haben je eine (unbekannte) Erfassungswahrscheinlichkeit pro Unfall (Annahme 4). Konkret entsprechen diese Annahmen dem sogenannten M_{t} -Modell [46,54]. In unserem Datenbeispiel ergibt sich mit dem M_{t} -Modell ein Total von 49 613 Fällen (95%-Vertrauensintervall: [45 637; 54 147]). Der wahre Wert von 50 000 wird somit erfolgreich erfasst und die Schätzungenauigkeit kann quantifiziert werden. Dank der Capture-Recapture-Methode konnten obige Schwierigkeiten (Unterschätzung) bezüglich der Erfassung von KVG-Versicherten und Personen mit Wohnsitz im Ausland gelöst werden. Intuitiv lässt sich die deutliche Verbesserung so verstehen: Aus der SU lernen wir, wie hoch der Anteil polizeilich nicht erfasster Fälle ist. Dieses Wissen können wir als Hochrechnungsfaktor auf die polizeilichen Daten verwenden, die auch Unfälle von KVG-Versicherten und übrige Unfälle umfasst. Hierbei ist entscheidend, dass die polizeiliche Erfassung unabhängig von der Versicherungsart ist. Eine detaillierte Herleitung, die zeigt, dass das M_{t} -Modell in diesem Beispiel zutreffend ist, findet man in Anhang III. Im gleichen Anhang sind auch die Details für die Simulation zu finden.

Ein weiteres konkretes Beispiel, bei dem die Capture-Recapture-Methode zur Schätzung der Sterblichkeitsrate bei Verkehrsunfällen verwendet wird, ist in Entezami et al. (2018) [61] zu finden.

Obige Beispiele zeigen, dass mit einer Methodik, bestehend aus drei Schritten (Filter, Linkage, Erweiterung), aus bestehenden Datenquellen auf die Grundgesamtheit und somit in Relation zum Hellfeld auch auf die Dunkelziffer geschlossen werden kann. Die Methodik kann problemlos auf eine Verwendung von mehr als zwei Datenquellen erweitert werden (z. B. P, SU und KH), solange die Schritte Filter und Linkage erfolgreich durchgeführt werden können. Das Verwenden von mehr als zwei Datenquellen hat folgende Vorteile: Zum einen wird die Schätzgenauigkeit erhöht. Zum anderen ist der Einsatz von Modellen möglich, die weniger Annahmen machen.

5.3.3 Mögliche Probleme

Im Folgenden werden mögliche Probleme bei den einzelnen Schritten diskutiert.

Filter: Falls man eine Gesamtzahl für eine Untergruppe eines Attributs (z. B. «E-Bike» oder «schwerverletzt») berechnen will, so muss es in jeder Datenquelle möglich sein, nach diesen Attributen zu filtern. Während das bei gewissen Variablen problemlos möglich sein sollte (z. B. «Unfalljahr», «Unfallland»), gibt es bei manchen Variablen Probleme mit der Vergleichbarkeit (beispielsweise «Verletzungsschwere»). Wenn eine Vergleichbarkeit bezüglich eines Attributs nicht erzielt werden kann, so kann dieses Attribut und allfällige Untergruppen nicht untersucht werden. Wird es dennoch verwendet, so wird die Schätzung der Gesamtzahl ungenauer.

Linkage: Für eine Verknüpfung von zwei oder mehr Datenquellen wäre eine eindeutige ID zu bevorzugen. Diese kann künstlich erzeugt werden oder schon vorhanden sein (z. B. AHV-Nummer und Datum des Unfalls). Es muss abgeklärt werden, ob eindeutige IDs aus rechtlicher Sicht verwendet werden dürfen oder kryptographische Verfahren verwendet werden können (z. B. Private set intersection, siehe Morales et al. (2023) [62]). Falls keine ID verfügbar ist, kann Linkage anhand von verfügbaren Informationen in den Datensätzen versucht werden. Je nach Anzahl und Art der beteiligten Variablen kann so unter Umständen eine sehr gute Linkage erzielt werden. Bei fehlerbehafteter Linkage wird die Schätzung

der Gesamtzahl ungenauer. Verschiedene Methoden stehen hier zur Verfügung und müssten im Einzelfall anhand der Datensätze ausgewählt werden [54]. Eine konkrete Methode, die für grosse Datensätze geeignet ist, ist im R-Paket «fastLink» implementiert [63].

Erweiterung: Wie bei jeder statistischen Methode werden auch bei der Capture-Recapture-Methode Annahmen gemacht, die im Anwendungsfall kaum abschliessend geprüft werden können, aber zumindest plausibel erscheinen sollten. Capture-Recapture-Methoden sind ein aktives Forschungsgebiet, sodass mehrere Schätzmethoden mit teils leicht unterschiedlichen Annahmen zur Verfügung stehen (siehe z. B. Das et al, 2023 [59]. Auch die Korrektur von fehlerhafter Linkage wird dabei untersucht [64]. Falls entsprechende Annahmen gemacht werden und die Schritte «Filter» und «Linkage» hinreichend gut gelöst werden können, so kann es von entscheidendem Vorteil sein, mit der Capture-Recapture-Methode die Grundgesamtheit zu schätzen und die Schätzungenauigkeit zu quantifizieren und eine systematisch fehlerbehaftete Schätzung zu korrigieren. In der Literatur gibt es mehrere Standard-Modelle, die die gemachten Annahmen formalisieren. Im Kontext der vorliegenden Problemstellung scheinen die M_t - und M_{th} -Modelle ein vernünftiger Ausgangspunkt zu sein [55]. Diese Modelle sind im R-Paket «Rcapture» gut unterstützt.

Eine Übersicht über verfügbare Softwarepakete für das Capture-Recapture Problem ist – teilweise mit Vergleich der Leistungsfähigkeit – in mehreren Veröffentlichungen zu finden, z. B. Bunge (2013) [65]. Für kompliziertere Modelle könnten neben dem R-Paket Rcapture folgende Implementationen herangezogen werden: Software MARK ([66]; es gibt eine Verknüpfung zum R-Paket RMark); R-Paket mra (insbesondere F.huggins.estim), R-Paket drpop [59], R-Paket VGAM [58], R-Paket dga (mit Bayesian Model Averaging), R-Paket LCMCR (Bayesianischer Ansatz). Für das Modellieren von Auswahlwahrscheinlichkeiten, die von gemessenen Kovariablen abhängen, scheint das R-Paket VGAM [58] gut geeignet zu sein.

5.4 Maschinelles Lernen

Beim maschinellen Lernen («Machine Learning») wird üblicherweise eine Zielgrösse von Interesse, im vorliegenden Fall beispielsweise die Anzahl der Unfälle einer gewissen Art in einem räumlichen und zeitlichen Fenster, durch mehrere, leicht verfügbare und erklärende Variablen modelliert («supervised learning»). Dabei kann eine Vielzahl von Methoden zum Einsatz kommen (z. B. lineare Regression, Random Forest, neuronale Netzwerke etc., s. Hastie et al (2009) [40]).

Unabhängig von der Methode muss ein Trainingsdatensatz verfügbar sein, mit dem das Modell trainiert werden kann. Im vorliegenden Kontext bedeutet das, dass in möglichst vielen repräsentativen Raum-Zeit-Segmenten (z. B. einer zufälligen Auswahl) sowohl die wahre Gesamtzahl (und somit auch die Dunkelziffer) der Unfälle als auch möglichst viele erklärende Variablen, die das entsprechende Raum-Zeit-Segment beschreiben, bekannt sein müssten.

Für die Intuition hilft der Vergleich mit einer Tabelle: Eine Spalte dieser Tabelle ist die Zielgrösse, die uns interessiert (z. B. die Anzahl der Kollisionen von PW auf einem Autobahnabschnitt in der nächsten Stunde oder Anzahl der E-Bike-Unfälle mit leichter Verletzung auf einem Streckenkilometer des Schweizer Strassennetzes). Weitere Spalten der Tabelle sind erklärende Variablen zum gleichen Raum-Zeit-Segment (z. B. Tageszeit, Temperatur, Verkehrsfluss, Information zu Ferien etc.). Falls eine solche Tabelle mit hinreichend vielen (typischerweise hunderten oder mehr) Zeilen erstellt werden kann, so kann mit maschinellem Lernen versucht werden, einen Zusammenhang zwischen Zielgrösse und erklärenden Variablen herzustellen. Ob dies gelingt, hängt von der Qualität und Beschaffenheit der erfassten Daten ab. Die Güte des geschätzten Modells kann mit statistischen Methoden (z. B. Kreuzvalidierung) beurteilt werden.

Falls ein gut passendes Modell gelernt werden kann, ist es möglich, Vorhersagen für die Zielgrösse zu machen, wenn Informationen zu erklärenden Variablen verfügbar sind. So kann beispielsweise abgeschätzt werden, wie viele Unfälle es auf einem bestimmten Autobahnsegment in der nächsten Stunde geben wird, wenn die aktuellen Daten zu Wetter, Verkehrsfluss etc. bekannt sind. Es muss betont werden, dass solche Vorhersagen in der Praxis immer fehlerbehaftet sind. Allerdings lässt sich die Genauigkeit häufig abschätzen.

Zudem gilt das Modell nur, solange die Grundsituation nicht substantiell verändert wird, z. B. durch eine Baustelle oder eine andere Geschwindigkeitsbegrenzung.

Zum besseren Verständnis wird ein konkretes Vorhersagemodell für Unfälle skizziert, das auf maschinellem Lernen beruht und die Schwierigkeiten bei einer möglichen Anwendung auf das vorliegende Problem der Bestimmung der Dunkelziffer aufzeigt.

In Zhang et al. (2024) [67] trainierten die Autoren Machine-Learning-Methoden (Random Forest, XGBoost, SVM etc.), um Unfälle von Fahrzeugen auf Freeways des Bundesstaats Alabama (US) in einem Zeitfenster (ca. 20 Minuten) und einem räumlichen Segment («Traffic management channels») vorhersagen zu können. Dabei verknüpfen die Autoren Informationen aus drei verschiedenen Datenbanken: CARE («Critical Analysis Reporting Environment»), eine Datenbank mit Unfalldaten der Universität Alabama) enthält vollständige Informationen über solche Unfälle in der Vergangenheit; die Datenbank von HERE Technologies enthält umfassende Informationen, die den Verkehrsfluss beschreiben; HPMS («Highway Performance Monitoring System») enthält Informationen bzgl. der Infrastruktur.

Während dieser Ansatz sinnvoll für Massnahmen der Verkehrssteuerung sein könnte, ist nicht klar, wie er für die Schätzung der Dunkelziffer eingesetzt werden könnte. Dies aus folgenden Gründen:

- Es wird eine Datenquelle verwendet (CARE), die für Unfälle auf Freeways als sehr zuverlässig gilt. Die Situation ist eventuell vergleichbar mit der Situation auf Schweizer Autobahnen. Für Unfälle in der Schweiz auf anderen Strassentypen oder im Langsamverkehr gibt es keine vergleichbare Datenquelle, die zum Trainieren des Modells herangezogen werden könnte.
- Die Vorhersagen sind auf kleine zeitliche und räumliche Segmente beschränkt. Weil es vermutlich Abhängigkeiten zwischen den Ereignissen in diesen Segmenten gibt, ist nicht klar, wie die einzelnen Ergebnisse über grosse Zeiträume und Flächen (z. B. das jährliche Total in der Schweiz) aggregiert werden können.
- Die Vorhersagen sind fehlerbehaftet. Das beste Modell hatte eine Genauigkeit («accuracy») von 58 %. Es ist nicht klar, wie sich diese Fehler auf aggregierte Werte über grosse Zeiträume und Flächen übertragen.

In einer systematischen Übersichtsarbeit untersuchen die Autoren Vorhersagemodelle für Unfälle, die auf maschinellem Lernen beruhen [68]. Es ist bemerkenswert, dass die Autoren in diesem Text das Problem der Dunkelziffer («underreporting») ansprechen und Linkage («data fusion») als mögliche Lösung erwähnen, also genau die Methode, die in Kapitel 5.3 vorgestellt wurde (s. Kapitel 5.1.2 in Ali et al (2024) [68]).

Entscheidend bleibt bei allen angedachten Herangehensweisen die zuverlässige Verfügbarkeit von Trainingsdaten: Wenn die wahre Anzahl der Unfälle nicht erfasst wurde, scheint es nicht möglich, ein brauchbares Modell zu schätzen. Ob die steigende Anzahl von Sensordaten die Situation entscheidend verbessern wird, bleibt abzuwarten. Die Anzahl der erklärenden Variablen mag dadurch ansteigen, aber entscheidend ist die zuverlässige, repräsentative und häufige Beobachtung der Zielgrösse (Unfall). Diese scheint nicht immer das direkte Ziel der Sensordaten zu sein, bzw. könnte aus rechtlichen Gründen nicht verfügbar sein (z. B. e-Call).

In Tab. 4 sind die Vor- und Nachteile der drei Ansätze «Stichprobe bzw. Befragungen», «Datenspuren» und «Maschinelles Lernen» aufgeführt.

Tab. 4: Vor- und Nachteile der drei vorgeschlagenen Ansätze		
Methode	Vorteil	Nachteile
Stichprobe bzw. Befragungen	• Bewährt, verbreitet • Einfache Auswertung • Kann gewünschte Details erfragen • Erlaubt Aussage bzgl. Schätzgenauigkeit	• «Einmalige» Momentaufnahme • Methodische Annahmen (kein Recall Bias u. Ä.) • Aufwendig • Stichprobenrahmen muss vorhanden sein
Datenspuren	• Nutzt bestehende Datenquellen • Wenig Aufwand und regelmässig machbar, falls gut aufgesetzt • Branchenüblich (s. IRTAD, WHO) • Erlaubt Aussage bzgl. Schätzgenauigkeit	• Evtl. kein Zugang zu relevanten Datenquellen in gewünschter Granularität (z. B. KVG) • Filter und Linkage muss zuverlässig in allen Datenquellen machbar sein • Methodische Annahmen (z. B. Unabhängigkeit der Listen u. Ä.)
Maschinelles Lernen	• Falls machbar, gute Unfallschätzungen in «Echtzeit» • Wenig Aufwand und regelmässig machbar, falls gut aufgesetzt	• Trainingsdaten müssen verfügbar sein • Unklar, wie man mit zeitlicher und räumlicher Abhängigkeit von punktuellen Vorhersagen umgeht (z. B. beim Aggregieren) • Unklar, ob Vorhersagegüte ausreichend ist

6 Potenziell nutzbare Datengrundlagen

Um die in Kapitel 5 dargestellten methodischen Ansätze zur Schätzung von differenzierten Dunkelziffern umsetzen zu können, braucht es entsprechend geeignete Datengrundlagen. Nachfolgend wird pro Ansatz eine Übersicht über vorhandene bzw. potenziell künftig vorliegende Datengrundlagen gegeben.

Der Fokus liegt dabei auf bestehenden und gut erschlossenen Datensätzen im Hinblick auf die kurzfristige Verwendung für die Umsetzung der drei methodischen Ansätze (Kapitel 5). Zudem werden Hinweise gegeben, welche derzeit nicht erschlossenen Datenquellen künftig weitere Beiträge leisten könnten und deshalb (je nach gewählter Methode) erschlossen werden sollten.

6.1 Strassenverkehrsunfallstatistik des ASTRA

Als Voraussetzung zur Herleitung der Dunkelziffer im Strassenverkehrsunfallgeschehen dienen die Daten der polizeilich registrierten Strassenverkehrsunfälle des ASTRA. Sie zeigen im Detail die in der Schweiz erfassten Unfälle nach dem Territorialitätsprinzip (Unfälle in der Schweiz).

Die Unfälle werden in einer relationalen Datenbank mit drei Tabellen erfasst: Unfallereignisse, beteiligte Objekte sowie Mitfahrende. Diese Tabellen sind über Fremdschlüssel verknüpft, wobei sich Mitfahrende (Personen) zu Objekten (Fahrzeugen) und beides zu den Unfallereignissen zuordnen lassen. Die Analyse- und Differenzierungsmöglichkeiten sind dabei sehr gross und breit.

Dennoch besteht Einigkeit darüber, dass es Strassenverkehrsunfälle gemäss Definition gibt, die nicht von der Polizei erfasst wurden und daher nicht in den Daten erscheinen. Insbesondere folgende Unfälle bzw. Unfallfolgen sind nicht erfasst bzw. unterrepräsentiert:

- Meldepflichtige, aber nicht gemeldete Unfälle
- Nicht meldepflichtige Unfälle mit «geringem» Sach- oder Personenschaden
- Unvollständige Erfassung: Fahrerflucht, Gefährdungen
- Gesundheitliche Spätfolgen durch Strassenverkehrsunfälle

Dabei ist auch denkbar, dass in der Praxis je nach Region bzw. erfassender Person die definierten Erfassungsschwellen bezüglich «geringem» Sach- und/oder Personenschäden unterschiedlich interpretiert oder vor Ort noch nicht adäquat eingeschätzt werden können. Tab. 5 fasst in aller Kürze die wichtigsten Merkmale des Datensatzes zusammen. Im Detail können die verfügbaren Merkmale pro registriertem Unfallereignis, beteiligten Objekt und beteiligter Person kann im aktuell gültigen Unfallerfassungsprotokoll 2018 nachvollzogen werden [1].

Tab. 5: Daten im Hellfeld: Strassenverkehrsunfallstatistik des ASTRA

Datenquelle	Datenqualität	Differenzierungsmöglichkeiten	Grundgesamtheit	Repräsentativität	Periodizität
ASTRA Strassenverkehrsunfälle (Datenbank)	Grundsätzlich hohe Qualität. Professionelle Einschätzung durch erfassende Behörde (Polizei) vor Ort und im Nachgang Teilweise abgestimmt mit Informationen der Rettungsdienste	<u>Gemäss UAP 2018:</u> • Verletzungsschwere (angelehnt an NACA-Kodierung) • Alter • Geschlecht • Fahrzeug (Objekt) • Hauptverursacher • Unfallhergang (Alleinunfall vs. Kollision) • Ort (Koordinate, Gemeindenummer) • Wetter / Fahrbahnzustand • Ursachen • Zustand Person / Fahrzeug (z. B. Alkoholeinfluss) • Infrastrukturtyp • usw.	Alle Strassenverkehrsunfälle in der Schweiz (Territorialitätsprinzip) gemäss Definition des ASTRA	Vollerhebung angestrebt gemäss Definition der Strassenverkehrsunfälle (siehe hierzu die weiter oben erwähnten Erläuterung zu nicht erfassten bzw. unterrepräsentierten Ereignissen)	Laufende Erfassung, jährliche Publikation

Die Unfälle werden über das Unfallerfassungsprotokoll 2018 erfasst, für das zusätzliche Instruktionen zum Ausfüllen vorliegen. Die Erfassung erfolgt in separaten Teilmasken für das Unfallereignis («Unfall»), die beteiligten Objekte («Objekt») sowie die «Mitfahrenden».

Beim Unfallteil werden Datum, Uhrzeit, Unfallort, Unfallhergang, Hauptursache und die Anzahl der beteiligten Objekte und Personen erfasst. Zudem werden Details zum Umfeld (Strassenart, Signalisation, Unfallstelle, Vortrittsregelung, Verkehrsaufkommen, Witterung, Strassenzustand etc.) erfasst. Der Ablauf des Unfalls wird in Prosa beschrieben und kann mit einer Unfallskizze ergänzt werden. Beim Unfallblatt wird zusätzlich ein grob bezifferter Sachschaden erfasst. Diese Variable wird jedoch vom ASTRA aufgrund von Bedenken zur Qualität nicht publiziert bzw. nicht als Teil des Datensatzes abgegeben.

Pro am Unfall beteiligtem Objekt wird ein Objektblatt erfasst. Die Personalien und Verletzungsschweren von Fahrzeuglenkern/-innen oder Fussgängern/-innen werden direkt im jeweiligen Objektblatt erfasst. Die Verletzungsschwere von anderen Beteiligten wird im Blatt «Mitfahrende» (vorne, hinten, unbekannt) unter Unfallfolgen erfasst. Dabei wird bei Verletzten zwischen «leicht verletzt», «erheblich verletzt» und «lebensbedrohlich verletzt» differenziert. Beim öffentlichen Ausweis der Zahlen werden «erheblich» und «lebensbedrohlich» Verletzte zu «Schwerverletzte» zusammengefasst. Einige Beispiele für die Zuordnung von Verletzungsmustern zu den Verletzungsschweren sind in den Instruktionen zum UAP enthalten und sollen bei der Bestimmung helfen. Eine weitere Möglichkeit besteht in der Abstimmung der Verletzungsschwere zwischen Polizei und dem Rettungsdienst vor Ort. Die Rettungsdienste der Schweiz nutzen den NACA-Score des «National Advisory Committee for Aeronautics», der Vorläuferorganisation der NASA, zur präklinischen Triage der Verletzten. Dieser Score reicht von unverletzt (0) bis tödlich verletzt (7). Die Erfassung des NACA Score im Unfallprotokoll ist möglich, jedoch nicht obligatorisch. Bei Todesfällen wird zwischen «Auf Platz gestorben» und «Innert 30 Tagen gestorben» unterschieden. Auch nicht verletzte beteiligte Personen werden erfasst. Sowohl für Lenkende als auch für Mitfahrende werden das Geburtsdatum und das Geschlecht sowie allfällige verwendete Schutzsysteme erfasst.

Bei Fahrzeugen (Objekten) kann aus über 50 Fahrzeugarten ausgewählt werden. Daneben stehen die separaten Objekt-Kategorien «Fussgänger/-in» und «fäG» (Fahrzeugähnliche Geräte) zur Verfügung. Bei Fahrzeugen ist die Erfassung der Kontrollschildnummer sowie die Stammnummer (gem. Fahrzeugausweis) für Schweizer Fahrzeuge vorgeschrieben, bei ausländischen Kontrollschildern nur die Erfassung des Landes. Über die Stammnummer können weitere Angaben über das Fahrzeug direkt aus der Datenbank IVZ-Fahrzeuge abgefragt und übernommen werden. Für jeden Unfall muss ein beteiligtes Objekt als mutmasslicher «Hauptverursacher» identifiziert werden. Standardmässig beim Objekt erfasst werden Fahrzweck, Angaben zum Führerausweis sowie zum Zustand der lenkenden Person (Alkohol/Arzneimittel/Betäubungsmittelinfluss).

6.2 Stichproben- bzw. Befragungsdaten

In diesem Kapitel werden Befragungen in der Schweiz behandelt, bei denen (auch) Unfälle Thema sind. Gleichzeitig wird analysiert, ob bestehende (wiederkehrende) Befragungen existieren, die aktuell diese Thematik noch nicht miteinschliessen, sich aber potenziell dafür eignen würden.

6.2.1 Übersicht erschlossene Daten

Bisherige Befragungen in der Schweiz, welche sich zumindest teilweise mit Strassenverkehrsunfällen beschäftigen, sind die Schweizerische Gesundheitsbefragung (SGB) des Bundesamts für Statistik BFS sowie die Haushaltsbefragung der BFU. Diese beiden Quellen sind in Tab. 6 charakterisiert.

Schweizerische Gesundheitsbefragung (SGB)

In der Schweizerischen Gesundheitsbefragung (SGB) des BFS werden alle 5 Jahre – letztmals 2022 – rund 13 000 bis 20 000 Personen mit einem schriftlichen Fragebogen zu ihrem gesundheitlichen Befinden befragt. Der Stichprobenrahmen umfasst alle in der Schweiz wohnhaften Personen ab 15 Jahren. Personen mit Wohnsitz im Ausland sowie Personen unter 15 Jahren sind damit nicht abgedeckt [69]. Die Befragung enthält den Themenblock «Unfälle» mit der Frage «Hatten sie in den letzten 12 Monaten eine Unfallverletzung?», wobei als Antwortmöglichkeit auch «Verkehrsunfall» angegeben werden kann. Der Begriff «Verkehrsunfall» wird allerdings nicht genauer definiert. Zusätzlich wird die Behandlungsart (keine, selbst, ambulant, Spital) sowie die darauffolgende Arbeitsunfähigkeit (Dauer in Tagen) abgefragt. Weitere Differenzierungen zum Verkehrsunfall (z. B. Verkehrsmittel, Ort, Anzahl Beteiligte, Ursache) werden nicht abgefragt. Dadurch ist eine genaue Identifizierung von Strassenverkehrsunfällen nach ASTRA-Definition in diesen Daten nicht möglich.

Haushaltsbefragung der BFU

Für die Haushaltsbefragung 2024/25 der BFU wurde eine Stichprobe von 60 000 Haushalten aus dem Stichprobenrahmen des BFS bereitgestellt. Die Stichprobe wurde als eine nach Haushaltszusammensetzung (Haushalte mit Kindern bis 14 Jahre, Haushalte mit Senioren über 64 Jahre und andere Haushalte) und Sprachregionen (Deutschschweiz, Romandie und Tessin) disproportional geschichtete Zufallsauswahl realisiert. Ein standardisierter Fragebogen wurde per Webinterview (CAWI) durch eine Person im Haushalt stellvertretend für alle Haushaltsmitglieder beantwortet. Der Fragebogen umfasst Fragen zu allen erlittenen Unfällen mit Verletzungsfolge in den 12 Monaten vor der Beantwortung der Befragung. Die Befragung wurde während 12 Monaten in sechs Wellen durchgeführt. Erhoben wurden der Unfallmechanismus, der Unfallort, die Tätigkeit zum Zeitpunkt des Unfalls und die am Unfall beteiligten Gegenstände, Fahrzeuge etc. Zusätzlich wurden Verletzungsfolgen erhoben. Der Fragenkatalog ermöglicht die Identifikation von Strassenverkehrsunfällen nach ASTRA-Definition. Ob und wie die Haushaltsbefragung in Zukunft wiederholt wird, ist aktuell noch offen.

Tab. 6: Befragungen mit Bezug zu Strassenverkehrsunfällen in der Schweiz

Datenquelle	Datenqualität	Differenzierungsmöglichkeiten	Grundgesamtheit	Repräsentativität	Periodizität
Schweizerische Gesundheitsbefragung (BFS)	Selbstauskünfte von Befragten Fragebogen 2022: https://dam-api.bfs.admin.ch/hub/api/dam/assets/28965074/master Bekannte Probleme von Befragungen (z. B. Non-Response, Recall Bias)	Verkehrsunfall ja/nein (letzte 12 Monate): • Unfallart • Behandlungsart • Arbeitsunfähigkeit (Dauer) Sozioökonomische Angaben: • Alter • Geschlecht • Wohnort • Anzahl Kinder im Haushalt (nach Altersklasse)	Schweizer Wohnbevölkerung ab 15 Jahren	Stichprobenerhebung Rund 13 000 bis 20 000 Personen, je nach Erhebungsjahr	Alle 5 Jahre (letzte Erhebung 2022)
BFU-Haushaltsbefragung 2024/25	Selbstauskünfte von Befragten. Teilweise Auskünfte über Drittpersonen. Bekannte Probleme von Befragungen (z. B. Non-Response, Recall Bias)	Bereich: Strassenverkehr, Sport, Haus- und Freizeit sowie Berufsunfälle Genutzte und beteiligte Fahrzeuge, Unfallart, Verletzungsart und -lokalisation, Spitalaufenthalt, Ausfalltage, Alter, Geschlecht	Schweizer Wohnbevölkerung, Unfälle im In- und Ausland	Stratifizierte Zufallsauswahl von Haushalten aus dem Stichprobenrahmen des BFS	Noch offen, maximal alle 5 Jahre

6.2.2 Weitere bisher nicht erschlossene Möglichkeiten / Varianten

Im Folgenden wird auf Erhebungen eingegangen, die sich potenziell eignen würden, Fragen zu Strassenverkehrsunfällen zu integrieren, aber bisher nicht für diesen Zweck eingesetzt wurden.

Mikrozensus Mobilität und Verkehr (MZMV)

Der Mikrozensus Mobilität und Verkehr (MZMV) des BFS und des ARE wird seit 1974 etwa alle fünf Jahre durchgeführt. Die letzte Erhebung war im Jahr 2021. Mittels Telefoninterviews (CATI) wurden rund 55 000 Personen unter anderem zu ihrem Mobilitätsverhalten am Tag vor der Befragung (Etappen, Distanzen, genutzte Verkehrsmittel) befragt. Grundsätzlich ist die Schweizer Wohnbevölkerung ab einem Alter von sechs Jahren. Die Zufallsstichprobe wird aus dem Stichprobenrahmen für Personen- und Haushaltserhebungen (SRPH) des BFS gezogen [70]. Fragen zum Strassenverkehrsunfallgeschehen könnten bei entsprechender Finanzierung als Zusatzmodul im MZMV integriert werden. Oft stehen neue Frageblöcke im MZMV aber in Konkurrenz zu anderen Themenblöcken. Es verbleiben zudem die generellen Herausforderungen der Verzerrung (z. B. durch «Nicht-Teilnahme») und des Problems, dass Unfälle durch Personen mit Wohnsitz im Ausland aufgrund des Stichprobenrahmens nicht abgedeckt werden können. Der Vorteil dieses Ansatzes wäre, dass nur die Grenzkosten des Fragemoduls berücksichtigt werden müssen.

Befragungen von Verletzten in Krankenhäusern/Arztpraxen; European Injury Database (EU-IDB)

Daten zu Strassenverkehrsunfällen, ihren Hergängen und ihren Verletzungsfolgen können auch durch gezielte Befragung von Verletzten in Krankenhäusern und/oder Arztpraxen gewonnen werden. Ein solcher Ansatz wird von einigen europäischen Ländern zur Speisung der «European Injury Database» (EU-IDB, [71]) eingesetzt. Dabei werden Daten zu Unfällen in allen Bereichen (nicht nur Strassenverkehrsunfälle) in Spitälern durch Interviews der Betroffenen selbst oder Angehörigen gesammelt. Die Erfassung der Unfälle richtet sich nach der ehemaligen WHO-Klassifikation «International Classification of External Causes of Injury (ICECI)» [72], einer umfangreichen Kodierliste, in der Transportmittelunfälle über ein eigenes Modul erhoben werden können. Zusammen mit der Information des Unfallorts (öffentliche Strassen) lassen sich Strassenverkehrsunfälle selektieren. In den meisten Ländern wird in einer Stichprobe von Spitälern erhoben und deren Ergebnisse auf das gesamte Land hochgerechnet. Erhoben wird oft durch eigene Teams, die in der Anwendung der Unfallklassifikation geschult sind. Für die Erhebung existieren zwei Fragebogenversionen: Eine Kurzfassung mit einer deutlich gekürzten Kodierliste («Minimum Data Set») und eine ungekürzte Version («Full Data Set»). Ziel der IDB ist das Unfall-Monitoring, um z. B. gefährliche Produkte zu identifizieren (nur über das «Full Data Set» möglich) und eine durch die standardisierte Erfassung internationale Vergleichbarkeit der Ergebnisse herzustellen. Für die Erhebung ist ein hoher Personaleinsatz mit entsprechend hohen Folgekosten notwendig, der über die Anzahl der Spitäler in der Stichprobe gesteuert werden kann. Durch die Erhebung in Spitälern werden vor allem mittelschwere und schwere Verletzungen erfasst. Problematisch erscheint die Validität der aus IDB-Daten erstellten Hochrechnungen: Nicht alle Spitäler eines Landes nehmen an der Datenerfassung teil. Auch innerhalb eines Spitals wird nicht zu jedem Zeitpunkt Befragungspersonal anwesend sein und es muss auch mit einer Auskunftsverweigerung seitens der Betroffenen gerechnet werden. Die Schweiz hat sich der IDB-Erhebung bisher nicht angeschlossen.

6.3 Daten für Datenspuren / Linkage

Beim Datenspuren / Linkage-Ansatz stehen bestehende Datenquellen im Fokus, die einen Bezug zu Strassenverkehrsunfällen im Helffeld aufweisen. Um die Dunkelziffer für ein bestimmtes Würfelement zu berechnen, müssen die verwendeten Datenquellen nach den entsprechenden Kriterien gefiltert werden können. Je feiner die Filterkriterien gesetzt werden können, desto differenzierter lassen sich Dunkelziffern schätzen. Ebenfalls muss die Anzahl der in beiden Quellen vorhandenen Beobachtungen bestimmt werden können (Linkage) (s. Kapitel 5.3).

Je besser die Filtersetzung in allen verwendeten Datenquellen funktioniert, je höher die Bestimmtheit der Zuordnung im Einzelfall (eindeutiger Ein- bzw. Ausschluss der Überschneidung) ist und je grösser das Würfelement, desto höher ist die Qualität des Ergebnisses.

6.3.1 Übersicht erschlossene Daten

Die im Folgenden aufgeführten Quellen enthalten Daten, die zur Erfüllung der in Kapitel 4 erwähnten Bedürfnisse eingesetzt werden könnten und bereits erschlossen sind:

- UVG-Statistik / Spezialstatistik, SSUV
- Medizinische Statistik der Krankenhäuser (MSK) des BFS
- Patientendaten Spital ambulant (PSA) des BFS
- Todesursachenstatistik des BFS

Tab. 7 zeigt für die verfügbaren und erschlossenen Daten einen Überblick über die Datenqualität, die Differenzierungsmöglichkeiten und einige weitere Angaben.

Tab. 7: Erschlossene Datenquellen, die Informationen zu Strassenverkehrsunfällen beinhalten und die für einen Linkage-Ansatz infrage kommen

Datenquelle	Datenqualität	Differenzierungsmöglichkeiten	Grundgesamtheit	Repräsentativität	Periodizität
Spezialstatistik der Sammelstelle für die Statistik der Unfallversicherung (UVG-Statistik)	Die UVG-Statistik besteht aus einer Stichprobenerhebung von 5% aller Unfälle sowie aus allen Rentenfällen, Berufskrankheiten und Todesfällen.	Strassenverkehrsunfälle Freizeit und Beruf: • Benütztes Transportmittel (ca. 6 Kategorien in guter Qualität) • Alter • Geschlecht • Vorhandensein Polizeirapport (nur bis 2022) • Gemeinde, PLZ Die Unfall- bzw. Verletzungsschwere wird als Masszahl abgeleitet aus der Zahl der entschädigten Tage infolge des Unfalls, die für jeden Schadenfall bekannt ist.	Unfälle von UVG-Versicherten	5%-Stichprobe	Jährlich
Medizinische Statistik der Krankenhäuser MSK (BFS) (Einzeldaten)	Gut; Einzeldaten zu Patienten	Verkehrsunfälle sind Teilmenge des ICD10-GM-Codes V99! «Transportmittelunfälle»	Patienten in Kranken- oder Geburtshaus bzw. Klinik mit (teil-)stationärem Aufenthalt	Vollerhebung	Jährlich
Patientendaten Spital ambulant (PSA), BFS (Einzeldaten)	Gut; Einzeldaten zu Patienten	• Pro Patient: • Ambulante Leistungen (ICPC) • Sozio-Demografie (Basis, Anonymisiert) • Wohnort • Leistungserbringer • Keine standardisierte Kodierung medizinischer Diagnosen oder von Unfallhergängen. Verkehrsunfälle nicht identifizierbar	Patienten mit fakturierten ambulanten Leistungen in Kranken- und Geburtshäusern in der Schweiz • Spitalaufenthalt unter 24 Stunden • Keine Bettenbelegung über Nacht • Kein Transfer in ein (anderes) Spital • Kein Todesfall	Vollerhebung	Jährlich
Todesursachenstatistik, BFS (Einzeldaten)	Gut; Einzeldaten zu Sterbefällen	Pro Todesfall: • Äussere Ursachen nach Kapitel 20, ICD-10-WHO • Todesursache in Haupt- und Nebendiagnosen (ICD-10 WHO) • Geschlecht • Geburtsdatum • Sterbedatum • Unfalldatum • Sterbeort • Sterbeland • Wohnort • Zivilstand • Beruf	Todesfälle der ständigen Schweizer Wohnbevölkerung im In- und Ausland	Vollerhebung	Jährlich

UVG-Statistik (SSUV)

Die Schweizerische Sammelstelle der Unfallversicherungen (SSUV) erstellt jährlich eine Spezialstatistik zu den gemeldeten Unfällen in der Schweiz. Dabei handelt es sich um eine 5%-Stichprobe aller von UVG-versicherten Personen gemeldeten Nichtberufsunfälle in der Schweiz. Fälle mit einem Rentenanspruch werden zu 100 % erhoben. Diese Fälle werden detailliert und nach vorgegebener Anleitung kodiert.

Gegen Berufsunfälle nach UVG sind sämtliche Arbeitnehmenden in der Schweiz obligatorisch versichert. Gegen Nichtberufsunfälle – hierzu zählen die Unfälle im Strassenverkehr abseits des direkten Arbeitswegs – sind zusätzlich jene Arbeitnehmenden versichert, welche mindestens acht Wochenstunden beim selben Unternehmen arbeiten. Auch Bezügerinnen von Arbeitslosenentschädigung (ALE) sind versichert. Nicht obligatorisch gegen Nichtberufsunfälle im Verkehr versichert sind somit selbstständig erwerbende und Nicht-Erwerbstätige (Hausfrauen und -männer, Kinder, Studierende sowie Rentnerinnen und Rentner). Diese müssen bei einer Krankenversicherung (nach KVG) eine Unfallversicherung abschliessen (Unfallzusatz).

Alle UVG-Versicherer sind verpflichtet, der SSUV sämtliche wichtigen Falldokumente zu übermitteln für Fälle, welche in diese Spezialstatistik einfließen. Aus diesen Unterlagen ermittelt und kodiert die SSUV detailliert die Unfallursachen, die beteiligten Gegenstände (z. B. Fahrzeuge bei Strassenverkehrsunfällen), medizinische Diagnosen (ICD-10) und Fallkosten.

Medizinische Statistik der Krankenhäuser (MSK)

Die Medizinische Statistik der Krankenhäuser (MSK) ist eine obligatorische Vollerhebung aller (teil-)stationären Fälle. Bei der Kodierung der Diagnosen und Unfallhergänge kommt die ICD-10 GM (German Modification) zur Anwendung. Im Unterschied zur WHO-Version der ICD-10 sind die Diagnosecodes des Kapitel 20: «Äussere Ursachen von Morbidität und Mortalität» in der GM-Version stark zusammengefasst. Während die WHO-Version mit den Codes V01 bis V89 eine differenzierte Kodierung des genutzten Fahrzeugs der Verunfallten und des Fahrzeugs der Kollisionsgegner erlaubt, wird in der GM-Version nur ein einziger Code (V99!) vergeben. In diesem sind alle Transportmittelunfälle, inkl. Unfälle mit Luft- und Wassertransportmitteln zusammengefasst. Eine zuverlässige Identifizierung von Strassenverkehrsunfällen ist damit nicht möglich.

Statistik der Patientendaten Spital ambulant (PSA)

Die Statistik der Patientendaten Spital ambulant (PSA) ist eine obligatorische Vollerhebung bei allen Spitälern und Geburtshäusern der Schweiz und ergänzt die MSK mit Daten ambulanter Spitalkonsultationen [73]. Erhoben werden primär die erbrachten ambulanten Leistungen und soziodemografische Angaben der Patientinnen und Patienten. Die Erfassung von Diagnosen ist nicht obligatorisch und nicht standardisiert, d. h., es kommen unterschiedliche Klassifikationen zum Einsatz [74]. Mit den aktuell verfügbaren Daten ist keine Differenzierung zwischen Leistungen aufgrund von Erkrankungen oder Unfällen möglich.

Todesursachenstatistik

In der Todesursachenstatistik des BFS werden alle Sterbefälle der ständigen Wohnbevölkerung der Schweiz seit 1969 elektronisch erfasst [75]. Neben Geburts- und Sterbedatum werden das Geschlecht, Zivilstand, Wohnort, der Beruf und die Sterbegemeinde bzw. das Sterbeland erfasst. Bei Unfällen wird – wenn bekannt – ebenfalls das Unfalldatum erfasst. Die Todesursachen werden über die ICD-10 in der WHO-Version kodiert [76]. Bei Unfällen kommt das Kapitel 20 der ICD-10, «Äussere Ursachen von Morbidität und Mortalität», zur Anwendung. Grundsätzlich ist die Definition von Strassenverkehrsunfällen in der ICD-10 der des ASTRA ähnlich: Am Unfall muss ein Transportmittel beteiligt sein und der Unfall muss sich auf einem öffentlichen Verkehrsweg ereignet haben. Landtransportmittelunfälle werden mit den ICD-10-Codes V01 bis V89 kodiert. Eine detaillierte Anwendung der Codes erlaubt die Identifizierung des genutzten Fahrzeugs, des Fahrzeugs des Kollisionsgegners,

sowie die Differenzierung, ob es sich um einen Verkehrsunfall handelt oder einen Unfall ausserhalb des Verkehrs.

6.3.2 Beurteilung der verfügbaren Datenquellen

Um beurteilen zu können, ob sich diese erschlossenen Quellen für einen Linkage-Ansatz und zur Erfüllung der genannten Differenzierungsbedürfnisse eignen, geht Tab. 8 im Detail auf diese Datenquellen ein und stellt diese direkt den Daten im Hellfeld (ASTRA Unfalldaten) gegenüber. Insbesondere ist für jede Datenquelle angegeben:

- Lassen sich Opfer von Strassenverkehrsunfällen identifizieren (im Einklang mit der Definition des ASTRA)?
- Ist eine Überschneidung mit dem Hellfeld feststellbar oder eine direkte oder indirekte Linkage möglich?
- Welche räumliche Abgrenzung gilt für die erfassten Unfallopfer bzw. Unfälle? Sind die Unfälle und Unfallfolgen nach Unfallort (In- bzw. Ausland) abgrenzbar?
- Für die in Kapitel 4 erwähnten Differenzierungsmerkmale: ist eine solche Differenzierung im Einklang mit den Differenzierungen im Hellfeld möglich? Mit welcher (maximalen) Ausprägung?

Tab. 8: Beurteilung der Anwendbarkeit der Datenquellen für die Ansätze Datenspuren und Linkage.
Farbkodierung: türkis = ja, gelb = teilweise/mit Einschränkungen, rot = nein, grau = nicht anwendbar

	Hellfeld (Strassenverkehrsunfälle – ASTRA)	UVG-Statistik (SSUV)	Medizinische Statistik der Krankenhäuser MSK (BFS)	Spitaldaten Spitalambulant (PSA)	Statistik der Todesursachen und Totgeburten (nur Todesfälle)
Identifikation Strassenverkehrsunfall (ja/nein)		Ja, mit Einschränkung: - Genutzte und andere am Unfall beteiligte Verkehrsmittel - Unfallort öffentliche Strasse	Unschärf (ICD-10-GM V99!, Strassenverkehrsunfälle sind eine Teilmenge aller Transportmittelunfälle)	Nein	Ja, mit Einschränkung
Direkte Verknüpfung mit Hellfeld / Polizeirapport		Kodierung «Polizeirapport vorhanden» (bis 2022): - Eindeutig Ja - Wahrscheinlich Ja - Wahrscheinlich Nein/Unbekannt	Nein	Nein	Nein
Andere Ansätze zur Verknüpfung	- Fz-Kennzeichen (nur CH) - Stammmummer Fahrzeugausweis (nur CH)	Probabilistische Verknüpfung (z. B. Datum des Unfalls, Alter, Geschlecht etc.)	Probabilistische Verknüpfung (z. B. Datum des Unfalls, Alter, Geschlecht etc.)	Probabilistische Verknüpfung (z. B. Datum des Unfalls, Alter, Geschlecht etc.)	Probabilistische Verknüpfung (z. B. Datum des Unfalls, Alter, Geschlecht etc.)
Territorialprinzip	Ja - Unfälle von Schweizern in der CH - Unfälle von Ausländern in der CH	Teilweise: (UVG-Versicherte) - Mit: Unfälle von Schweizern im Ausland (Ausschluss möglich) - Mit: Unfälle von UVG-Versicherten Grenzgängern in der CH - Ohne: Unfälle von nicht UVG-Versicherten (Personen mit Wohnsitz im Ausland) in der CH	Teilweise (bei Behandlung in der CH) Mit: Unfälle von Schweizern im Ausland (Ausschlussmöglichkeit unklar)	Teilweise (bei Behandlung in der CH) Mit: Unfälle von Schweizern im Ausland (Ausschluss nicht möglich)	Nein (Inländerprinzip)
Grundgesamtheit	Alle Personen mit Strassenverkehrsunfall in der Schweiz	UVG-Versicherte mit gemeldetem Unfall Ohne Unversicherte (insb. Kinder, Jugendliche, Rentner)	Alle Personen mit (teil-)stationärer Behandlung in einem Krankenhaus in der Schweiz	Alle Personen mit ambulanter Behandlung in einem Krankenhaus in der Schweiz	Alle Sterbefälle der ständigen Schweizer Wohnbevölkerung (ohne nicht ständige Bev. / Wohnsitz im Ausland)
Alter	Ja (Geburtsdatum)	Ja: Ab ca. 15 Jahre bis ca. 64 Jahre	Ja	Ja (anonymisiert)	Ja (Geburtsdatum)
Geschlecht	Ja	Ja	Ja	Ja (anonymisiert)	Ja

Tab. 8: Beurteilung der Anwendbarkeit der Datenquellen für die Ansätze Datenspuren und Linkage.

Farbkodierung: türkis = ja, gelb = teilweise/mit Einschränkungen, rot = nein, grau = nicht anwendbar

	Hellfeld (Strassenverkehrsunfälle – ASTRA)	UVG-Statistik (SSUV)	Medizinische Statistik der Krankenhäuser MSK (BFS)	Spitaldaten Spitalambulant (PSA)	Statistik der Todesursachen und Totgeburten (nur Todesfälle)
Verletzungsschwere	<u>Unfallschwere (~NACA)</u> - Unbekannt - Nicht verletzt (0) - Leichtverletzte: Leicht verletzt (1–2) - Schwerverletzte: Erheblich verletzt (3–4) und lebensbedrohlich verletzt (5–6) - Getötete: Auf Platz gestorben (7) Innert 30 Tage gestorben <u>NACA-Code:</u> nicht obligatorisch, rund 65 % nicht erfasst	- Anzahl entschädigte Ausfalltage - IV-Rente - Informationen zu den entschädigten Ausfalltagen und Rentenleistungen erst nach 4 Jahren annähernd vollständig - Kodierte ICD10 Diagnosekategorien	- ICD-10-GM - Anzahl Aufenthaltstage - Intensivmedizinische Behandlung	- Ambulante Leistungen (ICPC) - Teilweise Diagnoseinformationen	Nur Todesfälle
Fahrzeugart	Objektart (Verkehrsteilnahme): Feine Differenzierung möglich	Detaillierte Kodierung der genutzten und beteiligten Fahrzeuge Kategorien (in guter Qualität): - Velo / E-Bike - Fussgänger - PW - Motorrad / Mofa - Schwere Fahrzeuge - Andere	Keine Identifizierung / Differenzierung möglich	Keine Identifizierung / Differenzierung möglich	Über ICD10-WHO Code Erfassung möglich, jedoch in der Praxis oft als «nicht näher bezeichneter Transportmittelunfall» kodiert
Unfallhergang (Alleinunfall vs. Kollision)	- Anzahl erfasste Fahrzeuge und Mitfahrende - Unfalltyp (Schleuder-/Selbstunfall)	Unterscheidung Alleinunfall und Kollision über «genutztes» und «beteiligtes» Fahrzeug möglich	Nein	Nein	Über die V-Codes der ICD10 möglich, jedoch in der Erfassungspraxis meist nicht differenziert kodiert
Unfallort	Unfallort (Gemeindenummer, Koordinaten)	- Inland: Postleitzahl des Unfallortes (Gemeindezuordnung unscharf) - Ausland: Ländercode	Nein	Nein	Nur Sterbegemeinde oder Sterbeland
Unfallzeit	Datum und Uhrzeit	Datum / Zeit	Nein	Nein	Unfalldatum (nur teilw. verfügbar)
Unfallursache bzw. Hergang	- Ursache/Hauptursache - Strassenart - Strassenzustand - Witterung - Strassenbeleuchtung	Kodierung gemäss Wegleitung: Hergangskategorie 09	Nein	Nein	Der Hergang ist über Kapitel 20, ICD-10 kodiert.
Verursacher (-prinzip)	- Mutmasslicher Hauptverursacher (Objekt) - Mängel (Objekt)	Nein	Nein	Nein	Nein

Tab. 8: Beurteilung der Anwendbarkeit der Datenquellen für die Ansätze Datenspuren und Linkage.
Farbkodierung: türkis = ja, gelb = teilweise/mit Einschränkungen, rot = nein, grau = nicht anwendbar

	Hellfeld (Strassenverkehrsunfälle – ASTRA)	UVG-Statistik (SSUV)	Medizinische Statistik der Krankenhäuser MSK (BFS)	Spitaldaten Spital-ambulant (PSA)	Statistik der Todesursachen und Totgeburten (nur Todesfälle)
Ursache Infrastruktur (Gestaltung)	Evtl. Kombination aus Unfallstelle (Kreisel, Kurve, Platz), Strassenzustand (Zusatzfelder: z. B. Rollsplit, Schlaglöcher, Spurrinnen), Strassenbeleuchtung, Ursachen (z. B. Zustand der Infrastruktur: Spitzwinklige Gleisquerung)	Annäherung über Kodierung der am Unfall beteiligten Gegenstände (z. B. Kreisel, Tramschienen, Trottoirrand, Mauern)	Nein	Nein	Nein
Kanton / Raumtyp	Ja (Koordinate)	Ja, über Postleitzahl (Zuordnung PLZ zu Gemeinde unscharf)	Nein	Nein	Sterbegemeinde
Antriebsart	Über Verknüpfung mit Fahrzeugdaten aus IVZ-Auswertung (wird vom ASTRA bereits gemacht) [nur CH-Nummernschilder]	Nein	Nein	Nein	Nein

Die Beurteilung der Datenquellen lässt sich wie folgt zusammenfassen:

- Aktuell **gut erschlossen** und für Abschätzungen zur Dunkelziffer von Verletzten gemäss dem Linkage-Ansatz nutzbar sind nur die UVG-Daten der SSUV. Für die Todesfälle (bei denen in der Regel von einer minimalen Dunkelziffer ausgegangen wird), kann die Statistik der Todesursachen genutzt werden.
- **Opfer von Strassenverkehrsunfällen** gemäss ASTRA-Definition lassen sich nur in den Daten der SSUV ausreichend scharf abgrenzen. Die Abgrenzung gegenüber anderen Unfällen erfolgt teilweise über die «beteiligten Gegenstände», wobei die Beteiligung von Fahrzeugen oft Hinweis auf Strassenverkehrsunfälle sind. Insbesondere die räumliche Abgrenzung («auf öffentlichen Strassen») sowie der «ursächliche Zusammenhang mit den Gefahren des Strassenverkehrs» lassen sich oft nicht eindeutig bestimmen. Keine der anderen aktuell erschlossenen Quellen lässt derzeit eine ausreichende Identifikation von Strassenverkehrsunfällen zu.
- Angaben zum **genutzten Fahrzeug** bzw. zu den **beteiligten Fahrzeugen** sind nur in den UVG-Daten der SSUV kodiert. Diese sind jedoch deutlich stärker aggregiert als im Hellfeld. Die Qualität der erfassten Unfallhergänge sowie deren Kodierung in der UVG-Statistik bestimmt somit die maximal mögliche Differenzierung der Fahrzeugkategorien in der Dunkelziffer.
Die übrigen untersuchten Datenquellen erfassen entweder keine Angaben zur Fahrzeugart (ICD-10-GM) oder nutzen entsprechende Kodierungsmöglichkeiten (ICD-10-WHO) in der Praxis kaum (z. B. Todesursachenstatistik).
- Die **soziodemografischen Merkmale** Alter und Geschlecht sind in allen erschlossenen Datenquellen vorhanden. Diese könnten als Grundlage für die Linkage dienen, allerdings mit einer hohen Ungenauigkeit. Die Bereitstellung von datengeschützten personenbezogenen Angaben oder Identifikationsnummern (z. B. AHV-Nummer, Führerausweis etc.) könnte die direkte Zuordnung deutlich verbessern. Ebenfalls für eine Linkage beigezogen werden können Unfalldatum, -zeitpunkt und -ort, falls hierzu vergleichbare Angaben vorhanden sind.
- Bei der Abbildung der **Verletzungsschwere** gibt es erhebliche Unterschiede zwischen den Quellen: Alle untersuchten Datengrundlagen enthalten medizinische Angaben zur

Verletzung. Die Kompatibilität mit der Verletzungsschwere im Helffeld ist jedoch aufgrund der unterschiedlichen Kodierung der medizinischen Angaben nicht gegeben. Damit für den Linkage-Ansatz ein Filter für die Verletzungsschwere gesetzt werden kann, müssen diese kompatibel gemacht bzw. harmonisiert werden (mehr dazu s. Kapitel 0).

6.3.3 Nicht erschlossene Quellen

Noch nicht systematisch erschlossen sind insbesondere Daten zu den Unfallversicherungen der Krankenkassen (KVG-Versicherer), der Motorfahrzeugversicherer sowie von Behandlungen in Arztpraxen (ausserhalb der Spitäler). Die Krankenkassen (gemäss dem Bundesgesetz über die Krankenversicherung (KVG)) und Motorfahrzeugversicherungen unterliegen im Unterschied zur Unfallversicherung (UVG) nicht einer Pflicht, Unfallmeldungen systematisch zu erfassen und durch einen zentralen Akteur zusammenzuführen.

Unfalldaten der KVG-Versicherer

Jede in der Schweiz wohnhafte Person muss gegen Unfälle versichert sein. Dies geschieht entweder über den Arbeitgeber (UVG-Versicherung, für in der Schweiz berufstätige Personen), oder über den Unfallzusatz der obligatorischen Krankenversicherung (insbesondere für Kinder und Senioren/-innen).

Rund die Hälfte der Schweizer Wohnbevölkerung ist via UVG [77], die andere über das KVG gegen Unfälle versichert. Die Grundgesamtheit ist daher komplementär zu den UVG-Versicherten. Auf dem Schweizer Markt gibt es etwa 50 KVG-Versicherer. Dabei sind 90 % der KVG-Versicherten bei den grössten 20 Unternehmen versichert. Durch Zusammenschluss der UVG- und KVG-Unfallmeldungen zu einer «Versicherungsliste» ergäbe sich für in der Schweiz versicherte und ärztlich behandelte Unfallopfer ein weitgehend vollständiges Bild. Voraussetzung für einen Einsatz dieser Daten zur Schätzung der Dunkelziffer ist die Verwendung einer Kodierung von Unfällen und Unfallfolgen kompatibel mit dem Helffeld.

Daten der Motorfahrzeugversicherungen

Im Bereich der Motorfahrzeugversicherungen gibt es verschiedene Arten von Versicherungen, in denen Datenspuren zu Strassenverkehrsunfällen anfallen:

- Die Motorfahrzeug-Haftpflichtversicherung ist in der Schweiz obligatorisch. Bei einem Unfall mit motorisiertem Fahrzeug mit Verletzten (Personenschaden) oder Sachschäden, zahlt in der Regel die Motorfahrzeug-Haftpflichtversicherung des Verursachers oder der Verursacherin für Heilungskosten oder Einkommensausfall von geschädigten Dritten.
- Personenschäden beim Verursacher eines Strassenverkehrsunfalls sind nicht durch die Motorfahrzeug-Haftpflicht versichert, sondern werden über deren Unfallversicherung (UVG oder Unfallzusatz KVG) gedeckt.
- Insassenunfallversicherungen decken Personenschäden an Mitfahrenden, die bei selbstverschuldeten Unfällen entstehen.
- Vollkaskoversicherungen (auch Kollisionskasko) decken Sachschäden am eigenen Fahrzeug, die bei einer Kollision entstanden sind.

Strukturierte Daten zu den durch diese Versicherungen bearbeiteten Fällen (Anzahl Fälle nach Fahrzeugkategorie, Ausbezahlte Versicherungsleistungen für Personen- bzw. Sachschäden, Anzahl Fälle nach Höhe der Leistungen) könnten wertvolle Hinweise für durch motorisierte Fahrzeuge verursachte Unfälle und Schäden geben. Insbesondere Informationen zu Sachschäden sind in allen anderen Quellen nicht erfasst.

Nicht über eine Fahrzeugversicherung versichert sind in der Regel Schäden, die ohne Beteiligung eines Motorfahrzeugs von Velos, Pedelets oder Fussgängern verursacht wurden, da solche Fälle meist über deren Privathaftpflichtversicherung laufen. Für Mofas und E-Bikes mit Nummernschild gibt es obligatorische kantonale Kollektiv-Haftpflichtversicherungen («Vignette»).

Über eine Vollerhebung zu den ausbezahlten Versicherungsleistungen für Sachschäden (Kollisionskasko, Motorfahrzeug-Haftpflicht, Mofa/E-Bike-Haftpflicht) könnte für den motorisierten Verkehr eine gute Grundlage für die Abschätzung der entstandenen Sachschäden im Strassenverkehrsunfallgeschehen geschaffen werden. Dabei gibt es auch hier eine Dunkelziffer bzw. Sachschäden, die nicht der Versicherung gemeldet werden (z. B. wegen eines drohenden Malus oder Bonusverlusts bei der Versicherungsprämie). Sachschäden, die durch im Ausland versicherte Motorfahrzeuge verursacht wurden, sind mit dieser Grundlage nicht erfasst.

Voraussetzung für einen Einsatz dieser Daten zur Schätzung einer differenzierten Dunkelziffer bei Verletzten ist die Verwendung einer Kodierung von Unfällen und Unfallfolgen kompatibel mit dem Hellfeld. Derzeit sind die Akteure im Schweizer Motorfahrzeugversicherungsmarkt nicht dazu verpflichtet, Daten mit Bezug zum Strassenverkehrsunfallgeschehen systematisch zu erfassen. Eine Ad-hoc-Erhebung bei den Versicherungen wäre aufgrund des stark fragmentierten Marktes, der unterschiedlichen Datenerfassung sowie weil es sich aus Sicht der Versicherer meist um schützenswerte Daten handelt, sehr aufwendig und mit grossen Unsicherheiten bezüglich Datenqualität verbunden.

Exkurs: Dunkelziffer zu Sachschäden

Der vorliegende Forschungsbericht fokussiert auf die Dunkelziffer bei Strassenverkehrsunfällen mit Personenschäden (Verletzungen oder Todesfälle). Volkswirtschaftlich spielen jedoch auch Sachschäden im Zusammenhang mit Strassenverkehrsunfällen eine Rolle. Gemäss aktuellen Schätzungen sind diese für rund 5–7% der Unfallkosten des Strassenverkehrs verantwortlich [12].

Die Datenlage zum Helffeld ist dabei anders gelagert als bei den Verletzten. Erstens handelt es sich um monetäre Beträge, im Unterschied zu Fallzahlen bei Unfällen, Verletzten oder Getöteten. Zweitens erfasst die Polizei zwar im Unfallerfassungsprotokoll (UAP) die geschätzten Sachschäden und meldet diese dem ASTRA. Das ASTRA erachtet jedoch die Qualität der so erfassten Sachschäden als nicht ausreichend, um diese für weitergehende statistische Auswertungen einsetzen zu können. Es werden deshalb nur Daten zur erfassten Zahl der Unfälle «nur mit Sachschaden» publiziert, wobei auch bei Unfällen mit Personenschäden meist Sachschäden entstehen. Entsprechend ist es auch nicht möglich, analog zu den verletzten Personen einen Faktor zur Ermittlung der Dunkelziffer zu schätzen. Es muss direkt eine Schätzung des Gesamtvolumens über andere Quellen vorgenommen werden.

Die oben erwähnte Schätzung zu den Sachschäden im Strassenverkehr beruht deshalb auf einer Angabe zum durchschnittlichen Sachschaden pro Unfall, einer groben Abschätzung der Gesamtzahl der Unfallereignisse (inkl. Dunkelziffer) sowie der dabei involvierten Fahrzeuge, differenziert nach verschiedenen Fahrzeugkategorien. Der mittlere Sachschaden pro Unfall basiert auf einer Auswertung von freiwillig bereitgestellten und stark aggregierten Daten einiger Motorfahrzeug-Haftpflichtversicherungen (ausbezahlte Leistungen und Anzahl Versicherungsfälle mit Sachschäden). Bei der Hochrechnung auf das gesamte Unfallgeschehen wird berücksichtigt, dass nur durch Motorfahrzeuge verursachte Schäden erfasst sind, nur Schäden an Dritten durch die Haftpflicht gedeckt sind und gewisse Unfälle mit Sachschäden nicht den Versicherungen gemeldet werden (weil beispielsweise ein Malus/Bonusverlust bei der Prämie vermieden werden soll und/oder weil der Schaden unterhalb des Selbstbehalts liegt). Ergänzend werden bei Personenwagen die Leistungen der Kollisionskaskoversicherungen berücksichtigt, welche auch Schäden am eigenen Fahrzeug decken.

Als alternativer Ansatz wäre beispielsweise eine Erhebung der Sachschäden über Fahrzeug-Reparaturkosten (z. B. bei Reparaturwerkstätten oder über eine Haushaltsbefragung), ergänzt mit Angaben der Strasseneigentümer zu unfallbedingten Reparaturen an der Strasseninfrastruktur (in Abgrenzung zum normalen Unterhalt), denkbar. Aus Kosten-Nutzen-Überlegungen stellt sich dabei die Frage, ob der dadurch gewonnene Erkenntnis- bzw. Genauigkeitsgewinn den entstehenden zusätzlichen Erhebungsaufwand rechtfertigt.

Patientendaten aus Arztpraxen

Im Unterschied zu Patienten mit Behandlung in Spitälern (stationär und ambulant) liegen für Patienten, die in Arztpraxen (ausserhalb der Spitäler) behandelt wurden, keine strukturiert und zentral verwalteten Daten vor. Im Unterschied zur Behandlung im Spital ist bei Arztpraxen nicht von einer Überrepräsentation von bei Strassenverkehrsunfällen mittelschwer- bis schwerverletzten Personen auszugehen. Ein Einbezug von Daten aus Arztpraxen könnte daher insbesondere bei den Leichtverletzten ein vollständigeres Bild ergeben. Eine Erschliessung dieser Daten zur Erfassung der Dunkelziffer bei Strassenverkehrsunfällen erscheint jedoch unverhältnismässig, insbesondere verglichen zum Fall, dass neben den UVG- auch die KVG-Unfalldaten erschlossen würden, die beide auch Behandlungen in Arztpraxen umfassen.

Der Einsatz von Algorithmen bei der systematischen Erfassung und Kodierung von Unfallmeldungen oder Anamnesetexten (s. Kapitel 6.4) könnte die künftige Erschliessung dieser Quellen unterstützen.

6.4 Weitere Datenquellen und Daten für maschinelles Lernen

Folgende weitere Datenquellen wurden auf ihre Eignung zur Erfassung von Strassenverkehrsunfällen in Betracht gezogen: Anamnesetexte in Spitälern, Smartphone-/Smart-Device-Funktionen, Mobile-Apps, Meldesysteme, Videoaufnahmen und Fahrzeugsensoren. Diese werden im Folgenden kurz erläutert.

Anamnesetexte in Spitälern

Für jede medizinische ambulante oder stationäre Behandlung in einem Schweizer Spital wird ein freier Anamnesetext erstellt, der im Fall von Verletzungen auch den für die Verletzung kausalen Hergang beschreibt. Die Texte liegen elektronisch vor, sind aber nicht standardisiert und können damit in ihrem Informationsgehalt innerhalb oder auch Spitalübergreifend stark variieren. Im Frühjahr 2024 wurde von der BFU in Zusammenarbeit mit der Zürcher Hochschule für Angewandte Wissenschaften ZHAW und dem Kinderspital Zürich ein Pilotprojekt zur «Machine Learning»-basierten Auswertung von Unfallanamnesen gestartet. Hintergrund ist der Aufbau eines Monitoringsystems für Kinderunfälle. Unfalldaten für Kinder werden – bis auf die polizeilich registrierten Strassenverkehrsunfälle – nicht systematisch erfasst [78]. Zentral für den Erfolg des Projekts ist es, über möglichst viele Trainingsdaten zu verfügen.

Im Bereich Strassenverkehr könnte für eine automatische Kodierung der Unfalldaten der Unfallort problematisch sein: Im Projektverlauf wurde festgestellt, dass diese Information oft in den Anamnesetexten fehlt. Wird z. B. ein Sturz mit einem Velo beschrieben, ist nicht erwähnt, ob sich dieser auf der Auffahrt zu Hause oder auf einer öffentlichen Strasse ereignete. Somit können Strassenverkehrsunfälle nicht zuverlässig identifiziert werden.

Aufwendig gestaltet sich dieses Projekt wegen des Datenschutzes: Der Zugriff auf die Spitaldaten erfordert eine Bewilligung durch eine Ethikkommission. Bei einer Ausweitung des Projekts auf weitere Spitäler müssten zusätzliche Ethikanträge gestellt werden. Eine weitere Herausforderung ist das Trainieren der Algorithmen auf die verschiedenen Landessprachen. Durch die Nutzung von Spitaldaten werden eher schwerere Verletzungen erfasst. Für stationäre Fälle liegen die ICD-10-kodierten Diagnosen, für ambulante Fälle eine textliche Beschreibung vor.

Unfallerkennung per Smartphone-/Smart-Device-Feature

Apple hat in seinen Betriebssystemen iOS und WatchOS bereits zwei Unfallerkennungs-Features implementiert: Die Sturzerkennung und die Verkehrsunfallerkennung. Auch in Android in der Version 13 wurde inzwischen eine Unfall- und Sturzerkennung implementiert. Diese ist aber noch nicht für alle Smartphone-Modelle verfügbar. Mittels der Daten verschiedener Sensoren (Beschleunigungsmesser, Gyroskop und Mikrofon) und Software-Algorithmen, die teilweise durch «Machine Learning» trainiert werden, werden verschiedene Bewegungsabläufe als Unfall erkannt. Während sich in den Medien falsch positiv als Unfall bewertete Ereignisse in Vergnügungsparks verbreiteten, liegen über die generelle Zuverlässigkeit der Sturz- und Unfallerkennung (falsch Positive, falsch Negative) keine offiziellen Daten der Softwarehersteller vor.

Über einen GPS-Stempel zum Unfallzeitpunkt kann über entsprechendes Kartenmaterial bestimmt werden, ob der Unfallort auf einer öffentlichen Strasse liegt. GPS kann theoretisch und unter besten Bedingungen eine Ortsgenauigkeit im Zentimeterbereich bieten. In der Realität verringern verschiedenen Faktoren diese Genauigkeit deutlich: Die in Smartphones verbauten Antennen sind weniger leistungsstark und Chip-Sätze oft auf Energieeffizienz als auf Genauigkeit optimiert. Die Abstraten weisen – ebenfalls aus Effizienzüberlegungen – häufig hohe Intervalle auf. So werden Lücken im Routenverlauf verursacht. Auch die Umgebung, z. B. hohe Gebäude, kann einen starken Einfluss auf die GPS-Genauigkeit haben.

In einer Übersichtsarbeit wurden verschiedenen Studien zur Echtzeit-Sturzerkennung evaluiert: Die Spannweite der Sensitivität lag bei 60 % bis 99 %, die der Spezifität bei 75 % bis

100 % [79]. Die Autoren kritisieren aber die Evaluationsmethoden und vermuten eine Überschätzung der Testkriterien. Zur Unfallerkennung schreibt Apple «Die Unfallerkennung erkennt schwere Autounfälle (z. B. Frontal-, Seiten- und Heckaufprall oder Überschläge) verschiedener Personenkraftwagen, etwa Limousinen, Minivans, SUVs oder Pickup-Trucks.» Gleichzeitig wird als Fussnote die Warnung ausgegeben: «Die Unfallerkennung kann nicht alle Autounfälle erkennen» [80]. Ob auch Stürze, z. B. mit einem Velo, erkannt werden, ist nicht bekannt. Erkennungsalgorithmen existieren aber bereits für andere Produkte, wie die bis Ende 2023 verkaufte Velohelmalternative «Hövdning» (ein Airbag für den Kopf). Insgesamt werden Unfallerkennungsalgorithmen in Zukunft immer besser und zuverlässiger werden. Eine Nutzung dieser Daten steht aber vor grossen Datenschutzhürden. Gesundheitsdaten werden als besonders schützenswert deklariert und eine Nutzung durch Drittpersonen ist schwer bis unmöglich zu realisieren. Zudem werden zuverlässige Informationen zur Verkehrsmittelnutzung während des Unfalls und auch Informationen zu Verletzungen oder der Verletzungsschwere kaum vorhanden sein.

Unfallfassung per Mobile-App

Inzwischen existiert eine Handvoll Smartphone-Apps, die sich zur Aufgabe gemacht haben, Unfälle zu erkennen und anschliessend entsprechende Massnahmen einzuleiten. Diese Apps agieren ähnlich wie die bereits in den Betriebssystemen implementierten Unfallerkennungsalgorithmen. Es gelten ebenfalls die hohen Datenschutzhürden und die geäusserten Bedenken hinsichtlich Verkehrsmittelnutzung – eine Ausnahme wären spezialisierte Apps für z. B. das Velo- oder Motorradfahren und Verletzungen. Neben Bedenken hinsichtlich der Repräsentativität solcher Daten aufgrund der Freiwilligkeit solcher Apps wirft die Nutzung von Daten diverser Anbieter auch Fragen zur Datenkompatibilität auf.

Eine analoge Funktionalität zur Unfallerkennung könnte auch in bestehende GPS-gestützte Mobilitätspanels integriert werden, welche neben der laufenden Erfassung von Mobilitätsdaten die gezielte Nacherfassung von Unfällen und weiterer Unfallmerkmale wie die Verletzungsschwere oder die Angabe der Nummer eines allfälligen Polizeirapports erlauben. Die Angaben zum Unfallgeschehen könnten zudem mit den Daten zur Mobilität verknüpft werden. Diese Verknüpfung ermöglicht auch die Herleitung von Unfallraten. Einschränkungen in der GPS-Genauigkeit sind aber auch bei diesem Ansatz zu beachten.

Unfallfassung über Meldesystem

Es existieren bereits zahlreiche Mobilitätsplattformen, über die Gefahrenstellen (z. B. für Velofahrende Gefahrenstellen im Radwegenetz Mainz – uMap [openstreetmap.de] oder allgemein: Gefahrenstellen auf Strassen erkennen und melden – Gefahrenstellen), die versuchen, über eine Kombination von Mobilitätsdaten, Unfalldaten und Meldungen von Nutzern (Crowd-Science) Risikokarten, teilweise in Echtzeit, zu erstellen. Nutzende können hier subjektiv empfundene Gefahrenstellen oder teilweise auch erlebte oder beobachtete Unfälle melden. Der Vorteil eines solchen Meldesystems ist es, dass über die Eingabeformulare gesteuert werden kann, welche Informationen erfasst werden sollen. Gerade im Bereich einer Unfallmeldung ist aber davon auszugehen, dass nur Unfälle gemeldet werden, die sich vermeintlich aufgrund von Infrastrukturproblemen ereignet haben. Unfälle, die sich aufgrund eines Fehlverhaltens der Fahrzeugnutzer ereigneten, werden dort nicht erfasst. Zudem ist die Zuverlässigkeit der Angaben kaum prüfbar. Die Systeme sind von der Bereitschaft der Nutzer abhängig, entsprechende Meldungen zu erfassen. Dazu müssen die Systeme bekannt sein und vor allem den Nutzern einen Mehrwert bringen, indem z. B. die Betreiber auf gemeldete Mängel reagieren. Dies macht einen dauerhaften Betrieb aufwendig.

Unfallfassung über Videoaufnahmen

Insbesondere zur Konfliktanalyse wurden bereits zahlreiche Forschungsprojekte durchgeführt. Videoaufnahmen werden durch Computeralgorithmen automatisiert ausgewertet, um Konflikte («Beinaheunfälle»), zumeist an Knoten, über festgelegte Indikatoren («Time to collision» (TTC), «Post encroachment time» (PET)) zu identifizieren. Auch Unfälle können über solche Systeme erfasst werden. Während ein Schweizer Forschungsprojekt noch

zahlreiche Schwächen bei der Konfliktanalyse attestierte, ist in Zukunft damit zu rechnen, dass die Systeme durch Verbesserungen in Hard- und Software immer zuverlässiger werden [81].

Video- oder auch andere Aufzeichnungsmöglichkeiten wie Lidar sind aufwendig zu installieren und meist mit hohen Kosten für die Hardware verbunden. Zukünftig wird sich der Rechenaufwand auf die Kameras selbst verlagern, sodass grosse Mengen an Videodaten nicht mehr zentral gespeichert und bearbeitet werden müssen. Stattdessen übertragen die Kameras nur noch die Ergebnisse der Auswertung. Dadurch könnten auch bestehende, dauerhaft installierte Verkehrsüberwachungskameras mit solchen Systemen ausgerüstet werden.

Problematisch bleibt die räumliche Abdeckung: Verkehrskameras sind vor allem an kritischen Verkehrsinfrastrukturen (Tunnel, Knoten, Nationalstrassen) und in Ballungsräumen installiert. Eine flächendeckende Abdeckung oder eine zufällige Auswahl von Standorten ist aufgrund des hohen Aufwands und der begrenzten Akzeptanz in der Bevölkerung kaum realisierbar. Zudem lassen sich allein mit Videobildern noch keine Informationen zur Verletzungsschwere, zur polizeilichen Erfassung und zu anderen für die Differenzierung der Dunkelziffer relevanten Kriterien gewinnen.

Weiter zeigte das vom ASTRA in Auftrag gegebene Forschungsprojekt «Velo-Infrastruktur-Sicherheitsinstrumente V-ISSI», bei dem eine Mobility Observation Box eingesetzt wurde, dass Videoaufnahmen bei Dunkelheit und unzureichender Strassenbeleuchtung keine zuverlässigen Ergebnisse liefern. Zudem können je nach Bewertungsmethode (TTC und PET) Selbstunfälle nicht erfasst werden können.

Auch der Datenschutz ist bei Videoaufnahmen zu berücksichtigen, da personenbezogene Daten aufgenommen und optional gespeichert werden. Die KI-basierte Auswertung von Konfliktsituationen ist noch recht neu und weist verschiedene Unzulänglichkeiten bei der automatischen Erkennung auf. Die Fortschritte bei der Entwicklung solcher Systeme sind jedoch rasant und dürften in Zukunft valide Ergebnisse erzeugen.

Unfallerfassung über Fahrzeugsensoren

Moderne Motorfahrzeuge sind mit einer Vielzahl von Sensoren ausgerüstet. Diese reichen von Unfalldatenschreibern (Even Data Recorder, EDR), elektronische Notrufsystemen (Emergency Call, eCall) bis zu Daten, die für das automatisierte Fahren benötigt werden [53].

Problematisch bei diesem Ansatz sind wiederum die Datenverfügbarkeit, der Datenschutz und die fehlende Erfassung von Verletzungsschweren. So sind z. B. Daten der eCall-Systeme nach EU-Richtlinien ausschliesslich für die Abwicklung der Notfallsituation vorgesehen. Nach Abschluss sollten diese umgehend gelöscht werden. Daten von Unfalldatenschreibern lassen sich dagegen bereits heute von Fahrzeugherstellern käuflich erwerben. Die Hersteller bestimmen, welche Daten sie herausgeben. Selbstverständlich können Unfälle nur von Fahrzeugen erfasst werden, die mit entsprechenden Sensoren ausgestattet sind. Zudem erfassen diese Systeme keine Daten zu den involvierten Personen, der Unfallschwere und zu Verletzungen.

6.5 Zusammenfassung Datenquellen

Die Datenquellen wurden spezifisch für die Nutzung anhand der drei beschriebenen methodischen Ansätze Stichprobe/Befragung, Datenspuren/Linkage sowie Maschinelles Lernen untersucht.

Stichprobe bzw. Befragungen

Über Befragungen können von der Polizei nicht erfasste Strassenverkehrsunfälle erhoben und in ihrer Zahl abgeschätzt werden. Die Schweizerische Gesundheitsbefragung (SGB) des BFS, die einzige bestehende wiederkehrende Erhebung des Bundes, welche sich zumindest am Rande mit Unfällen beschäftigt, enthält dabei zu wenige Angaben zur Identifizierung von Strassenverkehrsunfällen. Selbst wenn eine solche Abgrenzung erreicht werden könnte, lassen die aktuell erhobenen Daten kaum weitere Differenzierungen zu.

Eine in regelmässigen Abständen durchgeführte und explizit auf diesen Anwendungsfall ausgerichtete Befragung eignet sich jedoch grundsätzlich zur Schätzung von differenzierten Dunkelziffern im Strassenverkehrsunfallgeschehen in der Schweiz. Ein Beispiel hierfür stellt die von der BFU durchgeführte Haushaltsbefragung dar, welche auf die Erfassung aller Arten von Unfällen ausgerichtet ist, aber die Identifizierung von Strassenverkehrsunfällen in der Schweiz und von polizeilich nicht erfassten Unfällen sowie eine für die meisten Bedürfnisse ausreichende, differenzierte Auswertung zulässt. Ob und wie es mit der Haushaltsbefragung der BFU weitergeht, d. h., ob sie auch in Zukunft als Datenquelle zur Verfügung stehen wird, ist zum jetzigen Zeitpunkt unklar. Aufgrund der bei solchen Befragungen gängigen Stichprobenrahmen, welche auf dem Einwohnerregister basieren, können Unfälle von nicht in der Schweiz wohnhaften Personen nicht erfasst werden.

Datenspuren / Linkage

Die Datenspuren zu Strassenverkehrsunfällen sind derzeit nur vereinzelt systematisch erschlossen. Die bereits erschlossenen Datenquellen bieten für sich allein noch kein vollständiges Bild des Unfallgeschehens. Für eine Linkage müssen die erhobenen Fälle zwischen den Datenquellen abgeglichen werden können (Linkage). Darüber hinaus müssen Strassenverkehrsunfälle von anderen Unfällen abgegrenzt werden können. Für eine weitere Differenzierung muss ein vergleichbarer Filter nach dem entsprechenden Kriterium in allen verwendeten Datenquellen gesetzt werden können.

Insgesamt stellt die differenzierte Ermittlung der Dunkelziffer hohe Anforderungen an die Qualität der Datenspuren, die derzeit nur wenige erschlossene Datenquellen erfüllen:

Die Todesursachenstatistik des BFS eignet sich gut für einen Abgleich mit dem Helffeld im Bereich der Todesfälle.

Die UVG-Statistik ist dank ihrem strukturierten Aufbau und ihrer Repräsentativität eine vielversprechende Datenquelle für den Linkage-Ansatz. Aufgrund der auf die UVG-Versicherten beschränkten Grundgesamtheit eignet sie sich jedoch nur zur Abschätzung der Dunkelziffer von unselbstständig erwerbstätigen Erwachsenen und Bezüger/-innen der Arbeitslosenentschädigung. Im Vergleich zum Helffeld sind die Verletzungsschweren anders kodiert und die kodierten beteiligten Verkehrsmittel sind stark aggregiert.

Für Kinder, Studierende, nicht arbeitstätige Erwachsene und Senior/-innen ist der Unfallzusatz bei der Krankenkasse (KVG) obligatorisch. Eine ideale Ergänzung zur UVG-Statistik wäre deshalb eine vergleichbare Erschliessung der den Krankenkassen gemeldeten Unfällen. Derzeit liegen solche Daten jedoch nicht vor, und aufgrund der starken Segmentierung des Krankenversicherungsmarktes ist ohne gesetzliche Grundlage nicht davon auszugehen, dass sie in absehbarer Zeit vorliegen werden.

Selbst bei erfolgreicher Erschliessung der KVG-Daten bleibt eine Datenlücke für Personen bestehen, die nicht in der Schweiz unfallversichert sind – insbesondere nicht in der Schweiz arbeitstätige Personen mit Wohnsitz im Ausland. Um deren nicht der Polizei gemeldeten

Unfälle über erschlossene Datenspuren zu erfassen, könnten Behandlungsdaten der Spitäler oder Arztpraxen herangezogen werden. Allerdings erfassen diese nur Fälle, die in der Schweiz ärztlich behandelt wurden.

Die verschiedenen systematisch erhobenen Daten des Bundesamts für Gesundheit (BAG) zu Spitalbehandlungen (ambulant und stationär) erfüllen die hohen Anforderungen jedoch nur bedingt. Insbesondere lassen sich Strassenverkehrsunfälle nur schwer identifizieren und mit dem Hellfeld abgleichen, zudem stehen kaum zusätzliche Informationen zum Unfallereignis oder zu den Beteiligten zur Verfügung. Dies schränkt die Differenzierungsmöglichkeiten selbst bei erfolgreicher Abgrenzung stark ein. Darüber hinaus ist bei Spitalbehandlungen von einer Überrepräsentation schwerer Verletzungen auszugehen. Behandlungen in Arztpraxen, die insbesondere leichte Verletzungen betreffen, werden derzeit nicht systematisch und zentral erfasst.

Weitere Datenquellen und Daten für maschinelles Lernen

Weitere Datenquellen wie die Erfassung von Strassenverkehrsunfällen über Smartphones, Fahrzeugtechnik oder Videokameras, sind derzeit noch zu wenig verbreitet und nicht ausreichend ausgereift, um zur Schätzung einer Dunkelziffer für Strassenverkehrsunfälle eingesetzt werden zu können. Die bloße Feststellung und Identifikation eines Unfalls genügt nicht, um die Daten für eine differenzierte Dunkelzifferberechnung nutzen zu können. Notwendige Detailangaben zum Unfall, zu verletzten Personen oder zu den involvierten Fahrzeugen fehlen oder unterliegen einschränkenden Datenschutzbestimmungen. In diesen Daten liegt aber ein grosses Potenzial für eine zukünftige Nutzung zur Dunkelzifferabschätzung. Für eine erfolgreichen Nutzung solcher Daten in Modellen des maschinellen Lernens müssen möglichst vollständige und qualitativ hochwertige Daten zu Strassenverkehrsunfällen und ihren Folgen vorhanden sein. Dies ist aktuell noch nicht der Fall.

7 Methodenvorschläge

Im Folgenden wird aufgrund der Überlegungen in den vorangehenden Kapiteln nur auf die Ansätze «Stichproben- bzw. Befragungen» und «Datenspur / Linkage» eingegangen. Mit der aktuellen Datenlage ist der Ansatz des maschinellen Lernens zur Schätzung der Dunkelziffer derzeit nicht anwendbar. Allerdings sollte diese Methode – insbesondere im Hinblick die zunehmende Menge und Diversität von Sensordaten und eine sich möglicherweise ändernde rechtliche Ausgangslage – zu einem späteren Zeitpunkt neu evaluiert werden.

Für die Methode «Datenspur / Linkage» besteht die Herausforderung, dass viele «bestehende» Datenquellen in ihrer aktuellen Form die hohen Anforderungen nicht erfüllen (s. auch Kapitel 6.3):

- Strassenverkehrsunfälle gemäss Definition des ASTRA lassen sich nicht eindeutig von anderen Unfällen und Verletzungsursachen abgrenzen (eine wesentliche Voraussetzung für die Nutzung eines Datensatzes für die Dunkelzifferberechnung), siehe medizinische Statistik
- Daten sind nicht erschlossen oder nicht verwendbar (z. B. KVG). Selbst wenn solche Daten existieren, ist zu beachten, dass auf diese – beispielsweise aus datenschutzrechtlichen Gründen – nicht ohne Weiteres zugegriffen werden kann.

Da die zur Ermittlung der Dunkelziffer verfügbaren Datenquellen für Getötete und Verletzte nicht dieselben sind, muss für Getötete eine andere Methodik angewendet werden als für Verletzte. Deshalb wird im Folgenden zwischen der Abschätzung der Dunkelziffer für Getötete und für Verletzte unterschieden. Auf einen ausführlichen Methodenvorschlag zur Schätzung der Dunkelziffer bei Sachschäden wird verzichtet (siehe hierzu Exkurs in Kapitel 6.3.3)

7.1 Dunkelziffer für Getötete

Neben den polizeilich erfassten Strassenverkehrsunfällen steht für Todesfälle die Todesursachenstatistik des BFS zur Verfügung (s. auch Kapitel 6.3.10).

Auf seiner Website publiziert das BFS die Sterbefälle nach Ursache [82]. Das aktuelle Jahr der Todesursachenstatistik war zum Zeitpunkt der Erarbeitung dieses Berichts 2022. Ein Vergleich der Sterbefälle durch Strassenverkehrsunfälle mit den vom ASTRA publizierten Daten zeigt über die Jahre sowohl positive als auch negative Abweichungen (Tab. 9).

Tab. 9: Getötete bei Strassenverkehrsunfällen in der ASTRA- und der Todesursachenstatistik des BFS, 2018-2022

Jahr	Todesursachenstatistik BFS	ASTRA	Quotient TUS/ASTRA
2018	241	233	1.03
2019	196	187	1.05
2020	204	227	0.90
2021	199	200	1.00
2022	233	241	0.97

Diese Abweichungen können mehrere Gründe haben:

- **Zuordnung zum Statistikjahr:** Das ASTRA zählt einen Getöteten in dem Jahr, in dem sich der Unfall ereignete, die Todesursachenstatistik im Jahr des Sterbedatums. Eine Person, die z. B. Mitte Dezember 2021 verunfallt und Anfang Januar 2022 verstirbt, wird beim ASTRA zum Jahr 2021, vom BFS dagegen zum Jahr 2022 gezählt.
- **Territorial- vs. Inländerprinzip:** Bei ASTRA werden alle Getöteten auf Schweizer Strassen unabhängig vom Wohnland der Verunfallten gezählt, bei BFS alle Todesfälle

der ständigen Wohnbevölkerung unabhängig vom Unfallland. Dazu muss angemerkt werden, dass die Information zur Todesursache bei im Ausland verunfallten Personen in der Todesursachenstatistik nur selten vorliegen. Bei allen Sterbefällen im Ausland ist in 80 % bis 90 % die Todesursache als «R99 Sonstige ungenau oder nicht näher bezeichnete Todesursachen» kodiert. Der Anteil der Sterbefälle im Ausland bei Strassenverkehrsunfällen ist folglich gering (0 % bis 3 % in den letzten fünf Jahren).

- **30-Tage-Regel:** Beim ASTRA werden nach internationaler Konvention Verstorbene als Strassenverkehrstote gezählt, wenn zwischen Unfalldatum und Sterbedatum maximal 30 Tage liegen. Bei der Todesursachenstatistik sind längere Zeitspannen möglich.
- **Publizierte V-Codes:** Zu den tödlichen Strassenverkehrsunfällen werden vom BFS nur Fälle mit den V-Codes V01 bis V79, V85 bis V87 und V89 gezählt. Unfälle mit Pferden und tierbespannten Fuhrwerken (V80), Unfälle mit Zügen oder Trams (V81 und V82) werden ebenso ausgeschlossen wie Unfälle mit landwirtschaftlichen Fahrzeugen oder Spezialfahrzeugen für Baustellen und Betriebsgelände (V83 bis V86).

Grundsätzlich lassen sich einige der Vergleichbarkeitsprobleme lösen. Die Zuordnung zum Statistikjahr liesse sich über das Unfalldatum in der Todesursachenstatistik der Systematik des ASTRA angleichen. Auch die «30-Tage-Regel» liesse sich über das Unfalldatum umsetzen. Jedoch fehlt in der Todesursachenstatistik das Unfalldatum in 10 % bis 15 % der Strassenverkehrsunfälle. Ein geringer Anteil des Unfalldatums enthält zudem falsche Werte (Unfalldatum liegt nach dem Sterbedatum, Unfalldatum entspricht dem Geburtsdatum).

Problematischer ist die Zuordnung der einzelnen V-Codes zum Bereich Strassenverkehrsunfall. Grundsätzlich können auch Fälle mit den vom BFS ausgeschlossenen Codes beim ASTRA als Strassenverkehrsunfall gewertet werden. Insbesondere Unfälle mit Trams oder landwirtschaftlichen Fahrzeugen. Bei den eingeschlossenen Codes ist es ebenso möglich, dass sich diese Unfälle abseits öffentlicher Verkehrswege ereigneten und daher nicht zum Strassenverkehrsunfallgeschehen gezählt werden sollten. Bei vielen V-Codes ist es möglich, mit einer vierten Stelle des Codes die Information zu kodieren, ob sich ein Unfall im oder ausserhalb des Verkehrs ereignete. Rund 6 % der vom BFS gezählten Strassenverkehrsunfälle ereigneten sich demnach «ausserhalb des Verkehrs». Bei rund 50 % der Fälle ist diese Information aber nicht vorhanden. Insgesamt muss auch infrage gestellt werden, ob jeder Fall aus der ASTRA-Statistik in der Todesursachenstatistik mit einem V-Code kodiert wird.

Mit der Todesursachenstatistik liegt eine vollständige Datensammlung aller Sterbefälle der ständigen Schweizer Wohnbevölkerung der Schweiz vor. Durch die Kodierung der äusseren Ursachen mittels ICD-10 wäre theoretisch eine nahezu perfekte Übereinstimmung mit den Getöteten in den ASTRA Daten möglich. Die Kodierungspraxis zeigt jedoch ein anderes Bild: Oftmals ist nicht bekannt, ob sich ein Unfall im oder abseits des Verkehrs ereignete. Auch andere Restkategorien («nicht näher bezeichnet») sind stark besetzt, was eine zuverlässige Zuordnung der Fälle zu den Strassenverkehrsunfällen erschwert. Zudem wird es in beiden Statistiken Fehler geben (unerkannte Suizide, Tod vor Kollision).

Einen pragmatischen Weg geht die World Health Organization (WHO): Für den bereits häufiger aktualisierten «Global Status Report on Road Safety» wird die Vollständigkeit der «offiziellen Landesangaben» (für die Schweiz die Zahlen des ASTRA) mit einem Korrekturfaktor multipliziert. Aktuell liegt für die WHO die Dunkelziffer der Getöteten bei 1,03 [83].

Angesichts der Unsicherheiten, der gemäss Literatur relativ kleinen Dunkelziffern, der relativ kleinen Abweichungen zwischen den Daten des BFS und dem ASTRA sowie dem erwarteten Aufwand wird vorgeschlagen, vereinfachend von einer Dunkelziffer von 1.0 (Faktor) ausgehen und die benötigten Differenzierungsinformationen direkt aus den Daten des ASTRA nehmen.

7.2 Dunkelziffer für Verletzte

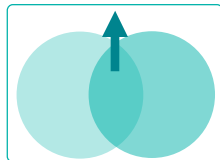
Um die Dunkelziffer für Leicht- und Schwerverletzte bei Strassenverkehrsunfällen abzuschätzen, kommen verschiedene Ansätze infrage. Basierend auf den Erkenntnissen aus den vorangehenden Kapiteln werden folgende Methoden vorgeschlagen:

- Stichprobe (Befragung) inkl. Hochrechnung



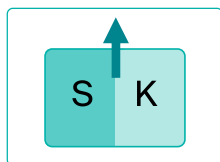
Das Piktogramm zeigt die befragte Stichprobe (türkis Kreis) sowie die Hochrechnung (Pfeil) auf die Gesamtzahl der verletzten Personen (Box) auf

- Linkage bestehender Datenquellen inkl. Capture-Recapture



Die beiden Kreise im Piktogramm repräsentieren zwei verschiedene Datenquellen. Mit der Überlappung sind Fälle, die in beiden Datenquellen vorhanden sind, dargestellt. Der Pfeil zeigt die Hochrechnung auf die Gesamtzahl der verletzten Personen (Box) mittels Capture-Recapture.

- Versicherungsliste (UVG-Daten der SSUV und KVG) (inkl. Hochrechnung)



Das Piktogramm zeigt die Kombination der Daten der SSUV («S») und der KVG-Versicherer («K») (diese beiden Datenquellen werden zusammengeführt und überschneiden sich nicht). Der Pfeil repräsentiert die Hochrechnung auf die Gesamtzahl der verletzten Personen (Box).

Methoden 1 und 2 können bereits kurz- bzw. mittelfristig umgesetzt werden. Um Methode 3, eine Versicherungsliste, bestehend aus den Daten der SSUV (UVG-Versicherer) sowie der KVG-Versicherer zu verwenden, müssten die Unfallmeldungen der KVG (bzw. eine Stichprobe davon) noch erschlossen werden. Da diese Erschliessung Zeit beansprucht, ist diese Methode allenfalls eine längerfristige Option. Im Folgenden wird auf die Ausgestaltung dieser drei Methoden genauer eingegangen und es werden Empfehlungen zum Vorgehen bei der Umsetzung gemacht.

7.2.1 Methode 1: Stichprobe (Befragung) inkl. Hochrechnung

Die Durchführung einer (Querschnitts-)Befragung ist eine bewährte und breit akzeptierte Methode. Wie bereits in Kapitel 1 und 3.1 erwähnt, wurden Befragungen bereits für eine Abschätzung des Strassenverkehrsunfallgeschehens in der Schweiz durchgeführt. Die generellen Vor- und Nachteile dieser Methode sind in Tab. 4 auf Seite 50 aufgeführt.

Stichprobendesign und Erhebungsmethode

Um eine aussagekräftige Dunkelziffer berechnen zu können, ist es zentral, dass die Ziehung der Stichprobe nach dem Zufallsprinzip aus einer gut definierten Grundgesamtheit erfolgt. Der SRPH des BFS bietet auf Basis der quartalsweise aktualisierten Daten der Einwohnerregister einen idealen Stichprobenrahmen.

Über die Grundlage des SRPH können verschiedene Stichprobendesigns angewendet werden, z. B. eine Stichprobe von Haushalten oder Einzelpersonen, die anhand diverser im SRPH vorhandener demografischer Merkmale (z. B. Haushaltsgrösse, Alter, Geschlecht und Wohnort) stratifiziert oder selektiert werden können. So könnten auch regionale Befragungen oder Befragungen bestimmter demografischer Gruppen realisiert werden: Bei einer spezifischen Erhebung zu Unfällen von Kindern bzw. wenn das Unfallgeschehen bei Kindern in die Erhebung eingeschlossen werden soll, würden Haushalte mit Kindern in die Stichprobe gelangen und eine auskunftsgibende Person bestimmt werden.

Als Erhebungsmethoden sind telefonische, schriftliche (postalische) und Online-Befragungen möglich, wobei die erste Variante mit deutlich höheren Kosten verbunden ist. Denkbar ist auch eine Kombination verschiedener Erhebungsmethoden («Methodenmix»), bei der z. B. bei Online-Befragungen Haushalte mit hochaltrigen Personen telefonisch befragt werden.

Neben Querschnittsbefragungen sind auch Panelbefragungen denkbar, bei denen Personen über einen längeren Zeitraum wiederholt befragt werden. Falls regelmässig (z. B. alle 3 Monate) nach Unfällen gefragt wird, dürfte mit diesem Vorgehen die Erinnerung an den Unfall besser und somit die Angaben zuverlässiger sein. Ein Nachteil ist jedoch, dass der Aufwand insbesondere auch für die Studienteilnehmenden grösser ausfällt als bei einer einmaligen Befragung, was sich negativ auf die Teilnahmebereitschaft und dadurch die Repräsentativität der Stichprobe auswirken könnte.

Sind die Inzidenzen in der Stichprobe für die interessierenden Würfelemente bestimmt, werden die Unfallzahlen (im jeweiligen Würfelement) auf die Schweiz hochgerechnet. Dabei muss eine Gewichtung erstellt werden, die das Studiendesign berücksichtigt und für «Nicht-Teilnahme» korrigiert.

Bedeutung der Stichprobengrösse

Ein Problem stellt die kleine Inzidenz von Strassenverkehrsunfällen dar, was zwar aus Public-Health-Sicht erfreulich, aber statistisch gesehen nicht ideal ist. Momentan rechnet die BFU auf Basis der Fortschreibung der Haushaltsbefragung der BFU aus dem Jahr 2010 mit jährlich ca. 79 000 im Schweizer Strassenverkehr verletzten Personen (Schweizer Wohnbevölkerung) (Durchschnitt 2017–2021). Bei einer Bevölkerung von 8,7 Millionen entspricht dies einer Inzidenz von knapp 1 % pro Jahr, also eine von 100 Personen pro Jahr verletzt sich einem Strassenverkehrsunfall. Die vorläufigen Resultate der Haushaltsbefragung der BFU aus dem Jahr 2024 weisen auf eine vergleichbare Inzidenz hin. Will man weiter differenzieren, z. B. nur Unfälle mit dem E-Bike berücksichtigen, verkleinert sich die Inzidenz weiter, und die Unsicherheiten der Resultate werden entsprechend grösser.

Dies verdeutlicht Abb. 4: Simuliert wurde eine Befragung von 5000, 10 000 und 50 000 Personen (obere 3 Balken) für ein Würfelement der Grösse $n = 50\,000$, was für die Schweiz eine Inzidenz von ca. 0,6 % bedeutet. Diese Zahl könnte ungefähr der Anzahl bei Strassenverkehrsunfällen leicht verletzter Erwachsener in der Schweiz (ohne weitere Differenzierung) entsprechen.

Bei den Fehlerbalken handelt es sich um die Wurzel des mittleren quadratischen Fehlers («Root mean squared error», RSME) zwischen dem vorgegebenen «wahren Wert» ($n = 50\,000$) und der Schätzung aus der 100-fach simulierten Stichprobe. Der Fehler setzt sich zusammen aus einem Bias (systematische Verzerrung) und der Standardabweichung.

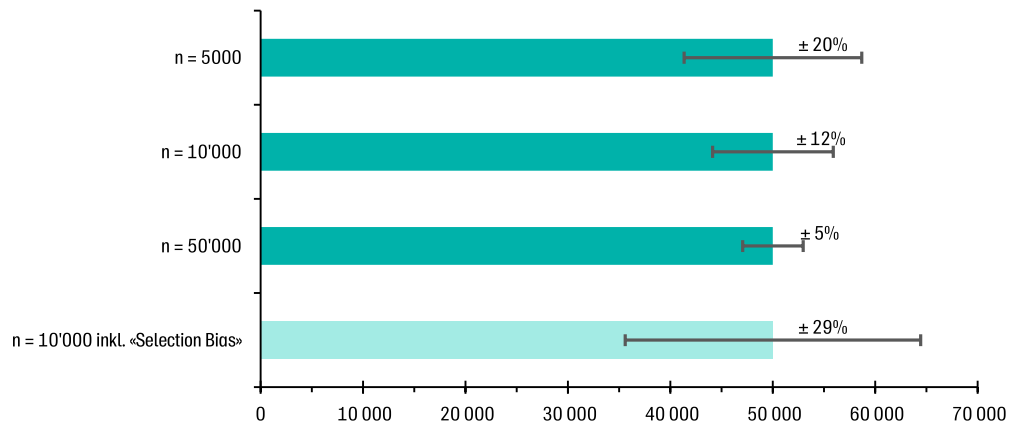


Abb. 4: Simulation des Fehlerbereichs (RSME) bei Befragungen einer zufälligen Stichprobe für ein Würfelement der Grösse $n = 50\,000$ für die Schweizer Bevölkerung (entspricht einer Inzidenz von ca. 0,6 %) mit unterschiedlicher Stichprobengrösse. Anhang III enthält die Details zu den Simulationen.

Je mehr Personen befragt werden, desto kleiner fällt erwartungsgemäss der Fehler aus. Besonders wenn verlässliche Aussagen zu kleinen Würfelementen gemacht werden sollen, muss die Stichprobe genügend gross sein, um zuverlässige Ergebnisse zu erzielen. Daher ist die Stichprobengrösse zentral. Im Rahmen der Konzeption einer neuen Befragung sollten daher für die Bestimmung der minimalen Stichprobengrösse u. a. Überlegungen dazu angestellt werden, wie stark differenziert werden soll und welches Mass an Schätzunsicherheit – bedingt durch den Stichprobenfehler – akzeptabel ist. Eine grosse Stichprobe geht jedoch auch mit höherem Aufwand und Kosten einher. Dies ist ein Nachteil dieser Methode.

Um das Problem der kleinen Inzidenzen anzugehen und potenziell mehr Differenzierungsmöglichkeiten zu haben, könnte der abgefragte Zeitraum erweitert werden (Unfälle nicht nur in den vergangenen 12 Monaten, sondern z. B. über die letzten 2 Jahre). Das Problem dabei ist allerdings, dass sich die Menschen tendenziell schlechter an den Unfall erinnern dürften, wenn er schon länger zurückliegt. Gerade bei leichten Verletzungen dürfte die Erinnerung schneller abnehmen als bei schweren Verletzungen. Deshalb ist ein Zeitraum von länger als 2 Jahren nicht empfehlenswert. Zudem beziehen sich die Zahlen nicht mehr auf ein Jahr, was dem Bedürfnis der Stakeholder entspricht, sondern würden einen 2-Jahres-Schnitt repräsentieren.

Umfang und Inhalt des Fragebogens

Obwohl man bei genügend grosser Stichprobe sehr detailliert befragen kann, sollte der Fragebogen nicht zu umfangreich werden, da sich dies negativ auf die Teilnahmequote auswirken kann [42]. Ausserdem sind gewisse Angaben möglicherweise schwierig oder heikel zu erfragen, insbesondere wenn es sich um Proxy-Interviews handelt, also wenn z. B. eine Person im Haushalt Auskunft zu einer dritten Person im Haushalt gibt.

Es gibt aber ein paar zwingend zu erhebende Daten, die dem Ziel eines möglichst kurzen Fragebogens entgegenlaufen:

- Für das Vorliegen eines Strassenverkehrsunfalls nach ASTRA-Definition müssen diverse – relativ komplex definierte – Kriterien erfüllt sein. Im Fragebogen oder Interview lediglich zu fragen, ob sich jemand bei einem «Strassenverkehrsunfall» verletzt hat, wird von den Befragten unterschiedlich interpretiert werden. Dies resultiert in nicht trennscharfen Filtern (s. auch Kapitel 5.3.1), was zu einer Verzerrung der Resultate führt. Stattdessen ist es wichtig, Fragen präzise zu formulieren, damit möglichst wenig Interpretationsspielraum vorhanden ist. Im Fall der Frage nach Strassenverkehrsunfällen bedeutet dies beispielsweise, dass mehrere Fragen gestellt werden müssen (z. B.

eine ergänzende Frage, ob sich der Unfall auf einem öffentlichen Verkehrsweg oder einer öffentlichen Verkehrsfläche ereignet hat).

- Weiter dürfte es schwierig sein, die Schwere der eigenen Verletzung oder insbesondere die einer angehörigen Person richtig einschätzen zu können. Auch hier reicht die Frage nach der Verletzungsschwere ohne zusätzliche Erklärung nicht.

Nicht zuletzt muss bedacht werden, dass bei gewissen Angaben auch die soziale Erwünschtheit oder möglicherweise eine (gewollte oder ungewollte) Fehlinterpretation der Situation eine Rolle spielen können, beispielsweise bei der Frage, wer den Unfall verursacht hat. Eine direkte Frage nach der/dem Unfallverursacher/-in dürfte wenig verlässlich sein und ist daher für eine Befragung nicht zu empfehlen.

Umgang mit «Nicht-Teilnahme»

Da die Teilnahme an Befragungen in der Regel freiwillig ist, stellt die Nicht-Teilnahme eine Herausforderung dar. Wenn sich Nicht-Teilnehmende systematisch von Teilnehmenden unterscheiden, kann dies ein Problem darstellen.

So ist es beispielsweise nicht ausgeschlossen, dass die Teilnahmebereitschaft davon abhängt, ob man einen Unfall erlebt hat oder nicht. Ist dies der Fall, liegt ein sogenannter «Selection Bias» vor. In Abb. 4 (S. 73) ist ersichtlich, dass eine solche Verzerrung einen beachtlichen Effekt auf den Fehlerbereich hat (unterster Balken): Dort wurde simuliert, dass alle Personen, die sich bei einem Unfall verletzt hatten, an der Befragung teilnehmen (100 % Teilnahmequote), jedoch «nur» 4 von 5 Personen, die keinen Unfall erlitten haben (80 % Teilnahmequote). Verglichen mit dem Fall ohne Selection Bias ist der entsprechende Fehler (RMSE) bei gleicher Stichprobengrösse mehr als doppelt so gross ($\pm 29\%$ vs. $\pm 12\%$). Ein solcher Fehler ist also möglichst zu vermeiden. In diesem Zusammenhang könnte es hilfreich sein, das Thema der Befragung (Unfälle) nicht von Anfang an preiszugeben. Wenn die Befragung keine anderen Themen abdeckt, ist dies allerdings nicht umsetzbar.

Anstrengungen zur Charakterisierung von nicht teilnehmenden Personen sind ebenfalls empfehlenswert. Solche Angaben sind z. B. im oben erwähnten Stichprobenrahmen des BFS vorhanden, was es ermöglicht, demografische und regionale Unterschiede zwischen teilnehmenden und nicht teilnehmenden Personen auszumachen. Wenn eine bestimmte Gruppe in der Stichprobe über- oder unterrepräsentiert ist, kann der Datensatz wo sinnvoll danach gewichtet werden.

Weiter sollten bei Befragungen Massnahmen zur Steigerung der Teilnahmequoten ergriffen werden: Solche können kurze und motivierende Anschreiben, mehrmalige Erinnerungsschreiben oder z. B. bei telefonischen Befragungen speziell durchgeführte Non-Responder-Befragungen sein, bei denen Personen, die die Teilnahme verweigerten, durch speziell geschultes Befragungspersonal erneut kontaktiert werden.

Neue Befragung vs. Integration in bestehende Befragung

Die Befragung kann entweder im Rahmen einer neuen, eigenständigen oder bereits bestehenden (periodischen) Stichprobenbefragung durchgeführt werden. Grosser Vorteil der Verwendung einer bestehenden periodischen Stichprobenbefragung sind die deutlich geringeren Kosten (Grenzkosten für zusätzliche Fragen). Andererseits ist man – sofern man die Stichprobe nicht erweitert, was wiederum mit Kosten verbunden wäre – an den Stichprobenrahmen gebunden. Dieser könnte in einer neuen Befragung für die vorliegende Fragestellung gezielter ausgewählt werden. In einer neuen Befragung ist man auch inhaltlich flexibler.

Als Grundlage für eine Berechnung der Dunkelziffer der Strassenverkehrsunfälle für das Jahr 2024 eignet sich die Haushaltsbefragung der BFU (für mehr Details s. auch Kapitel 6.2.1 auf Seite 52), die zum Zweck der Abschätzung des Gesamtunfallgeschehens (auch Arbeits-, Sport- und Freizeitunfälle) in der Schweizer Wohnbevölkerung durchgeführt wurde. Falls diese Haushaltsbefragung in Zukunft in ähnlicher/gleicher Form wiederholt wird, wird sie sich weiterhin als Datengrundlage eignen.

Die bestehenden periodischen Querschnittsbefragungen SGB [69] und MZMV [70,84] sind in der aktuellen Form für eine Abschätzung der Dunkelziffer nicht geeignet (s. Kapitel 6.2):

- In der SGB ist eine zuverlässige Identifizierung von Strassenverkehrsunfällen nach Definition des ASTRA nicht möglich (Problem Filter). Ebenso sind auch die Differenzierungsmöglichkeiten stark begrenzt (eine Auswertung nach Verkehrsteilnahme ist beispielsweise nicht möglich). Bei der SGB bräuchte es eine Optimierung und Erweiterung der bestehenden Frage nach Strassenverkehrsunfällen (und optimalerweise detailliertere Fragen z. B. zum benützten Fahrzeug).
- Der MZMV enthält bis einschliesslich 2025 keine Fragen zu Unfällen; entsprechende Fragen müssten ab 2030 integriert werden.

Von diesen beiden Optionen sollte die Variante MZMV weiterverfolgt werden. Erstens umfasst der MZMV eine grössere Stichprobe (MZMV: ca. 55 000 Befragte im Jahr 2021, SGB: ca. 22 000 Befragte im Jahr 2022), und zweitens schliesst er auch Kinder mit ein (MZMV ab 6 Jahren, SGB ab 15 Jahren). Konkret beinhaltete die MZMV-Stichprobe im Jahr 2021 etwas mehr als 7000 Kinder (6–16 Jahre). Bei einer Inzidenz von ungefähr jährlich 0,9 % in dieser Altersgruppe (berechnet aus den vorläufigen Daten der aktuellen BFU-Haushalts-erhebung für 6- bis 16-Jährige) wäre somit im MZMV mit ungefähr 60 bis 70 bei Strassenverkehrsunfällen verletzten Kindern zu rechnen. Eine Dunkelzifferberechnung für diese Altersgruppe sollte somit möglich sein. Die geringen Zahlen zeigen jedoch auch, dass die Differenzierungsmöglichkeiten innerhalb der Kinderunfälle beschränkt sind.

Um den MZMV für den vorliegenden Zweck optimal nutzen zu können, müsste er um einen «Unfallfragebogen» ergänzt werden, der durch eine Auskunftsperson im Haushalt beantwortet wird und idealerweise auch Kinder im Haushalt, die noch nicht 6 Jahre alt sind, miteinschliessen. Da davon auszugehen ist, dass sich bei einer durchschnittlichen Haushaltsgrosse von ungefähr 2 Personen [85] in ca. 98 % der Haushalte kein Strassenverkehrsunfall in den vergangenen 12 Monaten zugetragen hat, wäre für diesen Anteil der Befragten die zusätzliche Befragung zu den Unfällen nach einer Frage abgeschlossen. Für die 2 % der Haushalte, bei denen sich ein Strassenverkehrsunfall ereignet hat, müssten ein paar zusätzliche Fragen z. B. bzgl. verwendetem Fahrzeug, dem Unfallhergang und allenfalls der polizeilichen Erfassung gestellt werden. Die Differenzierungsmöglichkeiten hängen letztlich davon ab, wie viele zusätzliche Fragen gestellt werden.

Territorialitätsprinzip

Um die Unfälle von Personen, die nicht in der Schweiz wohnhaft sind (also nicht im SRPH enthalten sind), aber in der Schweiz verunfallt sind, mitzuberücksichtigen, können anhand des Verhältnisses von in der Schweiz wohnhaften Verunfallten und nicht in der Schweiz wohnhaften Verunfallten in den Daten des ASTRA, die auf dem Territorialitätsprinzip beruhen, Annahmen getroffen werden. Eine Befragung von Personen, die in der Schweiz verunfallt sind (und nicht in der ASTRA-Statistik gelandet sind), ist auf Grundlage der vorliegenden Stichprobenrahmen nicht durchführbar.

Periodizität der Befragung und Fortschreibung

Aufgrund des hohen Aufwands von Befragungen ist es nicht realistisch, eine solche jährlich durchzuführen. Eine «neue» Befragung findet aber idealerweise mindestens alle 5 bis 7 Jahre statt, sodass Trends (Entwicklung in den Mobilitätsformen und Expositionen) adäquat berücksichtigt werden können. Wenn eine bestehende Befragung verwendet wird, ist das Intervall der Befragung vorgegeben. Der MZMV wird in der Regel alle 5 Jahre durchgeführt.

Eine Fortschreibung für die Jahre zwischen den Befragungen kann dann auf Basis der Entwicklung im Hellfeld erfolgen. Für Erwachsene (17- bis 64-Jährige) können zu diesem Zweck zusätzlich die UVG-Daten der SSUV hinzugezogen werden (s. auch Kapitel 0).

7.2.2 Methode 2: Linkage bestehender Datenquellen inkl. Capture-Recapture

Eine Verknüpfung bestehender Datenquellen inklusive einer Hochrechnung mittels Capture-Recapture-Verfahren ist ein attraktives Verfahren, da keine Daten neu erhoben werden müssen und daher relativ geringe Kosten anfallen. Die Dunkelzifferberechnungen können daher mit dieser Methode im Vergleich zu Methode 1 ohne grossen Aufwand durchgeführt werden.

Geeignete Datenquellen-Kombinationen

Wie bereits in Kapitel 6.5 beschrieben, eignen sich die UVG-Daten der SSUV für eine Linkage mit den Daten des ASTRA (Hellfeld). Ebenfalls denkbar ist es, Daten einer Befragung mit dem Hellfeld zu verknüpfen. Da auch mehr als zwei Datensätze miteinander verknüpft werden können, kommen daher folgende Verknüpfungskombinationen infrage:

- ASTRA + SSUV
- ASTRA + Befragung
- ASTRA + SSUV + Befragung

Auswirkungen der Qualität von Filter und Linkage auf die Fehlergrösse: Simulation

Um die Genauigkeit dieser Verknüpfungsvarianten abzuschätzen, wurden ähnlich wie für verschiedene Stichprobengrössen bei Befragungen (s. Abb. 4, Kapitel 8.2.1) Simulationen durchgeführt. Dazu wurde wiederum jeweils der RSME für ein Würfelement der Grösse $n = 50\,000$ (z. B. leicht verletzte Erwachsene) berechnet. Für die Verknüpfung ASTRA-SSUV wurden zudem verschiedene Szenarien simuliert. Diese beziehen sich auf die Homogenität der Daten, die Güte der Verknüpfung und des Filters. Inhomogenität in einem Datensatz ergibt sich, wenn die Registrierungswahrscheinlichkeit eines Unfalls im betrachteten Würfelement nicht für alle verunfallten Personen gleich ist (Voraussetzung des Capture-Recapture-Schritts, s. auch Kapitel 5.3.1). Vom Prinzip her ist dies ähnlich wie unterschiedliche Teilnahmequoten innerhalb eines Würfelements in einer Befragung. Eine Inhomogenität besteht beispielsweise zwischen verschiedenen Verletzungsschweren: je schwerer verletzt, desto grösser ist die Wahrscheinlichkeit, dass die Polizei vor Ort ist und den Unfall registriert. Dies kann anhand der Daten der SSUV verdeutlicht werden, wo diesbezüglich keine grosse Inhomogenität zu erwarten ist. In Abb. 5 ist nach Verletzungsschwere dargestellt, ob die Polizei den Unfall registriert hat oder nicht. Bei 62 % der Schwerverletzten hat die Polizei den Unfall registriert (Ausprägungen «eindeutig ja» und «wahrscheinlich ja» kombiniert für die Variable «Polizeirapport vorhanden»). Bei Leichtverletzten liegt dieser Anteil lediglich bei 19 %.

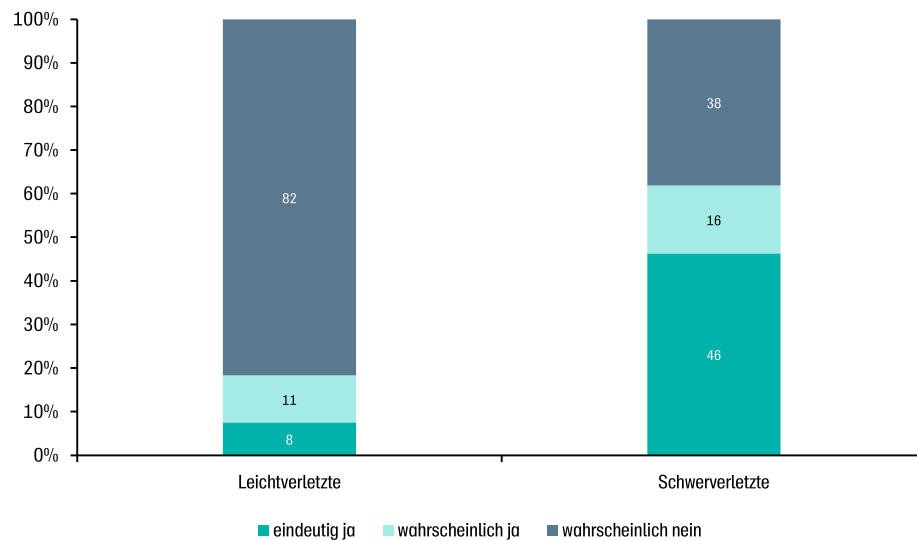


Abb. 5: Vorhandensein eines Polizeirapports nach Verletzungsschwere in den Unfalldaten der SSUV für das Unfalljahr 2017. Schwerverletzt (Ausprägungen «Mittelschwerverletzt», «Schwerverletzt» und «Invalidität») = Ausfall von mind. 1 Monat oder Invalidenrente).

Für die Simulation wurden folgende grundlegenden Annahmen getroffen: Die Polizei (Unfalldaten des ASTRA) erfasst Unfälle im Schnitt mit einer Wahrscheinlichkeit von 20 %, falls keine Inhomogenität angenommen wird. 80 % der Unfälle sind Arbeitsunfälle (UVG), 10 % der Unfälle sind andere Unfälle von Inländern/-innen (KVG), 10 % der Unfälle betreffen Personen mit Wohnsitz im Ausland. Von diesen drei Gruppen werden nur die Arbeitsunfälle (UVG) erfasst und davon ist eine 5%-Stichprobe (simple random sampling) verfügbar (Daten der SSUV). Die polizeiliche Erfassung erfolgt unabhängig vom Versicherungsstatus. Es muss betont werden, dass in Simulationen nur konkrete, plausible Situationen stellvertretend für die Realität beleuchtet werden können. Die Realität kann aber von diesen angenommenen Situationen abweichen. Das heisst, die Güte der Methoden kann in der Realität von der Güte abweichen, die in der Simulation ermittelt wurde.

Der RSME wurde für folgende Szenarien berechnet:

- Linkage ASTRA-SSUV
 - i. Keine Inhomogenität oder Fehler: sowohl die Unfalldaten des ASTRA als auch der SSUV sind homogen (gleiche Erfassungswahrscheinlichkeit für jede Person, s. auch Kapitel 5.3.1), die Verknüpfung verläuft ideal und der Filter kann in beiden Datensätzen korrekt gesetzt werden.
 - ii. Daten des ASTRA inhomogen: Es wurde angenommen, dass 80 % der Unfälle zu leichten Verletzungen führen und von der Polizei mit einer Wahrscheinlichkeit von 20 % erfasst werden. Bei den übrigen 20 % der Unfälle waren Schwerverletzte beteiligt; solche Unfälle werden von der Polizei mit 80% Wahrscheinlichkeit erfasst.
 - iii. Die ASTRA- und SSUV-Daten sind inhomogen: Zusätzlich zu der inhomogenen Erfassungswahrscheinlichkeit der Polizei (s. Szenario 1.ii) wird für die Daten der SSUV eine inhomogene Erfassungswahrscheinlichkeit gemäss Unfallschwere auf eine ähnliche Weise simuliert.
 - iv. Linkage-Fehler: 10 % der gematchten Fälle werden in ungematchte Fälle umgewandelt
 - v. Filter-Fehler: Zur Polizeiliste werden fälschlicherweise 10 % weitere Fälle hinzugenommen, die nicht aus der Unfallpopulation stammen
 - vi. Linkage- und Filter-Fehler: Kombination der Fehler iv. und v.

- vii. Linkage- und Filter-Fehler, Daten des ASTRA inhomogen: Kombination der Fehler ii., iv. und v.
- Linkage ASTRA-Befragung: unter der Annahme, dass die beiden Datensätze homogen sind und keine (groben) Linkage- oder Filter-Fehler passieren; 10 000 Befragungsteilnehmende mit einer Rücklaufquote von 100 % und ohne Recall Bias.
- Linkage ASTRA-SSUV-Befragung: unter der Annahme, dass alle Datensätze homogen sind und keine (groben) Linkage- oder Filter-Fehler passieren; 5000 bzw. 10 000 Befragungsteilnehmende mit einer Rücklaufquote von 100 % und ohne Recall Bias.

Abb. 6 zeigt, dass die Fehler je nach betrachtetem Szenario unterschiedlich gross ausfallen. Besonders wenn sowohl die Daten des ASTRA als auch der SSUV inhomogen sind (Szenario 1.iii), führt dies zu einem grossen Fehlerbereich. Aber auch Linkage- und Filter-Fehler, insbesondere in Kombination, können beträchtliche Auswirkungen auf die Genauigkeit haben. Bemerkenswert ist, dass der Fehler relativ klein bleibt, wenn nur in den Daten des ASTRA (nicht aber in denjenigen der SSUV) eine Inhomogenität vorliegt (Szenario 1.ii).

Der Fehlerbereich für Szenario 1.vii (inhomogene Daten des ASTRA, Linkage- und Filter-Fehler) ist kleiner als für Szenario 1.vi («nur» ein Linkage- und Filter-Fehler), was nicht intuitiv ist. Dies ist damit zu erklären, dass manche Fehler zu einer Überschätzung, andere zu Unterschätzung führen können, was sich in diesem Fall für Szenario 1.vii gegenseitig aufhob. Dies muss im Allgemeinen aber nicht so sein.

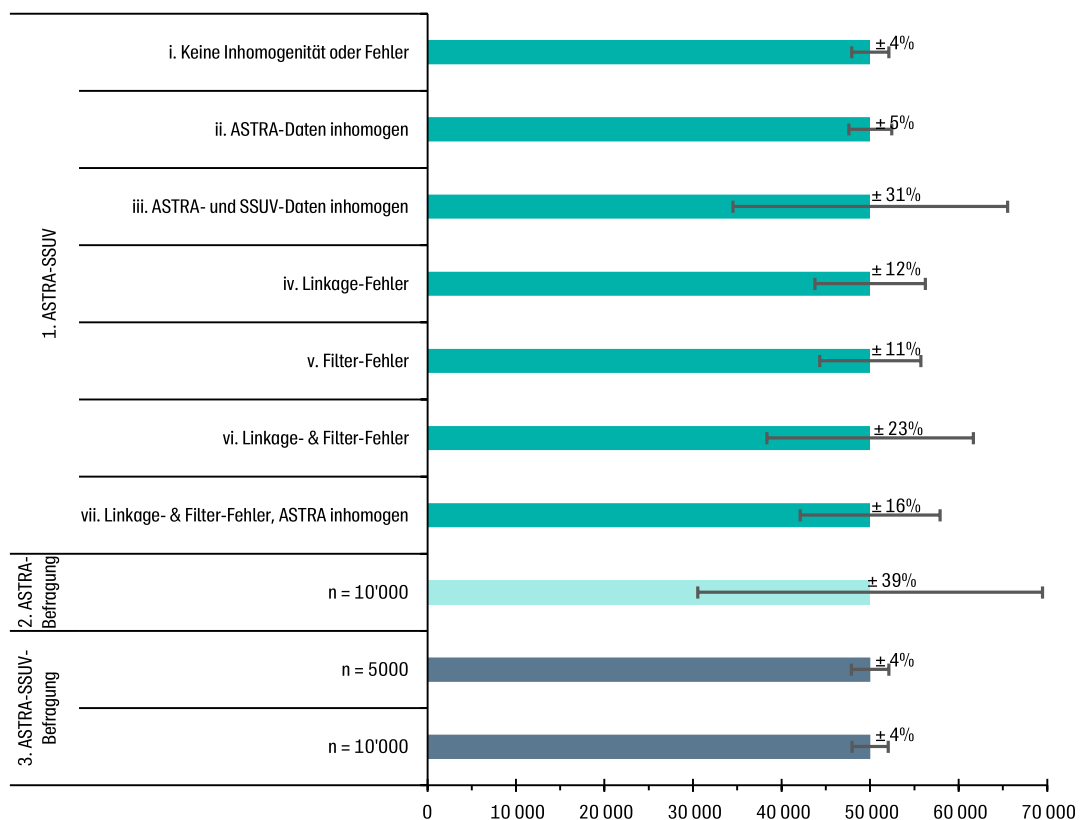


Abb. 6: Simulation des Fehlerbereichs (RMSE als Anteil des wahren Totals) für ein Würfelement der Grösse $n = 50\,000$ für unterschiedliche Szenarien.

Die Simulation der Linkage ASTRA-SSUV (Szenario 1) zeigt auf, wie wichtig korrekte Filter und funktionierende Linkage sind. Inhomogenität in einem Datensatz hat jedoch wenig Einfluss auf die Genauigkeit. Insgesamt kann davon ausgegangen werden, dass bei den meisten Würfelementen eine bestimmte Inhomogenität in den Daten des ASTRA zu finden ist

und auch die Filter und die Linkage in beiden Datensätzen nicht perfekt sind. Die Daten der SSUV dürften dagegen einigermaßen homogen sein. Somit entspricht insbesondere Szenario 1.vii wohl für viele Würfelemente aktuell dem realistischsten Szenario. Wenn in Zukunft die technischen Probleme von Filter und Linkage gelöst werden können, wäre Szenario ii realistisch.

Eine Linkage der ASTRA- mit Befragungsdaten ist mit sehr grossen Fehlern verbunden (Abb. 6, Szenario 2). Dies liegt daran, dass aufgrund der kleinen Inzidenz von Strassenverkehrsunfällen nur wenige Unfälle in der Schnittmenge landen. In der Simulation ist zudem nicht berücksichtigt, dass die Linkage «ASTRA-Befragung» wohl immer mit einer Unsicherheit verbunden wäre: Diese würde realistischerweise nicht über eine eindeutige ID, die in beiden Datensätzen vorhanden ist, passieren, sondern über eine Angabe im Fragebogen, ob die Polizei vor Ort war und den Unfall registriert hat. Dies dürfte nicht allen Verunfallten bekannt sein. In der Haushaltsbefragung 2024 der BFU gaben auf die Fragen, ob die Polizei über den Unfall informiert wurde und ob sie den Unfall aufgenommen oder protokolliert hat, 4 % der Verunfallten «Weiss nicht» an (vorläufige Resultate). Diese Unsicherheit dürfte den Fehlerbereich noch weiter vergrössern.

Die Abbildung zeigt ebenfalls, dass eine Linkage der drei Datenquellen ASTRA-SSUV-Befragung punkto Genauigkeit keine merkliche Verbesserung zur Linkage ASTRA-SSUV (inkl. Capture-Recapture) bringt. Dies rechtfertigt den hohen Aufwand zur Durchführung der Befragung nicht. Die beiden Linkage-Optionen ASTRA-Befragung und ASTRA-SSUV-Befragung werden deshalb nicht mehr weiterverfolgt.

Im Folgenden ist das Vorgehen zur Berechnung der Dunkelziffer bei einer Verknüpfung der Daten des ASTRA und der SSUV beschrieben. Es besteht aus den drei Schritten «Filter», «Linkage» und «Capture-Recapture» (s. auch Kapitel 5.3).

Filter

Eine Grundvoraussetzung bei diesem Ansatz ist, dass die gleichen Filter (Grundgesamtheit, für die die Dunkelziffer berechnet werden soll («Würfelemente»)) in beiden Datensätzen gesetzt werden können (s. auch Kapitel 5.3.1).

Dass ein fehlerhafter Filter grosse Auswirkungen auf die Genauigkeit der Resultate hat, ist in Abb. 6 ersichtlich. Möchte man die Dunkelziffer für alle Strassenverkehrsunfälle berechnen, sollte der Filter «Strassenverkehrsunfall» korrekt in beiden Datensätzen gesetzt werden können. Da die Definition von Strassenverkehrsunfällen in beiden Datensätzen vergleichbar ist, können solche Unfälle somit einigermaßen zuverlässig identifiziert werden.

Weiter sind in beiden Datensätzen verschiedene weitere ähnliche Variablen vorhanden, wonach differenziert werden kann. Es kann allerdings nicht stärker differenziert werden als in der Datenquelle mit der weniger starken Differenzierung. Dies bedeutet beispielsweise für den vorliegenden Fall, dass eine Unterscheidung zwischen Velo- und E-Bike-Unfällen nicht möglich ist, da diese heute in der UVG-Statistik nicht zuverlässig unterschieden werden können. Eine Dunkelziffer kann daher nur für beide Verkehrsteilnehmergruppen kombiniert berechnet werden.

Ein weiterer Nachteil dieser Methode ist, dass in den UVG-Daten keine Informationen zu Kindern und Senioren/-innen und erwachsenen Personen, die zum Zeitpunkt des Unfalls keine obligatorische Unfallversicherung (UVG) hatten (s. Kapitel 6.3.1), vorhanden sind. Für Kinder und Senioren/-innen können dadurch mit dieser Methode keine sinnvollen Abschätzungen vorgenommen werden. Für (alle) Erwachsenen (17- bis 64-Jährige) und Unterkategorien von Erwachsenen (z. B. erwachsene Velo-/E-Bike-Fahrende) können aber über Homogenitätsannahmen Aussagen gemacht werden.

Die möglichen Filter/Würfelemente, die mit dieser Methode abgedeckt (grün) bzw. nicht abgedeckt (grau, gestrichelte Umrandung) werden können, sind in Abb. 7 dargestellt.

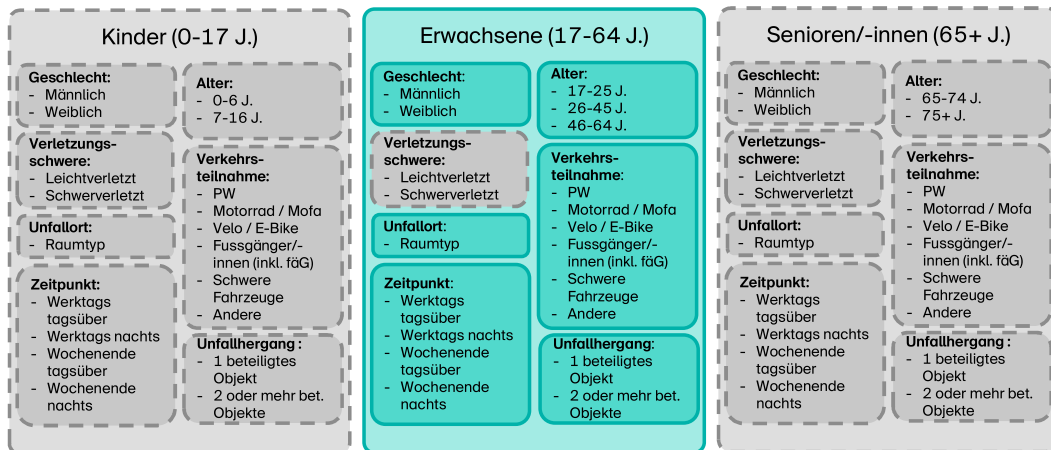


Abb. 7: «Würfelemente», die mit Methode 2 abgedeckt (grün) werden können. In grau (und gestrichelter Umrandung) sind die Elemente, die mit dieser Methode nicht abgedeckt werden können.

Herausforderung Verletzungsschwere: Um nach Verletzungsschwere differenzieren zu können, müssen in beiden Datensätzen Leicht- und Schwerverletzte anhand einer übereinstimmenden Definition differenziert werden können. Wie in Kapitel 6.1 beschrieben, wird in den Daten des ASTRA die Verletzungsschwere einerseits durch die Polizei in drei Kategorien («leicht verletzt», «erheblich verletzt» und «lebensbedrohlich verletzt») erfasst und kann optional mit dem NACA-Score des Rettungsdienstes vor Ort ergänzt werden.

Bei den UVG-Daten ist die Verletzungsschwere über die Anzahl der durch die Versicherung entschädigten Tage bei Arbeitsunfähigkeit definiert. Als schwer verletzt gelten Arbeitnehmende, die zumindest drei Monate Taggelder beziehen. Es ist anzunehmen, dass die durch die Polizei eingeschätzte Verletzungsschwere oder der NACA-Score mit der Anzahl entschädigter Tage korrelieren wird, die Konkordanz aber sehr gering ist. In den UVG-Daten werden aber auch Verletzungsdiagnosen nach ICD-10 kodiert. Die ICD-10 Codes lassen sich wiederum mittels Übersetzungstabellen in den AIS übersetzen. Entsprechende Tabellen werden von den Entwicklern des AIS kostenpflichtig angeboten [86,87]. Im Forschungsprojekt Vespa wurden solche Verfahren zur ICD10-AIS-Konvertierung mit den MSK-Daten bereits durchgeführt [88].

Die Angabe eines NACA-Codes im Unfallerfassungsprotokoll 2018 ist nicht obligatorisch und wird längst nicht für jede verletzte Person kodiert: Im Jahr der Einführung 2015 wurde für 93 % der Verletzten kein NACA-Code erfasst. Bis zum Unfalljahr 2023 sank der Anteil nicht kodierter Fälle auf 64 %, bleibt damit aber immer noch hoch. Im Unfalljahr 2023 wurde der NACA von 18 der 21 Datenlieferanten bei keinem Verletzten kodiert, von zwei Lieferanten für alle Fälle. Teilweise weicht der NACA-Code von der Zuordnung der durch die Polizei eingeschätzten Verletzungsschwere ab. Der NACA-Code gibt dann eine höhere Verletzungsschwere an als die Polizeieinschätzung oder umgekehrt. Dies trifft aber nur auf rund 1 % der in 2023 NACA-kodierten Fälle zu.

Die Einführung der differenzierten Kategorien für die Schwerverletzten und die Aufnahme der NACA-Codes in das UAP wurde durch das vom ASTRA beauftragte Projekt «SIS – Simplified Injury Scale» initiiert [89]. Neben einer differenzierteren Kodierung der Schwerverletzten und der Aufnahme des NACA-Scores sollte das Projekt in Feldversuchen klären, ob der AIS direkt durch die Polizei kodiert werden kann.

Ein Vergleich von stationären Fällen in Spitälern und Daten von Rettungsdiensten ergab, dass ein NACA-Code von 4+ am ehesten einem MAIS-Code von 3+ entspricht [90].

Tab. 10: Zuordnung der Verletzungsschwere zu den NACA-Codes laut UAP und entsprechende MAIS-Werte

Einschätzung Polizei (UAP)	NACA (UAP)	MAIS
Leicht verletzt	1	<3
	2	
Erheblich verletzt	3	3+
	4	
Lebensbedrohlich verletzt	5	3+
	6	

Vor dem Hintergrund, dass sich international durchgesetzt hat, dass eine «Schwere Verletzung» über einen MAIS-Wert von drei oder höher (MAIS3+) definiert wird [23], müsste für eine trennscharfe Einordnung der Verletzungsschweren in den ASTRA- und UVG-Daten der NACA-Code flächendeckend über das UAP erfasst werden. Eine Approximation über die von der Polizei beurteilte Verletzungsschwere ist nicht möglich, da der Übergang von NACA 3 auf NACA 4 in die Kategorie «Erheblich verletzt» fällt und nicht differenziert werden kann.

Verknüpfung der Datensätze / Linkage

Ist der Filter gesetzt, werden die beiden Datensätze verknüpft (Linkage). Dabei geht es darum, die Mengengrössen in den einzelnen Datenquellen sowie die Schnittmenge zu ermitteln. Bei zwei Datensätzen, wie im vorliegenden Fall, müssen somit die drei Mengengrössen n_{10} (nur Daten des ASTRA), n_{01} (nur Daten der SSUV) und n_{11} (Schnittmenge ASTRA-SSUV) bestimmt werden (s. Abb. 8).

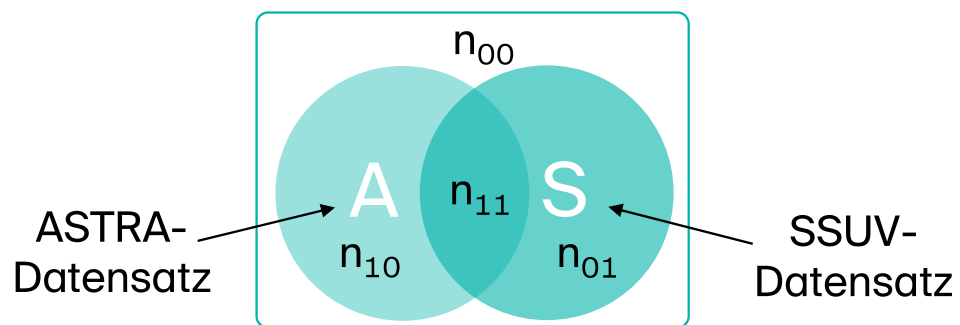


Abb. 8: Schematische Darstellung des Linkage- (inkl. Capture-Recapture) Ansatzes mit zwei Datenquellen (ASTRA und SSUV).

n_{10} : Anzahl Verletzte, die in den Unfalldaten des ASTRA, aber nicht der SSUV vorkommen; n_{01} : Anzahl Verletzte, die in den Daten der SSUV aber nicht in den Daten des ASTRA vorkommen; n_{11} : Anzahl Verletzte, die in beiden Datensätzen vorkommen; n_{00} : Anzahl Verletzte, die in keinem der beiden Datensatz vorkommen. Adaptiert von WHO, 2022, S. 19 [24].

Eine möglichst genaue Bestimmung der Mengengrössen n_{10} , n_{01} und n_{11} ist bei diesem Schritt zentral, da wie auch beim Filter bereits kleine Linkage-Fehler beachtliche Auswirkungen auf die Genauigkeit der Dunkelzifferschätzung haben können (Abb. 6, Szenario 1.iv, vi und vii). Bis anhin war es möglich, die Mengengrössen n_{01} und n_{11} über die in den Unfalldaten der SSUV vorhandenen Angabe zu bestimmen, ob ein Polizeirapport vorhanden ist oder nicht und die Mengengrösse n_{10} aus den Daten des ASTRA. Allerdings ist die Angabe über das Vorhandensein des Polizeirapports im SSUV-Datensatz dadurch, dass

sie die Ausprägungen «wahrscheinlich ja» und «wahrscheinlich nein» (und vor 2014 zusätzlich «ungewiss, unbekannt, keine Anhaltspunkte») annehmen kann, mit einer Unsicherheit behaftet. Zudem wird diese Variable seit 2022 nicht mehr erhoben. Für zukünftige Abschätzungen der Dunkelziffer müssen somit andere Linkage-Methoden zum Einsatz kommen, auf die im Folgenden eingegangen wird.

Idealerweise werden die Beobachtungen in beiden Datensätzen 1:1 deterministisch verknüpft, d. h. über (eine) Angabe(n) in beiden Datensätzen, mit denen ein Unfall eindeutig identifiziert werden kann. Eine eindeutige ID in beiden Datensätzen wäre optimal, existiert aber nicht. Eine weitere elegante Lösung wäre es, die AHV-Nummer der verletzten Person mit dem Unfalldatum zu kombinieren. Die AHV-Nummer ist aber in keinem der beiden Datensätze vorhanden. Während man beim ASTRA daran ist, diese Angabe im Datensatz zu integrieren, ist dies bei der SSUV nicht absehbar. Realistischer ist es also, die Verknüpfung über mehrere Variablen wie Unfalldatum/-zeit, Geburtsdatum, Geschlecht, Region etc. vorzunehmen (probabilistische Verknüpfung).

Erste Abklärungen mit der SSUV und dem ASTRA haben ergeben, dass eine probabilistische Verknüpfung der Datensätze aus rechtlicher Sicht möglich ist. Grundlage dafür liefert Art. 13.2b der Verordnung über das Informationssystem Strassenverkehrsunfälle (ISUV). Die Verknüpfung könnte beispielsweise über das ASTRA erfolgen. Dazu stellt das ASTRA bei der SSUV eine offizielle Anfrage, in der u. a. beschrieben wird, welche Merkmale im SSUV-Datensatz ans ASTRA geliefert werden sollen. Der Aufwand der SSUV muss gemäss der Verordnung über die Statistiken der Unfallversicherungen (SR 431.835, Art. 12 Abs. 3) durch die beauftragenden Bundesorgane getragen werden. Der Aufwand wird von der SSUV auf zwei oder drei Arbeitstage für die Initialisierung und ein oder zwei Arbeitstage pro Update (z. B. jährlich) geschätzt.

Falls sich das Teilen der Daten dieser beiden Anbieter dennoch als Problem herausstellt – z. B., weil die Koordinationsgruppe für die Statistik der Unfallversicherung (KSUV) einer Lieferung der Daten an das ASTRA nicht zustimmt –, könnten allenfalls für die Bestimmung der Schnittmengen der beiden Datensätze die bereits in Kapitel 5.3.3 erwähnten modernen technischen (kryptografischen) Methoden zur Anwendung kommen, bei denen solche Berechnungen möglich sind, ohne dass Daten geteilt werden, z. B. die «Private set intersection»-Methode [62,91]. Diese ermöglicht es zwei Parteien, die gemeinsamen Elemente ihrer jeweiligen Datensätze zu finden, ohne weitere Informationen über andere Elemente in den Datensätzen offenzulegen. Das heisst, jede Partei erfährt nur die Schnittmenge ihrer Datensätze und nichts über die Elemente, die ausschliesslich im Datensatz der anderen Partei enthalten sind. Zudem wäre es denkbar, dass nur eine vertrauenswürdige Instanz volle Einsicht in die Daten erhält und zur weiteren Berechnung lediglich die Informationen bzgl. Schnittmengen weiterleitet bzw. eine Anonymisierung vornimmt.

Capture-Recapture

Falls die Schritte «Filter» und «Linkage» zuverlässig durchgeführt werden können, ist mit geringem Aufwand ein Überblick über die tatsächliche Anzahl Unfälle möglich. Dieser Überblick kann unter gewissen Annahmen mit der Capture-Recapture-Methode verfeinert werden. Damit kann anhand der bekannten Mengengrössen in den beiden Datensätzen (n_{01} und n_{10}) sowie in der Überschneidung (n_{11}) die Anzahl geschätzt werden, die in keinem der beiden Datensätze zu finden ist (n_{00}) (Abb. 8 auf Seite 81). Letztlich ergibt die Summe der vier Mengenwerte n_{00} , n_{01} , n_{10} sowie n_{11} das Gesamtausmass der im betrachteten Würfelement verletzten Personen.

Die Annahmen 1–3 der Capture-Recapture-Methode, also eine geschlossene Grundgesamtheit, die Möglichkeit der Zuordnung der Individuen zwischen den Listen und die Unabhängigkeit der Individuen (s. Kapitel 5.3.1), dürften im vorliegenden Fall ausreichend erfüllt sein. Bezüglich Annahme 4 (Homogenität) kann wie weiter oben erwähnt davon ausgegangen werden, dass die Erfassungswahrscheinlichkeit je nach Verletzungsschwere unterschiedlich ist. Da aber von einer relativ homogenen Verteilung der Erfassungswahrscheinlichkeit in den Unfalldaten der SSUV ausgegangen wird, ist die Verzerrung minimal, selbst wenn die Daten des ASTRA heterogen sind [56]. Dies ist auch in Abb. 6 ersichtlich (Szenario 1.ii vs. 1.i). Zudem gibt es zahlreiche Erweiterungen der einfachen Capture-Re-

capture-Methodik (z. B. Berücksichtigung von Inhomogenitäten durch beobachtete Kovariablen wie z. B. Verletzungsschwere), auf die je nach konkreter Anwendung zurückgegriffen werden kann (z. B. Vector Generalized Linear Models / Vector Generalized Additive Models [58]). In diesem Fall reicht es dann allerdings nicht mehr, nur die entsprechenden Mengengrößen zu kennen (Abb. 8).

Bei kleinen Würfelementen (kleine Mengengrößen, kleine Überschneidungen) können die Unsicherheiten gross werden und teilweise die Vertrauensintervalle nicht bestimmt werden. In Abb. 9 wurde beispielhaft die Gesamtzahl inkl. Vertrauensintervalle für ein kleines Würfelement berechnet: Es handelt sich um die Zahlen aus dem Jahr 2021 für bei Kollisionen verletzten Frauen nach Verkehrsteilnahme. Die entsprechenden Mengengrößen sind in Tab. 11 zu finden. Es wird ersichtlich, dass die Vertrauensintervalle recht gross ausfallen. Für die Kategorie «Andere» konnte kein oberes Vertrauensintervall bestimmt werden.

Tab. 11: Beispiel für kleine Mengengrößen basierend auf realen Daten: Anzahl von bei Kollisionen verletzten erwachsenen Frauen (17–64 Jahre) nach Verkehrsteilnahme im Jahr 2021 in den Unfalldaten der SSUV und des ASTRA. Die Details zu den Simulationen finden sich in Anhang III.

Verkehrsteilnahme	Nur SSUV	Überschneidung SSUV + ASTRA	Nur ASTRA
Personenwagen	238	112	2770
Motorrad	19	7	412
Velo / E-Bike	75	18	767
Fussgängerin (inkl. fäG)	25	13	529
Andere	12	2	84

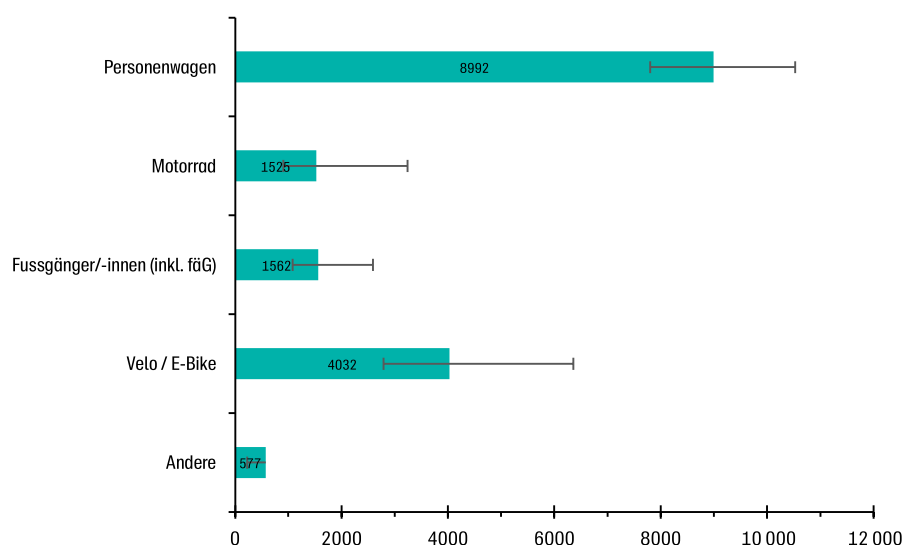


Abb. 9: Auf Basis der Zahlen in Tab. 11 mittels CR berechnete Gesamtzahl inkl. 95%-Vertrauensintervalle für bei Kollisionen verletzten, erwachsenen Frauen im Jahr 2021, nach Verkehrsteilnahme. Für die Kategorie «Andere» konnte aufgrund der kleinen Schnittmenge kein oberes Vertrauensintervall berechnet werden. Die Details zu den Simulationen finden sich in Anhang III.

Unfälle von erwachsenen Personen, die in der Schweiz verunfallen, aber nicht in der Schweiz UVG-versichert sind, sind in den Unfalldaten der SSUV nicht enthalten. Um diese mitzubetrachten, können anhand des Verhältnisses der Personen mit Wohnsitz Schweiz vs. Personen mit Wohnsitz Ausland in den Daten des ASTRA Annahmen getroffen werden.

Mögliche Verbesserungen

Die vorliegende Methode kann aktuell nicht allein für die Abschätzung der Dunkelziffer des Gesamtunfallgeschehens in der Schweiz eingesetzt werden, da dadurch nur Erwachsene abgedeckt werden können. Eine Dunkelziffer für Nicht-UVG-Versicherte – vorwiegend Kinder und Senioren – kann über diese Methode mit den Daten der SSUV nicht bestimmt werden. Nach unserer Einschätzung wäre auch über Analogie-Annahmen keine sinnvolle Schätzung möglich, da sich die Dunkelziffern dieser drei Gruppen systematisch unterscheiden dürften.

Damit die Methode allein für die Dunkelziffern aller Würfelemente infrage kommt, müssten daher zusätzlich zu den Daten der SSUV auch noch Daten der KVG-Versicherer erhoben und verfügbar gemacht werden. Dadurch könnten auch Kinder, Senioren und nicht UVG-versicherte Erwachsene (z. B. selbstständig Erwerbende) miteinbezogen werden, was die Verwendung zusätzlicher Methoden wie z. B. eine Befragung (Kapitel 7.2.1) überflüssig machen würde. Für mehr Details zur KVG-Erschliessung siehe auch Methode 3 (Kapitel 7.2.3).

Weiter würde eine Erweiterung der Stichprobe der UVG-Statistik (aktuell 5 %) die Resultate robuster machen (kleinere Vertrauensintervalle) und es könnten (sofern kodiert) auch die Dunkelziffern für kleinere Würfelemente, bei denen mit der aktuellen Stichprobe zu kleine Mengengrössen im Datensatz bzw. in der Überschneidung SSUV-ASTRA vorhanden sind, berechnet werden. Die SSUV strebt mittels «Machine Learning» (KI-unterstützte Textanalyse und Klassifizierung) eine solche Vergrösserung der Fallzahlbasis an, das (Fern-)Ziel wäre eine Vollerhebung (100 %). Aktuell ist allerdings noch nicht klar, ob und wann dies umgesetzt wird.

Insbesondere bei den Daten der SSUV wäre es auch sinnvoll, die Zuverlässigkeit der Angaben im Datensatz zu überprüfen oder gar Schritte einzuleiten, mit denen die Zuverlässigkeit der Angaben verbessert werden kann (z. B. durch eine stärkere Systematisierung des Unfallhergangs im Unfallerfassungsformular). Dadurch wäre in Zukunft eventuell auch eine genauere Unterscheidung der benützten Fahrzeuge möglich, insbesondere auch die von den Stakeholdern ausdrücklich gewünschte Unterscheidung zwischen Velo und E-Bike.

Pilotberechnung Methode 2 (Linkage ASTRA-SSUV)

Im Rahmen eines Pilots wurde mit Methode 2 (Linkage ASTRA-SSUV inkl. CR) das Gesamtausmass der Verletzten bei Strassenverkehrsunfällen für verschiedene Würfelemente berechnet und den Resultaten der Hochrechnung der BFU-Haushaltsbefragung aus dem Jahr 2010 gegenübergestellt.

Für die Pilotberechnung wurde ein Würfelement gewählt, das in allen Datensätzen (Hochrechnung, ASTRA, SSUV) anwendbar ist. Folgende Filter wurden angewendet:

- Jahr 2010
- Nichtberufsunfälle
- Erwachsene Personen (17- bis 64-Jährige), die in der Schweiz wohnhaft sind
- Alle Verletzungsschweren (ohne Todesfälle)

Die Dunkelziffern sind nach Verkehrsteilnahme (Fahrzeug im Monitoringprinzip) differenziert, wobei die maximale Differenzierungsmöglichkeit durch die UVG-Statistik vorgegeben war.

Abb. 10 vergleicht die Resultate der Berechnungen mit Methode 2 (blaugrüne Balken) mit der Hochrechnung der Haushaltsbefragung (graue Balken). Die Überschneidung n_{11} der beiden Datensätze ASTRA-SSUV wurde über die Variable «Polizeirapport vorhanden» in den Daten der SSUV und die Mengengrössen n_{10} und n_{01} über das jeweilige Datenset bestimmt. Da bei der Variable «Polizeirapport vorhanden» aber in 46 % der Fälle (im Jahr 2010) die Ausprägung «ungewiss, unbekannt, keine Anhaltspunkte» kodiert wurde,

ist diese Angabe mit einer relativ grossen Unsicherheit behaftet. Es wurden dabei zwei verschiedene Varianten gerechnet:

- Variante 1: Alle Beobachtungen mit dieser Ausprägung wurden als «nein» kodiert
- Variante 2: 90 % der Beobachtungen wurden als «nein», 10 % als «ja» kodiert.

Abbildung zeigt einerseits, dass insgesamt die Resultate der Haushaltsbefragung und Methode 2 (beide Varianten) in der Grössenordnung vergleichbar sind. Variante 1 von Methode 2 (dunkelgrüne Balken) weist erwartungsgemäss – da eine kleinere Überschneidung angenommen wird – für alle Würfelemente leicht höhere Werte aus als Variante 2 von Methode 2 (hellgrüne Balken). Dies demonstriert einmal mehr, dass die Güte der Linkage für die Qualität des Ergebnisses von Methode 2 zentral ist. Bei gewissen Verkehrsteilnehmergruppen (Personenwagen, Velo/E-Bike, Motorrad) liegen die Ergebnisse von Methode 2 (Variante 1 und 2) über den Ergebnissen der Haushaltsbefragung, bei Fussgänger/-innen (inkl. fäG) und der Kategorie «Andere» leicht darunter. Die Vertrauensintervalle überschneiden sich dabei. Die Hochrechnung der Haushaltsbefragung geht tendenziell mit der grössten Spannweite in den Vertrauensintervallen einher. Das Total von Methode 2 liegt (leicht) über dem Total aus der Hochrechnung der Haushaltsbefragung und wurde aus der Summe der einzelnen Verkehrsteilnehmergruppen berechnet (ohne Vertrauensintervalle).

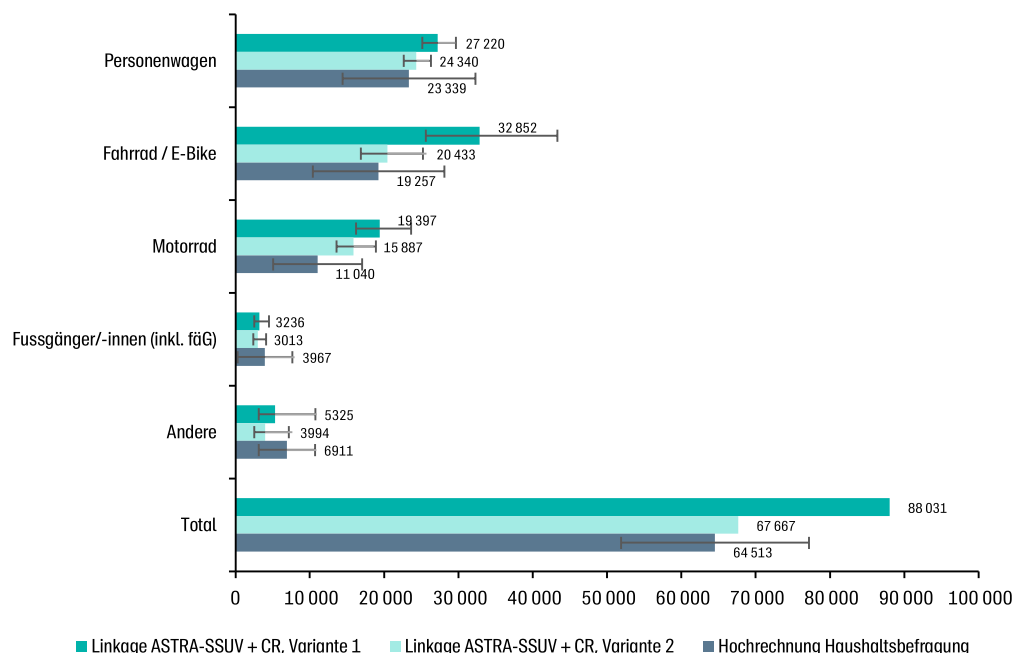


Abb. 10: Gesamtausmass inkl. 95%-Vertrauensintervalle der Verletzten bei nichtberufsbedingten Strassenverkehrsunfällen von Erwachsenen (Schweizer Wohnbevölkerung im Alter von 17–64 Jahren) in der Schweiz nach Verkehrsteilnahme für das Jahr 2010, berechnet mit Methode 2 (Linkage ASTRA-SSUV + CR) und mit der Hochrechnung der Haushaltsbefragung der BFU aus dem Jahr 2010. Variante 1: 100 % der Beobachtungen, bei denen die Variable «Polizeirapport vorhanden» die Ausprägung «ungewiss, unbekannt, keine Anhaltspunkte» annimmt, werden als «nein» kodiert. Variante 2: 90 % der Beobachtungen werden als «nein», 10 % als «ja» kodiert. Die Details zu den Simulationen (Methode 2) finden sich in Anhang III.

Tab. 12 zeigt die entsprechenden Zahlen (Gesamttotal) aus Abb. 10, die Zahlen im Hellfeld sowie Dunkelziffern für die beiden Varianten von Methode 2 und der Hochrechnung, aufgeteilt nach Fahrzeugart. Die Dunkelziffer liegt je nach Fahrzeugart und Variante zwischen einem Faktor 2,2 (Personenwagen) und knapp 15 (Velo/E-Bike). Wie zu erwarten, sind die Dunkelziffern für Variante 1 höher als für Variante 2. Die Dunkelziffern von Variante 2 und der Hochrechnung sind ähnlich hoch.

Tab. 12: Anzahl Fälle im Hellfeld, berechnete Gesamttotals sowie Dunkelziffern mit den beiden Varianten für Methode 2 und die Hochrechnung der BFU für das Jahr 2010

	Anzahl im Hellfeld	Linkage ASTRA-SSUV + CR, Variante 1		Linkage ASTRA-SSUV + CR, Variante 2		Hochrechnung BFU	
		Gesamttotal (95 %-VI)	Dunkelziffer (95 %-VI)	Gesamttotal (95 %-VI)	Dunkelziffer (95 %-VI)	Gesamttotal (95 %-VI)	Dunkelziffer (95 %-VI)
Personenwagen	10 414	27 220 (25 133–29 637)	2.6 (2.4–2.8)	24 340 (22 639–26 296)	2.3 (2.2–2.5)	23 339 (14 407–32 270)	2.2 (1.4–3.1)
Velo / E-Bike	2257	32 852 (25 618–43 298)	14.6 (11.4–19.2)	20 433 (16 867–25 196)	9.1 (7.5–11.2)	19 257 (10 403–28 111)	8.5 (4.6–12.5)
Motorrad	3728	19 397 (16 217–23 620)	5.2 (4.4–6.3)	15 887 (13 571–18 883)	4.3 (3.6–5.1)	11 040 (5060–17 020)	3 (1.4–4.6)
Fussgänger	1276	3236 (2499–4480)	2.5 (2–3.5)	3013 (2363–4092)	2.4 (1.9–3.2)	3967 (280–7653)	3.1 (0.2–6)
Andere	561	5325 (3102–10 729)	9.5 (5.5–19.1)	3994 (2517–7174)	7.1 (4.5–12.8)	6911 (3117–10 705)	12.3 (5.6–19.1)
Total	18 236	88 031	4.8	67 667	3.7	64 513 (51 887–77 140)	3.5 (2.8–4.2)

7.2.3 Methode 3: Versicherungsliste (SSUV und KVG)

Wie in Kapitel 0 beschrieben, werden die Unfälle von UVG-Versicherten in einer zufälligen Stichprobe von 5 % systematisch kodiert (UVG-Daten der KVG). Die Unfälle der bei den KVG-Versicherern versicherten Personen werden jedoch nicht systematisch erfasst. Würden diese – z. B. nach dem Vorbild der UVG-Statistik – registriert, so stünde nach Zusammenführen dieser beiden Datenquellen eine Versicherungsliste zur Verfügung, mit der die (bzw. eine zufällige Stichprobe der) Unfälle von allen in der Schweiz Unfallversicherten und wohnhaften Personen erfasst werden. Einzig bei den KVG-Versicherten kann bei leichten Verletzungen ein leichtes Under-Reporting erwartet werden, weil zu erwarten ist, dass einige Unfälle je nach Höhe der Franchise nicht gemeldet werden.

Eine Abklärung bei einigen KVG-Versicherungen ergab, dass Unfallmeldungen zwar mit ähnlichen Formularen wie bei den UVG-Versicherten erfasst werden, bei denen der Unfallhergang von den Versicherten als Freitext beschrieben wird. Jedoch unterscheidet sich der Umgang mit diesen Unfalldaten deutlich [78]. Teilweise werden die Formulare noch heute analog per Post abgewickelt und dann z. T. mit Informationsverlust digital erfasst. In anderen Fällen werden sie als Bildinformation gescannt, aber nicht weiter ausgewertet. Für eine systematische und standardisierte Erfassung sowie die Bereitstellung der Daten für die Präventionsarbeit müsste – wie auch für die UVG-Statistik – zuerst eine gesetzliche Grundlage geschaffen werden.

Wie auch bereits bei den ersten beiden vorgeschlagenen Methoden sind auch bei dieser Methode Personen, die nicht in der Schweiz unfallversichert sind, aber in der Schweiz verunfallen, nicht abgedeckt. Wiederum können auch bei dieser Methode über das Verhältnis der in der Schweiz wohnhaften vs. nicht in der Schweiz wohnhaften Personen aus den Daten des ASTRA Annahmen getroffen werden.

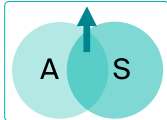
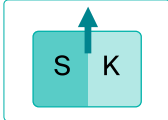
7.2.4 Übersicht / Beurteilung der Methoden zur Abschätzung der Dunkelziffer für Verletzte

Übereinstimmung mit Bedürfnissen der Stakeholder

In Tab. 13 werden die drei Methodenvarianten im Hinblick auf die Bedürfnisse der Stakeholder (Kapitel 4) beurteilt. Alle Methoden können grundsätzlich mindestens alle 5 bis 7 Jahre angewendet werden, und es ist eine Differenzierung nach Raumtyp (mit Fragezeichen bei der Qualität) möglich. Das Monitoringprinzip ist bei allen Methoden auswertbar.

Das Verursacherprinzip kann aufgrund fehlender Attributierung der Verursacher in bestehenden Datengrundlagen sowie Schwierigkeiten bei deren Erfassung in Befragungen mit keiner der Methoden umgesetzt werden.

Tab. 13: Beurteilung der Methodenvarianten nach den Bedürfnissen der Stakeholder (Stakeholder-Workshop I). Farbkodierung: grün: Anforderung erfüllt, gelb: Anforderung teils erfüllt, rot: Anforderung nicht erfüllt. Eine Erklärung zu den Piktogrammen findet sich in Kapitel 7.2 (ab S. 71) (Zusatzinformation zum mittleren Piktogramm: A = ASTRA-Daten, S = SSUV-Daten).

			
	Stichprobe (Befragung) + Hochrechnung	Linkage ASTRA-SSUV + CR	Versicherungsliste SSUV + KVG
Berechnung alle 5–7 Jahre	Ja	Ja	Ja
Zeitpunkt Erscheinung (12–18 Mt. nach ASTRA)	Ja	Aktuell nicht (Daten der SSUV 2 Jahre verzögert), in Zukunft voraussichtlich machbar	Unklar (hängt von KVG-Daten ab)
Territorialprinzip	Abschätzung mit Annahmen	Ja (Annahme Homogenität)	Abschätzung mit Annahmen
Differenzierung Raumtyp	Ja	Ja	Ja
Monitoringprinzip	Ja	Ja	Ja
Verursacherprinzip	Eher nein (Selbstausskunft wohl wenig verlässlich)	Nein	Nein
Differenzierungsmöglichkeiten / Filter	Theoretisch alles möglich. Limitiert durch Stichprobengrösse; Selbstausskunft	Variablen / Ausprägungen, die in beiden Datensets vorkommen; genügend grosse Gruppen	Variablen / Ausprägungen, die in beiden Datensets vorkommen

Bei den übrigen Kriterien weisen die drei Methoden spezifische Herausforderungen sowie Vor- und Nachteile auf. Auf diese wird im Folgenden eingegangen.

Werden die für Methode 2 und 3 aktuell unentbehrlichen UVG-Daten verwendet, so kann die Dunkelzifferberechnung nur mit einer gewissen Verzögerung erfolgen, da die Daten immer erst knapp zwei Jahre später vorliegen. Somit kann dem Bedürfnis, dass die Dunkelzifferberechnung max. 18 Monate nach Erscheinen der Unfalldaten des ASTRA vorliegen, aktuell nicht nachgekommen werden. Allerdings muss hierzu erwähnt werden, dass die SSUV daran arbeitet, diese Verzögerungen zu verringern. Methode 3 nutzt ebenfalls UVG-Daten, kann bezüglich Verzögerung jedoch erst beurteilt werden, wenn zusätzlich zu UVG-Daten auch die KVG-Daten erschlossen sind. Für Methode 1 bestehen verschiedene Möglichkeiten, die über die Befragung alle 5 bis 7 Jahre ermittelte Dunkelziffer jährlich fortzuschreiben. Die resultierende Verzögerung hängt dabei von den dafür verwendeten Datengrundlagen ab (z. B. Daten des ASTRA oder der SSUV, Daten zum Verkehrsaufkommen etc.).

In einer Befragung (Methode 1) sind theoretisch alle benötigten Differenzierungsmerkmale abfragbar. In der Praxis sind die Möglichkeiten jedoch durch eine maximale Länge des Fragebogens/Interviews beschränkt, weil eine zu detaillierte Abfrage negative Auswirkungen auf die Teilnahmequote haben kann. Ob die in der Befragung differenziert erhobenen Daten auch für eine Schätzung der jeweiligen Dunkelziffern mit akzeptablem Fehler verwendbar sind, hängt massgeblich von der Grösse der Stichprobe ab. Je tiefer die angestrebte Differenzierung und je kleiner die Würfelemente, desto grösser muss die Stichprobe gewählt werden. Dies wiederum hat Auswirkungen auf den Umsetzungsaufwand der Methode. Gewisse Angaben können sowohl in Befragungen als auch in vorhandenen Daten unzuverlässig sein, wenn sie auf Selbstausskünften der Unfallopfer basieren. Dies betrifft beispielsweise die Frage nach der/-m Verursacher/-in des Unfalls (Erinnerung, soziale

Erwünschtheit) und nach der Identifikation von Strassenverkehrsunfällen (komplexe Abgrenzung) sowie die Einschätzung der Verletzungsschwere (Erinnerung, komplexe Abgrenzung).

Die Differenzierungsmöglichkeiten sind bei Methode 2 und 3 dadurch eingeschränkt, dass die entsprechenden Variablen und die Ausprägungen in allen kombinierten Datensets vorkommen und vergleichbar sein müssen, damit die entsprechenden Filter gesetzt werden können. Bei Verwendung von UVG-Daten kann aktuell nicht zwischen Velo- und E-Bike-Unfällen unterschieden werden, wie es von einem Teil der Stakeholder explizit gewünscht wurde. Den detaillierten Differenzierungswünschen der Stakeholder (Tab. 3) kann am ehesten noch mit der Befragung (aber auch eingegrenzt bzgl. Fragebogenlänge und Selbstauskunft) nachgekommen werden.

Das Territorialitätsprinzip (alle Strassenverkehrsunfälle in der Schweiz) kann mit keiner der Methoden vollständig umgesetzt werden. Alle Methoden eignen sich zur Ermittlung der Dunkelziffer für in der Schweiz wohnhafte Personen. Die Ansätze haben jedoch Mühe, auch Unfälle von im Ausland wohnhaften Personen in der Schweiz zu erfassen. Für Methode 1 fehlt ein Stichprobenrahmen zur Abdeckung der im Ausland wohnhaften Personen, es können aber Abschätzungen aufgrund der Verteilung in den Daten des ASTRA gemacht werden. Bei Methode 2 fehlen in den Daten der SSUV Angaben zu den nicht UVG-versicherten Personen. Dies wiegt weniger schwer, da Grenzgänger meist UVG-versichert sind und die SVU-Statistik selbst im Territorialitätsprinzip erstellt wird. Über Homogenitätsannahmen ist eine Hochrechnung daher gut möglich. Bei Methode 3 kann analog Methode 1 eine Abschätzung anhand der Daten des ASTRA vorgenommen werden.

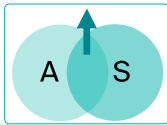
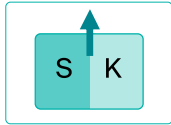
Weitere relevante Aspekte

In Tab. 14 werden weitere relevante Aspekte der Methoden zusammenfassend aufgeführt, die in den vorangehenden Kapiteln bereits aufgegriffen wurden.

Methoden 1 und 2 wären kurz- oder mittelfristig umsetzbar, Methode 3 jedoch nicht. Für Letztere muss zuerst eine KVG-Unfallstatistik erschlossen werden.

Die Identifikation von Strassenverkehrsunfällen ist bei Methode 1, 2 und – soweit in den noch nicht erschlossenen KVG-Daten gleich umgesetzt wie bei der UVG-Statistik – auch bei Methode 3 möglich. Bei Befragungen ist sie aber in der Qualität eingeschränkt durch die Selbstauskunft der verunfallten Person (bzw. über ein Familienmitglied bei einer Haushaltsbefragung). Oft unterliegen diese dabei bei weit zurückliegenden Ereignissen auch einem Recall Bias. Auch in den Daten der SSUV ist die geschulte Kodierungsstelle für die Abgrenzung auf eine korrekte und genügend ausführliche Selbstauskunft der verunfallten Person im Unfallmeldeformular angewiesen, wobei der Recall Bias aufgrund der Verschriftlichung im Protokoll kurz nach dem Unfall entschärft wird. Bei Methode 3 kann wie in Methode 2 von einer vorhandenen, aber in der Qualität schwierig überprüfbaren Abgrenzung ausgegangen werden. Es muss diesbezüglich bei allen Methoden von einer gewissen Unschärfe (Graubereich) in der Abgrenzung ausgegangen werden, die jedoch nicht quantifiziert werden kann.

Tab. 14: Beurteilung der Methodenvarianten nach weiteren, relevanten Aspekten. Eine Erklärung zu den Piktogrammen findet sich in Kapitel 7.2 (ab S. 71) (Zusatzinformation zum mittleren Piktogramm: A = ASTRA-Daten, S = SSUV-Daten).

			
	Stichprobe (Befragung) + Hochrechnung	Linkage ASTRA-SSUV + CR	Versicherungsliste SSUV + KVG
Kurz-/mittelfristig umsetzbar	Ja	Ja	Nein (Erschliessung KVG)
Identifikation Strassenverkehrsunfälle	(Ja) Selbstauskunft, Recall Bias	(Ja) (Graubereich)	(Ja) (Graubereich)
Ausweis Genauigkeit möglich	Ja	Ja	Ja
Genauigkeit bestimmt durch	- Stichprobengrösse - Kontrolle versch. Bias	Genauigkeit Linkage/ Filter, - Homogenität	
Beispiel Genauigkeit (RMSE)	Vgl. Abb. 4	Vgl. Abb. 6	-
Aufwand	++ (bestehend) bis ++++	+	+ bis ++++ (Initial)
Annahmen / Bias	Annahmen «überschaubar» Verschiedene Bias (Befragung)	Annahmen Capture-Recapture (CR)	

Die Genauigkeit der Resultate (95 %-Vertrauensintervalle) kann bei allen Methoden bestimmt werden. Bei Methode 1 ist sie von der Stichprobengrösse der Befragung abhängig sowie auch davon, wie gut verschiedene Bias kontrolliert werden können (s. dazu auch Abb. 4, Seite 73). Bei Methode 2 ist die Genauigkeit der Resultate von der Genauigkeit des Filters, der Linkage und von der Homogenität der Daten abhängig (veranschaulicht in Abb. 6 auf S. 78). Bei den Methoden 1 und 2 steigt die Genauigkeit (unter gleichbleibenden Bedingungen) mit der Grösse der geschätzten Würfelemente (höhere Inzidenz, grössere Überschneidung). Bei Methode 3 handelt es sich um eine repräsentative Erhebung sämtlicher in der Schweiz versicherten Unfallopfer, weshalb von einer hohen Genauigkeit ausgegangen werden kann.

Methode 1 (Befragung) geht mit dem höchsten Aufwand einher, insbesondere, wenn eine Befragung nur zum Zweck der Dunkelziffererhebung durchgeführt wird. Allerdings fallen die hohen Kosten nur alle 5 bis 7 Jahre an. Wird die Befragung in bestehende Erhebungsinstrumente integriert, kann der Aufwand auf die Grenzkosten reduziert werden. Bei Methode 2 ist der Aufwand einer Anwendung im Vergleich zu Methode 1 eher klein, wobei bei einer jährlichen Umsetzung der Aufwand jährlich anfällt. Mit zunehmender Automatisierung von Methode 2 (bei Linkage, Filter) kann der Aufwand voraussichtlich weiter gesenkt werden. Bei Methode 3 ist anfänglich mit einem bedeutenden – nicht bezifferbaren – Aufwand für die Erschliessung der KVG-Daten zu rechnen. Unter der Voraussetzung, dass diese analog zur UVG-Statistik erschlossen und auf jährlicher Basis gesammelt werden, wäre der Aufwand für die Dunkelzifferberechnung mit Methode 3 wiederum verhältnismässig gering.

8 Gesamtbeurteilung und Empfehlungen

8.1 Gesamtbeurteilung der Methoden zur Berechnung der Dunkelziffer für Verletzte

8.1.1 Bewertungshinweise der Stakeholder

Im Juli 2024 wurde ein zweiter Stakeholder-Workshop zu den beschriebenen Methodenvarianten durchgeführt. Es wurden die erarbeiteten Methodenvarianten vorgestellt, den am ersten Workshop genannten Bedürfnissen der Stakeholder gegenübergestellt und die vorhandenen Stellschrauben und Trade-Offs priorisiert.

Ein zentrales Ergebnis des Workshops war die Priorisierung der Bedürfnisse durch die Stakeholder. Dabei wurden die Trade-Offs zwischen einer starken Differenzierung der Dunkelziffer, einer hohen Qualität/Belastbarkeit der Zahlen und dem Erhebungsaufwand diskutiert:

- Fünf von sieben Teilnehmenden sprachen sich für eine gute Belastbarkeit der Zahlen aus, gewichteten dabei eine starke Differenzierung und einen möglichst geringen Aufwand weniger hoch.
- Zwei der Teilnehmenden sprachen sich für eine etwas stärkere Differenzierung bei gleichzeitig etwas geringerer Belastbarkeit aus, insbesondere eine Differenzierung zwischen E-Bike- und Velo-Unfällen wurde hoch priorisiert.

Wichtiger als eine (sehr) hohe Belastbarkeit der Zahlen ist dabei, dass die Grösse des Fehlers beziffert werden kann, mit dem Ziel der Abschätzung einer Bandbreite der Dunkelziffer. Die Stakeholder erachten es als wichtig, genügend Ressourcen bereitzustellen, den Aufwand gleichzeitig zu reduzieren oder auf verschiedene Interessierte zu verteilen. Dies durch eine pragmatische Erhebung/Bearbeitung – beispielsweise im Rahmen eines bestehenden Instruments – sowie durch Zusammenarbeit mehrerer an der Umsetzung beteiligter Stellen.

Die Aktennotiz des zweiten Stakeholder-Workshops ist in Anhang IV zu finden.

Im Folgenden wird eine Gesamtbeurteilung der verschiedenen Methoden zur Berechnung der Dunkelziffer für Verletzte mit Vor- und Nachteilen vorgenommen.

8.1.2 Methode 1: Befragung

- **Hauptvorteile:** Befragungen eignen sich am besten für die möglichst differenzierte Erfassung von Strassenverkehrsunfällen. Alle Altersgruppen können damit abgedeckt werden und es ist eine Differenzierung zwischen Velo und E-Bike möglich. Befragungen sind ein flexibel einsetzbares Instrument, das auf die spezifische Fragestellung adaptiert werden kann. Auch sind sie eine akzeptierte und bewährte Methode, und es konnten bereits umfassende Erfahrungen im Gebiet Strassenverkehrsunfälle gesammelt werden.
- **Hauptnachteile:** Vollständigkeit vom gewählten Stichprobenrahmen abhängig. Der Aufwand ist eher hoch, fällt aber auch nur alle 5 bis 7 Jahre an. Verfügbare Stichprobenrahmen enthalten keine im Ausland wohnhaften Personen (aber Abschätzung über die Daten des ASTRA möglich). Gewisse Angaben sind schwierig per Befragung zu erheben (z. B. Verletzungsschwere), Herausforderung Nicht-Teilnahme.
- **Zu beachten:** Belastbarkeit / Genauigkeit und Differenzierungsmöglichkeiten von Grösse des Samples (direkt beeinflussbar) und Bias (über Fragebogen kontrollierbar) abhängig. Bei der Auswertung der Dunkelziffern kann flexibel zwischen Genauigkeit und Differenzierung priorisiert werden. Für eine Berechnung der Dunkelziffer (Faktor) müssen die Befragungsdaten mit den Daten des ASTRA harmonisiert werden, insbesondere Abgrenzung Strassenverkehrsunfälle, Verletzungsschweren und Kategorien

der Verkehrsteilnahme. Für die jährliche Fortschreibung müssen weitere Datenquellen einbezogen werden (z. B. Daten des ASTRA).

- **Optimierungsmöglichkeiten:** Integration in bestehende Befragungen zur Senkung der Kosten. Allenfalls kann durch geeignete Sampling-Methoden der Standardfehler der Schätzungen reduziert werden.

8.1.3 Methode 2: Linkage ASTRA-SSUV + CR

- **Hauptvorteile:** Eher günstige, jährlich umsetzbare Methode. Die entsprechenden Daten werden ohnehin erhoben.
- **Hauptnachteile:** Keine Dunkelziffern für Kinder und Senioren mit den z. Zt. verfügbaren Datenquellen. Verzögerung von rund zwei Jahren (Daten der SSUV). Harmonisierung mit den Daten des ASTRA notwendig, insbesondere Abgrenzung Strassenverkehrsunfälle und Kategorien der Verkehrsteilnahme. Aktuell keine Differenzierung nach Verletzungsschwere möglich. Differenzierungsmöglichkeiten sind durch Kodierung in den Daten der SSUV beschränkt (z. B. keine Unterscheidung zwischen Velo und E-Bike möglich).
- **Zu beachten:** Belastbarkeit / Genauigkeit der Methode ist stark von der Qualität der Linkage, des Filters und von Grösse des Würfelements abhängig. Aktuell nicht vollständig harmonisierte Datensätze. Die Anzahl Personen mit Wohnsitz im Ausland, die in der Schweiz verunfallt sind, kann mittels Homogenitätsannahmen geschätzt werden. Pooling mehrerer Jahre möglich zur Erhöhung der Belastbarkeit.
- **Optimierungsmöglichkeiten:** Stichprobe der UVG-Statistik der SSUV erhöhen (bis zur Vollerhebung). Überprüfung der Qualität und Optimierung der Daten der SSUV (z. B. Anpassung Unfallformular). Harmonisierung Daten der SSUV und des ASTRA zur Steigerung der Linkage- und Filter-Qualität, insbesondere ist eine Harmonisierung bezüglich der Verletzungsschwere nötig. Eine Lösung wäre das obligatorische Erheben des NACA-Codes in den Daten des ASTRA.

8.1.4 Methode 3: Versicherungsliste

- **Hauptvorteile:** Relativ günstige Umsetzung der Methode, sobald die UVG- und KVG-Daten erschlossen und harmonisiert sind.
- **Hauptnachteile:** Heute nicht umsetzbar, weil Daten KVG fehlen. Diese müssen erschlossen werden. Personen, die nicht in der Schweiz unfallversichert sind, sind nicht abgedeckt, können aber anhand der Daten des ASTRA abgeschätzt werden. Die Differenzierungsmöglichkeiten sind durch die Kodierung und Harmonisierung der Datensätze beschränkt.
- **Zu beachten:** Für eine Berechnung der Dunkelziffer (Faktor) müssen die Versicherungsdaten mit den Daten des ASTRA harmonisiert werden, insbesondere Abgrenzung Strassenverkehrsunfälle, Verletzungsschweren und Kategorien der Verkehrsteilnahme. Die Belastbarkeit / Genauigkeit wird durch die Stichprobengrösse UVG/KVG sowie den Grad der Harmonisierung bestimmt.
- **Optimierungsmöglichkeiten:** Vergrösserung der Stichprobe UVG/KVG-Daten (bis zur Vollerhebung).

8.1.5 Weitere Datenquellen / Methoden für Verletzte

Im Projekt wurden verschiedene weitere mögliche Datenquellen für den Einsatz der Abschätzung der Dunkelziffer der Strassenverkehrsunfälle geprüft. Insbesondere in anderen Ländern wurden Spitaldaten hinzugezogen, um die Dunkelziffer zu schätzen, und mit dem Helffeld abgeglichen bzw. verknüpft. Aktuell kann dies für die Schweiz nicht empfohlen werden. Hauptproblem bei den verschiedenen vorhandenen Quellen (MSK, Spital ambulant, Anamnesetexte) ist, dass eine saubere Abgrenzung von Strassenverkehrsunfällen – eine Grundvoraussetzung für die Bestimmung der Dunkelziffer – und damit ein Abgleich mit dem Helffeld nicht möglich ist. Selbst wenn dies möglich wäre, wären die Differenzierungsmöglichkeiten stark eingeschränkt, weil nötige Zusatzinformationen fehlen (oder aufwendig über eine Befragung erhoben werden müssten): eine Differenzierung nach Fahrzeugart wäre bereits nicht möglich. So könnte man einem zentralen Bedürfnis der Stakeholder nicht nachkommen. Nicht zuletzt wären mehrheitlich nur schwere Verletzungen abgedeckt und Informationen zu Behandlungen in Arztpraxen fehlen gänzlich, wo aktuell keine strukturierten bzw. zentral verwalteten Daten vorliegen. Eine solche Erschliessung erscheint unverhältnismässig.

Weitere Datenquellen wie beispielsweise die Erfassung von Strassenverkehrsunfällen über das Smartphone, Fahrzeugtechnik oder Videokameras sind derzeit noch zu wenig verbreitet und zuverlässig, um sie für die Schätzung einer Dunkelziffer für Strassenverkehrsunfälle einsetzen zu können. Die Feststellung und Abgrenzung eines Unfalls reichen dabei noch nicht aus, um die Daten für eine differenzierte Dunkelziffer verwenden zu können. Notwendige Detailangaben zum Unfall, zu verletzten Personen oder zu den involvierten Fahrzeugen fehlen oder unterliegen einschränkenden Datenschutzbestimmungen. Zur erfolgreichen Verwendung solcher Daten in Modellen des maschinellen Lernens müssen möglichst vollständige und qualitativ hochwertige Daten zu Strassenverkehrsunfällen und deren Folgen vorhanden sein.

8.2 Empfehlungen

Im Folgenden werden konkrete Empfehlungen abgegeben, wie die Dunkelzifferberechnung für bei Strassenverkehrsunfällen verletzte oder getötete Personen mittelfristig umgesetzt werden kann. Dabei müssen insbesondere für die differenzierte Berechnung der Dunkelziffer für Verletzte (s. Kapitel 8.2.2) bereits kurzfristig Schritte eingeleitet werden, damit die Methode wie empfohlen umgesetzt werden kann.

Aufgrund der Komplexität der Herangehensweise beim Umsetzungsvorschlag der Dunkelzifferberechnung für Verletzte empfiehlt es sich, dass eine einzelne Organisation den Lead in der Koordination übernimmt und auf Basis der Resultate dieses Forschungsprojekts konkrete Strategieempfehlungen formuliert. Diese Rolle könnte insbesondere das ASTRA einnehmen. Das ASTRA verfügt über vertiefte Kompetenzen im Bereich der Strassenverkehrsunfälle und ist verantwortlich für die Erfassung und Verwaltung der Strassenverkehrsunfallstatistik (ASTRA VU). Die Umsetzung der Empfehlungen steht zudem in Harmonie mit den Zielen 2030 gemäss der Teilstrategie Verkehrssicherheit des ASTRA. Diese erwähnt u. a. die Schaffung eines neuen Standards zur Messung der Verkehrssicherheit (Kapitel 6.1 in der Teilstrategie Verkehrssicherheit des ASTRA [92]) und betont, wie wichtig der Transfer der Ergebnisse aus der Forschung in die Praxis ist (Kapitel 7.1 der Teilstrategie).

8.2.1 Dunkelziffer für Getötete

Angesichts der Unsicherheiten, der gemäss Literatur relativ kleinen Dunkelziffern, der relativ kleinen Abweichungen zwischen den Daten der BFU und dem ASTRA wird bezüglich Todesfällen empfohlen, von einer Dunkelziffer von 1,0 (Faktor) auszugehen. Die Zahl der Todesfälle bei Strassenverkehrsunfällen in der Schweiz wird somit direkt aus den Daten des ASTRA abgeleitet. Dies gilt auch für alle Differenzierungen innerhalb der Todesfälle, z. B. nach Verkehrsteilnahme oder Alter.

8.2.2 Dunkelziffer für Verletzte

Abb. 11 zeigt eine schematische Übersicht über die Empfehlungen im Bereich Dunkelziffer für Verletzte.

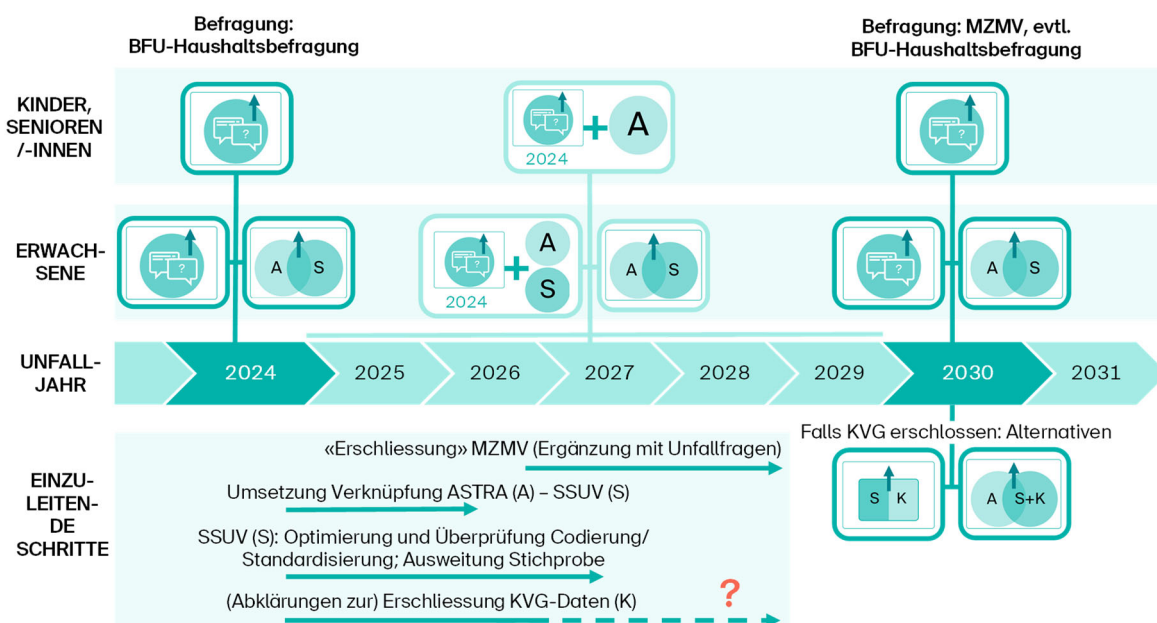


Abb. 11: Schematische Darstellung des vorgeschlagenen mittelfristigen Vorgehens zur Berechnung der Dunkelziffer für verletzte Personen bei Strassenverkehrsunfällen (ohne Todesfälle). Eine Erläuterung der verwendeten Piktogramme findet sich in Kapitel 7.2 (ab S. 71).

Befragung (Methode 1) als flexibel und gezielt einsetzbares Instrument

Kurz- und mittelfristig empfehlen wir, die Dunkelziffer für Strassenverkehrsunfälle mittels Befragung zu bestimmen. Befragungen sind eine breit akzeptierte Methode und haben sich in diesem Zusammenhang bereits bewährt. Grosser Vorteil gegenüber der Methode 2 «Linkage ASRA-SSUV inkl. CR» ist, dass alle Altersgruppen einbezogen werden können und dass eine Unterscheidung zwischen Velo und E-Bike möglich ist. Für die Dunkelzifferberechnung für das Jahr 2024 sollte die bereits durchgeführte Haushaltsbefragung der BFU für eine Dunkelzifferberechnung verwendet werden. Diese bietet eine saubere Abgrenzung von Strassenverkehrsunfällen sowie relevante Differenzierungsmöglichkeiten. Für die darauffolgenden Jahre empfiehlt es sich, eine Fortschreibung des Gesamtunfallgeschehens über die Entwicklung in den Daten des ASRA (und für Erwachsene allenfalls zusätzlich mit den Daten der SSUV) zu machen. Um jährliche Schwankungen in der Dunkelziffer der Erwachsenen auffangen zu können oder auf kurzfristige Trends einzugehen, kann gleichzeitig Methode 2 umgesetzt werden.

Danach sollte eine Befragung alle 5 bis maximal 7 Jahre durchgeführt werden. Wird die Haushaltsbefragung der BFU in diesem Intervall weitergeführt, eignet sie sich als Grundlage für die Dunkelzifferberechnung. Der bestehende Fragebogen ist nicht zu umfangreich, könnte aber für eine zukünftige Verwendung flexibel angepasst werden. Da aktuell aber noch unklar ist, ob und in welchem Intervall die BFU-Haushaltsbefragung weitergeführt wird, ist gleichzeitig zu empfehlen, bereits jetzt die notwendigen Schritte einzuleiten, den MZMV 2030 mit Fragen zu Unfällen zu ergänzen. Der MZMV eignet sich aufgrund der Stichprobengrösse und der Erhebungsform (Haushaltsbefragung) ideal für diesen Zweck, und weil er eine bestehende Befragung ist, ist das Vorgehen effizient. Die zusätzlichen Fragen zum Unfallgeschehen, in Kombination mit der Stichprobengrösse für das jeweilige Würfelement, werden die Differenzierungsmöglichkeiten bestimmen. Wird eine neue Befragung konzipiert, ist zu empfehlen, diese mittels einer Haushaltsbefragung zu gestalten, um auch Kinder zu erreichen, und für die Auswahl der Haushalte den SRPH-Stichprobenrahmen zu verwenden. Je nach gewünschter Differenzierung muss die Stichprobe grösser

gewählt werden und müssen allenfalls gewisse Gruppen überrepräsentiert werden. Es sollte darauf geachtet werden, die verschiedenen Bias (insb. Non-Responder) zu kontrollieren. Für die Berücksichtigung von Unfällen von Personen mit Wohnsitz im Ausland kann das entsprechende Verhältnis in den Daten des ASTRA herangezogen werden.

Linkage ASTRA-SSUV + Capture-Recapture (Methode 2) als Ergänzung und Kontrollinstrument

Gleichzeitig sollte als Ergänzung zur Befragung jährlich Methode 2 (Linkage ASTRA-SSUV inkl. CR) für das Würfelement Erwachsene und die möglichen Unterdifferenzierungen (s. Abb. 7) angewendet werden. Diese Methode kann relativ günstig jährlich umgesetzt werden und kann dazu dienen, die Resultate der Befragung zu validieren. Da die Variable «Polizeirapport vorhanden» in den Daten der SSUV seit 2022 nicht mehr existiert, muss in einem ersten Schritt die Verknüpfung dieser beiden Datenquellen angegangen werden. Zudem können dadurch bereits jetzt Erfahrungen im Linkage der beiden Datenquellen gesammelt werden. Liegen in Zukunft auch die Daten der KVG-Versicherer vor, würde diese Methode an Gewicht gewinnen und die wichtigen Schritte für eine erfolgreiche Umsetzung der Methode wären bereits eingeleitet.

Eine Harmonisierung der beiden Datenquellen ist ebenfalls prioritär zu behandeln, weil sich im Verlauf des Projekts herausgestellt hat, dass die Genauigkeit der Resultate stark von der Güte des Filters und der Linkage abhängt. Dies betrifft insbesondere die Verletzungsschwere, die in den beiden Datensets unterschiedlich definiert ist. Allgemein wird empfohlen, die Qualität der Angaben in den Daten der SSUV zu prüfen und allenfalls Schritte einzuleiten, um die Datenqualität zu erhöhen (z. B. Anpassung des Unfallformulars insbesondere bezüglich Unfallhergang).

Eine Ausweitung der Stichprobe der UVG-Statistik auf einen grösseren Anteil als die aktuellen 5 % oder gar alle Unfälle (100 %) wäre gewinnbringend und würde (falls in guter Qualität kodiert) auch die Analyse kleiner «Würfelemente» innerhalb der Altersgruppe der Erwachsenen erlauben.

Erschliessung KVG-Daten

Damit die Linkage-CR-Methode (Methode 2) strikt für alle Würfelemente angewendet werden kann, ist eine Erschliessung der Unfalldaten der KVG-Versicherer unumgänglich. Hier gilt es zeitnah zu klären, wie diese Daten erschlossen werden könnten.

Der Zugriff auf die KVG-Unfalldaten ist Grundvoraussetzung für eine Umsetzung von Methode 3 («Versicherungsliste»). Die Erschliessung dieser Daten wäre zwar aufwendig, danach ist aber mit einem tiefen Aufwand zur Auswertung der Versicherungslisten zu rechnen. Ähnlich wie bei Methodik 1 («Befragung») ist es allerdings auch hier notwendig, im Sinne des Territorialitätsprinzips eine Abschätzung der Unfälle von Personen mit Wohnsitz im Ausland mithilfe der Daten des ASTRA vorzunehmen.

Mögliche Differenzierungen

Die vorgeschlagene Methodik erlaubt eine Differenzierung nach Verkehrsteilnahme, Hergang (Alleinunfall vs. Kollision) und Verletzungsschwere (s. Ziele in Kapitel 1.2), wobei die Harmonisierung der Verletzungsschwere in einer Befragung mit den Daten des ASTRA eine Herausforderung darstellt und in den SSUV-Daten aktuell nicht möglich ist. Ebenfalls kann nach Alter und Geschlecht, Zeitfenster und Raumtyp differenziert werden. Die Differenzierung nach diesen Merkmalen entspricht dem Bedürfnis relevanter Stakeholder, ist kongruent mit Differenzierungen in anderen Ländern – was einen Vergleich ermöglicht – und es ist für die meisten Merkmale mit einem deutlichen Unterschied in der Erfassungswahrscheinlichkeit zwischen den Ausprägungen zu rechnen. Diese Differenzierungen sind daher bei der Umsetzung der Methodik zu empfehlen.

Je nach Fragestellung können (bzw. sollten) auch Differenzierungsmerkmale kombiniert werden, beispielsweise die Verletzungsschwere und Verkehrsteilnahme bei einer Abschätzung der Kosten von Strassenverkehrsunfällen. Dies ist mit der vorgeschlagenen Methodik möglich, vorausgesetzt, die Zahlen in den jeweiligen Würfelementen sind genügend

gross, sodass sie nicht mit einer zu grossen Schätzunsicherheit (95%-Vertrauensintervalle) einhergehen. Auch weitere Differenzierungen über die genannten Merkmale hinaus sind möglich. Voraussetzung ist auch hier, dass die Zahlen ausreichend gross sind. Zusätzlich ist wichtig, dass die Differenzierungsmerkmale in guter Qualität in der/den verwendeten Datenquelle(n) vorhanden sind, d. h., sie müssten in einer Befragung (Methode 1) zuverlässig erfragt werden können und bei Methode 2 oder 3 (Linkage ASTRA-SSUV + CR oder Versicherungsliste UVG-KVG) in allen verwendeten Quellen in guter Qualität vorhanden und vergleichbar sein.

9 Forschungsbedarf

Im Projektverlauf wurde in Zusammenhang mit der Quantifizierung der Dunkelziffer von Strassenverkehrsunfällen in der Schweiz weiterer Forschungsbedarf identifiziert:

Ein naheliegender nächster Schritt wäre die Durchführung eines Pilotprojekts zur Bestimmung der Dunkelziffer von Verletzten, in dem die praktische Umsetzung der empfohlenen methodischen Ansätze getestet wird. Ein solches Projekt könnte insbesondere die Verknüpfung der SSUV- und ASTRA-Daten in den Fokus rücken sowie die Erschliessung und Integration von KVG-Daten eingehend prüfen.

Für alle Methoden besteht Bedarf, Unfallopfer mit Wohnsitz im Ausland in die Erhebung der Unfalldaten einzubeziehen. Für keine der vorgeschlagenen Methoden können derzeit Personen einbezogen werden, die zwar in der Schweiz verunfallt sind, ihren Wohnsitz aber im Ausland haben und nicht in der Schweiz arbeiten. Es braucht methodische Ansätze, um ihre Berücksichtigung zukünftig zu ermöglichen. Bis dahin kann eine Abschätzung ihres Anteils in den ASTRA-Daten vorgenommen werden. Weiter müssen die datenschutzrechtlichen Rahmenbedingungen immer genau geprüft werden, um den Schutz der persönlichen Informationen zu gewährleisten und sicherzustellen, dass die Nutzung solcher Daten rechtskonform erfolgt, da es sich in diesem Thema oft um besonders schützenswerte Personendaten handelt (z. B. Verletzungsschwere). Im Folgenden wird auf den methodenspezifischen Forschungsbedarf eingegangen.

9.1 Methode 1: Befragung

Bei der Befragungsmethode handelt es sich um ein bewährtes und erprobtes Erhebungsinstrument. Mit dem SRPH steht für die Schweiz bereits ein hochwertiger Stichprobenrahmen zur Verfügung. Daher besteht bei dieser Methode wenig Forschungsbedarf.

Neben der bereits in Kapitel 9 erwähnten Herausforderung, Personen einzubeziehen, die ihren mit Wohnsitz im Ausland haben, aber in der Schweiz verunfallt sind, besteht ein weiteres Problem dieser Methode in der verlässlichen Erhebung der Verletzungsschwere. Hier ist zu prüfen, welche Fragen diesbezüglich die verlässlichsten Ergebnisse liefern können, damit eine ausreichende Qualität erreicht wird und die erhobenen Informationen mit der amtlichen Unfallstatistik kompatibel werden.

9.2 Methode 2: Linkage ASTRA-SSUV + CR

Bei dieser Methode besteht weiterer Forschungsbedarf in der Optimierung der Linkage-Prozesse. Während technologische Entwicklungen diesen Prozess fortlaufend verbessern (z. B. kryptografische Methoden wie die «Private set intersection»-Methode), bleibt die Herausforderung bestehen, die ASTRA- und SSUV-Einzeldaten trotz Anonymisierung konsistent zu verknüpfen und dabei den Datenschutz zu gewährleisten.

Ein besonderes Augenmerk sollte dabei auf die Harmonisierung der unterschiedlichen Klassifikationen zur Verletzungsschwere gelegt werden. Derzeit existieren verschiedene Kodierungssysteme wie die ASTRA-Verletzungsskala, der NACA-Score, (M)AIS oder ICD-10, die nicht ohne Weiteres miteinander vergleich- und verknüpfbar sind. Eine systematische Umrechnung oder Harmonisierung dieser Skalen könnte dazu beitragen, eine stärker belastbare Berechnung der Dunkelziffer nach Verletzungsfolgen vorzunehmen.

Ein weiteres relevantes Forschungsfeld betrifft die Qualität und Optimierung der SSUV-Daten. Eine effizientere elektronische Erfassung und Vereinheitlichung von Unfallberichten (z. B. durch eine grössere Anzahl geschlossener Fragen) würde dazu beitragen, die Datenqualität zu steigern. Dies könnte beispielsweise dadurch erreicht werden, dass Unfallmeldungen von Versicherungen standardisiert und direkt durch die meldende Person kodiert werden, insbesondere mit Blick auf die detaillierte Erfassung der beteiligten Fahrzeugkategorien. Eine Unterscheidung zwischen Velo- und E-Bike-Unfällen wird dadurch möglich. Parallel dazu wäre es sinnvoll, den NACA-Code als verpflichtendes Element (statt

optional wie heute) in die offizielle Erhebung des ASTRA zu integrieren. Dies würde die Kompatibilität der Verletzungsschweren zwischen den ASTRA- und SSUV-Daten erhöhen.

Um in der UVG-Statistik eine grössere Stichprobengrösse und damit eine verbesserte Aussagekraft zu erzielen, können zukünftig maschinengestützte Methoden zur Textanalyse und zur Klassifizierung (z. B. Sprachmodelle) eingesetzt werden. Dies wird zur Zeit von der Suva für ihre Versicherten umgesetzt. Möglichkeiten zur Ausweitung dieser Methodik auf die Daten aller UVG-Versicherer sollten identifiziert werden.

9.3 Methode 3: Versicherungsliste

Zusätzlich zu den in Kapitel 9.2 erwähnten Punkten ergibt sich hier insbesondere Forschungsbedarf im Hinblick auf die Erschliessung von Unfalldaten der Krankenversicherer. Dies würde die UVG-Statistik optimal ergänzen, indem auch nicht UVG-versicherte Personen erfasst werden. Eine zentrale Frage besteht darin, ob eine eigene KVG-Datensammelstelle, ähnlich der SSUV, eingerichtet werden könnte. Dies erfordert insbesondere umfassende rechtliche Abklärungen. Ohne eine gesetzliche Grundlage für die Schaffung einer solchen Datensammelstelle ist es unrealistisch, dass solche Daten in absehbarer Zeit vorliegen würden.

9.4 Weitere Datenquellen / Methoden

Über die hier vorgeschlagenen Methoden hinaus sollten kontinuierlich noch nicht erschlossene, aber theoretisch vorliegende sowie auch potenzielle Datenquellen identifiziert und für eine verbesserte Dunkelzifferberechnung evaluiert werden. Hierzu gehören insbesondere moderne Technologien wie Sensoren, Smartphone-Features und -Apps oder Videoanalysen, die eine automatisierte und detailliertere Erfassung von Unfällen ermöglichen. Fahrzeugdaten, insbesondere solche aus Unfalldatenschreibern (EDR), mit denen seit dem 1. April 2024 neue Fahrzeuge ausgerüstet sein müssen, bieten ebenfalls potenziell wertvolle Informationen zur Abschätzung der Dunkelziffer von Unfällen mit Motorfahrzeugen. Aber auch hier müssen immer datenschutzrechtliche Aspekte abgeklärt und mitberücksichtigt werden.

Anhänge

I	Resultate der IRTAD-Befragung	99
I.1	IRTAD-Befragung: Zusammenstellung der Antworten der Länder	99
II	Aktennotiz Stakeholder-Workshop I	103
III	Mathematische Details und Simulationen	119
III.1	Notation	119
III.2	Likelihood	119
IV	Aktennotiz Stakeholder-Workshop II	136

I Resultate der IRTAD-Befragung

I.1 IRTAD-Befragung: Zusammenstellung der Antworten der Länder

Tab. 15: IRTAD-Befragung: Zusammenstellung der Antworten der Länder

	DZ für Subgruppe /-population	Verwendete Datenquelle(n)	Methodik	Berücksichtigte Verletzungsschweren (inkl. Definition)	Weitere Differenzierungen	Anteil der nicht registrierten Unfälle (%) bzw. Dunkelziffer	Frequenz Berichterstattung	Quelle / Bemerkungen
Finnland		2012: Haushaltsbefragung 2014-2015: Vergleich Polizeidaten mit Spitaldaten	2012: 1. Phase Haushaltsbefragung bei 41'661 Personen -> Teilnahmerate 38%). 2. Phase: Fokus auf Haushalte, in denen sich jemand bei einem Unfall verletzt hatte (n=71). Dabei wurde geprüft, ob diese Personen in versch. Registern (Polizeistatistik, dem Register der Rettungsdienste, Versicherungsregister und Gesundheitswesenregister) auftauchen 2014-2015: Vergleich von Registern. Die Studie scheint keine systematische Verknüpfung der Daten aus den verschiedenen Registern durchgeführt zu haben.	2012: alle Verletzungsschweren. Definition einer schweren Verletzung: Verletzungen, die eine Behandlung von mindestens zwei Tagen im Krankenhaus erforderten (erwartet oder tatsächlich). 2014 - 2015: MAIS3+		- 2012: von 71 befragten bei Strassenverkehrsunfällen verletzten Personen wurden 70% nicht registriert - 2014-2015: Unfälle mit schweren Verletzungen werden um ca. 50% unterschätzt	-	Quelle Haushaltsbefragung: Kautiala et al, 2012 (Bericht auf Finnisch, engl. Abstract) [93] Information zur Befragung von 2014/2015 stammt aus der Antwort-E-Mail auf die IRTAD-Umfrage.
Frankreich	Rhône Bezirk	- Spitaldaten (Stationäre und Notfälle) (aus Bezirk Rhone) - Polizeidaten	Die Daten wurden mittels spezifischen Merkmalen oder Identifikatoren (Unfalldatum, -zeit, -ort, personenbezogene Informationen d. Verletzten (z.B. Name, Geburtsdatum) abgeglichen. Spitaldaten aus Bezirk Rhone, Abgleich mit Polizeidaten via Capture-Recapture. Hochrechnung auf ganz Frankreich	Alle Verletzungsschweren (MAIS1+ und MAIS3+)	Verkehrsteilnahme: - Auto - Motorisierte Zweiräder - Velo - Fussgänger/-in - Andere	Letzte Abschätzung im 2017: 318'000 MAIS (1-2) und 20'000 MAIS 3+, verglichen mit 73'000 polizeilich registrierten Verletzungen (entspricht 78%); je fast 7000 motor. Zweiräder und Auto, je fast 3000 Fussgänger/-innen und Velofahrende	Jährlich	Quellen: - Amoros et al, 2019 [94] - Amoros et al., 2020 [95] Bemerkungen: Im Paper sind andere Resultate, aber die sind anscheinend veraltet (neue AIS scale und ein refinement of the method): Gemäss dem Paper: 2015: alle Verletzten: 260'000 gegenüber 71'000 polizeilich registrierten (-> 73%), MAIS 3+: 21'000 (6000 Auto, 8300 motor. Zweiräder, 4600 Velo, 31000 Fussgänger/-innen, 600

	DZ für Subgruppe /-population	Verwendete Datenquelle(n)	Methodik	Berücksichtigte Verletzungsschweren (inkl. Definition)	Weitere Differenzierungen	Anteil der nicht registrierten Unfälle (%) bzw. Dunkelziffer	Frequenz Bericht-erstattung	Quelle / Bemerkungen
Griechenland - > versch. Europäische Länder		- Spitaldaten - Polizeidaten	Als Grundgesamtheit wird die Vereinigung von Polizei (P) - und Krankenhausfällen (KH) angenommen. Probabilistische Verknüpfung über Unfalldatum, -zeit, -ort, beteiligten Fahrzeugen und Personenmerkmalen der Verletzten). Es wird argumentiert, dass Fälle, die hier nicht erfasst werden, nicht relevant sind. Es wird der "under-reporting coefficient" (= Dunkelziffer) definiert: (Anzahl P vereinigt mit KH) / (Anzahl P) Im Ländervgl. wird verglichen: underreporting vs. country, road user type, police severity, AIS score Modell: Poissonregression mit Offset: $\log(\text{total}) \sim \log(P) + \text{Variablen inkl. Wechselwirkung ("log-rate analysis")}$ Bis zur 4-fach Wechselwirkung signifikant und wird interpretiert	Alle: - Polizei: "leicht" und "schwer" - Spital: Umrechnung ICD-9/10 zu MAIS-Score 1-6). Einteilung MAIS 1-2 und MAIS 3+	Land Verkehrsteilnahme: - Auto - Motorrad - Velo - Fussgänger/-in - Andere	Dunkelziffern ("under-reporting correction coefficients"): - Tschechien: 1.58 - Frankreich: 2.75 - Griechenland: 6.51 - Niederlande: 1.05 - UK: 1.05 - Ungarn: 1.31 - Spanien: 1.07	-	Quellen: - Yannis et al., 2014 [34] - Broughton et al., 2010 [51] - SafetyNet, 2004 [96]
Indien	Nur Highways	In-Depth Unfalldatensatz: RASSI = "Road Accident Sampling System - India". http://www.rassi.org.in/	Bosch Accident Research hat die Untersuchung in Indien durchgeführt. RASSI sammelt seit 2011 Daten zu Unfällen. Die Unfälle werden von verschiedenen Quellen gemeldet: Polizei, medizinische Rettungsdienste, Krankenhäuser, eigenen Patrouillen sowie Zeugen eines Unfalls. -> Bessere Auflösung als polizeilich gemeldete Fälle, daher werden diese Daten für politische Entscheide verwendet "Verknüpfung" mit Polizeidaten: im RASSI-Datensatz ist die Information enthalten, ob ein Unfall polizeilich registriert wurde oder nicht.	- Alle Unfälle - Tödliche Unfälle - Unfälle mit Schwerverletzten - Unfälle mit Leichtverletzten (inkl. Spitalaufenthalt von <24h) - Unfälle mit "minimal" Verletzten oder Sachschaden	Verkehrsteilnahme (involviert in Unfall): Auto, Lastwagen, Motorrad, Fussgänger/-in, Bus Kombinationen Verkehrsteilnahme: Auto vs. Auto, Lastwagen vs. Lastwagen, Lastwagen vs. Auto, Auto vs. Motorrad, Lastwagen vs. Motorrad Unfalltyp: Kontrollverlust, Abbiegen, Einbiegen / Kreuzen, Fussgängerübergang, Parkierte Fahrzeuge, lateraler Verkehr, Anderer Strassentyp: National Highway, State Highway, Expressway Strassenstruktur: Gerade Strecke, Knoten, Kreuzung, Einfahrt, U-Turn, Brücke Unfallmeldung durch: Untersuchungsteam, Polizei, Ambulanz, Andere	- Alle Unfälle: 32% - Tödliche Unfälle: 0.3% - Unfälle mit Schwerverletzten: 29% - Unfälle mit Leichtverletzten: 73% - Unfälle mit "minimal" Verletzten oder Sachschaden: 88%	-	Quelle: Moennich et al., 2017 [49] PROBLEM: RASSI wird als "Ground-truth" angesehen; das scheint aber unrealistisch. Daher entsprechen die Zahlen eher einer Untergrenze für die Unterschätzung

	DZ für Subgruppe /-population	Verwendete Datenquelle(n)	Methodik	Berücksichtigte Verletzungsschweren (inkl. Definition)	Weitere Differenzierungen	Anteil der nicht registrierten Unfälle (%) bzw. Dunkelziffer	Frequenz Bericht-erstattung	Quelle / Bemerkungen
Niederlande	Getötete	- Todesursachenstatistik - Strafregister - Polizeidaten	Die drei Datenquellen werden verknüpft und verglichen (statistische Details sind allerdings nicht angegeben). Doppeleingaben, Todesfälle die keine Strassenverkehrsunfälle sind (z.B. Suizide) werden gelöscht. Es wird davon ausgegangen, dass jeder Todesfall im Strassenverkehr in mindestens einem der Register erfasst ist.	Nur Getötete	Verkehrsteilnahme: - Fussgänger/-in - Velo - Moped - Motorrad - Auto / Lieferwagen - Lkw / bus - Andere	Jahre 2011-2020: - Total: 14.4% - PW und Lieferwagen: 2% - Motorrad: 4% - Velo und E-Bike: 30% - Fussgänger/-in: 4% - Mofo: 15% - Lastwagen / Bus: 0% - Andere (inkl. E-Trottinette): 43%	Jährlich	Quellen: - Institute for Road Safety Research, 2024 [97] - Institute for Road Safety Research [99]
	Schwer-verletzte	Spitaldaten (Datenbank der Spitalentlassungen) Datenbank registrierter Strassenverkehrsunfälle (BRON): Polizeiregister, Meldungen von Strasseninspektoren der niederl. Strassenverkehrsbehörde, Medienberichte	Verknüpfung der Datenbank der Spitalentlassungen und der BRON-Datenbank (Datum, Geschlecht, Provinz, Zeit) beiden Datenquellen (nicht genau beschrieben wie die Verknüpfung gemacht wird) Machen auch eine Abschätzung, wie viele Verletzte in beiden Datenquellen fehlen	Schwerverletzte (MAIS3+)	Verkehrsteilnahme: - Fussgänger/-in - Velo - Moped - Motorrad - Auto / Lieferwagen - Andere	Total (2023): 37% der bei Strassenverkehrsunfällen mit Beteiligung Motorfahrzeug Schwerverletzten und 88% der bei Strassenverkehrsunfällen ohne Beteiligung Motorfahrzeug Schwerverletzten werden registriert	Jährlich	Quelle: Institute for Road Safety Research, 2024 [98]
	Leicht- und Schwer-verletzte	Spitaldaten	Stichprobe (12%) von Notaufnahmen niederländischer Krankenhäuser, bei denen detaillierte Infos erhoben werden. Die Fälle erhalten ein Gewicht von ca. 8.	Leicht- und Schwerverletzte (Schwere Verletzung: MAIS2+)	Verkehrsteilnahme: - Velo - PW - Motorrad - Roller - Fussgänger/-in - Andere/unbekannte Mobilitätsformen Unfallhergang: - Alleinunfall - Kollision mit anderem VT - Kollision mit Hindernis	Velo: 4x mehr Unfälle in Spitälern E-Bike: 3-4x mehr Unfälle in Spitälern	Jährlich	Quelle: Veiligheid (Niederländisch): https://www.veiligheid.nl/onderzoek/onze-onderzoeken
	Sachschäden: Motor-fahrzeuge	Versicherungsdaten	Vermutlich Vergleich der Anzahl Fälle im Polizeiregister mit den entsprechenden Fälle aus Versicherungsdaten (Details zur Methodik fehlen)	Sachschäden		Total: 86%	?	Bemerkung: Die Information stammt lediglich aus dem Antwortmail auf die IRTAD-Befragung. Keine weiteren Informationen gefunden.
Österreich	Velo und E-Bikes	Spitaldaten (IDB, Injury Database Austria): Für Velo und E-Bike: Polizeidaten)	Alle Verkehrsteilnahmen: jährlich ca. 15000 Interviews in ausgewählten Spitälern (im Jahr 2021: 13'400 Interviews stellvertretend für rund 600000 Personen, die jährlich in Krankenhäusern in Österreich wegen einem Unfall behandelt werden) -> Hochrechnung auf Österreich. Gewichtung nach Aufnahmedatum, Behandlungsart (stationär/ambulant), Altersgruppen und Covid-Effekte). Für Velo und E-Bike erfolgte ein Vergleich mit den Polizeidaten (unklar, welche Methodik hier verwendet wurde).	Verletzt (nicht klar definiert)	Für Hochrechnung: Verkehrsteilnahme (detailliert erhoben aber berichtet wird folgendes): - (E-)Fahrrad - Moped/Motorrad - Fussgänger+ - PW - Sonstige Alter: in 5-Jahresschritten Geschlecht	Velo: 4x mehr Unfälle in Spitälern E-Bike: 3-4x mehr Unfälle in Spitälern	Jährlich	Quelle: Donabauer & Pommer, 2021 (IDB Jahresbericht) [100]

	DZ für Subgruppe /-population	Verwendete Datenquelle(n)	Methodik	Berücksichtigte Verletzungsschweren (inkl. Definition)	Weitere Differenzierungen	Anteil der nicht registrierten Unfälle (%) bzw. Dunkelziffer	Frequenz Bericht-erstattung	Quelle / Bemerkungen
Schweden		- Polizeidaten - Spitaldaten (Notfall) (Swedish Traffic Accident Data Acquisition (STRADA))	Direkte Vergleiche der Gesamtanzahl von polizei- und krankenhausgemeldeten Unfällen sowie Capture-Recapture-Methoden. Die Verknüpfung erfolgte über die persönliche Identifikationsnummer der Unfallbeteiligten.	Polizeidaten: Anweisung zu schwerer Personenschaden: Personen mit Knochenbrüchen, stumpfen Traumata, Brüchen, Stichverletzungen, Gehirnerschütterung oder anderen inneren Verletzungen. Ebenso Personen, die wahrscheinlich in ein Krankenhaus eingeliefert werden. Spitaldaten: AIS/ISS	Verkehrsteilnahme (involviert in Unfall): Auto, Lastwagen, Bus, Motorrad, Moped, Velo, Fussgänger/-in Alle Kombinationen v. Unfällen: Auto - Lastwagen, Auto - Bus, Auto - Motorrad, ... Spezielle Untergruppe: Velo-Alleinunfälle (mind. 96% Underreporting) Unfallmonat Verletzungsschwere: Schwere Unfälle, Leichte Unfälle Tageszeit: tagsüber (7:00 - 18:59 Uhr), nachts (19:00 - 6:59 Uhr)	2003-2013: Underreporting mind. 48%	-	Quelle: Held, 2016 [50]
Tschechien -> s. unter "Griechenland"			SafetyNet Report (2007); Spitaldaten aus dem einzigen Spital am Ort werden mit den Polizeidaten gelinkt; Linkage-Variablen: Datum Unfall, Geburtsjahr, Geschlecht, Typ Strassenutzer, Unfallschwere					Quellen: - Yannis et al., 2014 [34] - Broughton et al., 2010 [51] - SafetyNet, 2004 [96]
Ungarn	Velo und E-Bikes	- Polizeidaten - Spitaldaten - Befragung unter Velofahrenden	Der englische Bericht ist noch nicht verfügbar, wird aber zur Verfügung gestellt sobald der wissenschaftliche Artikel veröffentlicht ist. Zusätzlich war Ungarn auch Teil des SafetyNet Report (s. "Griechenland")	Verletzt (leichte Verletzung; schwere Verletzung)	Unfallhergang: - Alleinunfall - Velo - Velo und Fussgänger/-in - Velo - Motorisiertes Fahrzeug - Velo Verletzungsschwere: - Leichte Verletzung - Schwere Verletzung	Total (Velo und E-Bikes): 91.7%		Quellen: - Yannis et al., 2014 [34] - Broughton et al., 2010 [51] - SafetyNet, 2004 [96]

II Aktennotiz Stakeholder-Workshop I

Aktennotiz

Projekt	MEDUSA: Methodik Dunkelziffer Strassenverkehrsunfälle
Sitzung	Stakeholderworkshop I – Bedürfnisse
Datum, Zeit	22.1.2024, 09:30-12:00
Ort	Beratungsstelle für Unfallverhütung BFU, Hodlerstrasse 5a, 3011 Bern
Moderation	– Matthias Amacher, Ecoplan (Leitung) – Patrizia Hertach, BFU – Markus Kalisch, ETH
Teilnehmende Stakeholder	– Christian Kamenik, ASTRA – Alexandra Quandt, BFS – Josephine Leuba, ARE – Wernher Brucks, Stadt Zürich – Nicole Felder, ASTRA Langsamverkehr – Stefan Scholz, SSUV – Steffen Niemann, BFU – Christian Egeler, Verkehrspolizei Kanton BL
Entschuldigt	– Michael Pfäffli, axa
geht an	alle Teilnehmenden und Entschuldigten sowie –
Unterlagen	Präsentation
Verfasser	Matthias Amacher, Ecoplan
Datum	31.1.2024

Programm

Zeit	Traktandum	Inhalt
09:30	1 Start / Einleitung	Programm / Vorstellung + Erwartungen (1 Satz) Vorgehen / Abgrenzungen
09:50	2 Bedürfnisse	Fragen / Ziele, Lücken
10:25	3 Differenzierung (Teil 1)	Differenzierungsmerkmale
10:45	Pause (ca. 25 Min)	
11:10	3 Differenzierung (Teil 2)	
11:35	4 Periodizität, räumliche und zeitliche Auflösung	Vorschlag Forschungsstelle
12:00	5 Würdigung / Weiteres Vorgehen	
12:15	Schluss	

Inhalt

1	Start / Einleitung	3
1.1	Ziel des Forschungsprojekts	3
1.2	Ziel des Workshops	3
1.3	Abgrenzungen / Begriffe	4
2	Bedürfnisse	5
2.1	Einleitung	5
2.2	Ziele / Fragen	6
2.3	Lücken	8
2.4	Einleitung	9
2.4.1	Bisherige Differenzierung («BFU Hochrechnung»)	9
2.4.2	Mögliche Differenzierungen	9
2.4.3	Vorgehen	10
2.5	Priorisierung der Differenzierungsmerkmale	10
2.5.1	Personen	10
2.5.2	Objekte / Fahrzeuge	11
2.5.3	Ereignis / Umfeld	12
2.5.4	Hergang / Ursache	13
2.6	Zusammenfassung	14
3	Periodizität, Auflösung	15
4	Würdigung / Weiteres Vorgehen	16
4.1	Würdigung	16
4.2	Weiteres Vorgehen / Nächste Schritte	16

1 Start / Einleitung

1.1 Ziel des Forschungsprojekts

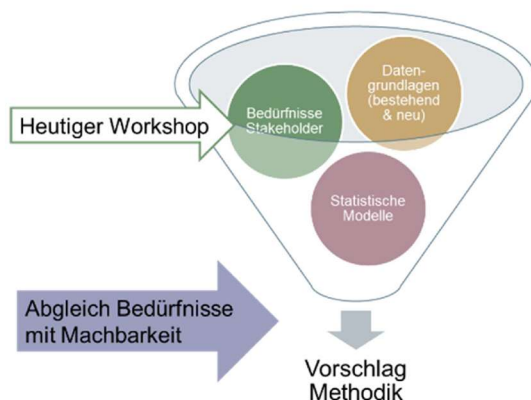
- Das Projekt hat zum Ziel, eine Methodik zur Herleitung einer differenzierten «Dunkelziffer» für das Unfallgeschehen auf Schweizer Strassen zu erarbeiten.
- Die «Dunkelziffer» beschreibt dabei auf aggregierter Ebene (nicht auf Ebene von Einzelpersonen oder einzelnen Unfällen) jenen Teil des Unfallgeschehens, welcher nicht in den offiziellen Statistiken der Strassenverkehrsunfälle enthalten ist.
- Die vorgeschlagene Methodik soll es ermöglichen, das Gesamtunfallgeschehen zu beschreiben, und zwar in einer Qualität und Differenzierung welche den **Bedürfnissen** der verschiedenen Nutzer dieser Informationen (Stakeholder) möglichst gut entspricht.
- Rahmenbedingungen für die Machbarkeit dieser Differenzierung sind dabei die Qualität und Verfügbarkeit der benötigten Datengrundlagen, die verfügbaren Ressourcen zur Datenaufbereitung und Anwendung der Methodik (Finanzen/Personal) sowie die generellen Herausforderungen der Schätzung unbekannter Grundgesamtheiten mittels statistischer Methoden.
- Im ersten Arbeitspaket AS 1 werden als Grundlage für die weiteren Arbeiten die Bedürfnisse der Nutzer / Stakeholder an einem Workshop aufgenommen.

1.2 Ziel des Workshops

Aufnahme der Bedürfnisse der Stakeholder hinsichtlich der notwendigen Differenzierung des «Dunkelfelds» bzw. der Dunkelziffer

- Breite und offene Herangehensweise
- Abgleich mit Machbarkeit erfolgt in nächstem Schritt

Abbildung 1-1: Vorgehen / Einordnung des Workshops

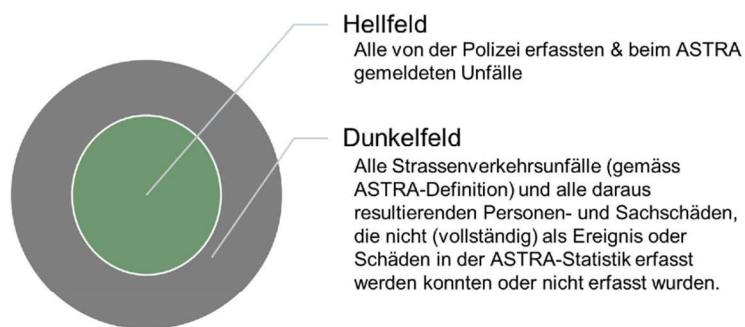


1.3 Abgrenzungen / Begriffe

Grundsätzlich gemäss ASTRA VU-Definition:

- Als **Strassenverkehrsunfall** gilt ein unvorhergesehenes Ereignis auf einer öffentlichen Verkehrsfläche, das einen Sach- und/oder Personenschaden zur Folge hat und an dem mindestens ein Fahrzeug oder ein fahrzeugähnliches Gerät beteiligt ist.
- Es gibt viele Abgrenzungsfragen (bei unterstrichenen Teilen).
- Gemäss Vorgehensvorschlag im Forschungsprojekt:
 - Fokus auf Personenschäden (Verletzte)
 - Verletzungsschweren gemäss Definition ASTRA UAP (NACA-Codierung)
 - Es gibt auch Unsicherheiten im Hellfeld --> Korrektur des Hellfelds ist nicht das Ziel des Projekts

Abbildung 1-2: Hellfeld vs. Dunkelfeld



Dunkelziffer = Verhältnis zwischen Anzahl im Hellfeld und Gesamtzahl (Summe aus Hell- und Dunkelfeld) z.B.

Hellfeld	Dunkelfeld	Gesamt	Dunkelziffer
20	10	30 = 20 + 10	1.5 = 30 / 20

2 Bedürfnisse

2.1 Einleitung

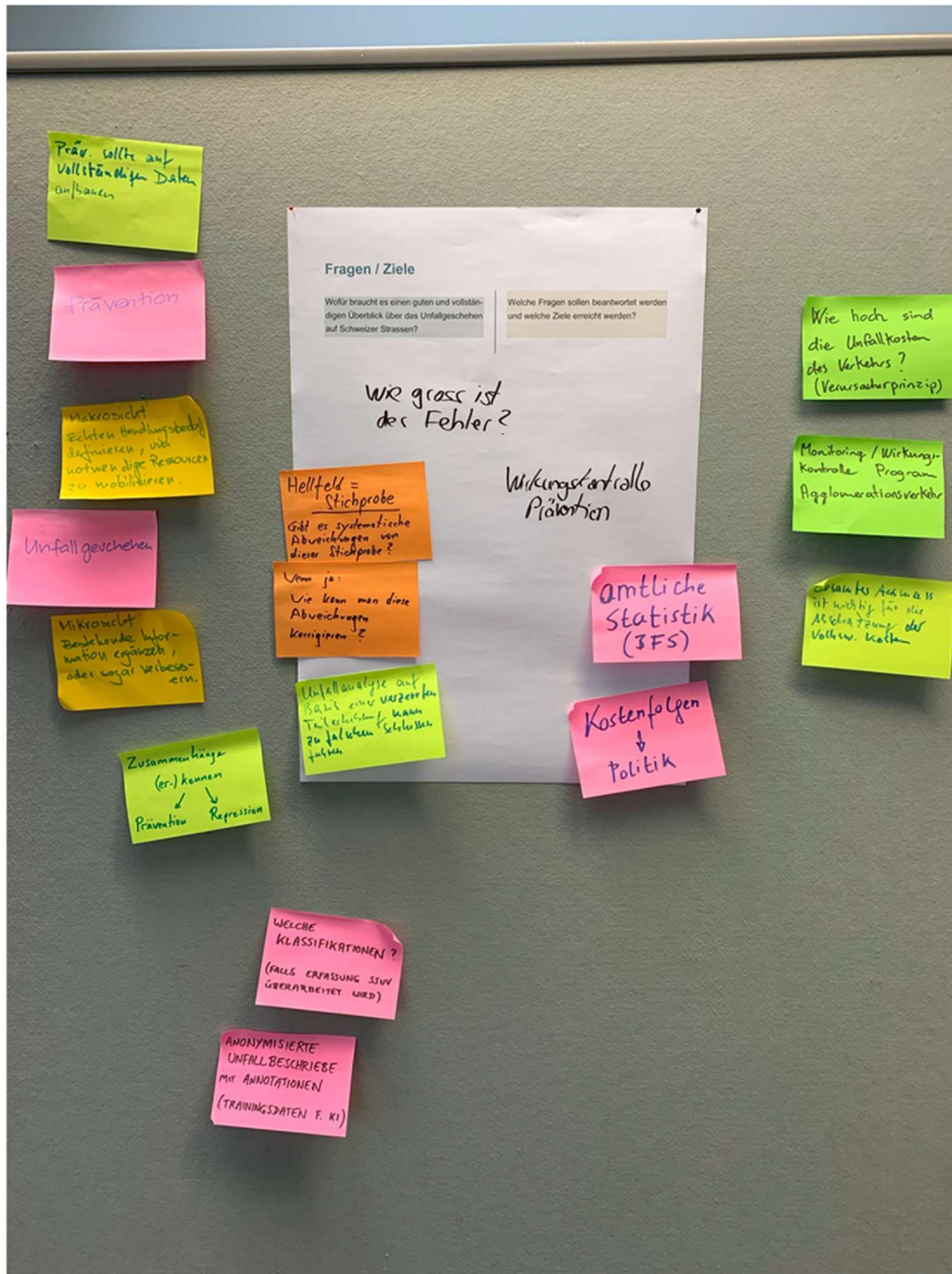
Fragen:

- In welchem für Sie wichtigen Teil gibt es grosse Wissenslücken?
- Wofür braucht es einen guten und vollständigen Überblick zum Unfallgeschehen? Welche Fragen sollen mit den Resultaten beantwortet werden können?

Vorgehen:

- Jede/r TN erhält 3 Kärtchen
- Stichwortartiges notieren der Antworten auf obige Fragen
- Zusammentragen und Diskussion im Plenum

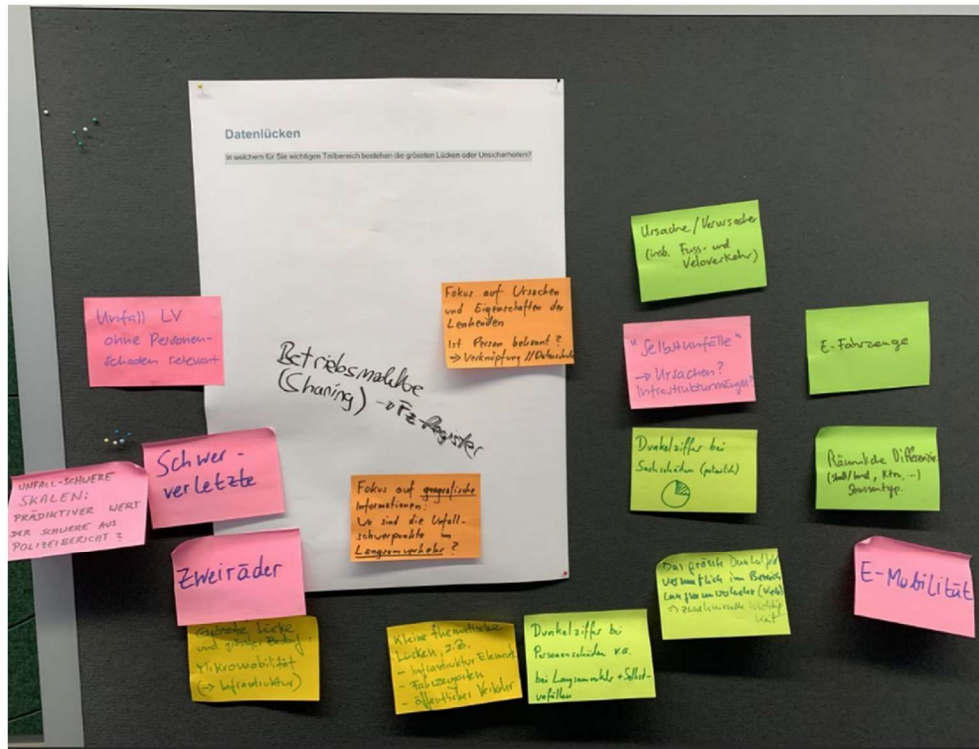
2.2 Ziele / Fragen



Die mit einer Abschätzung der Dunkelziffer zu beantwortenden Fragen bzw. die zu erreichenden Ziele lassen sich grob in folgende Kategorien einteilen:

- **Prävention** von Verkehrsunfällen (auch Repression: wo/wann muss die Polizei präsent sein). Hier sind auch relative Betrachtungen wichtig: sind die Verhältnisse im Hellfeld repräsentativ?
- **Amtliche Statistik** / Externe Kosten des Verkehrs (Kostenfolgen von Unfällen, Grundlagen für Politikentscheide)
- **Verkehrsplanung**: Analyse von Schwachstellen / Unfallanalyse, Planung von sicherer Verkehrsinfrastruktur
- **Wissen** / Abbildung des Unfallgeschehens: Gibt es systematische Abweichungen zwischen Hell- und Dunkelfeld bzw. für verschiedene Teilmengen des Hellfelds?
- **Wirkungskontrolle** von planerischen Massnahmen und Prävention

2.3 Lücken



Lücken bestehen insbesondere bei:

- Ursachen:
 - im Fuss- und Veloverkehr (wer ist Verursacher der Unfälle?)
 - Generell bei Selbstunfällen: Welche Rolle spielt die Infrastruktur? Ursache wird bei Infrastruktur vermutet, aber Informationen zum Unfallort fehlen --> Zusammenhang mit geografischen Schwerpunkten
- Mikromobilität (neue Fahrzeugtypen, verschiedene Betriebsmodelle wie Sharing/Leasing etc.) und Elektromobilität
- Geografische Schwerpunkte: Unfallorte / Unfallstellen und Infrastruktur am Unfallort
- Veloverkehr: Viele Unfälle ohne (bedeutende) Personenschäden werden nicht erfasst. Auch kleine Verletzungen oder Sachschäden sind in der Masse relevant. Unklare, evtl. uneinheitliche und schwierige Abgrenzung von bedeutenden zu unbedeutenden Verletzungen.
- Schwere Verletzungen: Qualität der heutigen Grundlagen ist unklar
- Öffentlicher Verkehr auf der Strasse: Führt eigene Statistiken, die teilweise nicht mit ASTRA-Daten übereinstimmen bzw. z.T. mehr/weniger erfassen.

2.4 Einleitung

2.4.1 Bisherige Differenzierung («BFU Hochrechnung»)

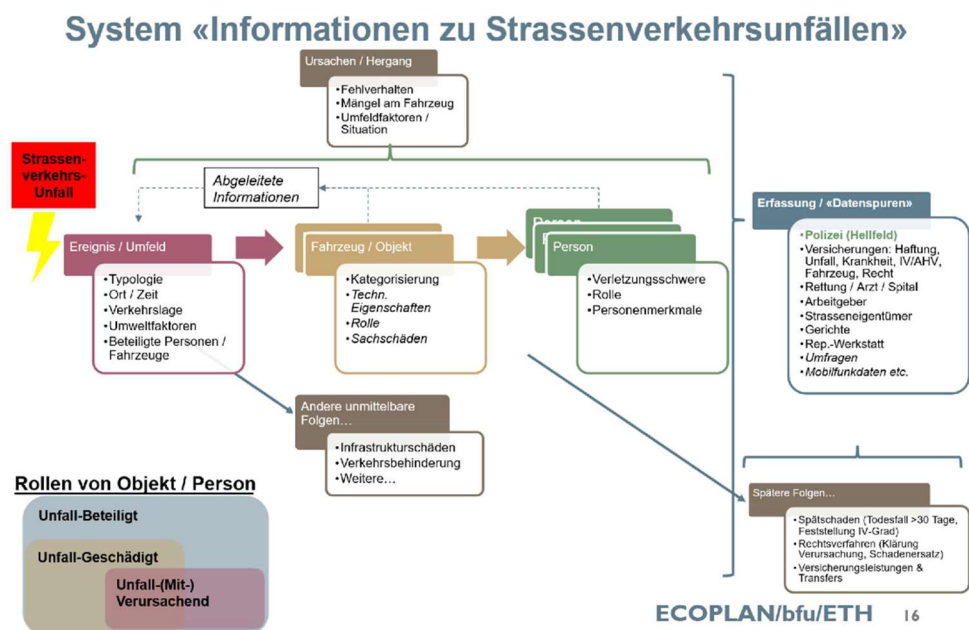
Anzahl Verletzte nach ...

- 4 Verletzungsschweren (SUVA-Kategorien, Ausfalltage)
- 4-8 Fahrzeugkategorien
- 3 Altersklassen
- Geschlecht

→ Publikation erfolgt in vereinfachter / aggregierter Form

2.4.2 Mögliche Differenzierungen

Welche Differenzierungen könnten von Interesse sein?



2.4.3 Vorgehen

- Es sind 4 Plakate mit möglichen Differenzierungsmerkmalen (Personen, Objekte, Ereignis/Umfeld, Hergang/Ursachen) aufgehängt
- Jede/r TN erhält 4 rote Punkte. 5-10 Minuten Zeit um Punkte bei notwendigen Differenzierungsmerkmalen zu setzen. Bei Bedarf können zusätzliche Zeilen ergänzt werden.
- Setzen der **roten** Punkte (hohe Priorität) bei **notwendigen** Differenzierungsmerkmalen
- Zusätzliche **grüne** Punkte (geringe Priorität) werden bei Bedarf abgegeben

2.5 Priorisierung der Differenzierungsmerkmale

2.5.1 Personen

Personen nach...	Differenzierung notwendig
Verletzungsschwere / Unfallfolgen	● ● ● ● ●
Rolle (z.B. geschädigt und/oder verursachend)	●
Demografische Merkmale (z.B. Alter, Geschlecht, Zivilstand)	● ● ●
Sozio-ökonomische Merkmale (z.B. Erwerbsfähigkeit, Versicherungsstatus, Einkommen)	
Zustand zum Zeitpunkt des Unfalls (z.B. Fahrfähigkeit, Sicherheitsvorkehrungen: Tragen von Gurten oder Helm) <i>ist zum Teil in der Ursache enthalten</i>	● ●

Ursache

→ selbstwert.

Prioritäten → wer?

2.5.2 Objekte / Fahrzeuge

Fahrzeuge nach...	Differenzierung notwendig
Fahrzeugart / Verkehrsmittel / Fahrzeugkategorie	● ● ● ● ● ● ●
Vereinfachte, zusammengefasste Art der Verkehrsteilnahme / Verkehrsart (z.B. Fussverkehr, Veloverkehr, MIV, Schwerverkehr, öff. Personenverkehr)	
Rolle (z.B. beteiligt, geschädigt und/oder verursachend)	● ● ●
Technische Eigenschaften (z.B. Antriebsart, Grösse, Alter, Assistenzsystem, Betriebszustand etc.)	● ● ●

Velo

E-Trotti's

Pedelec

E-Bike

mot. Leichtfahrzeug

Cargo-Bikes

Ursacher?
+ Monitoring

Besonders interessieren die kleineren bzw. langsameren Fahrzeuge wie Velos, E-Trottis, Pedelecs/E-Bikes/Cargo-Bikes sowie andere motorisierte Leichtfahrzeuge.

2.5.3 Ereignis / Umfeld

Ereignis / Umfeld nach...	Differenzierung notwendig
Unfallort (z.B. Kanton, Gemeinde etc.)	●
Raumtyp (z.B. Stadt-Land-Intermediär, Dichteklassierung, innerorts-ausserorts)	● ●
Strassenart (z.B. Autobahn, Hauptstrasse, Nebenstrasse etc.)	●
Zeitpunkt des Unfalls (z.B. Monat, <u>Wochentag</u> , <u>Tageszeit</u> etc.)	● ●
Signalisation (z.B. Höchstgeschwindigkeit)	
Verkehrsaufkommen / Verkehrsfluss	
Unfallstelle (z.B. Parkplatz, Kurve, gerade Strecke, Kreisverkehr, Verzweigung, Platz etc.)	● ● ●
Infrastruktur (z.B. Trottoir, Verkehrsstruktur, etc.)	● ●
ABGRENZUNG zu SPORT & SPIEL (Skatepark, MTB, ...)	●

→ selbsttest
"Sonntag 2 Uhr"

* Komplex
Daten nicht gut

2.5.4 Hergang / Ursache

↑ Alleinunfall (Gefährdung) → Vortritt genommen

Hergang / Ursachen nach...	Differenzierung notwendig
Typ: Selbstunfall oder Kollision	● ● ● → selbstversch.
Anzahl beteiligte Fahrzeuge (z.B. auch: ein beteiligtes Fz. vs. mehrere beteiligte Fz.)	● ●
Anzahl beteiligte Personen (z.B. auch: eine beteiligte Person vs. mehrere beteiligte Personen)	
Unfallursachen (z.B. einfache aggregierte Kategorien)	● ● ● ● ● ●

↳ multikausal Warum?

↳ Infrastruktur Unfallstelle (gegenseitig)

2.6 Zusammenfassung

Folgende Differenzierungen wurden (ohne gegenteilige Voten) als selbstverständlich angenommen und deshalb von vielen nicht mehr bepunktet:

- nach **Alter**
- nach **Geschlecht**
- nach **Unfalltyp** (Selbstunfall/Kollision): inkl. allenfalls Aussagen zu als «Selbstunfall» erfassten Unfällen mit Dritteinwirkung (Unfall nach einer gefährlichen Situation, bei dem der «Gefährder» nicht im Unfallprotokoll erfasst wurde: z.B. Unfall nach einem abrupten Ausweich- oder Bremsmanöver aufgrund eines nicht gewährten Vortritts oder «zu nahem Vorbeifahren»)
- nach **Zeitpunkt** (Tageszeit und Wochentag): Lücken im Hellfeld bestehen insbesondere zu Randzeiten bzw. in der Nacht am Wochenende («Sonntag 2 Uhr morgens»)

Mit den meisten Punkten bewertete Differenzierungsmerkmale sind:

- **Personen:** Der differenzierte Ausweis nach **Verletzungsschwere** wird als sehr wichtig erachtet, wobei eine kombinierte Aufschlüsselung nach Alter und Geschlecht insbesondere für die Prävention wichtig ist: Wer soll angesprochen werden?
- **Fahrzeuge/Objekte:** Sehr wichtig ist auch die **Fahrzeugkategorie** / das Verkehrsmittel: Auch hier steht die Frage im Vordergrund, wer im Rahmen der Prävention angesprochen werden soll, wobei die Zuweisung der verletzten Personen auf Fahrzeugkategorien Idealfall jeweils für unterschiedliche Bedürfnisse erfolgen sollte: im Monitoringprinzip (wer ist gefährdet?) und im Verursacherprinzip (wer gefährdet andere?). Ein Schwerpunkt sollte dabei auf weniger geschützten «kleineren» Fahrzeugkategorien («Mikromobilität») sowie die Elektromobilität gesetzt werden, für die es noch grössere Wissenslücken gibt.
- **Ursachen/Hergang:** Warum kam es zum Unfall? Gemeint sind dabei tieferliegende **Unfallursachen** (meist nicht in ASTRA VU abgebildet), wobei die Ursache multikausal sein kann. Eine Rolle kann dabei auch die Infrastruktur spielen (siehe nachfolgender Punkt).
- **Ereignis/Umfeld:** Unfallort, Unfallstelle, **Infrastruktur:** Es wird ein Zusammenhang zwischen Infrastruktur (Gestaltung, z.B. mit Randstein, Trottoir / Velowege) an der Unfallstelle und den Ursachen eines Unfalls angenommen. Einzelne Teilnehmende sind auch auf eine Differenzierung nach Kanton und/oder **Raumtyp** angewiesen.

3 Periodizität, Auflösung

Die Teilnehmenden können zum Vorschlag der Forschungsstelle im Plenum Stellung nehmen.

Frage	Vorschlag Forschungsstelle	Diskussion / Voten der TN
Wie häufig sollen Dunkelziffern berechnet und publiziert werden?	Jährlich (1x pro Jahr)	– Eine Grundlagenerhebung der Dunkelziffern reicht alle 5-7 Jahre. – Dazwischen ist eine laufende Fortschreibung notwendig, damit jährliche Gesamtzahlen und Mobilitätstrends aufgenommen werden können.
Wie lange sollte es maximal dauern, bis Daten zur Dunkelziffer für das Jahr 2024 publiziert werden?	1-2 Jahre	– Bei der Grundlagenerhebung ist die Publikationsfrist nicht so wichtig. – Für die jährliche Fortschreibung sollte 1 Jahr (nach Publikation des Helffelds) ausreichen.
Welche räumliche Differenzierung braucht es?	Ganze Schweiz im Territorialprinzip (alle Unfälle in der Schweiz)	– Eine Differenzierung nach Raumtyp oder Kanton wäre hilfreich. – Die räumliche Differenzierung muss nicht mit allen anderen Merkmalen kombiniert werden können.
Sind Angaben zu Verletzten den Fahrzeugkategorien im Monitoring - oder im Verursacherprinzip zuzuordnen?	«Monitoringprinzip» = in welchem Fahrzeug befand sich die verletzte Person?	– Es braucht sowohl das Monitoring- (wer ist betroffen?) als auch das Verursacherprinzip (wer gefährdet andere?).

4 Würdigung / Weiteres Vorgehen

4.1 Würdigung

- Die Teilnehmenden konnten ihre Bedürfnisse an die Methodenentwicklung der Dunkelziffer platzieren.
- Kenntnis der Dunkelziffer wird v.a. als wichtig für die Unfallprävention, Planung/Unterhalt der Infrastruktur und für die Abschätzung der Unfallfolgekosten erachtet
- Häufig genannt wurde die Problematik des Langsamverkehrs und neuer Mobilitätsformen wie die Elektro- und Mikromobilität (grosse Dunkelziffer vermutet/erwartet)
- Bzgl. Differenzierung sind Verletzungsschwere, Mobilitätsform, Unfalltyp, Alter, Geschlecht, Zeitfenster und Ursachen/Umwstände am wichtigsten. Ein Bedürfnis scheint es auch beim verbesserten Verständnis der Unfallentstehung zu geben, insb. bzgl. Infrastruktur.
- Einige der genannten Bedürfnisse gehen darüber hinaus, was im Hellfeld bekannt ist. Insbesondere tieferliegende Ursachen von Unfällen wie die Gestaltung der Infrastruktur, das Verhalten vor einem Unfall (Ablenkung) oder auch ob Gefährdungen oder Fehlverhalten durch Dritte vorlagen.
- Die Möglichkeiten zur Umsetzung der Bedürfnisse gilt es nun in einem weiteren Schritt zu prüfen.

4.2 Weiteres Vorgehen / Nächste Schritte

- AP 2: Analyse bestehender und neuer Datenquellen
- AP 3: Statistische Modelle zur Auswertung / Verknüpfung der Datenquellen (auch aus anderen Bereichen ausserhalb Unfallthema)

→ BK-Sitzung: 15. März 2024

- AP 4: Methodik erarbeiten
 - → Abgleich zwischen Möglichkeiten mit Daten / Modellen vs. Bedürfnisse

→ Stakeholderworkshop II: Juni/Juli 2024 --> Es folgt eine Terminumfrage

III Mathematische Details und Simulationen

III.1 Notation

Es werden folgende Notation, bzw. Abkürzungen verwendet:

Abkürzung	Bedeutung
p	Polizei
u	UVG
k	KVG
a	Ausland, d.h. kein Versicherungsstatus in der Schweiz
v	Versicherung, d.h. entweder UVG oder KVG.

Es sei p_p die Wahrscheinlichkeit, dass ein Individuum auf der Polizeiliste landet, analog bei p_v etc.

Annahme: Bei der Polizeiliste ist die Selektionswahrscheinlichkeit unabhängig vom Versicherungs-Status (scheint plausibel in der Realität).

$$p_{p|u} = p_{p|k} = p_{p|a} = p_p$$

Bei uns ist momentan die Versicherungsliste nur UVG, d.h. $p_{v|k} = p_{v|a} = 0$, also

$$\begin{aligned} p_v &= p_{v|u} \cdot p_u + p_{v|k} \cdot p_k + p_{v|a} \cdot p_a \\ &= p_{v|u} \cdot p_u. \end{aligned}$$

D.h. beim capture-recapture Problem mit den beiden Listen p und v gibt es nur die beiden Parameter p_p und p_v .

III.2 Likelihood

Auflistung der Möglichkeiten für capture-history und entsprechende Wahrscheinlichkeiten:

p	v	Wahrscheinlichkeit	Anzahl
1	1	$p_p \cdot p_v$	n_{11}
1	0	$p_p \cdot (1 - p_v)$	n_{10}
0	1	$(1 - p_p) \cdot p_v$	n_{01}
0	0	$(1 - p_p) \cdot (1 - p_v)$	n_{00}

Wir definieren $N = n_{11} + n_{10} + n_{01} + n_{00}$ und $D = n_{11} + n_{10} + n_{01}$.

Die Likelihood ist dann

$$\begin{aligned} L(n_{00}, p_p, p_v | n_{11}, n_{10}, n_{01}) &= \frac{N!}{n_{11}! \cdot n_{10}! \cdot n_{01}! \cdot n_{00}!} (p_p \cdot p_v)^{n_{11}} \cdot (p_p \cdot (1 - p_v))^{n_{10}} \cdot \\ &\quad ((1 - p_p) \cdot p_v)^{n_{01}} \cdot ((1 - p_p) \cdot (1 - p_v))^{n_{00}} \\ &= \frac{N!}{n_{11}! \cdot n_{10}! \cdot n_{01}! \cdot n_{00}!} p_p^{n_{10} + n_{11}} \cdot (1 - p_p)^{n_{00} + n_{01}} \cdot p_v^{n_{01} + n_{11}} \cdot \\ &\quad (1 - p_v)^{n_{00} + n_{10}} \\ &\propto \frac{N!}{n_{00}!} p_p^{R_p} \cdot (1 - p_p)^{N - R_p} \cdot p_v^{R_v} \cdot (1 - p_v)^{N - R_v}, \end{aligned}$$

wobei $R_p = n_{10} + n_{11}$ und $R_v = n_{01} + n_{11}$. Dies entspricht gerade der Gleichung (4) auf Seite

14 in King et al. (2019) [55]). Technisch gesehen entspricht dies dem sogenannten M_t Modell im Package Rcapture, weil wir zwei Listen haben, die jeweils ihre eigenen Wahrscheinlichkeiten haben.

Bemerkung: Wenn in der Polizeiliste nicht alle die gleiche Wahrscheinlichkeit haben, dann kann man die Likelihood nicht mehr so einfach aufschreiben: Der Vorfaktor ist dann anders, bzw. gruppenspezifisch.

Simulations for Report

Markus Kalisch

2025-03-21

Session info

```
sessionInfo()
```

```
## R version 4.4.3 (2025-02-28 ucrt)
## Platform: x86_64-w64-mingw32/x64
## Running under: Windows 11 x64 (build 26100)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=German_Switzerland.utf8 LC_CTYPE=German_Switzerland.utf8
## [3] LC_MONETARY=German_Switzerland.utf8 LC_NUMERIC=C
## [5] LC_TIME=German_Switzerland.utf8
##
## time zone: Europe/Zurich
## tzcode source: internal
##
## attached base packages:
## [1] stats      graphics  grDevices utils      datasets  methods   base
##
## loaded via a namespace (and not attached):
## [1] digest_0.6.35   R6_2.6.1        fastmap_1.1.1    xfun_0.48
## [5] cachem_1.0.8    knitr_1.48      htmltools_0.5.8.1 rmarkdown_2.28
## [9] lifecycle_1.0.4 cli_3.6.2       sass_0.4.9       jquerylib_0.1.4
## [13] compiler_4.4.3  rstudioapi_0.16.0 tools_4.4.3      evaluate_1.0.0
## [17] bslib_0.8.0     yaml_2.3.10     rlang_1.1.3      jsonlite_1.8.8
```

Kap 5.3.1: Kleines Beispiel

```
library(Rcapture)
## M_t Modell
X <- rbind( c(1,1,850),
            c(1,0,350),
            c(0,1,550))
closedpCI.t(X, dfreq = TRUE, m = "Mt")
```

```
##
## Number of captured units: 1750
##
## Poisson estimation and model fit:
## abundance stderr deviance df AIC BIC infoFit
## Mt 1976.5 22.9 0 0 30.428 46.83 OK
##
## Multinomial estimation, 95% profile likelihood confidence interval:
## abundance infCL supCL infoCI
## Mt 1975.4 1933.1 2023.1 OK
```

```
## sanity check mit Lincoln-Peterson
1400*1200/850 ## 1976
```

```
## [1] 1976.471
```

Kap. 5.3.2: Realistischeres Bsp

Zunächst werden Hilfsfunktionen definiert:

```
## Hilfsfunktionen ###
createMat <- function(captList, n) {
  mat <- matrix(FALSE, nrow = n, ncol = length(captList))
  colnames(mat) <- names(captList)

  for (i in 1:length(captList)) {
    mat[captList[[i]], names(captList)[i]] <- TRUE
  }

  empty <- rowSums(mat) == 0 ## which(apply(!mat, 1, all))

  mat.use <- mat[!empty,, drop = FALSE]
  mat.use
}

linkageError <- function(mat, prob){
  stopifnot(ncol(mat) == 2, prob < 1, prob > 0)
  sel.ind <- which(apply(mat, 1, sum) == 2)

  if(length(sel.ind) > 0 & round(prob * length(sel.ind)) > 0){
    mat.single <- mat[-sel.ind,]
    mat.double <- mat[sel.ind,]

    n.double.orig <- nrow(mat.double)
    n.double <- round((1 - prob) * n.double.orig)
    mat <- rbind(mat.single,
                 matrix(TRUE, nrow = n.double, ncol = 2),
                 matrix(c(FALSE, TRUE), nrow = n.double.orig - n.double, ncol = 2, byrow = TRUE),
                 matrix(c(TRUE, FALSE), nrow = n.double.orig - n.double, ncol = 2, byrow = TRUE))
  }
  mat
}

addPoll <- function(n, N, p=NULL, v=NULL, onlyPoll = FALSE){
  ## Ziehe n Personen aus der CH Bevölkerung
  tmp <- sample(x = 1:9e6, size = n, replace = FALSE)
  poll <- tmp[tmp <= N] ## Personen mit Unfall in Umfrage

  out <- if(onlyPoll){
    x <- sum(poll <= N) ## Anzahl befragter Personen mit Unfall
    (x / n) * 9e6
  }else{
    captList.poll <- list(p = p, v = v, poll = poll)
    mat.use.poll <- createMat(captList.poll, n = N)
    closedp(mat.use.poll)$results["Mt", "abundance"]
  }
  out
}
```

Für die Berechnung des "realistischeren Beispiels" wird der erste Schleifendurchlauf der folgenden grösseren Simulation verwendet. Im Folgenden wird die grosse Simulation nur im ersten Schleifendurchlauf ausgeführt:

```

set.seed(124)

## Anzahl Unfälle total
## Szenario: Leichtverletzte, Erwachsene (rel. homogene capture prob.)
N <- 50000

## Grundidee der Simulation:
## - Die Personen der CH haben eine ID 1, 2, ..., 9*10^6 (=9e6)
## - Wir sortieren die Personen so, dass die ersten 1, 2, ..., N die Personen
##   mit Unfall sind
## - Von diesen sind die ersten 40% UVG versichert, die nächsten 40% KVG versichert
##   und der Rest Ausländer ohne Versicherung
## - Aus diesen Grundpopulationen werden Stichproben gezogen;
##   Polizei und Versicherungen ziehen nur aus den UNFÄLLEN
##   Umfrage zieht aus der GANZEN BEVÖLKERUNG (enthält also wenige Unfälle)

## Polizei: capture probability
p.p <- 0.2

## Versicherungen: capture probabilities
p.u <- 0.8 ## UVG
p.k <- 0.1 ## KVG
p.a <- 0.1 ## Ausländer

p.v_u <- 0.05 ## Spezialstatistik
p.v_k <- 0 ## Wird nicht erhoben
p.v_a <- 0 ## Wird nicht erhoben

u <- 1:(N * p.u) ## Population UVG
k <- (N * p.u + 1):(N * p.u + N * p.k) ## Population KVG
a <- tail(1:N, N * p.a) ## Population Ausländer
Nu <- length(u)
Nk <- length(k)
Na <- length(a)

## Simulation approach (copy-pasted from above) #####

nreps <- 100
results.all <- matrix(NA, nrow = nreps, ncol = 13)
colnames(results.all) <- c("CR", "CRInhomP", "CRInhomPV",
                          "CR.errL", "CR.errF", "CR.errFL", "CR.errFLinhom",
                          "CR.poll.small", "CR.poll.large", "CR.poll.police",
                          "poll.small", "poll.large", "poll.vlarge")

## prepare vector for injury severity per person
nL <- round(0.8*N)
nS <- N - nL
injSev <- c(rep("L", nL), rep("S", nS))
injProb <- ifelse(injSev == "L", 0.2, 0.8)
## Alternative 1: strong noise
## runif(N, min = 0.1, max = 0.9)
## Alternative 2: Two diffuse clusters
## c(runif(nL, min = 0.1, max = 0.3), runif(nS, min = 0.7, max = 0.9))

for(i in 1:nreps){
  if (i%10 == 0) cat("i=", i, "\n")
  ## Daten generieren #####
  ## Versicherungsliste: Zufallsauswahl aus entsprechender Population
  nvu <- round(N*p.u*p.v_u)
  nvk <- round(N*p.k*p.v_k)
  nva <- round(N*p.a*p.v_a)
  v <- c(sample(u, size = nvu, replace = FALSE),
         sample(k, size = nvk, replace = FALSE),
         sample(a, size = nva, replace = FALSE))

  ## Polizeiliste: Zufallsauswahl aus entsprechender Population
  np <- round(N*p.p)
  p <- sample(1:N, size = np, replace = FALSE)

```

```
## Polizeiliste mit inhomogener Auswahl
sampledInjProb <- sample(injProb)
pInhom <- sample(1:N, size = np, replace = FALSE, prob = sampledInjProb)

## Versicherung mit inhomogener Auswahl
vInhom <- c(sample(u, size = nvu, replace = FALSE, prob = sampledInjProb[1:Nu]),
  sample(k, size = nvk, replace = FALSE, prob = sampledInjProb[(Nu+1):(Nu+Nk)]),
  sample(a, size = nva, replace = FALSE, prob = sampledInjProb[(Nu+Nk+1):N]))

## Mit Capture-Recapture
captList <- list(p = p, v = v)
captListInhomP <- list(p = pInhom, v = v)
captListInhomPV <- list(p = pInhom, v = vInhom)

mat.use <- createMat(captList, N)
mat.useInhomP <- createMat(captListInhomP, N)
mat.useInhomPV <- createMat(captListInhomPV, N)

## CR
results.all[i, "CR"] <- closedp(mat.use, m = "Mt")$results["Mt", "abundance"]

## CR Polizei mit inhomogener Auswahl
results.all[i, "CRInhomP"] <- closedp(mat.useInhomP, m = "Mt")$results["Mt", "abundance"]

## CR Polizei und Versicherung mit inhomogener Auswahl
results.all[i, "CRInhomPV"] <- closedp(mat.useInhomPV, m = "Mt")$results["Mt", "abundance"]

## CR mit Linkage Error
## Gematchte Fälle werden in ungematchte Fälle umgewandelt
## D.h. eine Zeile TRUE-TRUE wird zu TRUE-FALSE und FALSE-TRUE
results.all[i, "CR.errL"] <- closedp(linkageError(mat.use, prob = 0.1),
  m = "Mt")$results["Mt", "abundance"]

## CR mit Filter Error: Zu "p" werden 10% weitere Fälle hinzugenommen,
## die nicht aus der Unfallpopulation stammen
schrott <- N+seq(1,round(0.1*length(p)))
captListF <- list(p = c(p, schrott), v = v)
mat.useF <- createMat(captListF, N+length(schrott))

results.all[i, "CR.errF"] <- closedp(mat.useF, m = "Mt")$results["Mt", "abundance"]

## CR mit Linkage und Filter Error
results.all[i, "CR.errFL"] <- closedp(linkageError(mat.useF, prob = 0.1),
  m = "Mt")$results["Mt", "abundance"]

## CR P und V INHOMOGEN mit Filter & Linkage Error wie oben
captListFinhom <- list(p = c(pInhom, schrott), v = vInhom)
mat.useFinhom <- createMat(captListFinhom, N+length(schrott))

results.all[i, "CR.errFLinhom"] <- closedp(linkageError(mat.useFinhom, prob = 0.1),
  m = "Mt")$results["Mt", "abundance"]

## Umfrage gelinkt mit Polizei und Versicherung
## n befragte Personen (100% Rücklauf angenommen)
results.all[i, "CR.poll.small"] <- addPoll(n = 5000, N = N, p = p, v = v)
results.all[i, "CR.poll.large"] <- addPoll(n = 10000, N = N, p = p, v = v)

## Umfrage und nur polizei linken
tmp <- sample(x = 1:9e6, size = 10000, replace = FALSE)
poll <- tmp[tmp <= N] ## Personen mit Unfall in Umfrage
captList.poll <- list(p = p, poll = poll)
mat.use.poll <- createMat(captList.poll, n = N)
results.all[i, "CR.poll.police"] <- closedp(mat.use.poll)$results["Mt", "abundance"]

## Nur Umfrage
results.all[i, "poll.small"] <- addPoll(n = 5000, N = N, onlyPoll = TRUE)
results.all[i, "poll.large"] <- addPoll(n = 10000, N = N, onlyPoll = TRUE)
results.all[i, "poll.vlarge"] <- addPoll(n = 50000, N = N, onlyPoll = TRUE)
}
```


Nun werden die benötigten Grössen aus dieser Simulation ausgelesen und im Bericht verwendet ("Multinomial estimation" wird verwendet):

```
## Für "realistisches Bsp", Kap. 5.3.2: i <- 1 ##
N ## unbekanntes Total
```

```
## [1] 50000
```

```
np ## P
```

```
## [1] 10000
```

```
nvu ## SU
```

```
## [1] 2000
```

```
table(p=mat.use[,1], v=mat.use[,2])
```

```
##          v
## p      FALSE TRUE
## FALSE      0 1597
## TRUE     9597 403
```

```
closedpCI.t(mat.use, m = "Mt")
```

```
##
## Number of captured units: 11597
##
## Poisson estimation and model fit:
## abundance stderr deviance df    AIC    BIC infoFit
## Mt    49627.8 2164.1      0 0 34.058 56.134    OK
##
## Multinomial estimation, 95% profile likelihood confidence interval:
## abundance infCL supCL infoCI
## Mt    49613.4 45637 54146.7    OK
```

Kap 7.2.1: Bedeutung der Stichprobengrösse

Im Folgenden werden die Abb. 5 in Kap. 7.2.1 zu Grunde liegenden Zahlen reproduziert. Zunächst wird eine Hilfsfunktion definiert:

```
## Hilfsfunktion
filter_vector <- function(vec, N, prob = 0.9) {
  # Keep all numbers smaller than N
  small_numbers <- vec[vec < N]

  # For numbers >= N, keep with probability prob
  large_numbers <- vec[vec >= N]
  keep_large <- as.logical(rbinom(length(large_numbers), 1, prob) == 1)
  filtered_large <- large_numbers[keep_large]

  # Combine and return the result
  return(c(small_numbers, filtered_large))
}
```

n = 5000

Es wird für alle angefragten Personen eine 100% Rücklaufquote angenommen.

```
set.seed(123)
N <- 50000 ## Anzahl tatsächlicher Unfälle
n <- 5000 ## Anzahl angefragter Personen
popSize <- 9e6 ## Grösse Bevölkerung
res <- rep(NA, 100)
for (i in 1:100) {
  ## Ziehe Personen, die befragt werden; Pers 1:N hatten tatsächlich Unfall
  tmp0 <- sample(x = seq(1, popSize), size = n, replace = FALSE)
  ## Pers ohne Unfall: Perfekte Rücklaufquote
  poll <- filter_vector(tmp0, N, prob = 1)
  npoll <- length(poll)

  x <- sum(poll <= N) ## Anzahl befragter Personen mit Unfall
  res[i] <- (x / npoll) * popSize
}

sqrt(mean( (res - N)^2 ))/N ## 19.5% von 50'000
```

```
## [1] 0.1952922
```

n = 10000

Es wird für alle angefragten Personen eine 100% Rücklaufquote angenommen.

```
set.seed(123)
N <- 50000 ## Anzahl tatsächlicher Unfälle
n <- 10000 ## Anzahl angefragter Personen
popSize <- 9e6 ## Grösse Bevölkerung
res <- rep(NA, 100)
for (i in 1:100) {
  ## Ziehe Personen, die befragt werden; Pers 1:N hatten tatsächlich Unfall
  tmp0 <- sample(x = seq(1, popSize), size = n, replace = FALSE)
  ## Pers ohne Unfall: Perfekte Rücklaufquote
  poll <- filter_vector(tmp0, N, prob = 1)
  npoll <- length(poll)

  x <- sum(poll <= N) ## Anzahl befragter Personen mit Unfall
  res[i] <- (x / npoll) * popSize
}

sqrt(mean( (res - N)^2 ))/N ## 12% von 50'000
```

```
## [1] 0.1199667
```

n = 50000

Es wird für alle angefragten Personen eine 100% Rücklaufquote angenommen.

```
set.seed(123)
N <- 50000 ## Anzahl tatsächlicher Unfälle
n <- 50000 ## Anzahl angefragter Personen
popSize <- 9e6 ## Grösse Bevölkerung
res <- rep(NA, 100)
for (i in 1:100) {
  ## Ziehe Personen, die befragt werden; Pers 1:N hatten tatsächlich Unfall
  tmp0 <- sample(x = seq(1, popSize), size = n, replace = FALSE)
  ## Pers ohne Unfall: Perfekte Rücklaufquote
  poll <- filter_vector(tmp0, N, prob = 1)
  npoll <- length(poll)

  x <- sum(poll <= N) ## Anzahl befragter Personen mit Unfall
  res[i] <- (x / npoll) * popSize
}

sqrt(mean( (res - N)^2 ))/N ## 5% von 50'000
```

```
## [1] 0.05028121
```

n=10000 inkl. "Selection Bias"

Personen mit Unfall haben eine Rücklaufquote von 100% (für sie sind Unfälle wichtig); Personen ohne Unfall haben eine Rücklaufquote von 80%.

```
set.seed(123)
N <- 50000 ## Anzahl tatsächlicher Unfälle
n <- 10000 ## Anzahl angefragter Personen
popSize <- 9e6 ## Grösse Bevölkerung
res <- rep(NA, 100)
for (i in 1:100) {
  ## Ziehe Personen, die befragt werden; Pers 1:N hatten tatsächlich Unfall
  tmp0 <- sample(x = seq(1, popSize), size = n, replace = FALSE)
  ## Pers ohne Unfall: kleinere Rücklaufquote
  poll <- filter_vector(tmp0, N, prob = 0.8)
  npoll <- length(poll)

  x <- sum(poll <= N) ## Anzahl befragter Personen mit Unfall
  res[i] <- (x / npoll) * popSize
}

sqrt(mean( (res - N)^2 ))/N ## 29% von 50'000
```

```
## [1] 0.288356
```

Kap 7.2.2: Simulation verschiedener Szenarien mit Capture-Recapture Methode

```
set.seed(124)

## Anzahl Unfälle total
## Szenario: Leichtverletzte, Erwachsene (rel. homogene capture prob.)
N <- 50000

## Grundidee der Simulation:
## - Die Personen der CH haben eine ID 1, 2, ..., 9*10^6 (=9e6)
## - Wir sortieren die Personen so, dass die ersten 1, 2, ..., N die Personen
##   mit Unfall sind
## - Von diesen sind die ersten 40% UVG versichert, die nächsten 40% KVG versichert
##   und der Rest Ausländer ohne Versicherung
## - Aus diesen Grundpopulationen werden Stichproben gezogen;
##   Polizei und Versicherungen ziehen nur aus den UNFÄLLEN
##   Umfrage zieht aus der GANZEN BEVÖLKERUNG (enthält also wenige Unfälle)

## Polizei: capture probability
p.p <- 0.2

## Versicherungen: capture probabilities
p.u <- 0.8 ## UVG
p.k <- 0.1 ## KVG
p.a <- 0.1 ## Ausländer

p.v_u <- 0.05 ## Spezialstatistik
p.v_k <- 0 ## Wird nicht erhoben
p.v_a <- 0 ## Wird nicht erhoben

u <- 1:(N * p.u) ## Population UVG
k <- (N * p.u + 1):(N * p.u + N * p.k) ## Population KVG
a <- tail(1:N, N * p.a) ## Population Ausländer
Nu <- length(u)
Nk <- length(k)
Na <- length(a)
```

```
## Simulation approach (copy-pasted from above) ###

nreps <- 100
results.all <- matrix(NA, nrow = nreps, ncol = 13)
colnames(results.all) <- c("CR", "CRInhomP", "CRInhomPV",
  "CR.errL", "CR.errF", "CR.errFL", "CR.errFLinhom",
  "CR.poll.small", "CR.poll.large", "CR.poll.police",
  "poll.small", "poll.large", "poll.vlarge")

## prepare vector for injury severity per person
nL <- round(0.8*N)
nS <- N - nL
injSev <- c(rep("L", nL), rep("S", nS))
injProb <- ifelse(injSev == "L", 0.2, 0.8)
## Alternative 1: strong noise
## runif(N, min = 0.1, max = 0.9)
## Alternative 2: Two diffuse clusters
## c(runif(nL, min = 0.1, max = 0.3), runif(nS, min = 0.7, max = 0.9))

for(i in 1:nreps){
  if (i%%10 == 0) cat("i=", i, "\n")
  ## Daten generieren ###
  ## Versicherungsliste: Zufallsauswahl aus entsprechender Population
  nvu <- round(N*p.u*p.v_u)
  nvk <- round(N*p.k*p.v_k)
  nva <- round(N*p.a*p.v_a)
  v <- c(sample(u, size = nvu, replace = FALSE),
    sample(k, size = nvk, replace = FALSE),
    sample(a, size = nva, replace = FALSE))

  ## Polizeiliste: Zufallsauswahl aus entsprechender Population
  np <- round(N*p.p)
  p <- sample(1:N, size = np, replace = FALSE)

  ## Polizeiliste mit inhomogener Auswahl
  sampledInjProb <- sample(injProb)
  pInhom <- sample(1:N, size = np, replace = FALSE, prob = sampledInjProb)

  ## Versicherung mit inhomogener Auswahl
  vInhom <- c(sample(u, size = nvu, replace = FALSE, prob = sampledInjProb[1:Nu]),
    sample(k, size = nvk, replace = FALSE, prob = sampledInjProb[(Nu+1):(Nu+Nk)]),
    sample(a, size = nva, replace = FALSE, prob = sampledInjProb[(Nu+Nk+1):N]))

  ## Mit Capture-Recapture
  captList <- list(p = p, v = v)
  captListInhomP <- list(p = pInhom, v = v)
  captListInhomPV <- list(p = pInhom, v = vInhom)

  mat.use <- createMat(captList, N)
  mat.useInhomP <- createMat(captListInhomP, N)
  mat.useInhomPV <- createMat(captListInhomPV, N)

  ## CR
  results.all[i, "CR"] <- closedp(mat.use, m = "Mt")$results["Mt", "abundance"]
  ## Für "realistischeres Bsp", Kap. 5.3.2: i <- 1 ##
  # N ## unbekanntes Total
  # np ## P
  # nvu ## SU
  # table(p=mat.use[,1], v=mat.use[,2])
  # closedpCI.t(mat.use, m = "Mt")

  ## CR Polizei mit inhomogener Auswahl
  results.all[i, "CRInhomP"] <- closedp(mat.useInhomP, m = "Mt")$results["Mt", "abundance"]

  ## CR Polizei und Versicherung mit inhomogener Auswahl
  results.all[i, "CRInhomPV"] <- closedp(mat.useInhomPV, m = "Mt")$results["Mt", "abundance"]
}
```

```
## CR mit Linkage Error
## Gematchte Fälle werden in ungematchte Fälle umgewandelt
## D.h. eine Zeile TRUE-TRUE wird zu TRUE-FALSE und FALSE-TRUE
results.all[i, "CR.errL"] <- closedp(linkageError(mat.use, prob = 0.1),
                                   m = "Mt")$results["Mt", "abundance"]

## CR mit Filter Error: Zu "p" werden 10% weitere Fälle hinzugenommen,
## die nicht aus der Unfallpopulation stammen
schrott <- N+seq(1,round(0.1*length(p)))
captListF <- list(p = c(p, schrott), v = v)
mat.useF <- createMat(captListF, N+length(schrott))

results.all[i, "CR.errF"] <- closedp(mat.useF, m = "Mt")$results["Mt", "abundance"]

## CR mit Linkage und Filter Error
results.all[i, "CR.errFL"] <- closedp(linkageError(mat.useF, prob = 0.1),
                                   m = "Mt")$results["Mt", "abundance"]

## CR P und V INHOMOGEN mit Filter & Linkage Error wie oben
captListFinhom <- list(p = c(pInhom, schrott), v = vInhom)
mat.useFinhom <- createMat(captListFinhom, N+length(schrott))

results.all[i, "CR.errFLinhom"] <- closedp(linkageError(mat.useFinhom, prob = 0.1),
                                   m = "Mt")$results["Mt", "abundance"]

## Umfrage gelinkt mit Polizei und Versicherung
## n befragte Personen (100% Rücklauf angenommen)
results.all[i, "CR.poll.small"] <- addPoll(n = 5000, N = N, p = p, v = v)
results.all[i, "CR.poll.large"] <- addPoll(n = 10000, N = N, p = p, v = v)

## Umfrage und nur polizei Linken
tmp <- sample(x = 1:9e6, size = 10000, replace = FALSE)
poll <- tmp[tmp <= N] ## Personen mit Unfall in Umfrage
captList.poll <- list(p = p, poll = poll)
mat.use.poll <- createMat(captList.poll, n = N)
results.all[i, "CR.poll.police"] <- closedp(mat.use.poll)$results["Mt", "abundance"]

## Nur Umfrage
results.all[i, "poll.small"] <- addPoll(n = 5000, N = N, onlyPoll = TRUE)
results.all[i, "poll.large"] <- addPoll(n = 10000, N = N, onlyPoll = TRUE)
results.all[i, "poll.vlarge"] <- addPoll(n = 50000, N = N, onlyPoll = TRUE)
}
```

```
## i= 10
## i= 20
## i= 30
## i= 40
## i= 50
## i= 60
## i= 70
## i= 80
## i= 90
## i= 100
```

```
## RMSE
sqrt(apply((results.all - N)^2, 2, mean))/N
```

##	CR	CRInhomP	CRInhomPV	CR.errL	CR.errF
##	0.04194671	0.04845410	0.31001014	0.12478824	0.11438447
##	CR.errFL	CR.errFLinhom	CR.poll.small	CR.poll.large	CR.poll.police
##	0.23311318	0.15809922	0.04223643	0.04095310	0.38938022
##	poll.small	poll.large	poll.vlarge		
##	0.17339250	0.11794812	0.05976150		


```
## Bias
apply(results.all - N, 2, mean)
```

```
##          CR      CRInhomP      CRInhomPV      CR.errL      CR.errF
##    216.9954    151.0227   -15459.1370    5795.3430    5238.6950
##    CR.errFL CR.errFLinhom CR.poll.small CR.poll.large CR.poll.police
##    11374.8773   -7784.1077    215.0379    261.3472    5622.9296
##    poll.small poll.large poll.vlarge
##    2380.0000    562.0000    122.8000
```

```
## SD
apply(results.all - N, 2, sd)
```

```
##          CR      CRInhomP      CRInhomPV      CR.errL      CR.errF
##    2096.589    2430.174    1137.426    2323.412    2306.248
##    CR.errFL CR.errFLinhom CR.poll.small CR.poll.large CR.poll.police
##    2555.754    1383.916    2111.428    2041.140    18733.243
##    poll.small poll.large poll.vlarge
##    8378.544    5900.142    3000.591
```

Kap. 7.2.2, Tab. 11 und Abb 10: Kollisionen Frauen 2021

```
dat2010 <- data.frame(nurSSUV = c(238,19,75,25,12),
                      schnitt = c(112,7,18,13,2),
                      nurASTRA = c(2770,412,767,529,84))
dat2010$verkehrsteilnahme <- factor(x = c("PW", "MRad", "Fahrrad", "Fussg", "Andere"))
vn <- levels(dat2010$verkehrsteilnahme) # "Andere"
res <- matrix(0, length(vn), 3)
rownames(res) <- vn
colnames(res) <- c("abundance", "infCL", "supCL")
for (i in vn) {
  vnIdx <- dat2010$verkehrsteilnahme %in% i

  mat.use <- rbind(unlist(c(1, 1, sum(dat2010[vnIdx, "schnitt"]))),
                  unlist(c(1, 0, sum(dat2010[vnIdx, "nurSSUV"]))),
                  unlist(c(0, 1, sum(dat2010[vnIdx, "nurASTRA"]))))

  ## 95%-CI
  res[i,] <- unlist(closedpCI.t(X = mat.use, m = "Mt", dtype = "hist",
                              dfreq = TRUE, alpha = 0.05)$CI[,c("abundance", "infCL", "supCL")])
}
res
```

```
##          abundance      infCL      supCL
## Andere "577.497809956787" "223.284332131157" ">1806"
## Fahrrad "4031.94385893603" "2787.10541200186" "6358.19050706452"
## Fussg "1562.49796762126" "1077.65657004747" "2589.55412069372"
## MRad "1524.99780908626" "908.177492900319" "3242.42338380905"
## PW "8992.32137797656" "7800.31062219626" "10525.9089406144"
```

Kap 7.2.2: Pilotberechnung Methode 2 (Linkage ASTRA-SSUV)

Variante 1: 100% “ungewiss” zu “nein”

```
## Daten der HHB ###
data_astra_ssuv_hhb <- read.csv("data_astra_ssuv_hhb.csv",
                                stringsAsFactors = TRUE)

## Daten aufteilen (Code von Patrizia)
library(tidyr)
library(dplyr)

##
## Attache Paket: 'dplyr'

## Die folgenden Objekte sind maskiert von 'package:stats':
##
##   filter, lag

## Die folgenden Objekte sind maskiert von 'package:base':
##
##   intersect, setdiff, setequal, union

astra_2010 <- data_astra_ssuv_hhb %>%
  filter(quelle == "ASTRA" & unfalljahr==2010 & altersgruppe == "17-64" & berufsfahrt == 0) %>%
  group_by(verkehrsteilnahme) %>%
  summarise(anzahl = sum(nofcases))

library(forcats)

# We will make two little dataframes out of the "unknown" cases, one with
# some fraction of "No", and the other with remaining cases as "Yes".

## If fraction_unknown_to_no set to 1, then all unknowns become "No"
## (results same as before, confirmed by sanity check test)
fraction_unknown_to_no <- 1

tmp_ssuv_2010_focus <- data_astra_ssuv_hhb %>%
  filter(quelle == "SSUV" & unfalljahr == 2010 & altersgruppe == "17-64" &
         rentenfall == "Nein" & berufsfahrt == 0)

tmp_ssuv_2010_unknowns <- tmp_ssuv_2010_focus %>%
  filter(polizeirapport_binaer_ssuv == "Ungewiss, unbekannt, keine Anhaltspunkte")
tmp_ssuv_2010_unknowns_as_no <- tmp_ssuv_2010_unknowns %>%
  mutate(polizeirapport = "Nein",
         nofcases = ceiling(nofcases * fraction_unknown_to_no))
tmp_ssuv_2010_unknowns_as_yes <- tmp_ssuv_2010_unknowns %>%
  mutate(polizeirapport = "Ja",
         nofcases = nofcases - ceiling(nofcases * fraction_unknown_to_no))

## Add column to the known "yes" / "no" in order to merge dfs
tmp_ssuv_2010_knowns <- tmp_ssuv_2010_focus %>%
  filter(polizeirapport_binaer_ssuv != "Ungewiss, unbekannt, keine Anhaltspunkte") %>%
  mutate(polizeirapport = policeirapport_binaer_ssuv)

tmp_ssuv_2010_combined <- rbind(
  tmp_ssuv_2010_knowns,
  tmp_ssuv_2010_unknowns_as_no,
  tmp_ssuv_2010_unknowns_as_yes)
```

```
## Desired aggregate data frame as before, with the correct fraction of unknowns
## marked as "no" or "yes"
ssuv_2010 <- tmp_ssuv_2010_combined %>%
  group_by(verkehrsteilnahme, polizeiRapport) %>%
  summarise(anzahl = sum(nofcases)) %>%
  pivot_wider(names_from = polizeiRapport, values_from = anzahl)
```

```
## `summarise()` has grouped output by 'verkehrsteilnahme'. You can override using
## the `.groups` argument.
```

```
ssuv_2010
```

```
## # A tibble: 5 x 3
## # Groups:   verkehrsteilnahme [5]
##   verkehrsteilnahme      Ja  Nein
##   <fct>              <dbl> <dbl>
## 1 Andere              9    77
## 2 Fahrrad / E-Bike    51   692
## 3 Fussgänger (inkl. fäG) 27    42
## 4 Motorrad           86   362
## 5 Personenwagen     338   546
```

```
astra_ssuv_2010 <- ssuv_2010 %>%
  left_join(astra_2010, by = c("verkehrsteilnahme")) %>%
  rename(astra = anzahl) %>%
  mutate(nurAstra = astra - Ja)
```

```
astra_ssuv_2010
```

```
## # A tibble: 5 x 5
## # Groups:   verkehrsteilnahme [5]
##   verkehrsteilnahme      Ja  Nein astra nurAstra
##   <fct>              <dbl> <dbl> <int>   <dbl>
## 1 Andere              9    77   561    552
## 2 Fahrrad / E-Bike    51   692  2257   2206
## 3 Fussgänger (inkl. fäG) 27    42  1276   1249
## 4 Motorrad           86   362  3728   3642
## 5 Personenwagen     338   546 10414  10076
```

```
# fraction_unknown_to_no <- 1
# > astra_ssuv_2010
## # A tibble: 5 x 5
## # Groups:   verkehrsteilnahme [5]
##   verkehrsteilnahme      Ja  Nein astra nurAstra
##   <fct>              <dbl> <dbl> <int>   <dbl>
## 1 Andere              9    77   561    552
## 2 Fahrrad / E-Bike    51   692  2257   2206
## 3 Fussgänger (inkl. fäG) 27    42  1276   1249
## 4 Motorrad           86   362  3728   3642
## 5 Personenwagen     338   546 10414  10076

vn <- levels(astra_ssuv_2010$verkehrsteilnahme) #"Andere"
res <- matrix(0, length(vn), 3)
rownames(res) <- vn
colnames(res) <- c("abundance", "infCL", "supCL")
for (i in vn) {
  vnIdx <- astra_ssuv_2010$verkehrsteilnahme %in% i

  mat.use <- rbind(unlist(c(1, sum(astra_ssuv_2010[vnIdx, "Ja"]))),
                  unlist(c(1, 0, sum(astra_ssuv_2010[vnIdx, "Nein"]))),
                  unlist(c(0, 1, sum(astra_ssuv_2010[vnIdx, "nurAstra"]))))

  ## consider 99% (!) - CI => used in Bonferroni correction later
  res[i,] <- closedCI.t(X = mat.use, m = "Mt", dtype = "hist",
                       dfreq = TRUE, alpha = 0.05)$CI[,c("abundance", "infCL", "supCL")]
}
res
```



```
##          abundance    infCL    supCL
## Andere          5325.222  3102.465 10729.493
## Fahrrad / E-Bike 32852.482 25618.354 43298.013
## Fussgänger (inkl. fäG) 3236.480 2498.661 4479.928
## Motorrad         19396.500 16217.342 23619.635
## Personenwagen    27220.402 25132.687 29637.230
```

```
colSums(res)
```

```
## abundance    infCL    supCL
## 88031.09    72569.51 111764.30
```

```
## OHNE BERUFSUNFÄLLE ####
# fraction_unknown_to_no <- 1
# abundance    infCL    supCL
# Andere          5325.222  3102.465 10729.493
# Fahrrad / E-Bike 32852.482 25618.354 43298.013
# Fussgänger (inkl. fäG) 3236.480 2498.661 4479.928
# Motorrad         19396.500 16217.342 23619.635
# Personenwagen    27220.402 25132.687 29637.230
# > colSums(res)
# abundance    infCL    supCL
# 88031.09    72569.51 111764.30
```

Variante 2:

```
## Daten der HHB ####
data_astra_ssuv_hhb <- read.csv("data_astra_ssuv_hhb.csv",
                                stringsAsFactors = TRUE)

## Daten aufteilen (Code von Patrizia)
library(tidyr)
library(dplyr)
astra_2010 <- data_astra_ssuv_hhb %>%
  filter(quelle == "ASTRA" & unfalljahr==2010 & altersgruppe == "17-64" & berufsfahrt == 0) %>%
  group_by(verkehrsteilnahme) %>%
  summarise(anzahl = sum(nofcases))

library(forcats)

# We will make two little dataframes out of the "unknown" cases, one with
# some fraction of "No", and the other with remaining cases as "Yes".

## If fraction_unknown_to_no set to 1, then all unknowns become "No"
## (results same as before, confirmed by sanity check test)
fraction_unknown_to_no <- 0.9

tmp_ssuv_2010_focus <- data_astra_ssuv_hhb %>%
  filter(quelle == "SSUV" & unfalljahr == 2010 & altersgruppe == "17-64" &
         rentenfall == "Nein" & berufsfahrt == 0)

tmp_ssuv_2010_unknowns <- tmp_ssuv_2010_focus %>%
  filter(polizeirapport_binaer_ssuv == "Ungewiss, unbekannt, keine Anhaltspunkte")
tmp_ssuv_2010_unknowns_as_no <- tmp_ssuv_2010_unknowns %>%
  mutate(polizeirapport = "Nein",
         nofcases = ceiling(nofcases * fraction_unknown_to_no))
tmp_ssuv_2010_unknowns_as_yes <- tmp_ssuv_2010_unknowns %>%
  mutate(polizeirapport = "Ja",
         nofcases = nofcases - ceiling(nofcases * fraction_unknown_to_no))

## Add column to the known "yes" / "no" in order to merge dfs
tmp_ssuv_2010_knowns <- tmp_ssuv_2010_focus %>%
  filter(polizeirapport_binaer_ssuv != "Ungewiss, unbekannt, keine Anhaltspunkte") %>%
  mutate(polizeirapport = policeirapport_binaer_ssuv)
```

```
tmp_ssuv_2010_combined <- rbind(
  tmp_ssuv_2010_knowns,
  tmp_ssuv_2010_unknowns_as_no,
  tmp_ssuv_2010_unknowns_as_yes)

## Desired aggregate data frame as before, with the correct fraction of unknowns
## marked as "no" or "yes"
ssuv_2010 <- tmp_ssuv_2010_combined %>%
  group_by(verkehrsteilnahme, polizeiRapport) %>%
  summarise(anzahl = sum(nofcases)) %>%
  pivot_wider(names_from = polizeiRapport, values_from = anzahl)
```

`summarise()` has grouped output by `verkehrsteilnahme`. You can override using
the `.groups` argument.

ssuv_2010

```
## # A tibble: 5 x 3
## # Groups:   verkehrsteilnahme [5]
##   verkehrsteilnahme      Ja  Nein
##   <fct>              <dbl> <dbl>
## 1 Andere              12    74
## 2 Fahrrad / E-Bike     82   661
## 3 Fussgänger (inkl. fäG) 29    40
## 4 Motorrad           105   343
## 5 Personenwagen       378   506
```

```
astra_ssuv_2010 <- ssuv_2010 %>%
  left_join(astra_2010, by = c("verkehrsteilnahme")) %>%
  rename(astra = anzahl) %>%
  mutate(nurAstra = astra - Ja)
```

astra_ssuv_2010

```
## # A tibble: 5 x 5
## # Groups:   verkehrsteilnahme [5]
##   verkehrsteilnahme      Ja  Nein astra nurAstra
##   <fct>              <dbl> <dbl> <int>   <dbl>
## 1 Andere              12    74   561     549
## 2 Fahrrad / E-Bike     82   661  2257    2175
## 3 Fussgänger (inkl. fäG) 29    40  1276    1247
## 4 Motorrad           105   343  3728    3623
## 5 Personenwagen       378   506 10414   10036
```

```
# fraction_unknown_to_no <- 0.9
# > astra_ssuv_2010
## # A tibble: 5 x 5
## # Groups:   verkehrsteilnahme [5]
##   verkehrsteilnahme      Ja  Nein astra nurAstra
##   <fct>              <dbl> <dbl> <int>   <dbl>
## 1 Andere              12    74   561     549
## 2 Fahrrad / E-Bike     82   661  2257    2175
## 3 Fussgänger (inkl. fäG) 29    40  1276    1247
## 4 Motorrad           105   343  3728    3623
## 5 Personenwagen       378   506 10414   10036

vn <- levels(astra_ssuv_2010$verkehrsteilnahme) #"Andere"
res <- matrix(0, length(vn), 3)
rownames(res) <- vn
colnames(res) <- c("abundance", "infCL", "supCL")
for (i in vn) {
  vnIdx <- astra_ssuv_2010$verkehrsteilnahme %in% i

  mat.use <- rbind(unlist(c(1, 1, sum(astra_ssuv_2010[vnIdx, "Ja"]))),
    unlist(c(1, 0, sum(astra_ssuv_2010[vnIdx, "Nein"]))),
    unlist(c(0, 1, sum(astra_ssuv_2010[vnIdx, "nurAstra"]))))
```

```
## consider 99% (!) - CI => used in Bonferroni correction later
res[i,] <- closedpCI.t(X = mat.use, m = "Mt", dtype = "hist",
                      dfreq = TRUE, alpha = 0.05)$CI[,c("abundance", "infCL", "supCL")]
}
res
```

```
##          abundance    infCL    supCL
## Andere          3994.041  2517.091  7174.214
## Fahrrad / E-Bike 20432.830 16867.322 25196.489
## Fussgänger (inkl. fäG) 3013.309 2363.181 4092.308
## Motorrad        15886.748 13571.412 18882.796
## Personenwagen   24339.989 22638.749 26296.428
```

```
colSums(res)
```

```
## abundance    infCL    supCL
##  67666.92  57957.76 81642.23
```

```
# fraction_unknown_to_no <- 0.9
# abundance    infCL    supCL
# Andere          3994.041  2517.091  7174.214
# Fahrrad / E-Bike 20432.830 16867.322 25196.489
# Fussgänger (inkl. fäG) 3013.309 2363.181 4092.308
# Motorrad        15886.748 13571.412 18882.796
# Personenwagen   24339.989 22638.749 26296.428
# > colSums(res)
# abundance    infCL    supCL
#  67666.92  57957.76 81642.23
```

IV Aktennotiz Stakeholder-Workshop II

MEDUSA: Stakeholderworkshop II

ECOPLAN/BFU/ETHZ

MEDUSA: Stakeholderworkshop II – Aktennotiz

Datum	3. Juli 2024, 09:20 – 12:15 Uhr
Anwesende Personen	– Christian Kamenik (ASTRA) – Josephine Leuba (ARE) – Werner Brucks (Zürich) – Stefan Scholz (SUVA) – Alexandra Quandt (BFS) – Nicole Felder (ASTRA) – Markus von Runkel (Kapo BL – Stv. für Christian Egeler) – Steffen Niemann (BFU) – Patrizia Hertach (BFU), Leitung – Esther Luttrupp (BFU) – Lukas Meier (ETHZ) – Markus Kalisch (ETHZ) – Matthias Amacher (Ecoplan), Moderation – Martin Wagenbach (Ecoplan)
Entschuldigt	– Christian Egeler (Kapo BL)
Ort	BFU – Hodlerstrasse 5a
Traktanden	1. Start / Einleitung 2. Methodenvarianten 3. Bedürfnisse – Differenzierung Dunkelziffer nach Variante 4. «Stellschrauben» der Detailumsetzung und Trade-offs 5. Ausblick

Zeit (Dauer)	Traktandum
09:20	Eintreffen der Teilnehmenden (TN)
09:30 (5')	1. Start / Einleitung
09:35 (45')	2. Methodenvarianten
10:25 (45')	3. Bedürfnisse - Differenzierung Dunkelziffer nach Variante
11:10 (20')	Pause (20 Min)
11:30 (45')	4. «Stellschrauben» der Detailumsetzung und Trade-offs
12:15 (15')	5. Ausblick
12:30	Schluss

1 Start / Einleitung

- P. Hertach begrüsst die Teilnehmende des Workshops und stellt den Stand des Projekts vor. Es nehmen alle am 1. Stakeholderworkshop Anwesenden teil oder konnten eine Stellvertretung organisieren.
- M. Amacher stellt die Ziele und den Workshop-Ablauf vor.

2 Methodenvarianten

Input Forschungsteam (vgl. Präsentationsfolien für weitere Details)

- P. Hertach stellt nach einer kurzen Einleitung (Vorstellung der vier Varianten, relevante Dimensionen / Würfelemente) die Methodenvarianten einzeln vor. Dabei nennt sie die wichtigsten Eigenschaften und Annahmen.
- Ebenfalls erwähnt werden die aktuell verworfenen Methoden (basierend auf Patienten-/Spitaldaten sowie reinem «Machine Learning»)

Diskussion

Die TN werden gebeten, offene Fragen zu den vorgestellten Methodenvarianten zu formulieren und zu notieren [vgl. Abbildung 2-1 und Abbildung 2-2]. Am Plakat werden die Fragen entgegengenommen und kurz diskutiert / eingeordnet.

Zusammenfassung der Diskussionsergebnisse:

- Die Dunkelziffer soll **einfach kommunizierbar** sein. Journalisten interessiert in erster Linie das Endresultat und weniger die Methodik. Bei Interesse oder spezifischen Nachfragen muss aber die Methodik auch klar und transparent kommunizierbar sein. Wenn das Resultat zu ungenau ist (z.B. zu grosses Vertrauensintervall), ist zu prüfen, ob die Zahl überhaupt kommuniziert werden soll (etwa bei einzelnen Würfelementen).
- Thema **Datenschutz**: Die Resultate sollten «so öffentlich wie möglich» sein. Herausforderungen bestehen insbesondere bei Datenverknüpfungen. Methodisch liegen mit kryptografischen Vorgehen bereits Lösungen vor, die im ASTRA auch schon angewendet werden. Zudem kann Art. 39 des Datenschutzgesetzes als eine Art «Ermöglicher» von solchen Datenverknüpfungen zu Forschungszwecken betrachtet werden.
- Bei den allgemeinen Themen wurde auch die **Kombination von Methoden** und die **Vergleichbarkeit** angesprochen. Bei gewissen (unterschiedlichen) Würfelementen können unterschiedliche Methoden oder Kombinationen sinnvoll sein. Bzgl. Vergleichbarkeit ist insbesondere wichtig, dass bei gleichem Filter die Methoden oder Datengrundlagen kompatibel sind. Bei der Linkage/CR-Methode ist dies zentral.
- **Stichprobe / Befragung**: Verfügbare Erhebungen wie der MZMV oder die SGB könnten sich für die Befragung eignen. Das Einbringen von Frageblöcken zu Strassenverkehrsunfällen in diese bestehenden Erhebungen kann den Aufwand reduzieren, der bei Befragungen hoch ist (insbesondere im Vergleich mit den prinzipiell weniger aufwändigen Linkage/CR-Ansätzen). Mögliche Synergien und andere Nutzen sind bei Befragungen ebenfalls zu prüfen. Die Stakeholder wünschen sich pragmatische Empfehlungen zur Umsetzung der Variante (neue Erhebung? Anknüpfen -> wie?). Auch relevant ist, wie ein gezieltes Sampling von spezifischen Würfelementen erreicht werden kann.
- **Linkage ASTRA-SSUV** (und evtl. weitere Quellen / KVG): Bei der SSUV laufen Arbeiten, um die Stichprobe von heute 5% zu vergrössern. Dies erhöht die Qualität der Methodenvariante.
- **Linkage ASTRA-SSUV kombiniert mit Stichprobe / Befragung**: Die Frage stellt sich, wie viel grösser der Aufwand mit der zusätzlichen Befragung ist, gegenüber dem Nutzen. Mit der Befragung fällt der grosse Vorteil der Linkage-Methode (sehr geringer Aufwand, wenn das Setting stimmt) teilweise weg. Um die Vergleichbarkeit sicherzustellen, müssen die Filter übereinstimmen. Es stellt sich die Frage, ob die Medizinische Statistik der Krankenhäuser (MSK) trotzdem verwendet werden könnte. Dabei wird darauf hingewiesen, dass in dieser Statistik Strassenverkehrsunfälle nicht zuverlässig identifiziert werden können, wodurch sie für eine Abschätzung der

Dunkelziffer nicht in Frage kommt. Ausserdem sind in dieser Statistik kaum Informationen zu den Umständen des Verkehrsunfalls vorhanden sind, die für einen Filter von Würfelementen genutzt werden könnten. Für eine genauere/bessere Kategorisierung der Verletzungsschwere in der ASTRA-Statistik und Abstimmung mit den Daten der SSUV wäre es aber denkbar, die MSK hinzuzuziehen. Es muss aber bedacht werden, dass eher schwerere Verletzungen in der MSK-Statistik zu finden sind (stationäre Aufenthalte). Aktuell können gemäss Aussagen ASTRA ca. 60% der Verletzten in der ASTRA-Statistik mit der MSK gelinkt werden. Eine weitere Info hierzu: in Zukunft soll in der MSK die AHV-Nr. aufgenommen werden. Auch bei der ASTRA-Statistik ist man hier dran, diese Information aufnehmen zu können. Dies wird einen positiven Einfluss auf die Güte des Linkage haben.

- **Versicherungsliste SSUV-KVG:** Auch die ergänzte «Versicherungsliste» könnte mit anderen Varianten kombiniert werden.

Abbildung 2-1: Offene Fragen der Stakeholder zu Methodenvarianten – alle Methoden

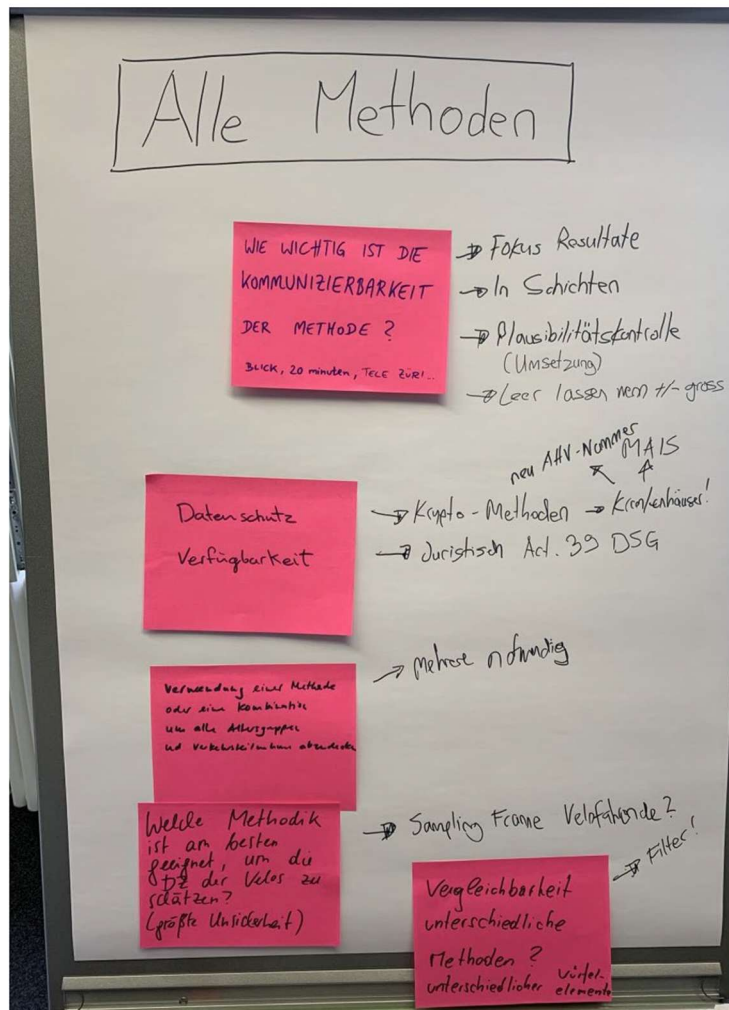


Abbildung 2-2: Offene Fragen der Stakeholder zu den Methodenvarianten - Variantspezifisch

Methodenvariante 1:
Stichprobe (Befragung)
vs. ASTRA (nur Filter)

Befragung:
Aufwand?
Andere Nutzung?
↳ Synergien?

Welche bestehende Erhebung eignet sich für die Befragung?
→ HEMV
→ SGB
↳ verfügbare Datenquellen

Methodenvariante 2:
Linkage ASTRA-SSUV
(Capture-Recapture CR, 2 Quellen)

(+ KVG)

*SSUV → Grössere Stichprobe
Prospektive Vollerhebung
↳ sich dran...*

Methodenvariante 3:
Linkage ASTRA-SSUV-Befragung
(CR, 3 Quellen)

(+ KVG)

Aufwand V3 ggü. V2 viel grösser?

Stichprobenziehung bei Befragungen (mit Infos aus anderen Datenquellen) 2 Bsp. med. Statistik Krankenhäuser

Gleiche Differenzierung / Kategorien? ↳ MUSS für Filter!

Methodenvariante 4:
Versicherungsliste (SSUV + KVG)
vs. ASTRA (nur Filter, ohne Linkage/CR)

V4 mit anderen Varianten kombinierbar?
→ Ja, M2+3 mit Linkage
→ spez. Würfelknoten

3 Bedürfnisse – Differenzierung Dunkelziffer nach Variante

Input Forschungsteam

- (vgl. Präsentationsfolien für weitere Details: Folien 15 bis 22)

Diskussion

- Bei der Verletzungsschwere gibt es Unterschiede zwischen den Datenquellen ASTRA und SSUV. Für eine Differenzierung der Dunkelziffer nach Verletzungsschwere müssten diese vergleichbar (Filter auf gleiche Verletzungsschwere) gemacht werden. Einzelne TN gehen davon aus, dass sich die ASTRA-Verletzungsschweren (Basis NACA) allenfalls über einen gemeinsamen Nenner AIS/MAIS in die SSUV-Schweren überführen lassen.
- Aus SSUV lassen sich E-Bikes und Velos nicht zuverlässig trennen. Eine solche Differenzierung entspricht aber einem Bedürfnis. Zudem wird festgehalten, dass sich die Mobilität und die eingesetzten Verkehrsmittel auch aufgrund von technologischen und gesellschaftlichen Entwicklungen immer wieder verändert. Die Statistik kann diese Entwicklung nicht immer sofort abbilden oder voraussehen. Vor 10 Jahren waren E-Trotts noch kein Thema.
- Die Simulationsresultate zeigen, dass Fehler auch beim Linkage-Ansatz zu grösseren Abweichungen bei den Schätzungen führen können. Hier gilt es festzuhalten, dass Fehler bei jedem der drei Schritte (Filter, Linkage, Capture-Recapture) auftreten können.
- Im Hinblick auf eine zukünftige DZ-Berechnung anhand der Linkage Methode ist eine Vereinheitlichung der Filter zwischen den Datenquellen sinnvoll.
- Der Linkage Ansatz ist zwar sehr attraktiv bei tiefen Fehlern, aber die Abweichungen nehmen bei Fehlern rasch zu. Im Vergleich zur Stichprobe kann der Schätzfehler zudem schlecht quantifiziert werden.
- Bei Verfeinerung der Sampling Frames (Fokus auf spez. Würfelemente bei gleicher Stichprobengrösse) können die auf Befragungen basierenden Methoden noch genauer werden. Stichproben / Befragungen sind daher (trotz hohem Aufwand) eine gangbare Lösung. Ein geeignetes Sampling Frame für Velofahrende oder Mikromobilität gibt es jedoch aktuell nicht.
- Schwer- VS Leichtverletzte: Man könnte sich auf Schwerverletzte mit hohen Kosten pro Fall konzentrieren. Aber: Leichtverletzte (erfasst wenn ein Arzt besucht wird) haben voraussichtlich eine deutlich höhere DZ, und haben in der Masse ebenfalls einen bedeutenden Anteil an den externen Unfallkosten des Verkehrs.

4 Stellschrauben der Detailumsetzung und Trade-Offs

Input Forschungsteam

Die TN werden vor der Pause gebeten, im Dreieck mit Tradeoffs «hohe Qualität / Belastbarkeit der DZ», «Geringer Aufwand» und «starke Differenzierung» einen Punkt zu setzen gemäss ihrer Prioritätensetzung bzw. Gewichtung. Es zeigen sich zwei Gruppen [vgl. Abbildung 4-1]:

- a) Priorität auf *hohe Qualität / Belastbarkeit* mit Berücksichtigung des Aufwands
- b) Priorität auf *hohe Qualität / Belastbarkeit* mit *starker Differenzierung*

Die grosse Mehrheit der Stakeholder spricht sich für a) aus. Hohe Priorität hat das Publizieren einer belastbaren Zahl. Auch wichtig zu wissen ist, wie hoch der Fehler ist. Der Aufwand soll dabei verhältnismässig sein.

Bei b) wird insbesondere die Differenzierung Velo und E-Bike angesprochen. Eine gewisse Differenzierung ist wichtig, damit die DZ auch effektiv verwendet werden kann. Der nötige Aufwand wurde weniger stark gewichtet.

Im Rahmen der weiteren Arbeiten ist zu prüfen, inwiefern beide Bedürfnisse (mit allenfalls unterschiedlichen Methoden) bestmöglich erfüllt werden können.

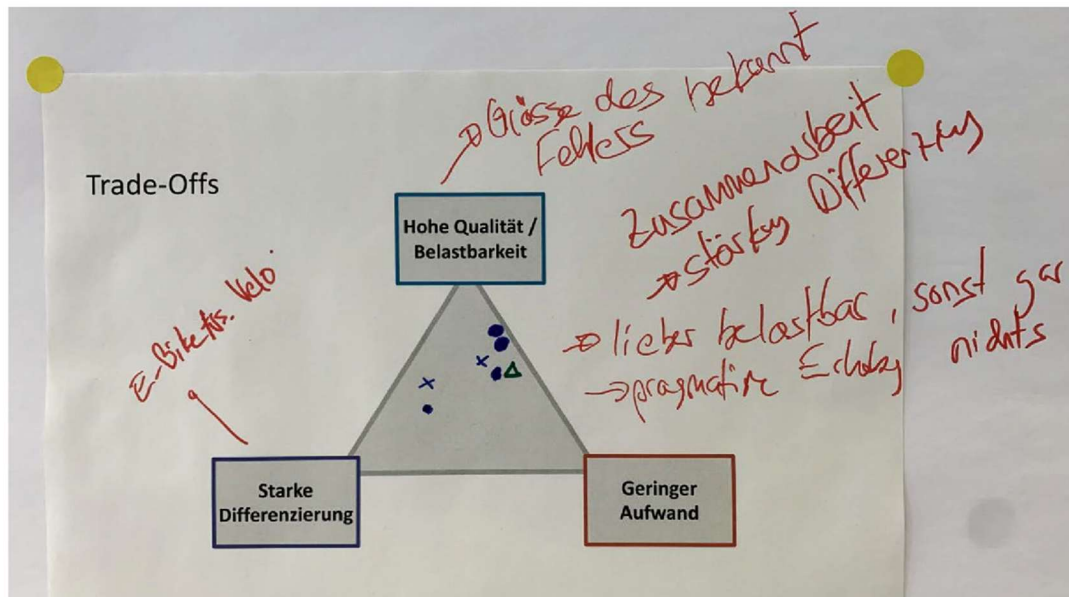
M. Amacher stellt mögliche Parameter vor, die in den Methodenvarianten angepasst werden können (vgl. Präsentationsfolien).

Diskussion

In der Diskussion wurden zudem folgende Punkte erwähnt:

- Die Erschliessung der KVG stellt eine grosse Hürde dar. Vergangene Arbeiten haben gezeigt, dass private Versicherer (u.a. Krankenversicherer) kein grosses Interesse haben, mitzuwirken und ihre Daten zu teilen. Um eine «Versicherungsliste» (UVG/KVG) erstellen zu können, braucht es voraussichtlich bei der KVG eine rechtliche Pflicht.
- Als weitere Stellschraube (für die Linkage Variante) gilt die Qualität des Filters bei der Verletzungsschwere. ASTRA Verletzungsschweren basierend auf NACA können in AIS/MAIS umkodiert werden, SSUV-Daten können über ICD nach AIS/MAIS umgerechnet werden, was die beiden Datensätze vergleichbar macht.
- Für eine ausreichende Belastbarkeit der DZ sind «kleine» Stichproben keine gute Lösung. In diesem Sinne ist die Reduktion der Stichprobengrösse zur Eindämmung der Kosten keine gute Stellschraube bei der Befragung.
- Eine geografische Abgrenzung wird verworfen. Bereits auf übergeordneter Ebene gibt es Schwierigkeiten, eine genügend grosse Grundgesamtheit zu erreichen. Zudem besteht das Problem, die Erkenntnisse aus einem geografischen Raum auf alle anderen zu Übertragen.
- Betreffend Empfehlungen soll in Betracht gezogen werden, differenzierte Aussagen zu machen (vgl. Trade-off Diskussion und die unterschiedlichen Prioritäten der Stakeholder). Auch bei der Datenerhebung kann eine Differenzierung sinnvoll sein (je nach Würfelement oder Priorität eine unterschiedliche Differenzierung festlegen).
- Durch Pooling mehrerer Jahresdatensätze (z.B. SSUV) könnte die Qualität erhöht werden.
- Ausblick / weiterer Forschungsbedarf: Bei den zentralen Datenquellen / beim Linkage (Erhöhung Stichprobe, Vereinheitlichung der Abgrenzungen bei Verletzungsschwere, Erfassung Fahrzeugkategorien und Unfallmerkmale) sollen in Zukunft Verbesserungen angestrebt werden (z. B. AHV-Nummer als ID prüfen).
- Wenn unterschiedliche Daten kombiniert/erhoben werden, ist es für die Umsetzung empfehlenswert, dass die Daten bei einer einzigen Institution «zusammenlaufen», sodass eine optimale Koordination ermöglicht ist (für Teil Empfehlungen).

Abbildung 4-1: Bedürfnisse – Priorisierung



5 Ausblick

- P. Hertach stellt die weiteren Schritte vor (vgl. Präsentationsfolien).
- Die Stakeholder haben keine weiteren Rückmeldungen.

Abkürzungsverzeichnis

Begriff	Bedeutung
AHV	Alters- und Hinterlassenenversicherung
AIS	Abbreviated Injury Scale (vereinfachte Verletzungsskala)
ARE	Bundesamt für Raumentwicklung
ASTRA	Bundesamt für Strassen
BFS	Bundesamt für Statistik
BFU	Beratungsstelle für Unfallverhütung
BK	Begleitkommission
CARE	Critical Analysis Reporting Environment
CAWI	Computer-Assisted Web Interview
CR	Capture-Recapture
EDR	Event Data Recorder (Unfalldatenschreiber)
EU-IDB	European Injury Database
IRTAD	International Traffic Safety Data and Analysis Group
ISS	Injury Severity Score
ISSI	Infrastruktur-Sicherheitsinstrumente
KVG	Bundesgesetz über die Krankenversicherung
MAIS	Maximum Abbreviated Injury Scale
MSK	Medizinische Statistik der Krankenhäuser
MZMV	Mikrozensus Mobilität und Verkehr
NACA	National Advisory Committee for Aeronautics
PET	Post encroachment time
RMSE	Root mean squared error
SGB	Schweizerische Gesundheitsbefragung
SRPH	Stichprobenrahmen für Personen- und Haushaltserhebungen
SSUV	Sammelstelle für die Statistik der Unfallversicherung UVG
Suva	Schweizerische Unfallversicherungsanstalt
SVG	Strassenverkehrsgesetz
SVU-Statistik	Statistik der polizeilich registrierten Strassenverkehrsunfälle des ASTRA
TTC	Time to collision
UAP	Unfallaufnahmeprotokoll
UVG	Bundesgesetz über die Unfallversicherung
VRV	Verkehrsregelnverordnung
WHO	World Health Organisation (Weltgesundheitsorganisation)

Literaturverzeichnis

- [1] Bundesamt für Strassen ASTRA. Unfallaufnahmeprotokoll 2018; 2018. <https://www.astra.admin.ch/astra/de/home/dokumentation/unfalldaten/grundlagen/unfallerfassung.html>. Zugriff am 25.01.2021.
- [2] Niemann S, Lieb C, Sommer H. *Nichtberufsunfälle in der Schweiz: Aktualisierte Hochrechnung und Kostenberechnung*. Bern: BFU, Beratungsstelle für Unfallverhütung; 2015. bfu-Report 71.
- [3] PIARC World Road Association. *Road safety manual: Data quality and under-reporting*. <https://roadsafety.piarc.org/en/road-safety-management-safety-data/quality-and-under-reporting>. Zugriff am 10.12.2022.
- [4] Amoros E, Martin J-L, Lafont S, Laumon B. Actual incidences of road casualties, and their injury severity, modelled from police and hospital data, France. *Eur J Public Health*. 2008; 18(4): 360–365. DOI:10.1093/eurpub/ckn018.
- [5] Uhr A, Hertach P. *Verkehrssicherheit von E-Bikes mit Schwerpunkt Alleinunfälle*. Bern: BFU, Beratungsstelle für Unfallverhütung; 2017. bfu-Report 75. DOI:10.13100/bfu.2.340.01.
- [6] Hertach P, Uhr A, Niemann S, Cavegn M. Characteristics of single-vehicle crashes with e-bikes in Switzerland. *Accid Anal Prev*. 2018; 117: 232–238. DOI:10.1016/j.aap.2018.04.021.
- [7] Bundesamt für Strassen ASTRA. *Instruktionen zum Ausfüllen des Unfallaufnahmeprotokolls UAP2018*. Bern: ASTRA; 2018.
- [8] Birkel C, Church D, Hummelsheim-Doss D et al. *Der Deutsche Viktimisierungssurvey 2017: Opfererfahrungen, kriminalitätsbezogene Einstellungen sowie die Wahrnehmung von Unsicherheit und Kriminalität in Deutschland*. Wiesbaden: Bundeskriminalamt BKA; 2019.
- [9] Eifler S, Pollich D, Hg. *Empirische Forschung über Kriminalität: Methodologische und methodische Grundlagen*. Wiesbaden: Springer Fachmedien; 2014.
- [10] Machu A, Renard A, Helmstetter P et al. *Situation en matière de sécurité du trafic cycliste sur les routes et dans les carrefours*; 2023. Rapport de recherche OFROU 1755.
- [11] Eberling P, Deublein M, Machu A et al. *Velo-Infrastruktur-Sicherheitsinstrumente VISSI*. Bern: Bundesamt für Strassen ASTRA; 2022. Forschungsbericht ASTRA 1730.
- [12] Lieb C, Amacher M, Sommer H et al. *Externe Effekte des Verkehrs 2021. Umwelt-, Unfall- und Gesundheitseffekte des Strassen-, Schienen-, Luft- und Schiffsverkehrs*; 2024.
- [13] Thoma J. *Das gesamte Ausmass der Strassenverkehrsunfälle*. Bern: BFU, Beratungsstelle für Unfallverhütung; 1990. Pilotstudie R 9032.
- [14] Niemann S, Achermann Stürmer Y, Ellenberger L, Meier D. *Status 2024: Statistik der Nichtberufsunfälle und des Sicherheitsniveaus in der Schweiz*. Bern: BFU, Beratungsstelle für Unfallverhütung; 2024. DOI:10.13100/bfu.2.533.01.2024.
- [15] Bundesamt für Raumentwicklung ARE. *Externe Kosten und Nutzen des Verkehrs in der Schweiz. Strassen-, Schienen-, Luft- und Schiffsverkehr 2020*. Bern; 2023.
- [16] Ecoplan, Infras AG. *Externe Kosten des Verkehrs in der Schweiz: Aktualisierung für das Jahr 2005 mit Bandbreiten – Zusammenfassung*. Bern: Bundesamt für Raumentwicklung ARE; Bundesamt für Umwelt BAFU; 2008.
- [17] Quandt A, Noirjean T, Pittet J-M, Biedermann F. *Kosten und Finanzierung des Verkehrs 2021*. Neuenburg: Bundesamt für Statistik BFS; 2024. 11 Mobilität und Verkehr 811-2100.
- [18] Noirjean T, Quandt A. *Kosten und Finanzierung des Verkehrs 2020*. Neuenburg: Bundesamt für Statistik BFS; 2023. 11 Mobilität und Verkehr 811-2000.
- [19] Schweizerischer Verband der Strassen- und Verkehrsfachleute VSS. *Kosten-Nutzen-Analysen im Strassenverkehr – Unfallraten und Unfallkostensätze*. Zürich: VSS; 2019. VSS 41 824.

- [20] Ringel L, Kielhauser C, Adey BT. Wider view over bicycle crashes: Complementing and extending bicycle crash statistics in urban areas using surveys. *J Safety Res.* 2023; 87: 217–231. DOI:10.1016/j.jsr.2023.09.018.
- [21] International Traffic Safety Data and Analysis Group IRTAD, Hg. *Special report: Under-reporting of road traffic accidents recorded by the police at the international level.* Norway: International Traffic Safety Data and Analysis Group IRTAD; 1994.
- [22] Derriks HM, Mak PM. *Underreporting of road traffic casualties: IRTAD Special Report.* International Traffic Safety Data and Analysis Group IRTAD; 2007.
- [23] International Traffic Safety Data and Analysis Group IRTAD. *Reporting on serious road traffic casualties: Combining and using different data sources to improve understanding of non-fatal road traffic crashes.* Paris: International Transport Forum ITF; 2011.
- [24] World Health Organization WHO. *Road traffic deaths: A toolkit for strengthening data systems.* Geneva: WHO; 2022.
- [25] Yadav SS, Edwards P, Porter J. Completeness of police records for injury surveillance: a systematic review. *Int J Med. Public Health.* 2021; 11(1): 1–9. DOI:10.5530/ijmedph.2021.1.1.
- [26] Pérez-Núñez R, Hidalgo-Solórzano E, Híjar M. Impact of Mexican road safety strategies implemented in the context of the UN's decade of action. *Accid Anal Prev.* 2021; 159: 106227. DOI:10.1016/j.aap.2021.106227.
- [27] World Health Organization WHO. *Global status report on road safety 2018.* Geneva: WHO; 2018. Zugriff am 11.01.2019.
- [28] Razzaghi A, Soori H, Abadi A, Khosravi A. World Health Organization's estimates of death related to road traffic crashes and their discrepancy with other countries' national report. *J Inj Violence Res.* 2020; 12(3): 39–44. DOI:10.5249/jivr.v12i3.1425.
- [29] Harris S. The real number of road traffic accident casualties in The Netherlands: a year-long survey. *Accid Anal Prev.* 1990; 22(4): 371–378. DOI:10.1016/0001-4575(90)90052-M.
- [30] Alsop J, Langley J. Under-reporting of motor vehicle traffic crash victims in New Zealand. *Accid Anal Prev.* 2001; 33(3): 353–359. DOI:10.1016/s0001-4575(00)00049-x.
- [31] Elvik R, Mysen A. Incomplete accident reporting: Meta-analysis of studies made in 13 countries. *Transp Res Rec.* 1999; 1665(1): 133–140. DOI:10.3141/1665-18.
- [32] Ward H, Lyons R, Thoreau R. *Under-reporting of road casualties – phase 1.* London: Department for Transport DfT; 2006.
- [33] Loo BPY, Tsui K. Factors affecting the likelihood of reporting road crashes resulting in medical treatment to the police. *Inj Prev.* 2007; 13(3): 186–189. DOI:10.1136/ip.2006.013458.
- [34] Yannis G, Papadimitriou E, Chaziris A, Broughton J. Modelling road accident injury under-reporting in Europe. *Eur Transp Res Rev.* 2014; 6(4): 425–438. DOI:10.1007/s12544-014-0142-4.
- [35] Janstrup KH, Hels T, Kaplan S et al. *Understanding traffic crash under-reporting: Linking police and medical records to individual and crash characteristics. Abstract from Transport Research Arena.* Paris; 2014.
- [36] Watson A, Watson B, Vallmuur K. Estimating under-reporting of road crash injuries to police using multiple linked data collections. *Accid Anal Prev.* 2015; 83: 18–25. DOI:10.1016/j.aap.2015.06.011.
- [37] Levy PS, Lemeshow S. *Sampling of populations: Methods and applications.* 4 ed. Hoboken, NJ: Wiley; 2008.
- [38] Cochran WG. *Sampling Techniques.* 3. Auflage. New York, NY: Wiley; 1977. Wiley Series in Probability and Mathematical Statistics.
- [39] Jahn K, Dreifuss D, Topolsky I et al. Detection and surveillance of SARS-CoV-2 genomic variants in wastewater. *medRxiv.* 2021: 1151–1162. DOI:10.1101/2021.01.08.21249379.

- [40] Hastie T, Tibshirani R, Friedman JH. *The elements of statistical learning: Data mining, inference, and prediction*. 2 ed. New York: Springer; 2009.
- [41] Kalton G, Anderson DW. Sampling rare populations. *J R Stat Soc Ser A*. 1986; 149(1): 65. DOI:10.2307/2981886.
- [42] Rolstad S, Adler J, Rydén A. Response burden and questionnaire length: Is shorter better? A review and meta-analysis. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2011; 14(8): 1101–1108. DOI:10.1016/j.jval.2011.06.003.
- [43] Abdalla S, Abdelgadir N, Shahraz S, Bhalla K. Respondents' recall of injury events: An investigation of recall bias in cross-sectional injury data from the Sudan Household Health Survey 2010. *Int J Inj Contr Saf Promot*. 2015; 22(3): 215–223. DOI:10.1080/17457300.2014.908222.
- [44] Tourangeau R, Yan T. Sensitive questions in surveys. *Psychol Bull*. 2007; 133(5): 859–883. DOI:10.1037/0033-2909.133.5.859.
- [45] Bird SM, King R. Multiple systems estimation (or capture-recapture estimation) to inform public policy. *Annu Rev Stat Appl*. 2018; 5(1): 95–118. DOI:10.1146/annurev-statistics-031017-100641.
- [46] Chao A, Tsay PK, Lin S-H et al. The applications of capture-recapture models to epidemiological data. *Stat Med*. 2001; 20(20): 3123–3157. DOI:10.1002/sim.996.
- [47] International Working Group for Disease Monitoring and Forecasting. Capture-recapture and multiple-record systems estimation I: History and theoretical development. *Am J Epidemiol*. 1995; 142(10): 1047–1058.
- [48] International Working Group for Disease Monitoring and Forecasting. Capture-recapture and multiple-record systems estimation II: Applications in human diseases. *Am J Epidemiol*. 1995; 142(10): 1059–1068.
- [49] Moennich J, Kumares G, Lich T, Georgi A. *Analysis of non-police reported accidents on Indian highways*: SAE International; 2017. SAE technical paper 2017-26-0005. DOI:10.4271/2017-26-0005.
- [50] Held F. *Investigation of under-reporting and the consistency of injury severity classifications in Swedish police crash data compared to hospital injury data based on the Swedish Traffic Accident Data Acquisition (STRADA)* [Masterarbeit]. Göteborg: Department of Applied Mechanics: Chalmers University of Technology; 2016.
- [51] Broughton J, Keigan M, Yannis G et al. Estimation of the real number of road casualties in Europe. *Safety Science*. 2010; 48(3): 365–371. DOI:10.1016/j.ssci.2009.09.012.
- [52] Short J, Caulfield B. Record linkage for road traffic injuries in Ireland using police hospital and injury claims data. *J Safety Res*. 2016; 58: 1–14. DOI:10.1016/j.jsr.2016.05.002.
- [53] Meyer N, Métillé S, Percassi M-L et al. *Sensor-based accident research and prevention: Exploring legal and technological opportunities*. Berne: Bundesamt für Strassen ASTRA; 2024.
- [54] Karimi S, Hosseinzadeh A, Kluger R et al. A systematic review and meta-analysis of data linkage between motor vehicle crash and hospital-based datasets. *Accid Anal Prev*. 2024; 197: 107461. DOI:10.1016/j.aap.2024.107461.
- [55] King R, McCrea R. Capture-recapture methods and models: Estimating population size. In: Srinivasa Rao, Arni S. R., Rao CR, Hg. *Handbook of statistics: Integrated population biology and modeling, part B*: Elsevier; 2019: 33–83 Volume 40. DOI:10.1016/bs.host.2018.09.006.
- [56] Chao A. Capture-recapture for human populations. In: John Wiley & Sons Ltd., Hg. *Wiley StatsRef: Statistics reference online*. New York, NY; 2015: 1–16. DOI:10.1002/9781118445112.stat04855.pub2.
- [57] Rivest L-P, Baillargeon S. Applications and extensions of Chao's moment estimator for the size of a closed population. *Biometrics*. 2007; 63(4): 999–1006. DOI:10.1111/j.1541-0420.2007.00779.x.

- [58] Yee TW, Stoklosa J, Huggins RM. The VGAM Package for Capture-Recapture Data Using the Conditional Likelihood. *J Stat Softw.* 2015; 65(5): 1–33. DOI:10.18637/jss.v065.i05.
- [59] Das M, Kennedy EH, Jewell NP. Doubly robust capture-recapture methods for estimating population size. *J Am Stat Assoc.* 2023; 119(546): 1–25. DOI:10.1080/01621459.2023.2187814.
- [60] Brittain S, Böhning D. Estimators in capture-recapture studies with two sources. *Adv Stat Anal.* 2009; 93(1): 23–47. DOI:10.1007/s10182-008-0085-y.
- [61] Entezami N, Hashemi-Nazari S-S, Soori H et al. Application of capture-recapture method to estimate traffic accident mortality rate. *Trauma Mon.* 2018; 23(3). DOI:10.5812/trauma-mon.57926.
- [62] Morales D, Agudo I, Lopez J. Private set intersection: A systematic literature review. *Comput Sci Rev.* 2023; 49: 100567. DOI:10.1016/j.cosrev.2023.100567.
- [63] Enamorado T, Fifield B, Imai K. Using a probabilistic model to assist merging of large-scale administrative records. *Am Polit Sci Rev.* 2019; 113(2): 353–371. DOI:10.1017/S0003055418000783.
- [64] AAP Lehrerwelt GmbH. *Ampelkarten: 84 rote, gelbe und grüne Karten zur Unterstützung der Kommunikation und Hilfe bei Abstimmungen.* Hamburg: Persen.
- [65] Bunge JA. A survey of software for fitting capture-recapture models. *WIREs Comput Stat.* 2013; 5(2): 114–120. DOI:10.1002/wics.1250.
- [66] White GC, Burnham KP. Program MARK: Survival estimation from populations of marked animals. *Bird Study.* 1999; 46(sup1): S120-S138. DOI:10.1080/00063659909477239.
- [67] Zhang Z, Nie Q, Liu J et al. Machine learning based real-time prediction of freeway crash risk using crowdsourced probe vehicle data. *Journal of Intelligent Transportation Systems.* 2024; 28(1): 84–102. DOI:10.1080/15472450.2022.2106564.
- [68] Ali Y, Hussain F, Haque MM. Advances, challenges, and future research needs in machine learning-based crash prediction models: A systematic review. *Accid Anal Prev.* 2024; 194: 107378. DOI:10.1016/j.aap.2023.107378.
- [69] Bundesamt für Statistik BFS. *Schweizerische Gesundheitsbefragung: Steckbrief – Erhebung / Statistik*; 2023. <https://www.bfs.admin.ch/asset/de/29045161>. Zugriff am 05.12.2024.
- [70] Bundesamt für Statistik BFS, Bundesamt für Raumentwicklung ARE. *Mikrozensus Mobilität und Verkehr (MZMV): Steckbrief – Statistische Aktivität und veröffentlichte Ergebnisse*; 2024. <https://www.bfs.admin.ch/bfs/de/home/statistiken/mobilitaet-verkehr/erhebungen/mzmv.assetdetail.32170858.html>. Zugriff am 05.12.2024.
- [71] European Commission. *IDB purpose and methods*. https://health.ec.europa.eu/other-pages/basic-page/idb-purpose-and-methods_en. Zugriff am 23.07.2024.
- [72] World Health Organization WHO. *International Classification of External Causes of Injury (ICECI)*. Zugriff am 23.07.2024.
- [73] Bundesamt für Statistik BFS. *Patientendaten Spital ambulant: Steckbrief*. <https://www.bfs.admin.ch/bfs/de/home/statistiken/gesundheit/erhebungen/psa.assetdetail.9306840.html>. Zugriff am 05.12.2024.
- [74] Wieser S, Stucki M, Kohler A et al. *Grundlagenbericht zur Klassifikation und Kodierung der Diagnosen und Prozeduren in der ambulanten Gesundheitsversorgung*; Bundesamt für Statistik BFS; 2023.
- [75] Bundesamt für Statistik BFS. *Statistik der Todesursachen und Totgeburten (CoD)*. <https://www.bfs.admin.ch/bfs/de/home/statistiken/gesundheit/erhebungen/ecod.html>. Zugriff am 22.07.2024.
- [76] Bundesinstitut für Arzneimittel und Medizinprodukte BfArM. *ICD-10-WHO Version 2019*. <https://klassifikationen.bfarm.de/icd-10-who/kode-suche/htmlamtl2019/index.htm>. Zugriff am 22.07.2024.

- [77] Koordinationsgruppe für die Statistik der Unfallversicherung UVG KSUV. *UVG-Statistik 2023: Unfälle und Berufskrankheiten in der Schweiz*. Luzern: KSUV; 2023.
- [78] Dratva J, Feer S, Zysset A. *Machbarkeitsstudie Datenerhebung Kinder- und Jugendunfälle: Schlussbericht*. Zürich: BFU, Beratungsstelle für Unfallverhütung; 2023. Forschung. DOI:10.21256/ZHAW-2439.
- [79] Stampfler T, Elgendi M, Fletcher RR, Menon C. Fall detection using accelerometer-based smartphones: Where do we go from here? *Front Public Health*. 2022; 10: 996021. DOI:10.3389/fpubh.2022.996021.
- [80] Apple-Support. *Mithilfe der Unfallerkennung deines iPhone oder deiner Apple Watch bei einem Unfall Hilfe rufen*. <https://support.apple.com/de-ch/104959>. Zugriff am 23.07.2024.
- [81] Steiner R, Benedek G, Doerfel M et al. *Alternative Methoden zur Messung lokaler Verkehrssicherheit*. Bern: Bundesamt für Strassen ASTRA; 2023. Forschungsbericht ASTRA 1746.
- [82] Bundesamt für Statistik BFS. *Todesursachenstatistik 2022*. <https://www.bfs.admin.ch/news/de/2023-0109>. Zugriff am 22.07.2024.
- [83] World Health Organization WHO. *Road safety Switzerland 2023 country profile*. <https://www.who.int/publications/m/item/road-safety-che-2023-country-profile>. Accessed on 01.12.2024.
- [84] Bundesamt für Statistik BFS, Bundesamt für Raumentwicklung ARE. *Mobilitätsverhalten der Bevölkerung: Ergebnisse des Mikrozensus Mobilität und Verkehr 2021*. Neuenburg: BFS; ARE; 2023. 11 Mobilität und Verkehr 840-2100.
- [85] Bundesamt für Statistik BFS. *Haushalte und Lebensformen*. <https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/familien/haushalte.html#:~:text=In%20mehr%20als%20einem%20Drittel,durchschnittliche%20Haushaltsgr%C3%B6sse%20betr%C3%A4gt%202.18%20Personen>. Zugriff am 30.04.2025.
- [86] Loftis KL, Price JP, Gillich PJ et al. Development of an expert based ICD-9-CM and ICD-10-CM map to AIS 2005 update 2008. *Traffic Inj Prev*. 2016; 17(Suppl 1): 1–5. DOI:10.1080/15389588.2016.1191069.
- [87] Association for the Advancement of Automotive Medicine AAAM. *Purchase of the AIS ICD ISS Map*. <https://www.aaam.org/abbreviated-injury-scale-ais/licensing/ais-icd-iss-map/>. Zugriff am 01.12.2024.
- [88] Schmitt K-U, Scholz S, Baumgartner L et al. *Medizinische Folgen des Strassenunfallgeschehens: Forschungspaket VeSPA Teilprojekt 5*. Bern: Bundesamt für Strassen ASTRA; 2014. Forschungsbericht ASTRA 1457.
- [89] Schmitt K-U, Baumgartner L, Furter K. *SIS-Simplified Injury Scale: Abschlussbericht*. Zürich: AGU Zürich; 2014.
- [90] Schmitt K-U, Furter K, Muser MH. Correlation between AIS-coded injury severity and injury severity classified by the NACA score: Short communication. In: *2016 IRCOB Conference Proceedings: 14.-16. September 2016 – Malaga*; 2016: 82–84.
- [91] Wikipedia. *Private set intersection*. https://en.wikipedia.org/wiki/Private_set_intersection. Accessed on 28.04.2025.
- [92] Bundesamt für Strassen ASTRA. *Teilstrategie Verkehrssicherheit*. Bern: ASTRA; 2024 V1.0.
- [93] Kautiala C, Seimelä K, Oy D. *Tieliikenteen onnettomuusrekistereiden peittävyystutkimus [Zusammenfassung in Englisch]*. Helsinki: LINTU-tutkimusohjelma; 2012. LINTU Report 7/2012.
- [94] Amoros E, Turbatu L, Soler G et al. *Estimation du nombre de blessés graves de la route, au sens MAIS3+, France métropolitaine, 2006-2016*: IFSTTAR - Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux; 2019 hal-02110133.
- [95] Amoros E, Casez C, Adelli M et al. Blessés de la route en France: combien? Quelle gravité? Quelles caractéristiques? Par capture-recapture sur deux sources de données. *Rev Epi Santé Pub*. 2020; 68: 56–57.

- [96] SafetyNet. *D.1.15. Final report on Task 1.5: Building the European road safety observatory*. Loughborough: Transport Research Laboratory TRL; 2004.
- [97] Institute for Road Safety Research SWOV. *Road deaths in the Netherlands*. The Hague, NL: SWOV; 2024. SWOV Fact sheet.
- [98] Institute for Road Safety Research SWOV. *Serious road injuries in the Netherlands*. The Hague, NL: SWOV; 2024. SWOV Fact sheet.
- [99] Institute for Road Safety Research SWOV. *Registraton rate fatalities and seriously injured*. <https://theseus.swov.nl/single/?appid=9f66b916-dff2-48c3-8146-a41e875ab887&sheet=dbd3ed6f-2d85-43e4-baeb-3d9d360cdea1&opt=cursel%2ccctx-menu&select=Injury%20severity,Dead&select=Jaar,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020>.
- [100] Donabauer M, Pommer A. *IDB Austria 2021: Verletzungsursachen in Österreich – Standarderhebungen des KfV*. Wien: Kuratorium für Verkehrssicherheit KfV; 2021.

Projektabschluss



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement für
Umwelt, Verkehr, Energie und Kommunikation UVEK
Bundesamt für Strassen ASTRA

FORSCHUNG IM STRASSENWESEN DES UVEK

Version vom 09.10.2013

Formular Nr. 3: Projektabschluss

erstellt / geändert am: 28.03.2025

Grunddaten

Projekt-Nr.: MFZ_20_06B_03
Projekttitel: MEDUSA: Methodik zur differenzierten Quantifizierung der Dunkelziffer der Strassenverkehrsunfälle in der Schweiz
Enddatum: 30.06.2025

Texte

Zusammenfassung der Projektergebnisse:

Das MEDUSA-Projekt hatte zum Ziel, eine Methodik zur differenzierten Bestimmung der Dunkelziffer der vom Bundesamt für Strassen ASTRA erfassten polizeilich registrierten Strassenverkehrsunfälle (ASTRA-Daten) zu entwickeln. Dabei sollten Verkehrsteilnehmergruppen, Unfallhergänge und Verletzungsschwere differenziert betrachtet werden können. Es wurden verschiedene Projektschritte durchlaufen: eine Literaturrecherche, zwei Stakeholderworkshops, eine Übersicht und Beurteilung potenziell nutzbarer Datenquellen und statistischer Methoden. Darauf basierend wurden ein Methodenvorschlag und Empfehlungen erarbeitet.

Als zum aktuellen Zeitpunkt vielversprechendste Datenquellen wurden die Daten der für in der Schweiz wohnhafte oder arbeitstätige Personen obligatorischen Unfallversicherungen (gemäss UVG und KVG) sowie Daten aus Haushaltsbefragungen identifiziert. Weitere Datenquellen wurden geprüft, sind jedoch in Kombination mit den identifizierten statistischen Methoden in der aktuellen Form nicht für eine differenzierte Dunkelzifferschätzung geeignet. Eine grundsätzliche Datenlücke besteht bei Unfällen von nicht in der Schweiz wohnhaften und/oder arbeitstätigen Personen (insbesondere Kinder und Senioren/-innen).

Für die Dunkelzifferberechnung von Getöteten wird aus Kosten-Nutzen-Überlegungen empfohlen, von einer Dunkelziffer von 1,0 (Faktor) auszugehen und damit die ASTRA-Daten mit polizeilich registrierten Strassenverkehrsunfällen zu verwenden.

Für Verletzte werden folgende Methoden vorgeschlagen:

1. Befragung mit Hochrechnung: entweder eigenständig oder integriert in bestehende Erhebungen (z.B. BFU-Haushaltsbefragung oder Mikrozensus Mobilität und Verkehr (MZMV)). Es handelt sich um eine bewährte, kurzfristig umsetzbare Methode, wobei die Differenzierungsmöglichkeiten zwar flexibel, aber auch abhängig von der Stichprobengrösse sind. Sie ist mit vergleichsweise hohen Kosten verbunden und anfällig für Erinnerungsfehler und sozial erwünschte Antworten.
2. Linkage ASTRA- und SSUV-Daten (Sammelstelle für die Statistik der Unfallversicherung UVG) mit Capture-Recapture (Hochrechnung): diese Methode ist kostengünstiger als Befragungen. Nachteile sind, dass damit Kinder und Senioren/-innen nicht abgedeckt werden können und bei nicht harmonisierten Daten nur begrenzte Differenzierungsmöglichkeiten bestehen.
3. Versicherungsliste (UVG- und KVG-Daten) mit Hochrechnung: Hier werden Daten der UVG-Unfallversicherer (SSUV) und der KVG-Versicherer kombiniert. Langfristig wäre dies eine kosteneffiziente Methode. Um diese Methode umzusetzen, müssen die KVG-Daten jedoch erst erschlossen werden, was eine gesetzliche Grundlage erfordert.

Maschinelles Lernen wurde als Methode ebenfalls geprüft, braucht aber gute Trainingsdaten zu Unfallereignissen und deren Verletzungsfolgen. Solche Trainingsdaten sind aktuell nicht verfügbar, weshalb der Ansatz derzeit nicht umsetzbar ist. Die im Zusammenhang mit dem Strassenverkehr erfassten Daten werden künftig in Menge und Qualität zunehmen, weshalb ihre Nutzbarkeit für den Zweck der Dunkelzifferberechnung regelmässig neu evaluiert werden sollte.



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement für
Umwelt, Verkehr, Energie und Kommunikation UVEK
Bundesamt für Strassen ASTRA

Zielerreichung:

Das Ziel, eine Methodik zur differenzierten Bestimmung der Dunkelziffer für Strassenverkehrsunfälle zu entwickeln, wurde erreicht. Die methodischen Ansätze wurden beschrieben und auf ihre Vor- und Nachteile untersucht. Eine Differenzierung von Verkehrsteilnehmergruppen, Unfallhergängen und Verletzungsschwere ist möglich. Die praktische Umsetzbarkeit der Methoden sowie die bei der Umsetzung zu beachtenden Herausforderungen konnten aufgezeigt werden. Der Bericht gibt Empfehlungen zu einer Implementierungsstrategie sowie einer sinnvollen Periodizität der Dunkelzifferabschätzung.

Folgerungen und Empfehlungen:

Es wird empfohlen, für die Dunkelziffer der im Strassenverkehr Getöteten von einem Faktor 1,0 auszugehen (vollständige Erfassung in den ASTRA-Daten).

Für die Berechnung der Dunkelziffer von Verletzten sollte eine Kombination der Methoden «Befragung» und «Datenverknüpfung» vorgenommen werden. Kurz- bzw. mittelfristig sind folgende Schritte einzuleiten:

- Ergänzung des MZMV mit Fragen zu Strassenverkehrsunfällen
- Harmonisierung und Verknüpfung von ASTRA- und SSUV-Daten (probabilistische Linkage)
- Optimierung, Überprüfung der Codierung sowie Standardisierung der SSUV-Daten; Ausweitung der SSUV-Stichprobe
- Sondieren der Möglichkeiten zur Erschliessung der Unfalldaten der KVG-Versicherer (Krankenversicherungsgesetz)

Eine zentrale Koordination, z.B. durch das ASTRA wird empfohlen, um eine einheitliche Strategie zur Dunkelzifferermittlung sicherzustellen.

Publikationen:

-

Der Projektleiter/die Projektleiterin:

Name: Hertach

Vorname: Patrizia

Amt, Firma, Institut: BFU, Beratungsstelle für Unfallverhütung

Unterschrift des Projektleiters/der Projektleiterin:

P. Hertach



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement für
Umwelt, Verkehr, Energie und Kommunikation UVEK
Bundesamt für Strassen ASTRA

FORSCHUNG IM STRASSENWESEN DES UVEK

Formular Nr. 3: Projektabschluss

Beurteilung der Begleitkommission:

Beurteilung:

MEDUSA liefert einen umfassenden methodischen Rahmen zur differenzierten Berechnung der Dunkelziffer von Strassenverkehrsunfällen in der Schweiz. Die Studie überzeugt durch eine strukturierte Herangehensweise, eine umfassende Literaturrecherche und eine systematische Analyse verfügbarer Datenquellen sowie verschiedener statistischer Modelle. Besonders hervorzuheben ist die Entwicklung dreier komplementärer methodischer Ansätze zur Abschätzung der Dunkelziffer (Befragungen, Datenverknüpfung mit Capture-Recapture, Versicherungsliste), die unterschiedliche Stärken und Anwendungsbereiche aufweisen. Die aktive Einbindung relevanter Stakeholder in Workshops sichert die Praxisrelevanz der Ergebnisse. Herausforderungen zeigen sich bei Einschränkungen aufgrund der Datenverfügbarkeit, etwa der fehlenden Differenzierung zwischen herkömmlichen Fahrrädern und E-Bikes, Datenlücken bei nicht UVG-versicherten Personen (insbesondere Kinder und Senioren), methodischen Herausforderungen bei der Harmonisierung unterschiedlicher Klassifikationen von Verletzungsschwere und unzureichenden Informationen zu im Ausland wohnhaften Unfallbeteiligten. Die Pilotberechnung demonstriert zwar die praktische Anwendbarkeit, verdeutlicht jedoch auch die Komplexität und die allfälligen Schwierigkeiten bei der vollständigen Umsetzung aller gewünschten Differenzierungen.

Umsetzung:

Inhaltlich entspricht die Umsetzung des Forschungsprojekts weitgehend dem Projektplan und liefert die angestrebten Ergebnisse. Die geplante Projektlaufzeit von 15 Monaten war allerdings zu ambitioniert; das Projekt musste um fünf Monate verlängert werden.

Die drei entwickelten methodischen Ansätze bieten einen flexiblen Implementierungsrahmen mit kurz-, mittel- und langfristigen Optionen. Positiv zu bewerten ist die phasenweise Implementierungsstrategie, die pragmatische Empfehlungen für den sofortigen Einsatz (Befragungsmethode) mit längerfristigen Optimierungen (Datenverknüpfung, Versicherungslisten) kombiniert. Die Einbindung der Stakeholder wurde wie geplant durch zwei Workshops realisiert, deren Ergebnisse direkt in die Methodenentwicklung einflossen. Die Simulationen (für welche der Programmcode vorliegt) und die Pilotberechnung demonstrieren die praktische Anwendbarkeit, verdeutlichen jedoch auch die Herausforderungen bei der Datenverknüpfung. Die zentrale Projektempfehlung, eine Kombination mehrerer Methoden unter Koordination einer federführenden Organisation einzusetzen, berücksichtigt die identifizierten Stärken und Limitationen der einzelnen Ansätze.

weitergehender Forschungsbedarf:

Der Forschungsbericht identifiziert mehrere prioritäre Forschungsfelder, die für eine verbesserte Dunkelzifferberechnung wesentlich sind: (a) Entwicklung optimierter Methoden zur probabilistischen Datenverknüpfung zwischen ASTRA- und SSUV-Daten unter Berücksichtigung datenschutzrechtlicher Anforderungen, (b) Harmonisierung der unterschiedlichen Klassifizierungssysteme für Verletzungsschwere (ASTRA, ARE, NACA, AIS, ICD-10) für konsistente Differenzierungen, (c) Erschliessung der KVG-Versicherungsdaten für eine umfassendere Abdeckung aller Altersgruppen, (d) Methodische Ansätze zur Erfassung von im Ausland ansässiger Unfallbeteiligter, (e) Weiterentwicklung von Machine-Learning-Ansätzen, sobald ausreichende Trainingsdaten verfügbar sind und (f) Verbesserte Datenerhebung zur zuverlässigen Unterscheidung zwischen Fahrrad- und E-Bike-Unfällen.

Einfluss auf Normenwerk:

Der Forschungsbericht liefert eine fundierte methodische Grundlage für eine präzisere Quantifizierung des Gesamtunfallgeschehens, was erhebliche Auswirkungen auf bestehende Normen und Standards haben könnte. Insbesondere die VSS-Normen für Verkehrssicherheitsbewertungen könnten von differenzierteren Unfallzahlen profitieren. Die methodischen Erkenntnisse ermöglichen genauere Berechnungen der externen Unfallkosten, präzisere Unfallraten für die Infrastrukturplanung sowie eine verbesserte Grundlage für Kosten-Nutzen-Analysen von Verkehrssicherheitsmassnahmen. Eine standardisierte Methodik würde zudem bessere internationale Vergleiche der Verkehrssicherheitssituation erlauben.

Der Präsident/die Präsidentin der Begleitkommission:

Name: Kamenik

Vorname: Christian

Amt, Firma, Institut: Bundesamt für Strassen ASTRA

Unterschrift des Präsidenten/der Präsidentin der Begleitkommission: