BRAIN-be 2.0

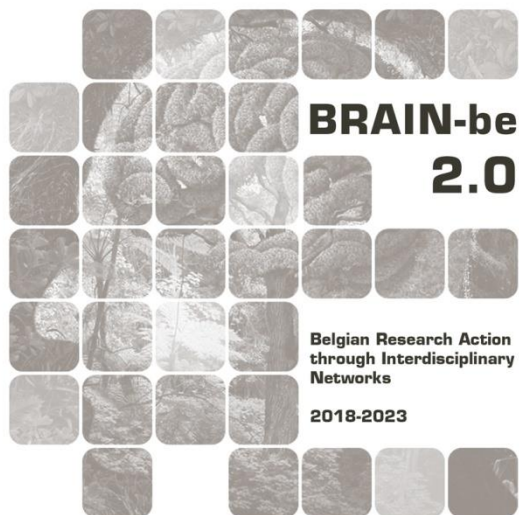**Belgian Research Action through Interdisciplinary Networks**

2018-2023

IIHA

It Is Happening Again: Digital criminal justice archives as building blocks for the study of recidivism

Benjamin Mine (INCC) – Patrick Jeuniaux (INCC) – Eric Maes (NICC) – Luc Robert (NICC) – Philippe Huynen (INCC) – Michaël Vande Velde (INCC)

Pillar 2: Heritage science

belspo .be

BRAIN-be
2.0

Belgian Research Action
through Interdisciplinary
Networks

2018-2023

NETWORK PROJECT

IIHA

It Is Happening Again: Digital criminal justice archives as building blocks for the study of recidivism

Contract – BR/202/P2/IIHA

FINAL REPORT

PROMOTORS: Benjamin Mine (NICC)
                    Patrick Jeuniaux (NICC)
                    Eric Maes (NICC)
                    Luc Robert (NICC)

AUTHORS: Benjamin Mine (NICC)
                    Patrick Jeuniaux (NICC)
                    Eric Maes (NICC)
                    Luc Robert (NICC)
                    Philippe Huynen (NICC)
                    Michaël Vande Velde (NICC)

NICC
INCC

belspo .be

Benjamin Mine, Patrick Jeuniaux, Eric Maes, Luc Robert, Philippe Huynen & Michaël Vande Velde. *It Is Happening Again: Digital criminal justice archives as building blocks for the study of recidivism*. Final Report. Brussels: Belgian Science Policy Office, 2024 – 75 p. (BRAIN-be 2.0 - (Belgian Research Action through Interdisciplinary Networks))

TABLE OF CONTENTS

ABSTRACT

Context

Recidivism is an important societal and political issue about which there is a clear knowledge gap in Belgium (Maes et al., 2021). The absence of systematically produced national recidivism statistics has important consequences, including the difficulties (or even impossibility) to scientifically evaluate the impact of criminal justice policies and interventions (including punishments). This implies that criminal justice interventions in Belgium take place without any scientific evidence about their effectiveness in tackling crime (be it through general or specific deterrence, incapacitation or rehabilitation).

However, recidivism remains a complex social phenomenon that is difficult to measure (Maes et al., 2021). Indeed, measuring recidivism comes up against numerous limitations such as the dark number, the grey number, and the impunity bias (Zara & Farrington, 2016), some of which are due to the very functioning of the criminal justice administration system. Despite their gaps and imperfections (P. Robert, 1977), official data is one of the sources that can potentially be mobilized to study recidivism.

The National Institute of Criminalistics and Criminology has access to a wealth of historical data from two criminal justice databases: the Central Criminal Record (over the 1995-2020 period) and the SIDIS-Greffe penitentiary database (over the 1974-2014 period) (hereafter: 'the Criminal Justice Databases'). Unfortunately, the optimal exploitation and valorisation of this heritage is made difficult due to lack of documentation on, and standardization of, the records.

The project aims to address societal questions through studies on recidivism and criminal careers by documenting and analysing these historical data. To achieve this, a computer platform (IHD) was developed containing a standardized version of multiple historical datasets extracted from these two criminal justice databases to make the data available for research. At the same time, analysis on recidivism and criminal careers were carried out on several themes (e.g. homicides, terrorism, sexual offences). Moreover, a prototype recidivism monitor was developed enabling the production of baseline recidivism rates for different cohorts.

Objectives

The research had four main objectives.

1. Documenting the Criminal Justice Databases

Expanding on a past project (De Blander et al., 2019), the first objective of this project is to thoroughly document the purpose, content, structure, limitations and potential for research of the two databases hereby solving the lack of an up-to-date centralized documentation.

2. Developing an Integrated Historical Database (IHD)

The second objective is to develop an integrated database to store a standardized version of historical records obtained from the Criminal Justice Databases. The purpose was to link an extraction of the Central Criminal Record (over the 1995-2020 period) with a full extraction of the

SIDIS-Greffe penitentiary database and to create an Integrated Historical Database (IHD). The goal of this integrated database is to make the datasets that are stored as separate files available within a common system that allow easy access for statistics production and research.

3. Analyses of recidivism and criminal careers

The third objective is to conduct several analyses. A first type of analysis concerns criminal careers and recidivism on year-based cohorts or the perpetrators of certain categories of offences based on conviction data and/or detention data. These results provide an immediate benefit for policy-makers and scientists.

4. Studying the feasibility of a recidivism module

Informed by the results from the previous steps and previous research (particulary Robert & Maes, 2012, Mine et al., 2015; De Blander et al., 2019), the feasibility of developing a computer-based module aimed at uncovering and monitoring recidivism is studied. The feasibility study includes collaborations with software developers as well as with the staff currently managing the two databases.

Conclusions

In a nutshell, we advocate for a culture of more and better scientific evidence that can serve to inform criminal justice policies and practices in Belgium, which implies a smooth and safe access to the Criminal Justice Administration System (CJAS) data, and data of better quality (in terms of internal, horizontal, vertical, and contextual integration).

Keywords

Historical records – Databases – Criminal justice – Criminal career – Recidivism

1. INTRODUCTION

This report presents the results of the research project "It Is Happening Again: Digital criminal justice archives as building blocks for the study of recidivism" (IIHA)[1].

The IIHA project has been funded for two years with two extensions of 6 months (2021-2024) by the Belgian Federal Science Policy Office (BELSPO) under Pillar 2 of the "Heritage Science" of the "BRAIN-be 2.0" research program. Pillar 2 is designed to support the conservation of State collections. Within the framework of this project, the databases of the Federal Public Service Justice are envisaged as collections that need to be preserved, so that they can be used for scientific, historical and statistical purposes. These data span decades of records. One of the databases even contains records that go back to 1922. These are the Central Criminal Record database and the former prison database (SIDIS-Greffe), two databases regularly used by the National Institute of Criminalistics and Criminology (NICC) as part of its research to address specific criminological research questions, e.g., pertaining to the study of recidivism or criminal careers.

In all those cases, preliminary steps had to take place in order to prepare the datasets for analysis. Such steps implied an effortful, repetitive and time-consuming process. This situation raised three types of issues.

1) Preparing and understanding the data was made difficult due to lacking and dispersed documentation describing the datasets (e.g., explaining the meaning of the variables or how the data was produced). Unfortunately, to this date, the documentation that had been reconstructed for past research purposes was not complete and did not meet strict quality standards that would facilitate future research.

2) The preparation of the data implied research-specific reorganization operations that were tight to the particular research questions at hand. In effect these operations transformed the original datasets (i.e., database extractions) into new datasets whose structure was not necessarily well suited to answer other research questions.

3) The datasets existed as isolated, and often duplicated, files scattered across a folder structure, which made it particularly difficult to not only find the necessary data but also to draw links between files although they originated from the same database.

In conclusion, based on this state of affairs, it was apparent that significant action was required to work towards proper conservation, preservation and management of this digital heritage.

The IIHA project therefore has four main objectives.

1) To document the two databases of interest. This documentation will help guide future criminological research based on these databases.

2) To develop a database integrating a dataset extracted from each of the two databases of interest.

---

[1] https://incc.fgov.be/IIHA

3) To exploit the integrated database for statistical analysis of recidivism and criminal careers in criminological studies.

4) Develop a prototype "recidivism monitor" to measure and track recidivism using data from the Central Criminal Records.

It is important to highlight the fact that recidivism is an important societal and political issue about which there is a clear knowledge gap in Belgium. The study of recidivism and criminal careers fits well within the priorities and strategies of the department of Criminology of the NICC, as they have been officially designated a priority this year by its stakeholders at the Federal Public Service Justice. Thus, the general idea of the IIHA project is therefore to increase the production of knowledge on recidivism and criminal careers in Belgium, but also to develop useful tools to support the achievement of this objective (i.e., documentation, a new database and an IT module).

We begin by presenting the state of the art in the study of recidivism and criminal careers, and then go on to explain more in details the objectives of this project, which organize the structure of the methodology and results sections of this report.

2. STATE OF THE ART AND OBJECTIVES

2.1. State of the art

Recidivism research can serve different purposes, including (i) the description of offender profiles, (ii) insights into the etiology of offending behavior (understanding the causes or factors responsible for reoffending), (iii) the study of criminal careers of offenders, (iv) an assessment of the impact of criminal justice interventions on offending (Wartna, 2009).

Recidivism research remains a challenge in many regards, not in the least because of a lacking consensus about an appropriate definition (e.g., Maltz, 1984; Zara & Farrington, 2016). Research that zooms in on recidivism is often constricted to the data that are available, which depends strongly on the specific context of the study. For instance, the definition of recidivism seen from a legal standpoint may differ from a definition used in social sciences and criminology (Robert et al., 2019). Moreover, depending on the definition of recidivism (the measure that is used), the cohort that is studied and the follow-up period, statistics on recidivism will vary greatly. A telling example in that regard comes from Andersen and Skardhamar (2017), who show the impact of the measure of recidivism (arrest, conviction, conviction to prison), cohorts and follow-up periods, with no less than 36 different national Norwegian recidivism prevalence statistics.

At the international level, the complexity is even further compounded. Countries often encounter difficulties when they attempt to compare their national statistics (Fazel & Wolf, 2015).This is often due to methodological issues, such as the type of the study cohort, the observation or follow-up period, and the definition of recidivism, alongside other underlying differences (the organization of the criminal justice system, the wider penal culture, victims' offence reporting behavior, …).

Both at the national and international levels, there is an urgent need to provide Belgium with the necessary tools to measure recidivism and model criminal careers (see De Blander et al., 2019; Robert et al., 2019). Belgium is one of the few countries in Western Europe and one of the few in the European Union that does not systematically produce national recidivism statistics (for a description of the issue, see e.g., Mine et al., 2015b). Until 2015, there were even no national recidivism statistics available. Until then, a number of studies existed that focused on specific (selective), and often small, groups of offenders, such as probationers in Antwerp (for an overview, see Robert et al., 2019).

The absence of systematically produced national recidivism statistics has important consequences, including the difficulties (or even impossibility) to scientifically evaluate the impact of criminal justice policies and interventions (including punishments). This implies that criminal justice interventions in Belgium take place without any scientific evidence about their effectiveness in tackling crime (be it through general or specific deterrence, incapacitation or rehabilitation).

Furthermore, Belgium is left aside in international (European) comparisons of recidivism rates (see, e.g., Wartna & Nijssen, 2006; Albrecht & Jehle, 2014; Yukhnenko et al., 2019).

A monitor that provides statistics on recidivism has several advantages (e.g., Wartna et al., 2005). First, it provides a better, more reliable description of recidivism, because it will draw on a consistent definition of recidivism and of other variables, on the basis of which it will provide base rates of

prevalence for offenders (including differences on a range of variables, such as sex, age, previous convictions and other important predictors for recidivism, e.g., Gendreau et al., 1996).

Second, a monitor will also provide the structured, cleaned data for national, large-scale, studies on criminal careers. Through the combined use of data related to all convictions and detentions of offenders, it will become possible to produce important insights about key parameters of criminal careers (e.g., Blumstein & Cohen, 1987; DeLisi & Piquero, 2011), including onset, duration, termination (or desistance), frequency of offences that led to convictions, intermittency, and crime mix (range of different types of offences). In Belgium, despite a rich criminological research tradition in the domain of criminal justice (e.g., Daems et al., 2013), researchers have only recently began to put to use large scale data for criminal career research (e.g., Jeuniaux et al., 2018; L. Robert et al., 2018; Spaan et al., 2020).

Third, the recidivism base rates (in comparison to groups subjected to an intervention) and information about criminal careers can be used for evaluation purposes, e.g., when new policy is introduced (e.g., a new type of punishment, such as electronic monitoring as a sentence). Belgium's criminal justice system lacks such types of evaluation research (but see Robert et al., 2017, for an example). Among the many issues involved in criminal career research are the impact of sentences on subsequent criminal careers (e.g., MacLeod et al., 2012), the impact of incarceration as a tool for incapacitation for prolific offenders (e.g., Blumstein & Cohen, 1987; Blumstein, 2016), and many other policy-relevant scientific questions.

Finally, at the national level, an in-depth description of the two criminal justice databases have never been reported to the best of our knowledge, except in a partial way in Daeninck et al. (2005) and Maes and Tange (2014) for detention data, and in Willems (2007) and Mine et al. (2015a, 2015b) for conviction data. There is therefore an important knowledge gap on these fundamental matters as well.

2.2. Objectives

This research has four objectives:

- **Objective 1: Documenting the Criminal Justice Databases**

Expanding on a past project (De Blander et al., 2019), this project documents thoroughly the functions, content, structure, limitations, and potential of the Central Criminal Record (CJCS) and SIDIS-Greffe (SIDIS) detention database.

This is important because the documentation describing the datasets (e.g., explaining the meaning of the variables or how the data was produced) is partially lacking and stay dispersed. Indeed, to this date, the documentation that has been reconstructed for past research purposes is not complete and does not facilitate future research.

Rich of the documentation constituted, this research examines the possibility of establishing reliable links between these two databases.

- **Objective 2: Developing an Integrated Historical Database (IHD)**

In the past, NICC researchers have made use of CJCS and SIDIS data. In addition to the limited documentation available, their task has been made difficult for two other main reasons.

First, available data existed in the form of isolated files, often duplicated, which had to be located, checked and reorganized for each analysis. Secondly, simultaneous analysis of conviction and detention data was made difficult by the absence of a unique personal identifier to identify these individuals within these two databases. This state of affairs was not conducive to long-term use of the data. Designed as a graph database, the IHD is intended to compensate these difficulties (scattered data and absence of unique personal identifier). It offers the advantage of being able to manage and analyze with a certain degree of flexibility data within which there are multiple opportunities to establish relationships.

We aim to create a computer-based repository (i.e., a data warehouse) of a standardized version of historical records obtained from the two Criminal Justice Databases. Contrary to the purpose of the first objective, the goal here is to offer a repository for the data itself (rather than merely the metadata). Our goal is to import the files extracted from these databases into a common system that will facilitate the access and exploration of the data, as well as the production of statistics for research and public policy purposes.

A key feature of the IHD is its pseudonymization. The names and other identification variables of the persons are to be replaced by random numbers and letters. This is an important characteristic of this project. First of all, it ensures a greater security in terms of data protection and safety. Second, it facilitates the process of authorization for data reuse in the future (i.e., beyond the end of the project). This therefore confers to the database a long-term advantage for research purposes.

Another particularly helpful feature of this integrated database is that it offers the possibility of drawing and storing new links between these records (hence generating new data). For instance, it is possible to create a link between the records of an individual registered in the conviction database to the records of that same individual registered in the detention database. The realization of this linkage is not trivial because there are no common personal identifiers to refer to individuals across the two databases. Instead, other variables (and combinations thereof) such as the national register identification number, the names, and the date of birth need to be used (when they are available).

Although there are many offenders in the conviction data that are not recorded in the detention data (i.e., there is only a limited overlap between the two databases because not all individuals are necessarily sentenced to imprisonment), in principle, all persons sentenced to imprisonment and included in the detention database should also appear in the Central Criminal Record.

It is really helpful to link the detention data to the conviction data because, for instance, the database of the Central Criminal Record tells us nothing about the effectiveness of the sentence handed down, neither how it was enforced. It would also facilitate finding the reason for which a person is incarcerated again (new conviction, non-compliance with conditions, etc.).

- **Objective 3: Analyses of recidivism and criminal careers**

We aim to conduct several analyses on the basis of the historical data from each database (CJCS and SIDIS) but also on data made available through the Integrated Historical Database (IHD[2]).

A first type of analysis concerns recidivism, hence the title of this proposal which sees "Digital criminal justice archives as building blocks for the study of recidivism". Recidivism being about the repetition of a particular event, the acronym of this proposal (IIHA) corresponds to the expression "It Is Happening Again". This analysis of recidivism is a key feature of this proposal as it informs Objective 4, particularly for conducting a business analysis and drafting the specifications for the development of a recidivism module.

A second type of analysis consists of the study of criminal careers (which can be considered as a natural extension of a study of recidivism). We will study criminal careers based on specific selection criteria (e.g., type of offences, age and gender of offenders, which allows to study the impact of continuity versus discontinuity in offending behavior).

- **Objective 4: Studying the feasibility of a recidivism module**

We aim to evaluate the feasibility of developing a computer-based module that will help uncover and monitor recidivism. The principle behind this module is to offer a computer interface to the manager of CJCS, on the basis of which he or she generates on the fly several figures and statistics (e.g., on reconviction rates) on the basis of some user-defined criteria (e.g., sex, age, type of offence, type of sentence, different follow-up periods).

The feasibility study (or business analysis) provides details about the architecture of this module (in terms of content and design), and supports the development of a demo. The role of the demo is to demonstrate to the authorities the need for investing resources in this area, and foster an evidence-based approach to sentencing and convictions, informed by results about a key outcome, recidivism.

Objective 4 is informed by the research done for the three previous objectives. However, the data exploited to make Objective 4 possible is not identical as the one used for Objective 2 and Objective 3 as the module would directly feed on the Criminal Justice Database being considered (i.e., CJCS) and not only on historical data.

---

[2] Due to the complexity and technical challenges of developing the IHD, most of the analyses were carried out in parallel with the creation of the IHD and are not based on it because the IHD was not fully operational on time. The exploitation of the IHD will, however, be carried out within the activities of the Research Unit on Recidivism and Criminal Careers (CReCC) and other projects carried out by members of this unit.

3. METHODOLOGY

The project is based on existing historical datasets originating from two databases: 1) the Central Criminal Record (CJCS) that deals with conviction data (concerning persons convicted over the 1995-2020 period), 2) the SIDIS-Greffe (SIDIS) detention database (persons detained over the 1974-2014 period). Both datasets cover activities taking place at the national level. This has enormous potential for criminological research and can provide important policy lessons about long-term offending behavior of nationally representative groups and national cohorts of offenders.

3.1. Documenting the Criminal Justice Databases

First of all, we identified and collected the data necessary to complete the project. Data on conviction and detention were contained in different database tables (i.e., detention history, prison regime, legal status, socio-biographical information, facts, punishments, etc.). Each table had been extracted from its database and exported as a separate file. These data were converted to a standard format (e.g., all files in CSV format, using UTF-8, all records on separate lines, lower-cased, using the same column separator).

Secondly, the research team conducted an investigation to identify and gather all the existing documents from the NICC's Criminology department that were mainly useful for understanding the structure and operation of the databases (i.e., entity-relationship diagram, nomenclatures, etc.) and recording practices (i.e., user manual).

Thirdly, several interviews (n=4) were then conducted with a number of resource persons from various departments of the FPS Justice (e.g., Criminal Policy Department, Central Criminal Records Department, ICT) in order to acquire knowledge on the two databases and collect any useful additional documents. The interviews were recorded and transcribed.

All the documents available to the NICC as well as those provided by the persons who were interviewed were sorted. The purpose of this sorting was to remove superfluous documents (duplicates, irrelevant versions of a text), documents of purely temporary interest (e.g., acknowledgements of receipt) and items that damaged the physical preservation of the archives (metal clips, filing cabinets, plastic folders, etc.). The remaining documents were then scanned and listed in a digital file within which a first classification was made (according to the institutional origin of the document and whether or not it is confidential).

Based on the examination of a heterogeneous material (e.g., interviews, previous research and technical reports, personal knowledge, databases explorations), a report for each database was drafted. These reports highlight the relevant tables to be taken into consideration (as well as the relationships between them) with a view to using them for scientific, statistical and historical purposes, particularly with regard to the production of knowledge on the issues of recidivism and criminal careers. A first version of the report on the Central Criminal Record was sent on 4 October 2022 to the members of the steering committee working for the FPS Justice, while the first version of the report on SIDIS was sent on 22 December 2022 to a former DGEPI staff member and SIDIS-Greffe expert. On the basis of the comments and remarks received from them, a second expanded and corrected version of these reports was drafted and finalized.

These two reports were resubmitted to the Central Criminal Records Department, to a former DGEPI staff member and to the DGEPI itself in order to validate their content and discuss the archiving of the documents inventoried. For the time being, it has only been possible to organize this discussion with DG EPI and the President's department of the FPS Justice.

In addition, a list of the grey and scientific literature relating to the two databases and the studies based on their scientific use was drawn up.

Finally, first contacts were made with the State Archives of Belgium to start the reflection about the preferred archiving procedure for both the products of the IIHA project, extractions and the historical documents inventoried in the course of the project.

These meetings led to taking several decisions. First, it was decided to store on the Social Sciences and Digital Humanities Archive (SODHA) website (https://www.sodha.be/) of the State Archives some documents from the documentation of the two databases considered to be publicly shareable by the representatives of the FPS Justice responsible for these databases (Objective 1), as well as databases metadata (Objective 2) and all IIHA public reports (Objective 1, Objective 2, Objective 3, Objective 4). However, the IHD and the extracts on which the analyses are based cannot be made available to the public, as the National Institute for Criminalistics and Criminology (NICC) does not own these data. Second, it was decided to open a dedicated electronic file on the SODHA project website for the NICC, through which the institution would easily reference its data (e.g., anonymized interview transcripts, databases containing aggregated data, list of values) and documents, not only for the benefit of the IIHA project but for all projects of the Institute.

3.2. Developing the Integrated Historical Database (IHD)

The purpose of the Integrated Historical Database (IHD) is two-fold. First, it aims at storing and preserving historical datasets. Second, it aims at facilitating their exploitation, notably by matching the records of individuals across different data sources. To achieve these objectives, a graph database technology was used[3]. Contrary to the relational databases that represent information as records in tables, graph databases represent information as nodes connected by relationships. For example, we can have nodes representing persons and detentions, and relationships specifying which person has been involved in what detention (i.e., which person node is related to which detention node).

A first version of the Integrated Historical Database (IHD) was produced in 2022, based on the software developed in the FAR project[4], a previously completed BELSPO-funded project[5] aimed a studying radicalization and using the same source of data, i.e., data from CJCS and SIDIS (Jeuniaux et al., 2022). This first version (IHD Version 1) was presented in the second follow-up committee meeting of the IIHA project on 20 October 2022.

To meet the specific needs of the IIHA project and satisfy future scientific needs it was decided to create a second version of the IHD (IHD Version 2), not only of the database itself, but of the

---

[3] More precisely that of Neo4j (https://neo4j.com/).
[4] https://incc.fgov.be/FAR
[5] Project # BR/175/A4/FAR

software used to create the IHD. The IHD Version 2 was presented at the IIHA study day on 14 November 2023 (see Jeuniaux et al., 2024).

The motivation behind IHD Version 2, and the differences with IHD Version 1 are shown in the Table 1 below.

Table 1 – Differences between Version 1 and Version 2 of the IHD

	IHD – Version 1	IHD – Version 2
Purpose	The purpose of the system is **intimately related to what was done in the FAR project**. It is oriented towards realizing specific analyses and providing a proof of concept of how links can be drawn between records originated from its two data sources.	The purpose of the system fully meets the **needs of the IIHA project**. It is designed to archive data, and allows their exploitation, including drawing links between records originating from different sources. It is designed towards the needs of future projects.
Software	The software of the IHD is a set of scripts that are specific to the selected data sources, and it not easily portable to other data sources. No documentation is available to explain its usage and functioning. The software is hard to **use and maintain**.	The software used to develop and use the IHD was completely rewritten and made more generic so that it could be adapted to different data sources. The usage and functioning of the code are documented. The software is **easier to use and maintain**.
Data selection	**Only a selection of data** from the data extractions, were stored in the IHD, those that were susceptible to be useful to compute statistics on recidivism and criminal careers. The most relevant tables were selected, and a selection of rows was chosen.	**All the data** extracted from the Central Criminal Records and SIDIS Greffe, were stored in the IHD, hereby satisfying the objective of archiving the State collections, and satisfying scientific objectives beyond the study of recidivism and criminal careers.
Data modeling	Some **data transformations** were performed on the raw data before it was modeled as a graph. Therefore, it was **difficult to distinguish the raw data from the new data** (e.g., the data from some tables were merged).	No data transformations were performed on the raw data. All **raw data was kept untouched and clearly distinguished from the new data**, hereby satisfying the objective of data preservation.
Data integration	The linking procedure used to match data from the different sources **did not allow an easy exploitation** of the data for the study of recidivism and criminal careers and needed to be improved.	The linking procedure has been **improved in terms of reliability and clarity**. A **new *person* node** has been created with links to its corresponding data sources. The system **facilitates the exploitation** of the data for the study of recidivism and criminal careers.

Like in the first version of the IHD, in the second version of the IHD, the goal is to find which persons records from SIDIS-Greffe (SIDIS) and which persons records from the Central Criminal Record (CJCS) concern the same persons. Each person record is represented by a node (see Figure 1).
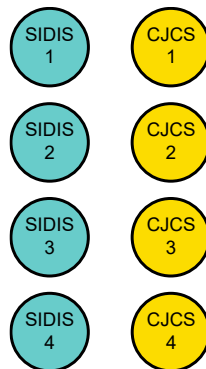
Figure 1 – nodes representing records in SIDIS and CJCS

To represent which person records are related to which actual persons, we create new person nodes which point towards their records in the original data sources, SIDIS or CJCS (see Figure 2). We therefore have person records nodes and person nodes. How these new person nodes are created is the subject of the integration procedure that has been designed and is described in Jeuniaux et al (2024). We nonetheless provide some hints about the procedure below.
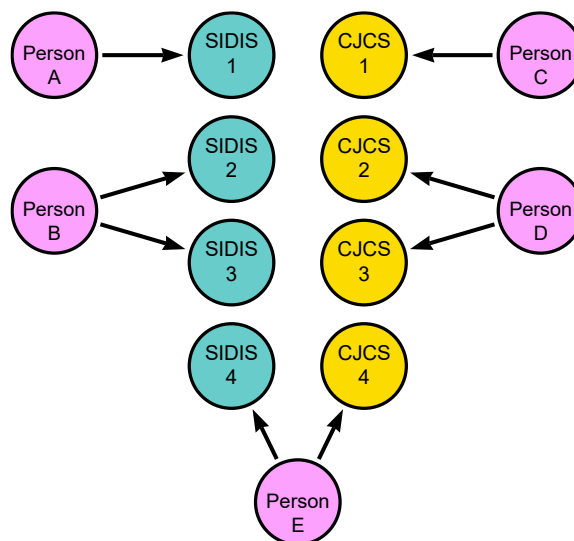
Figure 2 – Person nodes connected to records in SIDIS and CJCS

The nodes representing person records in Figure 1 – whether person records belonging to SIDIS or person records belonging to CJCS – are associated with integration nodes. Integration nodes contain a standardized version of information contained in person records (e.g., information on names, dates of birth).

These integration nodes may be used to connect two person records in the graph through the node that they share (see Figure 3). Figure 3 represents a fictional situation in which we try to determine whether a person called James Bond in SIDIS, is the same as a person called Gems Bont in CJCS. Is James Bond in SIDIS the same person as Gems Bont in CJCS, although their names are not identical? To answer this question, we must compare the two person records by using the available personal data.
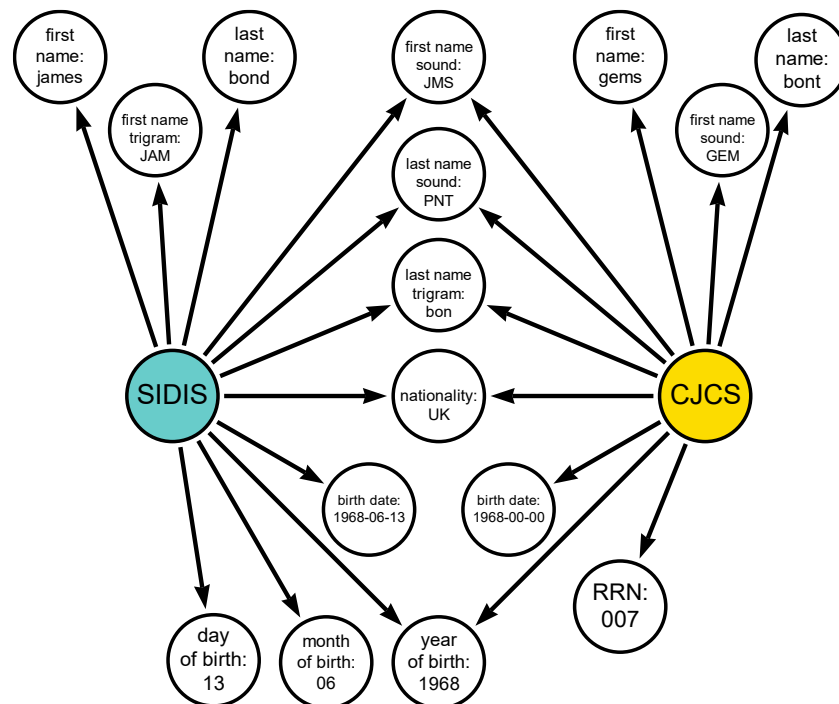
Figure 3 – (fictional) records in SIDIS and CJCS exhibiting similar personal information

As can be seen in Figure 3, integration nodes radiate from the SIDIS and CJCS person records nodes. Those person records nodes are connected to each other through the year of birth (1968), the nationality (British), the first three letters of the last name ('bon'), as well as the phonetic representation of the first name ('JMS') and the last name ('PNT').

Such integration nodes play a role to select candidates that we wish to compare. Proper selection procedures must be designed in order to reduce the number of candidates to compare. Indeed, it is not feasible to perform the trillion of comparisons of all possible pairs of person records that an uninformed, brute force, approach would entail. Candidates are determined across data sources (i.e., between SIDIS and CJCS) but also within each data source separately (within SIDIS alone and within CJCS alone) in order to detect person records duplicates.

Six types of methods have been designed to find the candidates that will be submitted to a comparison procedure. M1 selects candidates who share the same National register number (RRN), a unique identifier used to identify Belgian nationals and residents. In other words, if a SIDIS person record is connected to a CJCS person record through a RRN integration node with a particular value (e.g., 123456789), they will be compared to each other. M2 selects candidates who share the same first three letters of a first name and the same first three letters of a last name. M3 selects candidates who share the same phonetic representation of a first name and the same phonetic representation of a last name (the phonetic representation being based on the Double Metaphone algorithm). M4 and M5 are like M2 and M3, except that they take into account the possibility that in one record the first name has been mistakenly recorded in the last name field and vice versa. M6 selects candidates who share a same date of conviction.

For each selection method, the candidates will be compared to each other. A comparison uses the integration nodes that are available (as seen in Figure 3) and different mechanisms (e.g., textual similarity measurement, weighting scheme). When they are deemed to be sufficiently similar, an integration node will be drawn between the two candidates, while recording a normalized measure of similarity (a number between 0% and 100%) and the weight, i.e., the amount of evidence that has

been used to assess the similarity (a number between 0 and 40). These two values are helpful in assessing the quality of a link. For instance, if the amount of evidence being examined has a weight of 20, and that evidence that match between records weight 14, the normalized similarity measure is 14 / 20 = 70%. These integration links are then used to assess the added value of the six methods.

Next, these links are used to create components. A component is a set of nodes that are linked with one another by the links that have been specified. The normalized measure of similarity and the amount of evidence stored in every link will be used to define different scenarios. Other parameters will be used as well to define scenarios (for details, see Jeuniaux et al, 2024). Each scenario will therefore determine a specific set of components. Within each scenario, for its specific set of components, each component is then associated to a specific person node (like in the Figure 2).

For example, we could choose integration links that present a very high value of normalized similarity and a large quantity of evidence. This would likely reduce the number of false positives (the number of records that are linked but that do not belong to the same person) but decrease the number of false negatives (the number of records that are not linked but nonetheless do belong to the same person). Instead, we could reduce these values, which would likely increase the number of false positives, but decrease the number of false negatives. Using different configurations, we can measure the impact of some methodological choices on the results. In total, 24 configurations have been tested. We therefore have 24 distinct sets of person nodes (see Figure 4).
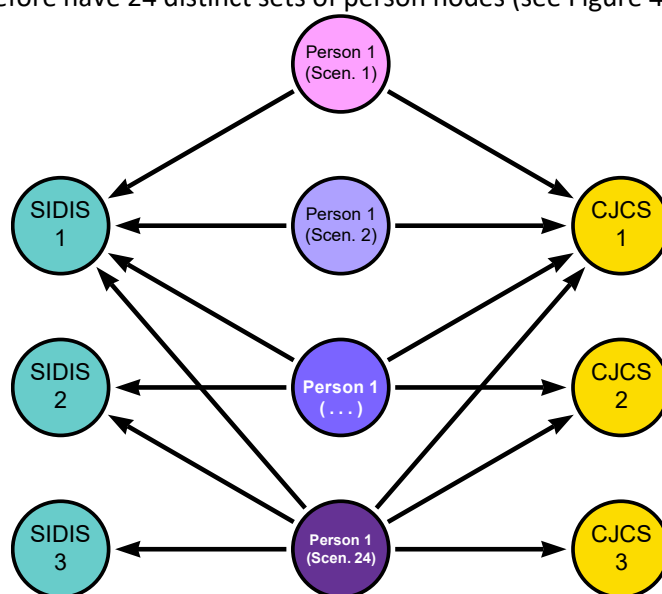
Figure 4 – different integration scenarios exhibiting different person nodes

Technologically speaking, the IHD was developed on a Tuxedo Book XUX7 (Gen 13. Intel Core i9-11900 K, 128 GB of RAM) with an iStorage diskAshur PRO2 encrypted SSD, the Tuxedo / Ubuntu 22.04 LTS operating system, Python 3.11.3 and Neo4j 4.4.4.

3.3. Studying recidivism and criminal careers

As the Integrated Historical Database (IHD) was only operational by November 2023, due to the complexity and high technicality of developing the system, it was not possible to complete all the analyses initially announced before the end of the project. However, these analyses have been initiated (e.g., see Jeuniaux, Mine et al, 2024) and will be continued as part of the research activities

of the Research Unit on Recidivism and Criminal Careers (CReCC)[6], and the Database on Offender Trajectories (DOT) project[7]. Moreover, as can be seen below, several other analyses were carried out on the basis of the Central Criminal Record. Most of these will be submitted for publication in 2024.

As already mentioned in the state of the art of this report, the study of recidivism can give rise to complex methodological discussions, if only because there is no scientific consensus on a definition of recidivism (Maltz, 1984; Zara & Farrington, 2016). The study of recidivism remains obviously eminently linked to the research questions that one proposes to address (Laub & Sampson, 2001) but also to the sources of the data (e.g., official justice databases, self-reported delinquency surveys, public documents such as press articles) that all may present biases, imperfections or errors.

The interpretation, scope, and limitation of recidivism statistic will depend also on the definition of recidivism that is chosen, and the way it is operationalized (Maes et al., 2021). The operationalization of recidivism statistic will depend on three essential components: the nature of the population of interest being studied, the type of measurement being taken, and the follow-up period during which the measurement is performed (Andersen & Skardhamar, 2017).

First, the population of interest refers to the category of offenders for which recidivism is measured, for example arrested persons, people for whom a case is opened at the public prosecutor's level, people convicted (by a court), prisoners, people on probation, etc. These may belong to specific subpopulations defined, for example, in terms of the type of offense they committed (as is the case here with terrorism offenders) or the type of sentence imposed on them, etc.

Second, the choice of a recidivism measure depends on how recidivism is defined. Do we define it as a new arrest or a new conviction? What type of recidivism is involved? Was the offender again involved in acts of a similar nature to that of the original offense[8] ?

Third, the follow-up period during which the person is monitored may be shorter or longer depending on when we begin to observe whether or not the person is reoffending and when we stop observing.

The choices made regarding these three components will greatly influence the reported results on recidivism (e.g., affect the estimate of the prevalence of recidivism or its interpretation). Moreover, because studies can vary widely in these components, it is often difficult to compare their results. For a more detailed discussion about the impact of methodological choices on results, see Maes et al. (2021)

---

[6] This research team was created at the end of 2021, as part of the Criminology Department of the NICC, thanks to structural funding provided by the Federal government. It is known as the "CReCC", as per its name in Dutch and French: Cel Recidive en Criminele Carrières (CReCC) / Cellule Récidive et Carrières Criminelles (CReCC). For more information, see https://incc.fgov.be/CReCC

[7] The DOT project was started in 2023, and was funded by Belspo as part of the INFRA-FED call. For more information, see https://incc.fgov.be/DOT

[8] In Belgian criminal law, recidivism consists of two terms. The first term of recidivism is a final conviction (and not a prosecution or dismissal) of a sentence for a felony, misdemeanor, or infraction. The second term is the fact of committing a new offense after this conviction for which the perpetrator is apprehended and brought before a court. We speak of special recidivism when the two terms of recidivism concern identical or similar offenses and general recidivism when the offenses are of different kinds. See Tulkens and Van de Kerchove (2007, p 472).

It should also be mentioned that certain biases (e.g., breaks in the time sequence, change of legislation) may affect the validity of the results proposed below. Contextualisation elements must be taken into account in order to fully understand the scope of the results. As the various work packages were conducted simultaneously, the analyses in this work package unfortunately do not benefit from all the lessons learned from the business analysis in work package 4. Studies on assessing the influence of breaks in continuity, change of legislation or the recording of new elements could therefore be relevant.

3.3.1. Survival analysis of persons convicted of terrorism

This study focused on individuals convicted of terrorism in Belgium between 2006[9] and 2020[10] and registered in the Central Criminal Record. The aim of the analysis is threefold: to describe the population of people convicted of terrorism (demographic information and criminal careers); to describe recidivism in this population (prevalence, time to recidivism, type of recidivism offense); and to study the factors that influence the risk of recidivism and the time to recidivism.

The first two points had already been addressed in a former study conducted by some members of our team using the same data (Mine et al., 2021). However, in this study, we defined recidivism differently from Mine et al. (2021). Here, recidivism is defined as an offense committed after the reference conviction, which is the first terrorist conviction. In Mine et al. (2021), recidivism was defined as any offense included in a judgment subsequent to the reference conviction, regardless of when the offense was committed.

To enhance our understanding of the time[11] until recidivism among individuals convicted of terrorism and to investigate potential factors associated with recidivism, we conducted a survival analysis (Cox, 1972). Survival analyses are statistical methods that primarily focus on the duration until a specific event occurs (such as recidivism). These techniques are particularly useful when dealing with censored data—instances where a portion of the sample does not experience the event during the observation period.

Specifically, we employed Cox Proportional Hazards models. These models center around the hazard function, which, in our context, represents the instantaneous probability of recidivism at a given time for individuals who have not yet reoffended. By using these models, we can estimate how predictor variables (such as gender and age) may impact the hazard of an event with the following assumption: the influence of a predictor is constant over time and implies constant hazard ratios.

In practical terms, we have run a series of Cox proportional hazards models to find which factors influence the instantaneous risk of – and thus the time to – recidivism. These models differed according to the predictors (i.e., the independent variables or covariates) included. Models could include the following predictors : demographic factors (e.g., gender, country of birth, age at the time

---

[9] The law punishing terrorist offences came into force in 2003, but the first conviction took place in 2006.
[10] This year corresponds to the date of extraction of the data from the Belgian Central Criminal Record: the 23rd October 2020.
[11] Defined here as the time between the sentence for the first terrorist offense (i.e., the reference offense) and the date of the recidivist offense, if any.

of the terrorist offense); criminal career indicators (onset age, prior convictions[12], crime mix index for prior convictions, year of the judgment[13] and seven binary indicators of specific categories[14] of prior offense); characteristics of the individual associated with the reference offense (three binary variables indicating if the individual was a leader of a terrorist group, a member of a terrorist group, had provided material assistance); and characteristics of the offenses committed contemporaneously[15] to the reference offense (crime mix index, three binary indicators of specific categories[16] of offense).

Our objective was to identify the optimal model in terms of both explanatory power and simplicity. To achieve this, we employed a stepwise variable selection approach, considering both forward and backward versions. In the forward version, we began with a null model and iteratively added factors that most significantly improved the model's performance. Model performance was evaluated using the Akaike Information Criterion (AIC), which balances model parsimony with its ability to explain the data. The process terminated when no additional factors enhanced the model. In the backward version, we started with a full model and successively removed the least influential factors based on model performance. Again, the procedure ceased when further factor elimination no longer improved the model. When the models obtained from these two procedures differed, we selected the model with the lowest AIC. Ultimately, the best model was determined by comparing AIC values across all model series. The rationale for running distinct model series was to account for the independence[17] of predictors.

Importantly, all the models focused on general recidivism, as the number of cases of special recidivism was too low to allow any statistical modelling of the data.

3.3.2. Analyses of the criminal trajectories of female sex offenders (with comparison groups).

Studies about criminal careers have increasingly turned towards those persons who have involved in sexual offending (e.g., Blokland & Lussier, 2015; Blokland, 2018). Most of the studies zoom in on males that committed sexual offenses. Few researchers have turned to the study of the criminal trajectories of women that committed sexual offenses (but see, amongst others, Blokland & Van Os, 2010; Wijkman et al., 2011). Furthermore, the samples of these studies remained relatively small.

In this study, the focus will be on the criminal careers of females. Two research questions will be tackled. First: what are the criminal trajectories of women convicted of sexual offenses? The second

---

[12] This variable was tested in two versions: as binary variable (i.e., had a prior conviction or not) and as a continuous variable (i.e., the number of prior convictions)

[13] This variable is used to assess whether there may be a temporal effect, in the sense that certain years reflect a certain jurisprudence, or a methodological bias induced by the time required to complete the procedure.

[14] These 7 categories were: drugs, property without violence, public order, traffic, violent crime, violent property crime and other.

[15] Defined as belonging to the same judgment.

[16] Namely: public order, property without violence and other. Other categories were too rare in the sample and therefore not tested in the models.

[17] So, for instance, we did not run models with both versions of the "prior convictions" variable, nor did we run models that included both the crime mix index and the binary variables associated with the offense categories.

research question is: To what extent do women convicted of sexual offenses differ from women convicted of non-sexual offenses in terms of the nature and sequencing of their criminal trajectory?

Data for this research are extracted from the national Central Criminal Records in Belgium. Data include all female offenders with at least one conviction in the period 1995-2020 (i.e. they are 'active' offenders), and include all convictions of the selected persons, also prior to 1995. Two large groups are composed. The Female Sex Offender (FSO) group consists of females with at least one conviction for a sex offense in their entire criminal past (n = 2342). A random sample of ca 10% of all females without a sex offense in their entire criminal past (up to the date of the data extraction) consists of 10.322 persons (FNSO). In this comparison group, a subgroup is identified of all females with at least one conviction outside of police court (i.e., not a misdemeanour or traffic offense) (n=2620) – FNSOPP or Female Non Sex Offenders - Police Plus.

In terms of the methodology, descriptives are provided about key criminal career dimensions. The career dimensions between the groups are compared and include both a p-value and the effect size of the difference (based on èta-squared). A next step consists in an analysis of the longitudinal panel data by means of a group-based trajectory model (GBTM). The trajectory models are based on all convictions (including any sex offense-related convictions) and are calculated with STATA 16. Cut-off ages for the trajectory model data are from 11 years old up to 80 years old. Models are calculated for 1 up to 10 trajectory groups per group (FSO, FNSO, FNSOPP). A weighting for year of birth is conducted based on the year of birth of the FSO. Models with the most advisable AIC and BIC (i.e. the AIC and BIC that are the lowest) are used to identify the optimal model, which in what follows is a model with 5 trajectories (and a spline at age 30). A 5-trajectory model is calculated for the three groups (FSO, FNSO, FNSOPP). Comparisons between trajectory groups of the FSO model on the one hand and FNSO or FNSOPP models on the other hand are calculated (t-tests).

3.3.3. Analysis of the duration of criminal careers of sex offenders

This study zooms in on the duration of criminal careers of sex offenders. Duration is a dimension of the criminal career that is rarely studied (e.g., Farrington, 2019). This is due to practical reasons, first and foremost a long follow-up period. Furthermore, studying the duration of the criminal careers of sex offenders is relevant, in that this population is often the focus of stereotypes that may trickle into policies and legislation (e.g., Robert, 2020).

This includes the stereotypical idea that sex offenders are highly specialised offenders that continue to commit sexual offenses until later in life. One way of looking at this is to not only study a criminal career that encompasses all types of offenses, but to distinguish between the sexual criminal career (i.e. all sex offenses) and the criminal career related to non-sexual offenses.

The central research question here is of a descriptive nature: how long is the criminal career of sex offenders? In the same vein, the duration of the sexual criminal career and the non-sexual criminal career will receive separate attention.

Data for this analysis are derived from the national Central Criminal Records. Data are thus limited to all offenses that have led to a conviction or any other type of judicial decision registered in the national Central Conviction Records.

For this analysis, a selection is made of all persons with at least one conviction (for any type of offense) in the period 1995 – 2020, that have at least one sex offense-related conviction in their entire criminal past (including before 1995) and that have reached or surpassed the age of 70 years at the end of the follow-up period. Persons are also included in case they have deceased before the end of the follow-up period if they had reached the age of 70 years prior to their death.

The descriptive information that is presented in this report serves as a springboard for further analyses that draw on more advanced statistical techniques. However, given the importance of disposing of such descriptive information, we believe it is valuable to report about descriptive information about several criminal career parameters, including the age of onset, the age of termination, duration of the criminal career, the frequency of offending (i.e. the frequency of convictions), the crime mix (based on 8 categories: sex offenses, violent offenses, traffic offenses, drug offenses, property offences, violent property offenses, public order offenses, and a rest category), the time between the last offense of any type (that led to a conviction) and the end of the follow-up (either end of 2020 or the date of death), and the time between the last conviction for any type of offense and the end of the follow-up.

Separate attention goes to the age of onset of the sexual criminal career, the termination, the duration of the sexual criminal career, the frequency in convictions that include a sex offense, the sexual crime mix (based on 7 categories: rape, sexual assault, sexual exploitation, child pornography, voyeurism, indecent exposure, other type of sex offense), the time between the last sex offense (that led to a conviction) and the end of the follow-up and the time between the last conviction for a sex offense and the end of the follow-up.

3.3.4. Continuity in sexual offending based on conviction data

One persistent stereotype about sex offenders relates to the age of onset in sex offending. The stereotype is that young sex offenders will evolve to become adult sex offenders, which suggests a high degree of continuity in sexual offending between different life phases. This stereotype flies in the face of empirical evidence in a number of countries (e.g., Blokland & Lussier, 2014).

Based on the national Central Conviction Records in Belgium, this question can be tackled for all sexual offenders that have reached the age of 70 years (see the analysis of duration). For this population, particular attention goes to the continuity in sexual offending across different life phases. Alongside the offense date, continuity can also be mapped based on the conviction dates for sexual offenses. The following is based on the offense dates only.

In this analysis, the continuity in sexual offending is done in two descriptive ways. Future analyses can develop these results more in detail. Here only descriptives are provided.

A first way relates to a check of two large life phases that are oftentimes used in the literature, i.e. the period of childhood, adolescence and young adulthood, up to the age of 25 years, as compared to the period of adulthood, from age 26 onwards up until old age. Continuity can be mapped based on the first offense date for a sexual offense across both time periods. Continuity in sexual offending is when persons have committed a sexual offense (for which a conviction is given) in both phases of life. A lack of continuity in sex offending would be that either a sexual offense was committed in the

first life phase, but not in the latter, or inversely, that a sexual offense is committed in the second life phase, but no such offense had been committed prior to the age of 26.

A second way distinguishes between four phases in life. A first period is the period of childhood up until adolescence, running up until 17 years of age. A second life phase is that of late adolescence and young adulthood, between the ages of 18 and 25. A third phase relates to the period of adulthood, between 26 years and 59 years of age. A fourth and last phase takes into consideration all offenses (and convictions) from the age of 60 and later. Continuity in this second approach takes on several possibilities, but will always include a sexual offense committed in at least two life phases, with a maximum of four phases.

3.3.5. Differentiating homicide offenders based on their criminal careers: A latent class

analysis

This study focused on offenders convicted of intentional homicide in Belgium between 1995 and 2020. The objectives were twofold: first, to describe the characteristics of the homicide offenders in terms of demographics (e.g., gender, age the time of the homicide) and criminal careers (e.g., onset age, nature of the prior offenses, crime mix) ; and second, to create an empirical (i.e. data-driven) typology based solely on their previous convictions. The rationale behind the second objective directly stemmed from a prior scoping review on empirical typologies of homicide offenders (Schils et al., submitted). Indeed, this review has revealed that criminal career information is severely underused in the literature for the classification of homicide offenders, both in terms of the number of publications and in terms of the quality and quantity of the criminal career information used, if any.

In this study, the unique source of data used was the Central Criminal Record. With regard to the first objective, this had two consequences: first, the information relating to the offender itself was limited; second, the criminal careers had to be reconstructed on the basis of the offenses officially recorded (i.e., those that led to convictions). For the criminal careers, we only considered the offenses that were committed before the homicide.

For the empirical classification of the homicide offenders, we applied a latent class analysis (LCA) on their criminal careers. The latent class analysis (LCA) is a model-based clustering technique that classify individuals into exhaustive and mutually exclusive categories, or latent classes, depending on how they typically respond to a collection of categorical indicators (i.e., manifest variables). In the criminological literature, LCA has been used to develop typologies of various offenders : sexual offenders (Spaan et al., 2020), burglars (Fox & Farrington, 2012) and homicide offenders (Khoshnood et al., 2023; Vaughn et al., 2009), among others. In our case, the LCA was carried out using a series of 16 binary indicators, each reflecting whether an individual had been convicted of a particular type of crime (8 categories: sexual crime, violent crime, violent property crime, property crime without violence, public order crime, drug-related crime, traffic crime or other crime) during a given period of their life (juvenile[18] or adult) *before* the homicide. To clarify, an offense was considered only if its date was anterior to the homicide, independently of the judgment date. These indicators were selected so

---

[18] Following the work of Spaan and colleagues (2020), the threshold between the juvenile and adult periods was set at 25 years of age. An offense was considered to have been committed in the juvenile period if the age of the offender at the time of the offense was less than 26 years; and in the adult period otherwise.

that the LCA could unfold different criminal careers patterns in terms of crime variety and career continuity. In practice, we generated a series of latent class models varying in the number of classes (from 1 to 7) via the poLCA package (Linzer & Lewis, 2011) in R. Then, we selected the best model on the basis of fit statistics, parsimony measures and classification statistics. The interpretability of the models was also considered in the model selection. The classes of the best model were described in terms of the 16 indicators used, as well as demographic variables and general career criminal indicators (e.g., onset age, crime mix).

3.4. Studying the feasibility of a recidivism module

This feasibility study began with meetings with the Business Intelligence Service and the strategic management of the FPS Justice (n = 2). Discussions led to a collaboration agreement on the development of a prototype recidivism monitor for future implementation in the 'Justice Dashboard' application.

The NICC was in charge of the business analysis (definition, choice of units of account and variables, etc.), drawing on its knowledge of the Central Criminal Record and its scientific expertise in recidivism, while the FPS Justice would be responsible for developing the tool, supplying it with information, maintaining it and making it accessible to the public by integrating it into the Justice Dashboard.

With a view to carrying out the business analysis, several meetings (n = 5) were held in conjunction with the Criminal Policy Department on how to envisage the production of statistics from the Central Criminal Record Department according to the needs of each service, as well as to improve the research team's knowledge of recording practices in the Central Criminal Record and its content.

At the end of these meetings (recorded and transcribed), a concept note on the development of a recidivism monitor was drafted. With a view to its operationalization, i.e. its transposition into specifications, two interviews were conducted with two international experts on the Swiss and Dutch experiences in developing a recidivism monitor.

A set of specifications based on the business analysis of the Central Criminal Record was drawn up and sent to the Criminal Policy Department, the Business Intelligence Department and the Strategic Directorate of the Federal Public Service Justice in order to begin working together on the development of the prototype for a recidivism monitor based on the Central Criminal Record database. The implementation of these specifications began in September 2022 thanks to the provision of a developer by the Business Intelligence Service and the Strategic Directorate of the FPS Justice.

Monitoring meetings involving several members of the support committee (every 14 days) and technical meetings with the developer (every other 14 days, alternately) were set up.

Several tests on the extracted data were carried out in the first quarter of 2023 and helped to refine the definitions in the specifications. A number of problems were identified in the data, and these were discussed at various meetings, which enabled solutions or workarounds to be considered. The first tests relating to the expected products (tables and graphs) were undertaken from April 2023. A

prototypical version of the recidivism monitor was available at the end of May 2023, as initially planned.

For more information on the methodology used and the implementation of the specifications in the development of the prototype recidivism monitor, see Huynen, Mine et al. (2024).

4. SCIENTIFIC RESULTS

4.1. Documenting the Criminal Justice Databases

The Criminal Justice Administration System (CJAS) comprises a multiplicity of recording systems (Mine & Vanneste, 2011; Detry et al., 2021), including the Central Criminal Record (which processes data relating to convictions, suspensions and internments) and the former SIDIS-Greffe database (which records data relating to detentions). The data recorded in these two systems in the form of digital files are considered by the IIHA project to be part of the Justice system's heritage, which has a number of functions.

Firstly, these files, which are national in scope, can play a role of institutional memory in terms of the evolution of recorded criminality in relation to representative groups of offenders, as well as the practices at work within the criminal justice administration system. Secondly, they can also be mobilized to enable public institutions to justify their actions to political representatives and citizens. Thirdly, as in the IIHA project, these files can be exploited for scientific, statistical and historical purposes, as part of research programs designed to study the operation of the Criminal Justice Administration System (CJAS) and its effects on society.

Indeed, in relation to this third point, the exploitation of these data represents an enormous potential for criminological research, and can provide important policy insights into, for example, (a) the repression of delinquency and decision-making behavior, (b) the execution of prison sentences, as well as (c) certain characteristics of incarcerated people and their backgrounds. And all this over a relatively long period of time.

Finally, the information recorded can be used as a basis for evaluating certain criminal policy decisions and developing new policies.

Over and above the legal obligation for public services to archive this type of digital production, for the reasons outlined above, it has to be said that documentation relating to the information contained in these different recording systems, their structuring and relationships, and the underlying encoding practices, is often scattered and incomplete.

This can be explained by the numerous demands and workload placed on staff, the limited human capacity and the rapid pace of technological change, which are not conducive to the documentation of this digital heritage. Moreover, it is not clear whether the FPS Justice has an archiving policy or a systematic documentation strategy for what it generates or receives in the course of carrying out its missions.

The 'institutional memory' is all the more fragile because expertise is usually held by one or two people, and dies out when they leave the institution or change jobs. Existing documentation may be lost or misplaced. However, documenting the structure of these recording systems and their contents, as well as archiving the documentation, means ensuring that the data remains accessible and usable, and thus preserving a certain continuity for future research.

The research activities carried out as part of this first workpackage of the IIHA project led to the production of four main results:

1) Firstly, the NICC obtained two complete extractions of raw data from the Central Criminal Record (CJCS), which contains conviction data (concerning people convicted

over the period 1995-2020), and the SIDIS-Greffe (SIDIS) detention database (concerning people detained over the period 1974-2014);

2) For each database, the research team has collected, sorted and organized a wide range of documents useful for understanding their structure and content (correspondence, technical reports, registration manuals, nomenclatures, interface training and presentations on the history of the databases, their architecture and content, working documents for the development of statistics);

3) A survey of the relevant grey and scientific literature relating, on the one hand, to each of the two databases and, on the other, to their scientific exploitation;

4) Based on the review of this heterogeneous material (interviews, technical reports, personal knowledge, database explorations) but also on previous research reports, one report on the Central Criminal Record (Huynen, Jeuniaux, et al., 2024) and a second on the SIDIS-Greffe database (Maes et al., 2024) have been written. These two reports provide a detailed description of each of the two databases, with a view to meeting the needs of criminological research. To this end, the data structure of both databases is described, as is the nature of their respective data. These reports highlight the relevant tables to be taken into consideration (and the relationships between them) with a view to their use for scientific, statistical and historical purposes, more specifically as regards the production of knowledge on the issues of recidivism and criminal careers. They also include a list of the documents indexed and the bibliographies established.

4.2. Developing the Integrated Historical Database (IHD)

The available data from the Central Judicial Record (CJCS) and the detention database (SIDIS) have been imported into the IHD in the form of a graph, amounting in the creation of multiple nodes and relationships. The CJCS data was modelled according to the original data structure (see Figure 5), as well as the SIDIS data (see Figure 6). This step in the development of the IHD satisfies the objective of storing and preserving the data.
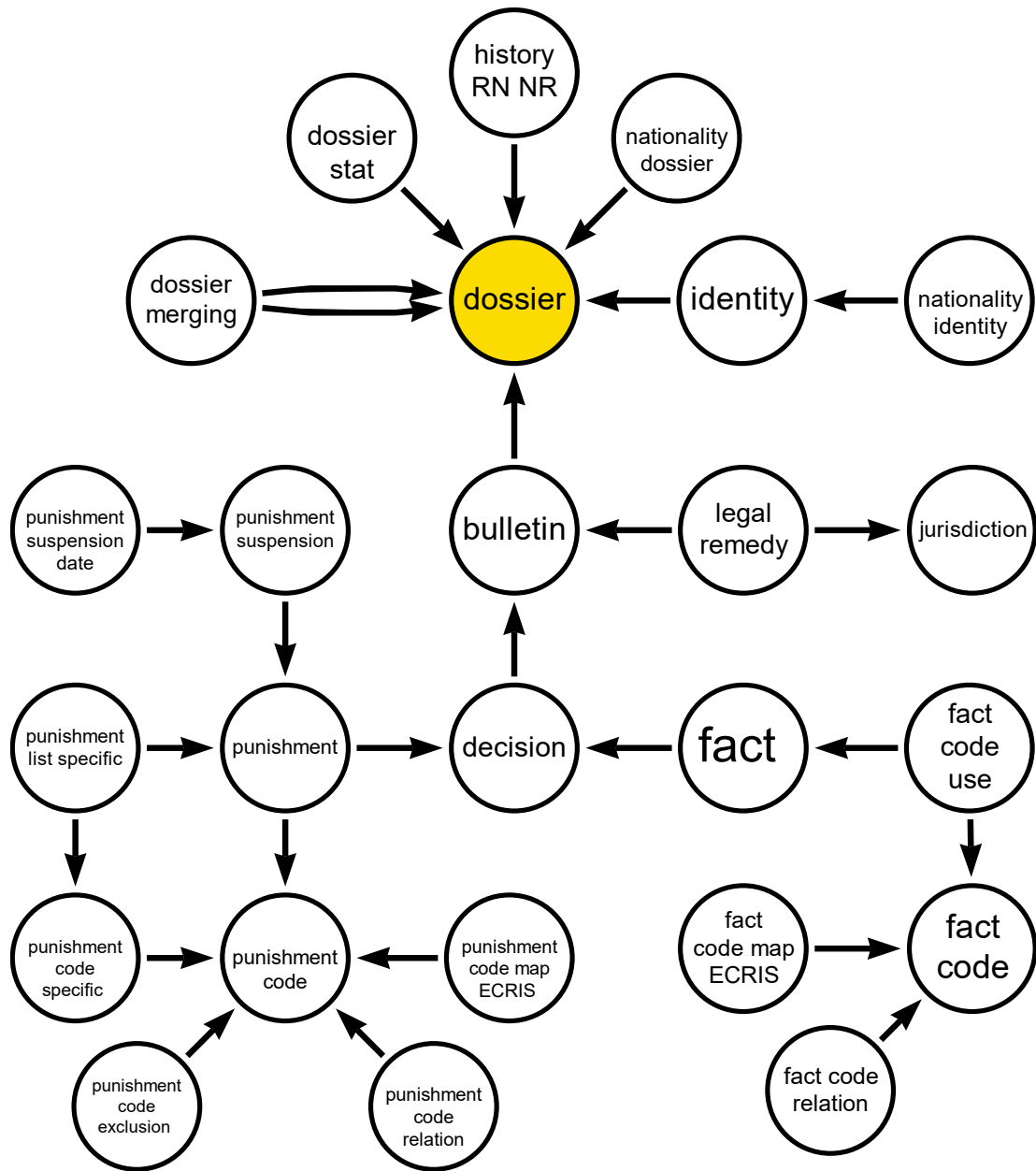
Figure 5 – key modelling structure of CJCS in the IHD
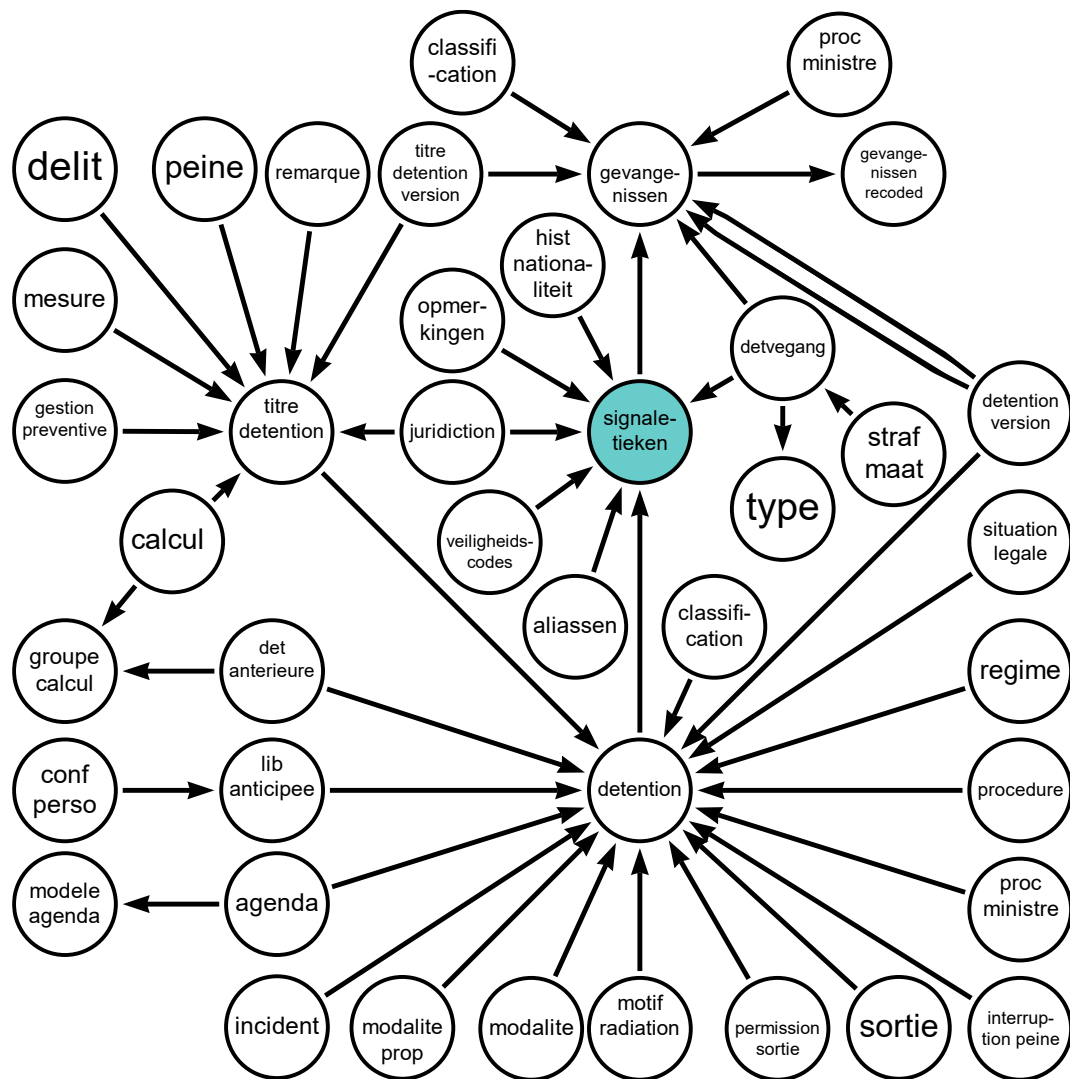
Figure 6 – key modelling structure of SIDIS in the IHD

The importation process led to the creation of 3.860.089 nodes representing person records in CJCS and 365.401 nodes representing person records in SIDIS (i.e., 4.225.490 person records nodes in total). To facilitate the exploitation of the data, it must be integrated at the personal level. Concretely speaking the goal of the integration step is to associate appropriate person nodes to these 4.225.490 person records nodes.

Integration nodes have been created for 20 types of elements: (1) National registration number (RRN), (2) gender, (3) nationality, (4-6) first name, (7-9) last name, (10-13) date of birth, (14-15) place of birth, (16-19) place of residence, (20) date of judgment. If CJCS records usually have a RRN (89% do), only 15% of the SIDIS records have this information, which will have a significant impact on the integration attempt.

On the basis of the six methods of integration (M1-M6), integration links have been created between the person records nodes, whether within SIDIS only (intra-SIDIS orientation), within CJCS only (intra-CJCS orientation) or between SIDIS and CJCS (inter-SIDIS-CJCS orientation).

The characteristics of the six methods (M1-M6) have been analyzed, while looking at the quality of the links being created (both in terms of weight of evidence, and filtering the links according to three types of thresholds: a normalized similarity of minimum 14 / 20 = 70% (Large threshold [L]), a minimum of 20 / 25 = 80 % (Medium threshold [M]) and a minimum of 27 / 30 = 90% (Narrow threshold [L]).

The usefulness of filtering out links according to a threshold is to potentially find a situation where we reduce both the number of false positives and the number of false negatives. Unfortunately, in this study we did not try to ascertain the presence or absence of these errors, as no principled way to identify these errors was designed. Instead, we merely checked a few cases of matches between records and evaluated whether they were false or true positives. Doing so it was apparent that the [L] filter would lead to a greater number of false positives than the [M] filter.

Looking at the links under the [M] filter, the six methods (M1-M6) can be ranked in terms of the number of links they have created: 287.835 links (M2: trigrams), 285.718 (M3: sounds), 52.351 (M1: RRN), 15.487 (M6: date of judgment), 1.545 (M4: inverted trigrams) and 1.507 (M5: inverted sounds). Because not all person records have a RRN, only a relatively modest amount of links can be drawn with M1 (n = 52.351). Therefore, relying on the other personal information (with M2 or M3) is essential.

Next, whereas connections can be established according to three orientations (Intra-CJCS, Intra-SIDIS or Inter-CJCS-SIDIS), 90% of these connections are performed between CJCS and SIDIS (Inter-CJCS-SIDIS orientation).

Because we have six methods (M1-M6), up to six links can be drawn between two person records, leading to 2^6 - 1 = 64 − 1 = 63 different ways of connecting two person records (e.g., two person records can be linked by only M1, or by both M1 and M2, etc.). Among the 63 different connectivity situations, only 46 showed up under the [M] filter. Moreover, some of them are much more frequent that others. Actually, 90% of the situations are of the type M2+M3 or M1+M2+M3. Finally, examining the 46 different types of connectivity situations, we can see that every method appears in a least a few cases, which shows that every method presents some "added value" or specificity compared to the other methods (i.e., none is strictly redundant).

On the basis of the available nodes and integration links, 24 scenarios of connectivity have been defined. Within each scenario, components have been created (one component per person), leading to 24 sets of persons, and therefore to 24 sets of statistics. A very basic statistics is the number of people that have been constructed. This number varies between 3.828.651 and 4.126.144 (this last number being the closest to the maximum number of person nodes that could be created within a scenario: 4.225.490). The number of persons that have a record in both SIDIS and CJCS varies between 52.717 and 295.659. In other words, recidivism rates including detention and conviction data could be calculated for a sample as small as 52.717 persons, and as large as 295.659 persons.

4.3. Studying recidivism and criminal careers

4.3.1. Survival analysis of persons convicted of terrorism

4.3.1.1. Sample

Between 2006 and 2020, a total of 482 individuals aged 16 and older were convicted of terrorism. However, we excluded 4 individuals due to the unavailability of the date of the terrorist offense (referred to as the "reference offense"). For the remaining 478 individuals, we obtained data from the Central Criminal Record regarding all offenses for which they were convicted throughout their criminal careers. This dataset included crucial information such as the date of the offense, the date of judgment, and the nature of the offense, among other details.

We encountered some offenses without a reliable date (approximately 3.3% of all offenses across all subjects). These cases could not be definitively categorized as either prior offenses or recidivist offenses in relation to the reference offense. To ensure an unbiased description of the criminal careers of the terrorist offenders and to facilitate the survival analysis, we retained participants whose criminal history was sufficiently comprehensive. Specifically, we excluded individuals with 30% or more of their offenses discarded based on this criterion. As a result, our final sample consisted of 463 persons convicted of terrorism.

The median age at the time of the reference offense (i.e., the first terrorist offense) was 24 years old (with a range of 16 to 72). The majority of the offenders were males (accounting for 83.6%), and they were predominantly born in Belgium (constituting 59.8% of the sample). The years during which the terrorist offenses occurred spanned from 1997 to 2019, but a significant proportion (87.4%) took place after 2010. As for the judgments, they were rendered between 2006 and 2020, with a notable concentration of judgments occurring after 2014, making them relatively recent (cf. Figure 7).

Based on the qualification of the reference offense as documented in the Central Criminal Record, the majority of the offenders (82.7%) were affiliated with a terrorist group. Additionally, approximately one out of ten individuals (10.8%) were identified as leaders within these terrorist groups. A smaller proportion, less than one out of twenty (4.1%), were involved in the terrorist offense solely by providing material support.
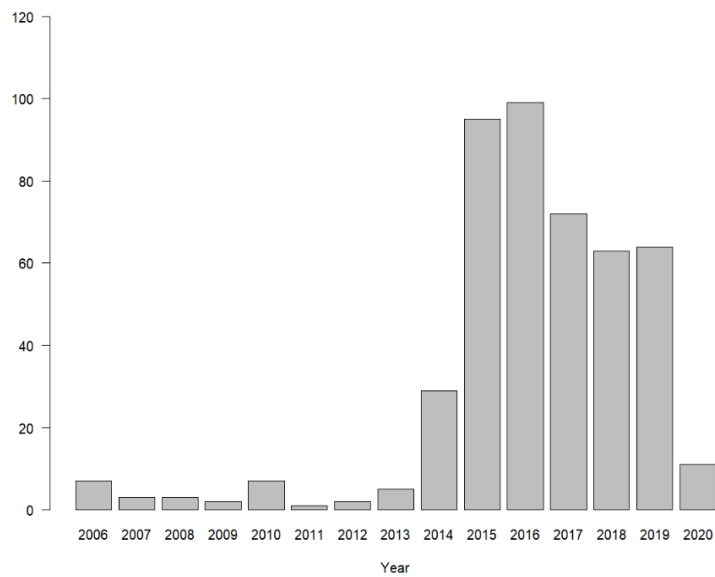
Figure 7 – Distribution of the (first) convictions of terrorism across the years (years of the judgments)

4.3.1.2. Criminal careers of persons convicted of terrorism

Based on information from the Central Criminal Record, nearly half of the sample (46.4%) had prior convictions before the reference offense. Among these individuals, the median age at the time of their first offense was 21 years (with a range spanning from 12 to 72, and an interquartile range of 18 to 26). For the entire sample, the average number of prior convictions[19] stood at approximately 2 (mean=2.3), although the distribution was highly skewed, ranging from 0 to 27 (with an interquartile range of 0 to 2).

The most prevalent category[20] of prior offenses was road traffic violations (accounting for 33.5% of the sample). Additionally, other relatively common offenses included property crimes without violence (19.4%), violent crimes (16.2%), and offenses categorized as 'other' (14.9%). Public order offenses were also notable, constituting 14.7% of the cases.

At the opposite end of the spectrum, sexual offenses accounted for less than 2% of convictions (1.3%), while drug-related offenses were less frequent, affecting less than 10% of the sample (8.6%). The crime mix index, which reflects the diversity of distinct offenses prior to terrorism convictions, averaged 1.2 (with an interquartile range of 0 to 2), indicating a relatively low variety of offenses prior to their involvement in terrorism.

4.3.1.3. Recidivism among terrorist offenders

The prevalence of general recidivism, which refers to the proportion of offenders convicted for an offense committed after the reference judgment, stands at 9.9% (n = 46) when considering all categories of offenses combined. However, when focusing specifically on special recidivism (i.e., reconviction for terrorism), this figure significantly drops to 1.1% (n = 5).

---

[19] To be precise, we refer here to a prior conviction as an offense committed before the reference conviction (i.e., the first conviction for terrorism).

[20] We have categorized the prior offenses in 8 distinct categories: drugs, other, property without violence, property with violence, public order, sex crime, (road) traffic, violence crime.

Interestingly, when compared to the work by Mine et al. (2021), our definition of recidivism—based on the date of the offense rather than the date of the sentence—results in a division of the prevalence of general recidivism by a factor of 2.3, and for special recidivism, by a factor of 5 (specifically, 23% for general recidivism and 5% for special recidivism).

The median time interval between the reference judgment and the recidivism offense was approximately 62.6 weeks (with a range spanning from 0.1 to 404 weeks, and an interquartile range of 25 to 129 weeks). Traffic offenses were overwhelmingly the most common type of recidivism offense, accounting for 63% of the recidivists. Additionally, 15% of recidivists were convicted for public order or violent offenses (n = 7), while 11% were involved in terrorist offenses (n = 5).

It's important to note that the recidivism figures presented above do not account for variations in the duration of the observation period across individuals. For each person, this period commenced with their first terrorist conviction and concluded either with the occurrence of the next offense, their demise, or the end of the study's observation period (i.e., the date of data extraction). The latter two cases correspond to censored data. To address this, the Kaplan-Meier hypothesis allows us to assess the probability of recidivism as a function of time for the entire sample, while considering censored data. Figure 8 visually depicts the probability of recidivism[21] over time since the reference conviction, under the Kaplan-Meier hypothesis.
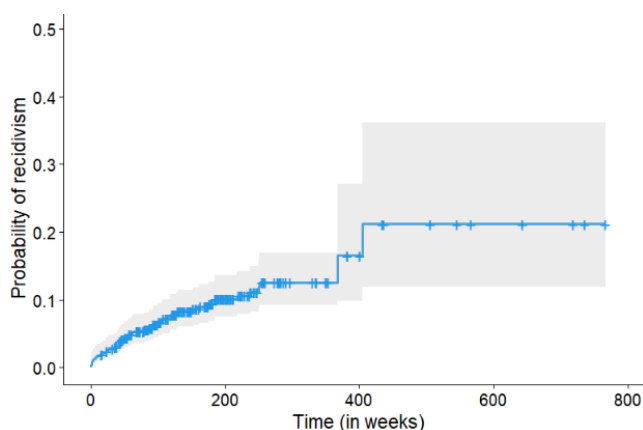
Figure 8 – Cumulative probability of recidivism as a function of time for individuals convicted of terrorism.

+ : Censored Data. Gray : 95% CI.

4.3.1.4. Factors influencing the risk of recidivism : the Survival Analysis

Upon analyzing the entire dataset, our modeling approach (as described in the Methodology) resulted in the identification of a best model that incorporates five predictors:

1. Number of prior convictions[22]
2. Year of judgment[23]
3. Presence (or absence) of a prior drug conviction
4. Leadership status within a terrorist group (or lack thereof)

---

[21] For a given value of x, the figure shows the probability of having reoffended in the interval [0-x].
[22] The best model includes the number of prior convictions as a continuous variable.
[23] Centered around the median judgement year (i.e., 2016) for the sake of convergence.

    5.   Country of birth[24]

Notably, this implies that other factors were not considered to significantly influence the risk of recidivism or the time until recidivism.

The findings reveal several noteworthy points regarding recidivism risk among individuals convicted of terrorism. Let's delve into the key observations:

1. Number of Prior Convictions:
    a. Having a higher number of prior convictions significantly increases the risk of recidivism compared to those with no prior convictions and significantly decreases the expected time to recidivism.
    b. The more previous convictions an individual has, the greater their likelihood of reoffending.
2. Year of Judgment:
    a. The year of judgment also significantly impacts recidivism risk.
    b. For each additional year, the instantaneous risk of recidivism decreases by 15%.
    c. However, this result should be interpreted cautiously due to potential methodological artifacts arising from the relatively short observation period.
    d. In particular, that offenses committed posterior to the reference sentence may have not been judged or encoded in the records yet at the time of the data extraction should not be neglected, especially for individuals sentenced in 2015 and later given the recency of their convictions.
3. Prior Drug Offense:
    a. Although not statistically significant (p=0.06), individuals with a prior drug offense exhibit a large estimated effect.
    b. These individuals are expected to have a 77% lower risk of instantaneous recidivism.
    c. The lack of significance likely stems from the small sample size of individuals with prior drug convictions.
4. Other Factors:
    a. Two remaining factors did not reach statistical significance, likely due to limited statistical power.
    b. However, their directions are intriguing:
        i. Being a Leader of a Terrorist Group: Surprisingly, this seems to have a protective effect on recidivism risk (p=0.13, HR: 0.48; CI: 0.18-1.26).
        ii. Being Born in Belgium: Conversely, being born in Belgium increases the risk of recidivism (p=0.13, HR: 1.65; CI: 0.86-3.17).
5. Illustration of Effects:
    a. Figure 9 and Figure 10 depict how the number of prior convictions and the year of judgment influence the probability of reconviction over time according to the model.
    b. Notably, as the number of prior convictions increases, the expected delay for a given proportion of the population to reoffend decreases significantly.

---

[24] Binary variable: Belgium vs not Belgium.

In summary, these findings shed light on the complex interplay of factors affecting recidivism risk among individuals convicted of terrorism.
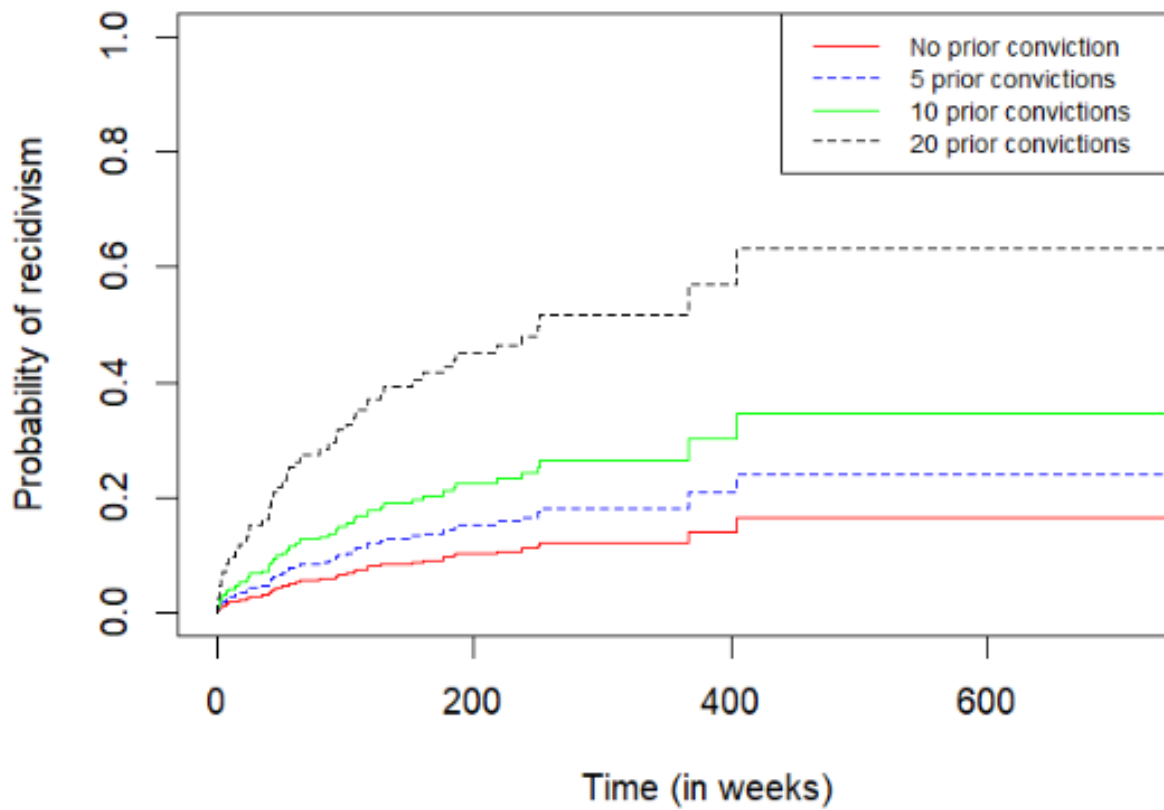
Figure 9 – Effect of the number of convictions on the cumulative probability of recidivism for an individual born in Belgium, convicted in 2014, with no prior drug conviction and not encoded as a leader of a terrorist group.
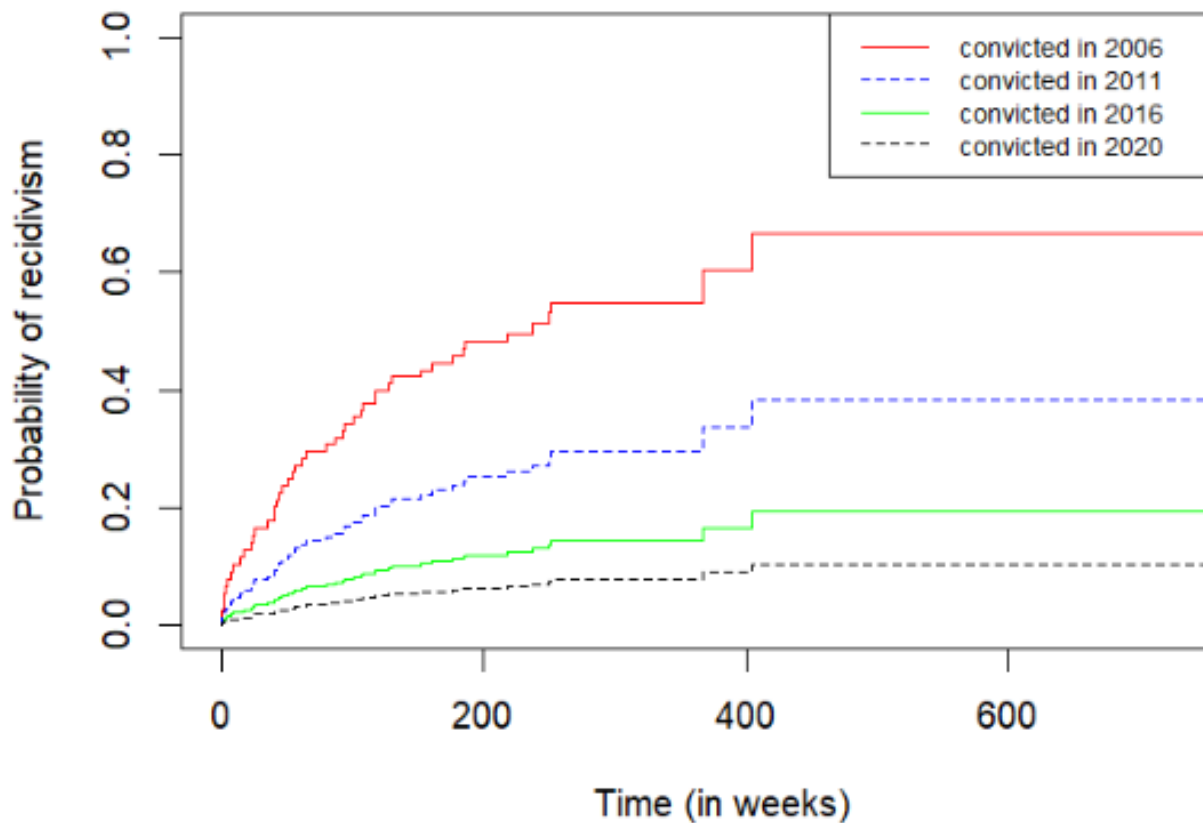
Figure 10 – Effect of the judgment year on the cumulative probability of recidivism for an individual born in Belgium, with no prior conviction and not encoded as a leader of a terrorist group.

In summary, our findings highlight the substantial impact of defining recidivism based on the offense date rather than the conviction date. This choice significantly affects the calculation of prevalence: general recidivism prevalence is divided by a factor of 2.3, while special recidivism (specifically reconviction for terrorism) prevalence is divided by a factor of 5 compared to our previous study (refer to Mine et al., 2021).

Our survival analysis offers a crucial advantage by accommodating the heterogeneity of follow-up periods and handling censored cases. Through this approach, we identified factors influencing both the risk and delay of recidivism.

Consistent with patterns observed in recidivism literature (with an exception for terrorists, as noted by Thijssen et al., 2023), we found that a higher number of prior convictions escalates the risk of recidivism. However, this result warrants further exploration, such as employing latent class analysis.

Interestingly, our results also suggest that recidivism is influenced by the type of offense committed. Notably, certain crime categories (such as drug offenses) appear to have a protective effect. However, all these findings should be interpreted cautiously, especially for individuals sentenced after 2015, considering the limited hindsight available.

For a more comprehensive discussion of these results, please refer to Mine et al. (submitted).

4.3.2. Female sex offenders: a trajectory analysis

4.3.2.1. Sample

The criminal career parameters that are included in the tables below relate to the age of onset, the duration, the age of termination, the frequency (of convictions) and the crime mix (in different offense types). For the female sex offenders, the parameters of the sexual criminal career are reported separately.

Descriptives include the median, the Inter Quartile Range (IQR), Mann-Whitney tests per comparison (including the mean rank, sum of ranks, the U-value, the significance and the effect size). Effect sizes are followed by an indication of their strength: s (small), i (intermediate), l (large).

Table 2 – Descriptive information criminal career dimensions FSO, FNSO, FNSOPP

	FSO	FNSO	FNSOPP
Onset	M: 27 (IQR: 14)	M: 33 (IQR: 19)	M: 29 (IQR: 17)
Termination	M: 40 (IQR: 19)	M: 37 (IQR: 20)	M: 38 (IQR: 20)
Duration	M: 7 (IQR: 21)	M: 0 (IQR: 3)	M: 0 (IQR: 12)
Frequency	M: 2 (IQR: 4)	M: 1 (IQR: 1)	M: 2 (IQR: 2)
Crime mix	M: 3 (IQR: 2)	M: 1 (IQR: 0)	M: 2 (IQR: 1)

For the FSO group, separate descriptives are given for the sexual criminal career, i.e. the onset, duration, termination, frequency and sexual crime mix.

Table 3 – Descriptive information of the sexual criminal career parameters

	FSO
Onset	M: 32 (IQR: 16)
Termination	M: 32 (IQR: 16)
Duration	M: 0 (IQR: 0)
Frequency	M: 1 (IQR: 1)
Crime mix	M: 1 (IQR: 1)

In terms of frequency of sex offences for which a conviction followed, 2147 (or 91.7%) had 1 conviction for a sex offense, 165 (or 7.0%) of the women had two convictions for a sex offense, 21 (or 0.9%) had 3 convictions for a sex offense, 7 (or .3%) had 4 convictions and 2 (or .1%) had 5 convictions for a sex offense. The median duration of the sexual criminal career is zero, which also confirms most of the women only had a single sex offense in their entire criminal career.

4.3.2.2. Comparisons of criminal career dimensions

In the table below, we compare the criminal career parameters of onset, termination, duration, frequency and crime mix of female sex offenders with those of female non-sex offenders.

Table 4 – Comparison descriptives FSO (n=2342) – FNSO (n=10332)

	Mean rank	Sum of ranks	MW U	P	Effect size (n²)

Onset	4891.36	11455567.00	8711914.000	***	.035 (s)
	6665.30	68865908.00			
Termination	6855.10	16054641.00	10886556.00	***	.005 (s)
	6220.17	64266834.00			
Duration	8350.56	19557003.00	7384194.000	***	.069 (i)
	5881.19	60764472.00			
Frequency	8339.04	19530035.50	7411161.500	***	.068 (i)
	5883.80	60791439.50			
Crime mix	9739.62	22810201.50	4130995.500	***	.196 (l)
	5566.33	57511273.50			

N² < .06 (small); n² ≥ .06 (intermediate); n² ≥ .14 (large)

Although the p-values are significant, the effect size (èta squared) shows that only for one parameter, a large effect size is observed. This relates to the crime mix, with a higher crime mix for female sex offenders. Two other parameters show an intermediate effect size (duration and frequency).

The table below zooms in on the comparison between criminal career parameters of female sex offenders with a selected group of female non-sex offenders (with at least one conviction outside of police court).

Table 5 – Comparison descriptives FSO (n=2342) – FNSOPP (n=2620)

	Mean rank	Sum of ranks	MW U	P	Effect size (n²)
Onset	2337.73	5474973.00	2731320.000	***	.009 (s)
	2610.01	6838230.00			
Termination	2617.36	6129851.50	2749841.500	***	.031 (s)
	2360.06	6183351.50			
Duration	2747.30	6434174.50	2445518.500	***	.008 (s)
	2243.90	5879028.50			
Frequency	2741.10	6421987.50	2457705.500	***	.03 (s)
	2248.56	5891215.50			
Crime mix	3031.95	7100829.50	1778863.500	***	.132 (l)
	1989.46	5212373.50			

N² < .06 (small); n² ≥ .06 (intermediate); n² ≥ .14 (large)

Again, p-values are significant, but only one parameter shows a large effect size and none of the others show an intermediate effect size; all are small. Only for the crime mix dimension, a large effect size is observed. The fact that no intermediate effect sizes are present indicates that both groups are more similar.

4.3.2.3. Trajectory results

In terms of the group-based trajectory models, the optimal model (according to BIC and AIC values) has 5 trajectories. The trajectory models are shown here.
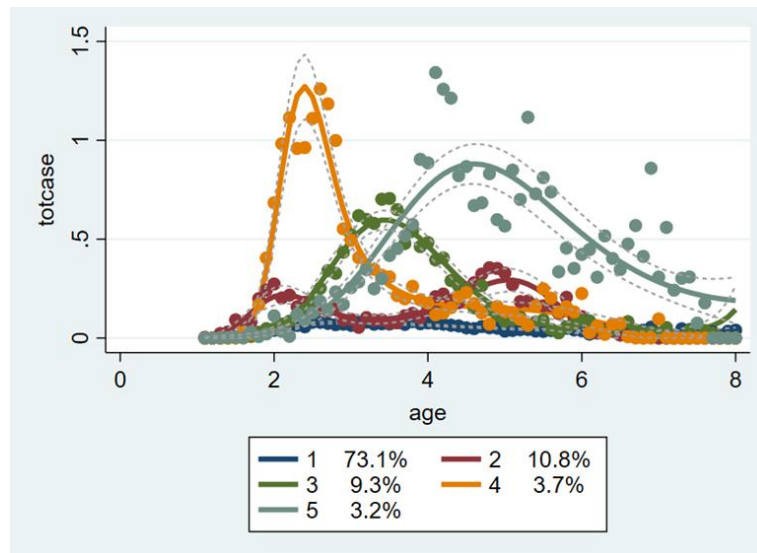
Figure 11 – Criminal trajectories of FSO
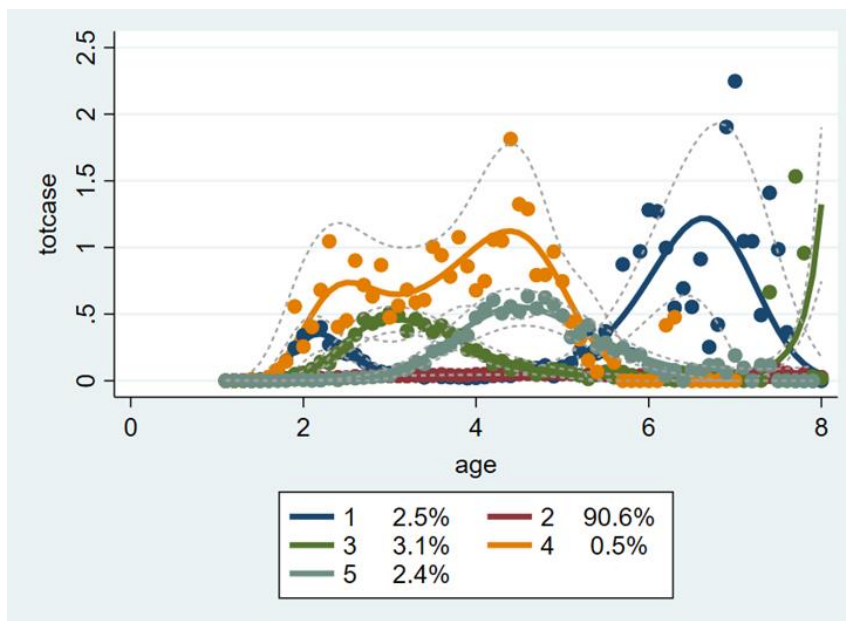
Figure 12 – Criminal trajectories of FNSO
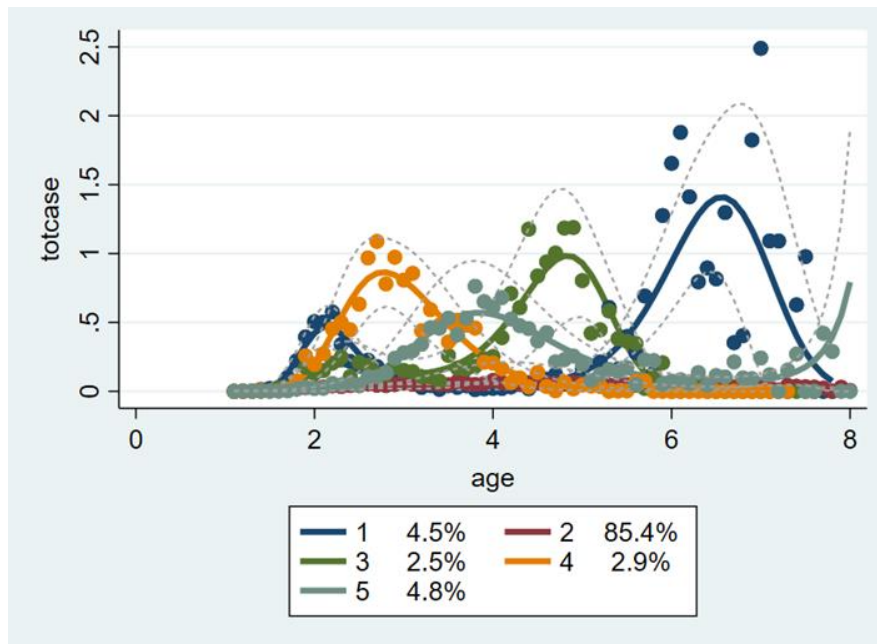
Figure 13 – Criminal trajectories of FNSOPP

Per (most) similar trajectory for the female sex offender group and the comparison group (either FNSO or FNSOPP), comparisons of the criminal career parameters are calculated (medians per trajectory, Mann-Whitney test significance and effect sizes).

Table 6 – Comparison trajectory groups FSO – FNSO

	Group comparison	FSO	FNSO	MW test	Effect size (èta squared)
Onset	1 (n=1792) 2 (n=9500)	Med: 29.00 Min-max: 13-78	Med: 34.00 Min-max: 11-91	U: 6449648.500; p= .000	.024 (s)
Duration		Med: 1.00 Min-max: 0-61	Med: 0.00 Min-max: 0-56	U: 5947438.500; p= .000	.036 (s)
Termination		Med: 39.00 Min-max: 13.86	Med: 39.20 Min-max: 11-91	U: 8260689.500; p= 0.047	0
Frequency		Med: 2.00 Min-max: 1-10	Med: 1.00 Min-max: 1-9	U: 5846696.500; p= .000	.039 (s)
Crime mix		Med: 2.00 Min-max: 1-7	Med: 1.00 Min-max: 1-5	U: 3388121.500; p= .000	.145 (l)
Onset	2 (n=197) 1 (n=255)	Med:19.00 Min-max:11-47	Med: 20.00 Min-max: 13-69	U: 24573.000; p= .691	0
Duration		Med:20.00 Min-max:0-48	Med: 4.00 Min-max: 0-43	U: 11211.500; p= .000	.226 (l)
Termination		Med: 48.00 Min-max:14-68	Med: 24.00 Min-max: 17-76	U: 12840.000; p= .000	.176 (l)
Frequency		Med:7.00 Min-max:1-19	Med: 4.00 Min-max: 2-39	U: 13494.000; p= .000	.158 (i)
Crime mix		Med:4.00 Min-max:1-7	Med: 2.05 Min-max: 1-6	U: 6935.500; p= .000	.386 (l)

Onset	3 (n=208) 3 (n=326)	Med:26.00 Min-max:16-39	Med: 25.00 Min-max: 17-34	U: 29557.5; p= .012	.012 (s)
Duration		Med:19.00 Min-max:0-50	Med: 11.00 Min-max: 0-59	U: 16309.5; p= .000	.192 (l)
Termination		Med:46.00 Min-max:27-76	Med: 36.00 Min-max: 24-77	U: 12965.5; p= .000	.272 (l)
Frequency		Med:10.00 Min-max:4-32	Med: 6.00 Min-max: 2-24	U: 15785.5; p= .000	.203 (l)
Crime mix		Med:4.00 Min-max:2-7	Med: 2.00 Min-max: 1-7	U: 7772.0; p= .000	.423 (l)
Onset	4 (n=79) 5 (n=198)	Med: 19.00 Min-max: 15-26	Med:35.00 Min-max:18-50	U: 219.5; p= .000	.576 (l)
Duration		Med: 21.00 Min-max: 4-48	Med:9.50 Min-max:0-44	U: 3912; p= .000	.152 (i)
Termination		Med: 40.00 Min-max: 22-69	Med:45.50 Min-max:35-74	U: 5113.5; p= .000	.073 (i)
Frequency		Med: 13.00 Min-max: 7-67	Med: 7.00 Min-max:3-28	U: 2811; p= .000	.25 (l)
Crime mix		Med: 5.00 Min-max: 2-8	Med: 1.00 Min-max:1-7	U: 849; p= .000	.484 (l)
Onset	5 (n=66) 4 (n=53)	Med:28.50 Min-max:16-24	Med:21.00 Min-max:14-33	U: 609.5; p= .000	.312 (l)
Duration		Med:27.00 Min-max:1-58	Med:15.00 Min-max:0-44	U: 925; p= .000	.163 (l)
Termination		Med:55.00 Min-max:36-76	Med: 38.00 Min-max:22-63	U: 406.5; p= .000	.433 (l)
Frequency		Med:20.50 Min-max:7-77	Med:18.00 Min-max:7-44	U: 1462; p= .125	.02 (s)
Crime mix		Med: 5.00 Min-max:2-8	Med:2.00 Min-max:1-7	U: 600.500; p= .000	.317 (l)

Only 6 comparisons show no or a small effect size, while 19 comparisons have an intermediate or large effect size. For one parameter, large effect sizes are present across all comparisons. The crime mix of female sex offenders per trajectory group systematically exceeds that of the female non-sex offender group.

Table 7 – Comparison trajectory groups FSO - FNSOPP

	Group comparison	FSO	FNSOPP	MW test	Effect size (èta squared)
Onset	1 (n=1792) 2 (n=2270)	Med: 29.00 Min-max: 13-78	Med: 31.00 Min-max: 11-78	U: 1854201.5; p= .000	.006 (s)
Duration		Med: 1.00 Min-max: 0-61	Med: 0.00 Min-max: 0-56	1776486; p= .000	.012 (s)
Termination		Med: 39.00 Min-max: 13.86	Med: 38.00 Min-max: 11-88	1954305.5; p= .032	.001 (s)
Frequency		Med: 2.00 Min-max: 1-10	Med: 1.00 Min-max: 1-9	1760561; p= .000	.013 (s)
Crime mix		Med: 2.00 Min-max: 1-7	Med: 1.00 Min-max: 1-5	1272726.5; p= .000	.104 (i)
Onset	2 (n=197) 1 (n=120)	Med:19.00 Min-max:11-47	Med: 19.00 Min-max: 13-50	11090; p= .353	.003 (s)
Duration		Med:20.00 Min-max:0-48	Med: 6.00 Min-max: 0-36	6401; p= .000	.148 (i)
Termination		Med: 48.00 Min-max:14-68	Med: 25.00 Min-max: 19-74	6488.5; p= .000	.143 (i)
Frequency		Med:7.00 Min-max:1-19	Med: 4.50 Min-max: 2-39	8650.5; p= .000	.051 (s)

Crime mix		Med:4.00 Min-max:1-7	Med: 3.00 Min-max: 1-6	5622; p= .000	.193 (l)
Onset	3 (n=208) 5 (n=109)	Med:26.00 Min-max:16-39	Med:29.00 Min-max:18-41	6855.5; p= .000	.124 (i)
Duration		Med:19.00 Min-max:0-50	Med:14.00 Min-max:0-59	7763; p= .000	.083 (i)
Termination		Med:46.00 Min-max:27-76	Med:43.00 Min-max:28-77	9803.5; p= .048	.02 (s)
Frequency		Med:10.00 Min-max:4-32	Med:8.00 Min-max:3-44	8609.5; p= .000	.052 (s)
Crime mix		Med:4.00 Min-max:2-7	Med:3.00 Min-max:1-7	4781; p= .000	.249 (l)
Onset	4 (n=79) 3 (n=41)	Med: 19.00 Min-max: 15-26	Med:25.00 Min-max:16-46	666; p= .000	.232 (l)
Duration		Med: 21.00 Min-max: 4-48	Med:23.00 Min-max:3-44	1518.5; p= .576	.003 (s)
Termination		Med: 40.00 Min-max: 22-69	Med:49.00 Min-max:35-63	851; p= .000	.151 (i)
Frequency		Med: 13.00 Min-max: 7-67	Med:14.00 Min-max:5-28	1538.5; p= 653	.002 (s)
Crime mix		Med: 5.00 Min-max: 2-8	Med: 3.00 Min-max:1-5	395.5; p= .000	.382 (l)
Onset	5 (n=66) 4 (n=78)	Med:28.50 Min-max:16-24	Med:22.00 Min-max:14-28	935; p= .000	.3 (l)
Duration		Med:27.00 Min-max:1-58	Med:13.50 Min-max:2-35	916.5; p= .000	.307 (l)
Termination		Med:55.00 Min-max:36-76	Med: 35.50 Min-max:27-56	231; p= .000	.613 (l)
Frequency		Med:20.50 Min-max:7-77	Med:10.00 Min-max:4-44	867.5; p= .000	.325 (l)
Crime mix		Med: 5.00 Min-max:2-8	Med:3.00 Min-max:1-7	1407; p= .000	.152 (i)

In 10 comparisons, the effect sizes are small, which suggest very limited differences between the groups. This is slightly more than with the general population of female non sex offenders (and thus this subgroup shows more similarities with the female sex offender group), but nonetheless, in 15 comparisons, intermediate (7) or large (8) effect sizes are found. In other words, the group of female non-sex offenders with at least one conviction outside of police court still seems to differ in important ways from the group of female sex offenders. Yet, in terms of the crime mix, not all trajectory groups show a comparison with a large effect size, which does suggest that some of the important overall differences dissipate.

4.3.3. The duration of criminal careers of convicted sex offenders

4.3.3.1. Sample

The criminal career information about male and female sex offenders that have reached the age of 70 years is presented separately. For the descriptive information about criminal career parameters, a further distinction between those still alive and those deceased is presented. We also distinguish between the onset, termination and duration based on the offense (each time the first offense) and the conviction.

Due to the heavily skewed data, for all the criminal career dimensions, the median is reported (i.e. the middle value, at the 50th percentile), as are the minimum and maximum values and the Inter Quartile Range (IQR), based on the difference between the values at the 75th and the 25th percentiles.

Overall, the dataset consists of 5119 sex offenders, of which 1821 are deceased at the date of extraction. 4809 are males (1746 are deceased) and 310 are females (of which 75 are deceased).

For both male and female sex offenders, three tables are presented. The first includes the descriptive information related to all criminal career parameters of the global criminal career (all offense types combined), a second table zooms in on the non-sexual criminal career and the third table presents the descriptives for the sexual criminal career. Due to the important number of offenders with only one conviction for a sex offense, results are also presented for those with at least two sex-offense related convictions.

4.3.3.2. Criminal career parameters for male sex offenders

Table 8 – Criminal career parameters male sex offenders (n = 4809) – all offense types combined (sex and non-sex)

Parameter	All (n = 4809)	Alive (n = 3063)	Deceased (n = 1746)
Onset (offense-based)	M: 29.00 (min: 13 – max: 89); IQR: 29	M: 29.00 (min: 13 – max: 88); IQR: 29	M: 28.00 (min: 13 – max: 89); IQR: 28
Onset (judgm-based)	M: 30.00 (min: 13 – max: 92); IQR: 30	M: 30.00 (min: 13 – max: 91); IQR: 31	M: 29.00 (min: 13 – max: 92); IQR: 29
Termination (off)	M: 63.00 (min: 44 – max: 92); IQR: 12	M: 61.00 (min: 44 – max: 92); IQR: 12	M: 67.00 (min: 46 – max: 92); IQR: 11
Termination (judgm)	M: 65.00 (min: 45 – max: 93); IQR: 12	M: 63.00 (min: 45 – max: 93); IQR: 12	M: 68.00 (min: 47 – max: 92); IQR: 11
Duration (off)	M: 32.00 (min: 0 – max: 69); IQR: 32.5	M: 29.00 (min: 0 – max: 69); IQR: 34	M: 36.00 (min: 0 – max: 68); IQR: 29
Duration (judgm)	M: 32.00 (min: 0 – max: 69); IQR: 34	M: 30.00 (min: 0 – max: 69); IQR: 34	M: 37.00 (min: 0 – max: 69); IQR: 29
Frequency (all)	M: 4.00 (min: 1 – max: 132); IQR: 9	M: 4.00 (min: 1 – max: 132); IQR: 8	M: 5.00 (min: 1 – max: 121); IQR: 11
Crime mix (max 8)	M: 3.00 (min: 1 – max: 8); IQR: 3	M: 3.00 (min: 1 – max: 8); IQR: 3	M: 3.00 (min: 1 – max: 8); IQR: 3
Years since last offense	M: 12.00 (min: 0 – max: 25); IQR: 11	M: 14.00 (min: 0 – max: 25); IQR: 11	M: 10.00 (min: 0 – max: 25); IQR: 9
Years since last conviction	M: 10.00 (min: 0 – max: 25); IQR: 11	M: 12.00 (min: 0 – max: 25); IQR: 12	M: 8.00 (min: 0 – max: 24); IQR: 9

The median duration of criminal careers here is 32 years (based on the offenses), with a maximum up to 69 years. For the deceased sex offenders, the median duration is 36 years. Remember, all persons included have at least reached the age of 70 years, which also implies some persons with a very active criminal career have continued their long criminal career up until old age. It is also very interesting that the median duration reaches up to a maximum of almost 7 decades of active offending.

Due to the data selection, the years since the last offence and the last conviction cannot be longer than 25 years (data extraction in 2020, active career between 1995 and 2020). This means the last two values per column are biased by the data selection process. Yet, most of them have a median number of years since the last offense of 12 overall, i.e. for 50% of all persons included, their last offense (that led to a conviction) was 12 years ago.

Table 9 – Criminal career parameters male sex offenders (n = 4809) – non-sex offenses only

Parameter	All (n = 3725*)	Alive (n = 2330)	Deceased (n = 1395)
Onset (offense-based)	M: 27.00 (min: 13 – max: 88); IQR: 20	M: 27.00 (min: 15 – max: 86); IQR: 20	M: 27.00 (min: 13 – max: 88); IQR: 21
Onset (judgm-based)	M: 28.00 (min: 13 – max: 89); IQR: 20	M: 28.00 (min: 16 – max: 87); IQR: 20	M: 28.00 (min: 13 – max: 89); IQR: 21
Termination (off)	M: 59.00 (min: 17 – max: 92); IQR: 16	M: 58.00 (min: 17 – max: 92); IQR: 16	M: 63.00 (min: 17 – max: 92); IQR: 15
Termination (judgm)	M: 61.00 (min: 18 – max: 93); IQR: 16	M: 59.00 (min: 18 – max: 93); IQR: 15	M: 64.00 (min: 18 – max: 92); IQR: 15
Duration (off)	M: 28.00 (min: 0 – max: 68); IQR: 34	M: 26.00 (min: 0 – max: 64); IQR: 33	M: 32.00 (min: 0 – max: 68); IQR: 34
Duration (judgm)	M: 28.00 (min: 0 – max: 69); IQR: 34	M: 27.00 (min: 0 – max: 66); IQR: 34	M: 32.00 (min: 0 – max: 69); IQR: 34
Frequency (all)	M: 5.00 (min: 1 – max: 130); IQR: 10	M: 5.00 (min: 1 – max: 130); IQR: 8	M: 5.00 (min: 1 – max: 116); IQR: 13
Crime mix (max 7)	M: 3.00 (min: 1 – max: 7); IQR: 2	M: 3.00 (min: 1 – max: 7); IQR: 3	M: 3.00 (min: 1 – max: 7); IQR: 3
Years since last offense	M: 16.00 (min: 0 – max: 72); IQR: 14	M: 17.00 (min: 1 – max: 72); IQR: 14	M: 13.00 (min: 0 – max: 65); IQR: 14
Years since last conviction	M: 14.00 (min: 0 – max: 71); IQR: 14	M: 16.00 (min: 0 – max: 71); IQR: 14	M: 12.00 (min: 0 – max: 65); IQR: 14

For this table, it is important to mention a difference in the number of included persons. This is due to the fact that a number of persons has one conviction only (i.e. a sex offense-related conviction), and thus no non-sex convictions are included[25].

The duration of the non-sexual criminal career has a median for the entire group of 28 years. For the group of deceased persons, the median is at 32 years, suggesting this group had a slightly more active criminal career (or that they may have been slightly older than those still alive at the time of the data selection, which could mean those alive still have time to 'catch up' in terms of committing further offenses).

Table 10 – Criminal career parameters male sex offenders (n = 4809) – sex offenses only

Parameter	All (n = 4809)	All with > 1 conviction (n= 938)	Alive (n = 3063)	Deceased (n = 1746)
Onset (offense-based)	M: 47.00 (min: 13 – max: 89); IQR: 34	M: 30 (min: 13 – max: 78); IQR: 21	M: 47.00 (min: 13 – max: 88); IQR: 32	M: 44.00 (min: 13 – max: 89); IQR: 37
Onset (judgm-based)	M: 49.00 (min: 13 – max: 92); IQR: 36	M: 31.00 (min: 13 – max: 80); IQR: 21.25	M: 50.00 (min: 13 – max: 91); IQR: 34	M: 46.00 (min: 14 – max: 92); IQR: 39
Termination (off)	M: 52.00 (min: 13 – max: 89); IQR: 31	M: 55.00 (min: 18 – max: 86); IQR: 26	M: 51.00 (min: 15 – max: 88); IQR: 29	M: 55.00 (min: 13 – max: 89); IQR: 36
Termination (judgm)	M: 54.00 (min: 14 – max: 92); IQR: 33	M: 57.00 (min: 18 – max: 86); IQR: 26	M: 53.00 (min: 15 – max: 91); IQR: 31	M: 57.00 (min: 14 – max: 92); IQR: 37
Duration (off)	M: 0.00 (min: 0 – max: 57); IQR: 0	M: 15.00 (min: 0 – max: 57); IQR: 23	M: 0.00 (min: 0 – max: 57); IQR: 0	M: 0.00 (min: 0 – max: 57); IQR: 0

---

[25] This analysis has a degree of telescoping, in that in cases where a single conviction is for several types of offenses including a sex offense, this is only counted once, in the table of the sexual criminal career. Future more detailed analyses could zoom in further on this.

Duration (judgm)	M: 0.00 (min: 0 – max: 58); IQR: 0	M: 15.50 (min: 0 – max: 58); IQR: 23	M: 0.00 (min: 0 – max: 57); IQR: 0	M: 0.00 (min: 0 – max: 58); IQR: 0
Frequency (all)	M: 1.00 (min: 1 – max: 23); IQR: 0	M: 2 (min: 2 – max: 23); IQR: 1	M: 1.00 (min: 1 – max: 13); IQR: 0	M: 1.00 (min: 1 – max: 23); IQR: 0
Crime mix (max 8)	M: 1.00 (min: 1 – max: 6); IQR: 1	M: 2 (min: 1 – max: 6); IQR: 1	M: 1.00 (min: 1 – max: 6); IQR: 1	M: 1.00 (min: 1 – max: 5); IQR: 1
Years since last offense	M: 23.00 (min: 0 – max: 72); IQR: 31	M: 20.00 (min: 0 – max: 63); IQR: 26.25	M: 23.00 (min: 0 – max: 71); IQR: 29	M: 20.00 (min: 0 – max: 72); IQR: 34
Years since last conviction	M: 20.00 (min: 0 – max: 71); IQR: 33	M: 18.00 (min: 0 – max: 63); IQR: 27	M: 20.00 (min: 0 – max: 71); IQR: 30	M: 18.00 (min: 0 – max: 71); IQR: 36

The median value of the duration of sexual criminal careers is zero, as 80.5% of all male sex offenders (n=3871) have one single conviction for a sex offense. For the 938 remaining persons, the criminal career parameters are reported here (all combined, without a distinction between alive and deceased). For this subgroup, the median duration of the sexual criminal career is 15 years (offense-based), with a maximum of 57 years.

4.3.3.3. Criminal career parameters for female sex offenders

Table 11 – Criminal career parameters female sex offenders (n = 310) – all offenses combined (sex and non-sex)

Parameter	All (n = 310)	Alive (n = 235)	Deceased (n = 75)
Onset (offense-based)	M: 28.00 (min: 16 – max: 74); IQR: 20	M: 29.00 (min: 16 – max: 73); IQR: 21	M: 27.00 (min: 16 – max: 74); IQR: 18
Onset (judgm-based)	M: 29.00 (min: 16 – max: 80); IQR: 20	M: 30.00 (min: 16 – max: 74); IQR: 11	M: 28.00 (min: 17 – max: 80); IQR: 19
Termination (off)	M: 61.00 (min: 45 – max: 86); IQR: 12	M: 60.00 (min: 45 – max: 84); IQR: 12	M: 64.00 (min: 52 – max: 86); IQR: 11
Termination (judgm)	M: 63.00 (min: 46 – max: 88); IQR: 12	M: 62.00 (min: 46 – max: 85); IQR: 13	M: 66.00 (min: 52 – max: 88); IQR: 13
Duration (off)	M: 31.00 (min: 0 – max: 61); IQR: 13.25	M: 30.00 (min: 0 – max: 61); IQR: 23	M: 39.00 (min: 0 – max: 58); IQR: 14
Duration (judgm)	M: 31.50 (min: 0 – max: 61); IQR: 24	M: 31.00 (min: 0 – max: 61); IQR: 23	M: 39.00 (min: 0 – max: 57); IQR: 15
Frequency (all)	M: 4.00 (min: 1 – max: 40); IQR: 6	M: 4.00 (min: 1 – max: 40); IQR: 6	M: 5.00 (min: 1 – max: 24); IQR: 6
Crime mix (max 8)	M: 3.00 (min: 1 – max: 8); IQR: 2	M: 3.00 (min: 1 – max: 8); IQR: 2	M: 3.00 (min: 1 – max: 7); IQR: 2
Years since last offense	M: 14.00 (min: 1 – max: 25); IQR: 11	M: 15.00 (min: 1 – max: 25); IQR: 13	M: 13.00 (min: 2 – max: 24); IQR: 7
Years since last conviction	M: 12.00 (min: 0 – max: 25); IQR: 10	M: 13.00 (min: 0 – max: 25); IQR: 11	M: 11.00 (min: 0 – max: 23); IQR: 8

For female sex offenders, the median duration of their criminal career is 31 years, with a maximum of 61 years. Here too, just as with the men, deceased female sex offenders have a median criminal career duration of 39 years, which is slightly longer than that of their counterparts that are still alive at the end of the data extraction. As far as the median time since the last offense is concerned, the median is at 14 years prior to the data extraction (or date of death).

Table 12 – Criminal career parameters female sex offenders (n = 310) – non-sex offenses only

Parameter	All (n = 280)*	Alive (n = 209)	Deceased (n = 71)
Onset (offense-based)	M: 34.50 (min: 16 – max: 86); IQR: 26	M: 35.00 (min: 16 – max: 84); IQR: 25.50	M: 31.00 (min: 17 – max: 86); IQR: 28
Onset (judgm-based)	M: 36.00 (min: 17 – max: 88); IQR: 26.75	M: 36.00 (min: 17 – max: 85); IQR: 26	M: 32.00 (min: 17 – max: 88); IQR: 29
Termination (off)	M: 61.50 (min: 27 – max: 86); IQR: 14	M: 60.00 (min: 31 – max: 84); IQR: 15	M: 64.00 (min: 27 – max: 86); IQR: 11
Termination (judgm)	M: 63.00 (min: 28 – max: 88); IQR: 14	M: 62.00 (min: 32 – max: 85); IQR: 15	M: 65.00 (min: 28 – max: 88); IQR: 13
Duration (off)	M: 22.00 (min: 0 – max: 61); IQR: 37	M: 21.00 (min: 0 – max: 61); IQR: 35	M: 25.00 (min: 0 – max: 54); IQR: 42
Duration (judgm)	M: 24.00 (min: 0 – max: 61); IQR: 36.75	M: 23.00 (min: 0 – max: 61); IQR: 35.50	M: 27.00 (min: 0 – max: 56); IQR: 41
Frequency (all)	M: 3.00 (min: 1 – max: 39); IQR: 2	M: 3.00 (min: 1 – max: 39); IQR: 5.50	M: 4.00 (min: 1 – max: 23); IQR: 5
Crime mix (max 8)	M: 2.00 (min: 1 – max: 7); IQR: 2	M: 2.00 (min: 1 – max: 7); IQR: 2	M: 2.00 (min: 1 – max: 6); IQR: 3
Years since last offense	M: 14.00 (min: 1 – max: 58); IQR: 13	M: 15.00 (min: 1 – max: 43); IQR: 14	M: 14.00 (min: 2 – max: 58); IQR: 10
Years since last conviction	M: 13.00 (min: 0 – max: 57); IQR: 12	M: 13.00 (min: 0 – max: 42); IQR: 13	M: 12.00 (min: 0 – max: 57); IQR: 10

As for male sex offenders, an important number of female sex offenders only had one single conviction. This explains why the non-sexual criminal career reports about 280 persons (instead of 310). The median duration of the non-sexual criminal career is 24 years.

Table 13 – Criminal career parameters female sex offenders (n = 310) – sex offenses only

Parameter	All (n = 310)	All with > 1 conviction (n= 33)	Alive (n = 235)	Deceased (n = 75)
Onset (offense-based)	M: 35.00 (min: 16 – max: 74); IQR: 27	M: 30 (min: 16 – max: 57); IQR: 19.50	M: 35.00 (min: 16 – max: 73); IQR: 26	M: 35.00 (min: 16 – max: 74); IQR: 28
Onset (judgm-based)	M: 37.00 (min: 16 – max: 80); IQR: 28.25	M: 33.00 (min: 16 – max: 58); IQR: 19.50	M: 37.00 (min: 16 – max: 75); IQR: 28	M: 36.00 (min: 17 – max: 80); IQR: 30
Termination (off)	M: 37.00 (min: 16 – max: 74); IQR: 27	M: 44.00 (min: 18 – max: 62); IQR: 17	M: 38.00 (min: 16 – max: 73); IQR: 27	M: 37.00 (min: 16 – max: 74); IQR: 27
Termination (judgm)	M: 38.00 (min: 17 – max: 80); IQR: 29	M: 46.00 (min: 20 – max: 64); IQR: 17	M: 38.00 (min: 17 – max: 75); IQR: 27	M: 38.00 (min: 17 – max: 80); IQR: 29
Duration (off)	M: 0.00 (min: 0 – max: 41); IQR: 0	M: 6.00 (min: 1 – max: 41); IQR: 15.50	M: 0.00 (min: 0 – max: 41); IQR: 0	M: 0.00 (min: 0 – max: 40); IQR: 0
Duration (judgm)	M: 0.00 (min: 0 – max: 44); IQR: 0	M: 7.00 (min: 1 – max: 44); IQR: 19.50	M: 0.00 (min: 0 – max: 44); IQR: 0	M: 0.00 (min: 0 – max: 44); IQR: 0
Frequency (all)	M: 1.00 (min: 1 – max: 4); IQR: 0	M: 2 (min: 2 – max: 4); IQR: 0.50	M: 1.00 (min: 1 – max: 4); IQR: 0	M: 1.00 (min: 1 – max: 4); IQR: 0
Crime mix (max 8)	M: 1.00 (min: 1 – max: 3); IQR: 0	M: 1.00 (min: 1 – max: 3); IQR: 1	M: 1.00 (min: 1 – max: 3); IQR: 0	M: 1.00 (min: 1 – max: 3); IQR: 0
Years since last offense	M: 38.00 (min: 6 – max: 70); IQR: 28	M: 29.00 (min: 15 – max: 59); IQR: 17	M: 38.00 (min: 6 – max: 70); IQR: 26	M: 39.00 (min: 7 – max: 70); IQR: 26
Years since last conviction	M: 36.50 (min: 1 – max: 70); IQR: 30	M: 28.00 (min: 11 – max: 59); IQR: 29.50	M: 36.00 (min: 2 – max: 67); IQR: 28	M: 38.00 (min: 1 – max: 70); IQR: 25

The median duration of the sexual criminal career is 0. This once again illustrates that most convicted female sex offenders only get one conviction for a sex offense in their entire criminal career. Only slightly more than 10% of the female sex offenders has more than one conviction for a sex offense. Focusing on female sex offenders with at least 2 sex offense-related convictions, the median duration for the 33 female sex offenders' sexual criminal career is 6 years, with a maximum of 41 years.

4.3.4. Continuity in sexual offending

4.3.4.1. Sample

The dataset that is used here is the same as for the study of duration (see section 3.2.1.). This has one inherent limitation: the data included always involve persons with an active criminal career in 1995 or later, with at least one conviction in or after 1995 (up to 2020). Since all persons in the dataset are at the very least 70 years, this means most will have received a conviction either in adulthood (between 26 and 59 years) or in late adulthood/older age (at 60 years or older). The follow-up period is a maximum of 25 years and the minimum age is 70, thus 45 years (when convicted) is the 'youngest' point of entry into the dataset. This eliminates the possibility of including persons that only committed offenses prior to that age, without a continuation of their criminal career later in life. This limitation is important when taking into consideration the results here. Future analyses could also be done on a younger population (e.g. 40 years as the minimum age for inclusion).

Here, the analysis is thus focused on persons who reached the age of 70 years and who had an active criminal career up until at least the age of 45 years.

For this group, the following results in terms of the continuity in sexual offending can be found. A distinction can be made between male and female sex offenders.

4.3.4.2. Continuity in sexual offending: descriptive results

In the table below, a first descriptive view is provided about the continuity in sex offending (based on offenses for which a conviction followed).

Table 14 – Continuity in sexual offending: two periods

		Age at sex offense – 26 years or older		Total
		No sex offense	Sex offense	
Age at sex offense – 25 years or younger	No sex offense	0	3981	3981
	Sex offense	860	**278**	1138
	Total	860	4259	5119

The population is selected based on the presence of at least one sex offense in the entire criminal career, which explains why there are no persons without sex offenses. In terms of continuity in sexual offending, this is limited to 278 out of 5119 (or 5,4%) of all persons. Of all young persons with a sex offense, 24.4% (or 278 out of 1138) committed a sexual offense after the age of 25 years. Of the persons over 25 years with a sexual offense in their criminal career after the age of 25 years, 6.5% also had committed a sexual offense before the age of 26 years.

Table 15 – Continuity in sexual offending (males only)

		Age at sex offense – 26 years or older		Total
		No sex offense	Sex offense	
Age at sex offense – 25 years or younger	No sex offense	0	3759	3759
	Sex offense	782	**268**	1050
	Total	782	4027	4809

For male sex offenders, continuity in sexual offending is ca 6% (5.6%), i.e. 268 of 4809 males committed both before the age of 26 and after the age of 25 a sexual offense. For young men who committed a sexual offense before the age of 26, ca a quarter (25.5%) also committed a sexual offense in adult life. For persons with a sex offense after the age of 25, 6.6% had committed a sexual offense earlier in life, up until the age of 25.

Table 16 – Continuity in sexual offending (females only)

		Age at sex offense – 26 years or older		Total
		No sex offense	Sex offense	
Age at sex offense – 25 years or younger	No sex offense	0	222	222
	Sex offense	78	**10**	88
	Total	78	232	310

For female sexual offenders, continuity in sexual offending is very low. 3.2% of all female offenders with a sexual offense at the age of 25 years or younger also committed a sexual offense later in life. Of all younger female offenders, 10 out of 88 (or 11,4%) go on to commit a sexual offense in adulthood (at the age of 26 years or later). Of all adult female sex offenders, only 4.3% also committed a sexual offense before the age of 26 years.

4.3.4.3. Continuity across four periods

This results clearly indicate that continuity depends upon how the numbers are looked at (and in that sense, there is some variation, between a low of 3.2% for female sex offenders overall and a high of 25.5% of male young sex offenders who go on to commit a sexual offense later in life (at the age of 26 years or later). All these statistics remain relatively small and rather indicate a very limited continuity in sexual offending.

The data are then further split up according to four phases in life. In case continuity refers to at least two periods in life with a sexual offense, this gives the following descriptive results.

Before the age of 18 years, 139 persons committed a sexual offense for which they received a conviction (or other type of judicial decision registered in the national Central Conviction Records).

At the age between 18 and 25 years, 1015 persons committed a sexual offense.

At the age between 26 and 59 years, 2869 persons committed a sexual offense.

For the group of 60 years or older, 1668 persons committed a sexual offense.

Per group, the continuity statistics are presented. Combinations that are included, are not reproduced in tables that follow.

To give one example that facilitates reading what follows: take for example the table below. This table contains the descriptive information related to persons with a sex offense committed prior to the age of 18 years (n = 139). Of these 139 persons (Group 1), 16 (11.5%) have also committed a sex offense in the age group 2 (i.e. between 18 and 25 years of age), 27 (19.4%) have committed a sex offense in the age group 3 (i.e. between 26 and 59 years of age) and 6 (4.3%) committed a sex offense in age group 4 (i.e. at the age of 60 years or older). Then, combinations are provided. For example, only those who committed a sex offense prior to the age of 18 and with a sex offense in the age group of 18 – 25 years (n = 16) are combined. For those, it is checked whether they also committed a sex offense in group 3 (n = 8 or 50%), or in group 4 (n = 2 or 12.5%) or in both groups 3 and 4 (n = 2, 25%).

Table 17 – Continuity descriptives – sex offense committed before the age of 18 years (n = 139)

Group	Group 1	Group 2	Group 3	Group 4
Below 18 years	139	16 (11.5%)		
	139		27 (19.4%)	
	139			6 (4.3%)
	G1&G2 (n=16)		8 (50%)	
	G1&G2 (n=16)			2 (12.5%)
	G1&G2&G3 (n=8)			2 (25%)
	G1&G3 (n=27)	8 (29.6%)	G1&G3 (n=27)	
	G1&G3 (n=27)		G1&G3 (n=27)	4 (14.8%)
	G1&G4 (n=6)	2 (33.3%)		G1&G4 (n=6)
	G1&G4 (n=6)		4 (66.7%)	G1&G4 (n=6)
	G1&G2&G4 (n=2)		2 (100%)	G1&G2&G4 (n=2)
	G1&G3&G4 (n=4)	2 (50%)	G1&G3&G4 (n=4)	

Of those who commit a sex offense prior to the age of 18, only a limited group goes on to recommit a sex offense later in life.

Table 18 – Continuity descriptives – sex offense committed between the ages 18 and 25 years (n = 1015)

Group	Group 1	Group 2	Group 3	Group 4
18 – 25 years	16 (1,6%)	1018		
		1018	234 (23.0%)	
		1018		53 (5.2%)
	8 (3.4%)	G2&G3 (n=234)		
		G2&G3 (n=234)		30 (12.8%)
	2 (3.8%)	G2&G4 (n=53)		G2&G4 (n=53)
		G2&G4 (n=53)	30 (56.6%)	G2&G4 (n=53)
	2 (6.6%)	G2&G3&G4 (n=30)		

In this table, of those with a sex offense committed between the age of 18 and 25, ca a quarter have a new sex offense at the age between 26 and 59 years.

Table 19 – Continuity descriptives – sex offense committed between the ages 26 and 59 years (n = 2869)

Group	Group 1	Group 2	Group 3	Group 4
26 – 59 years	27 (0.9%)		2869	
			2869	278 (9.7%)
	4 (1.4%)		G3&G4 (n=278)	
		30 (10.8%)	G3&G4 (n=278)	

This table shows a very limited continuity between those committing sex offenses at the age between 26 and 59 years and the age before 18 years. Barely 1% of those offenders has committed a sex offense (that led to a subsequent conviction) prior to the age of 18 years.

Table 20 – Continuity descriptives – sex offense committed at age 60 or later (n = 1668)

Group	Group 1	Group 2	Group 3	Group 4
Over 60 years	6 (0.4%)			1668
		53 (3.2%)		1668
			278 (16.7%)	1668

Again, for those who committed a sex offense at the age of 60 or older, a very limited degree of continuity can be observed in the prior age categories.

Counting the presence of sex offenses across these different life periods, the following figure shows the results (in percentages).
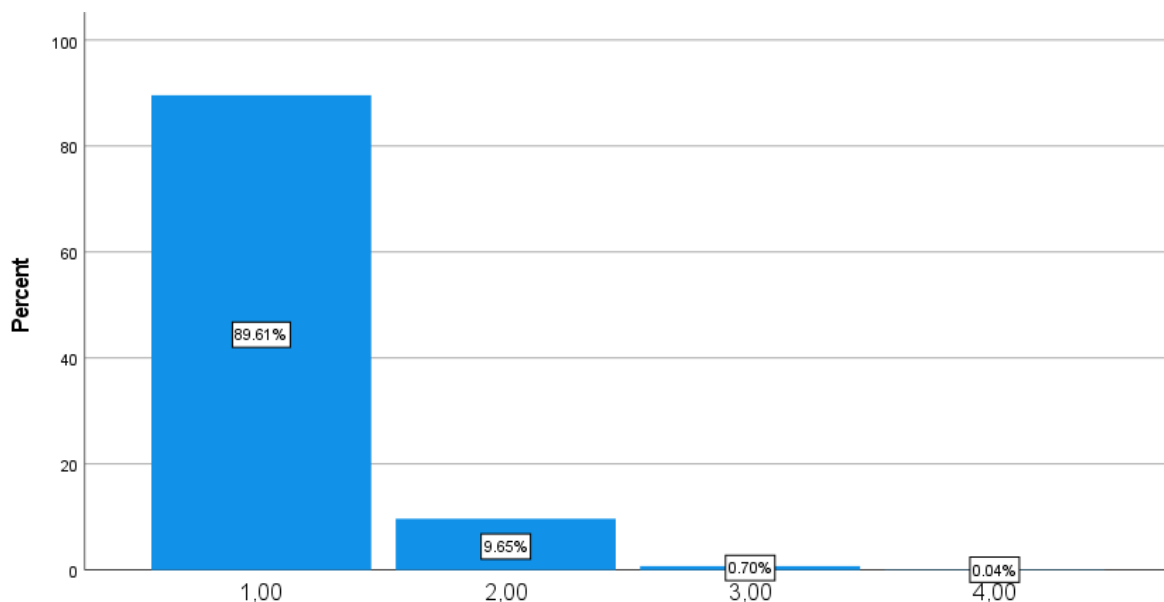

Figure 14 – continuity sex offence 4 groups [CAPTION TO ADAPT]

For men and women, the table below shows figures for continuity in sex offending across different life phases.

Table 21 – Continuity in sex offending for men and women across four periods

Periods in life with a sex offense	males		females		Total
	Frequency	% (rounded)	Frequency	%	
1 period	4290	89.2	297	95.8	4587
2 periods	481	10.0	13	4.2	494
3 periods	36	0.7	0	0	36
4 periods	2	0.04	0	0	2
Total	4809	100	310	100	5119

These descriptives show that continuity across all four time periods is very rare. Only 2 out of 5119 persons have committed a sexual offense in all four time periods in their life. This also extends to three time periods: even in this case, continuity seems to be very limited (0.7% overall). Next to the large difference in the numbers of men and women, it is also interesting to observe the difference in continuity in sex offending between women and men, where a very small group of men committed sex offenses in three different life phases, while 13 out of 310 women have committed sexual offenses in two different periods, but none in three or four periods.

4.3.5. Differentiating homicide offenders based on their criminal careers: A latent class analysis

4.3.5.1. Sample

Between 1995 and 2020, there were 6679 convictions for (intentional) homicide or attempted homicide according to the Central Criminal Record. The cases for which the date of the homicide was unavailable or unreliable (n=346) were removed from the original sample to allow us to clearly define what would constitute a criminal antecedent for each individual. For offenders with multiple homicide convictions, the first homicide was taken as the reference homicide. Subsequent homicides were not considered for the general description of the homicide data or in the latent class analysis. For this reason, we removed 384 homicide cases, committed by 354 individuals. Additionally, 474 individuals were excluded for unreliability[26] of their criminal histories, resulting in a final sample of 5475 homicide offenders.

4.3.5.2. Homicide offenders : Descriptives

The vast majority of the homicide offenders in our sample were male (88.3%). The median age at the time of the homicide was 33 years old with a large inter-individual variability (*sd*=12.7; range: 12-98). About two thirds of the sample (67.5%) had criminal antecedents to their homicide. This means than about one third of the sample started their criminal career by committing an homicide. On average, there was slightly less than 10 years between the criminal career onset (median=23.0, sd=11.4) and the homicide (median=33.0, sd=12.7) but again with a large inter-individual variability (sd=10.1; time

---

[26] The inclusion criterion was the following : 95% of the offenses of an homicide offender must have a reliable date in the CCR. The rationale was to prevent biases in the classification process both at the individual level and at the population level.

interval range: from 0 to 67 years). The average number of convictions for offenses committed prior to the homicide was about four (mean=4.0), but there was a huge inter-individual variability (*sd*=6.4; median=1; IQR:0-5; range:0-80). On average, the number of conviction was significantly higher for offenses committed during the adult period (mean=2.8) than for offenses committed during the juvenile period (mean=2.0). The variety of offenses was assessed by the crime mix index[27] (i.e., the number of different types of offense). The average crime mix index was significantly lower for the juvenile period (mean=1.3, *sd* = 1.8) than for the adult period (mean=1.6, *sd*=1.8). However, the increase in the variety of offenses committed between the two lifetime periods was not a universal pattern, as the LCA would demonstrate.

Table 22 provides a clearer picture of the types of offense committed by homicide offenders prior to their homicide. Criminal careers involving road traffic offense were common in our sample. Violence and property offenses were also relatively common. In contrast, there were very few homicide offenders with a criminal history involving sexual offense. Compared with the juvenile period, the prevalence of offending in the adult period was generally higher for each category of offense, except for the violent property crime.

Table 22 – Offenses committed before the homicide: proportion of convicts (%) by offense category and by lifetime period in the total sample

Offense at age 25 and below	
Drugs	10.2
Other	16.8
Property	23.1
Public order	15.3
Sexual crime	2.3
Traffic	32.9
Violence	22.3
Violent property	9.5
Offenses after age 25	
Drugs	10.5
Other	21.9
Property	23.2
Public order	21.0
Sexual crime	3.7
Traffic	42.9
Violence	29.4
Violent property	6.5

4.3.5.3. Latent class analysis

We estimated models from one up to seven latent classes. On the basis of fit indices, parsimony measures and diagnostic statistics, a model distinguishing six latent classes appeared to be optimal.

---

[27] As we used 8 categories to classify offenses, the individual crime-mix index necessarily ranged from 0 (n=1782) to 8 (n=28).

Furthermore, this model met the minimum class size conditions (min 5% of the sample for each class) and its classes were considered as interpretable. Homicide offenders were then assigned into the different classes on the basis of the maximum posterior probability rule. This allowed us to describe the shares, profiles and characteristics of the different classes. The labelling of the classes was done after examining the criminal careers patterns in terms of several crossed dimensions: time evolution (i.e., continuity between the two lifetime periods or not, growing or declining criminal trend); variety (i.e., large spectrum of offenses or criminal specialist); and level of activity (i.e., high or low probability of committing an offense; number of convictions).

4.3.5.4. Latent classes profiles

The Figure 15 shows the proportion of convicts per offense category and lifetime period for each of the six classes separately.

The first class (11.7% of the sample) is characterized by a clear dip between their juvenile and adult criminal activity. During their juvenile period, they have a relatively wide range of offenses for which they have a high or relatively high probability of being convicted. Non-violent property crime is particularly prevalent in this class during this period (78.4%). In contrast, their probability of being convicted for offenses committed in their adult period is low or relatively low for every type of offense, except for traffic offenses (39.0%). The distinction between the two lifetime periods is reflected in the mean crime indices (3.4 in the juvenile period vs 1.0 in the adult period) and in the mean number of convictions (4.5 and 1.8 respectively). In sum, the variety and the intensity of their criminal careers seem to decrease in adulthood. We labelled this pattern *juvenile delinquents.* Individuals assigned to Class 2 (13.8%) contrast drastically with those of the first class in terms of temporal aspects of their criminal careers. Indeed, they are unlikely to have been convicted for any type of offense in their juvenile period (mean crime index: 0.5; mean number of convictions: 0.6), whereas, by contrast, they can be convicted with a relatively high probability for a wide range of offenses in their adult period (mean crime index: 3.3). Thus, we labelled this pattern *adult onset.* The individuals of Class 3 (11.2% of the sample), contrary to the first two classes have a criminal pattern that is consistent across the two lifetime periods. Another characteristic is the very narrow range of criminal activity, mainly made up of traffic crimes, and to a lesser extent, violent crime. For these reasons, we labelled this class : Persistent criminals with a limited spectrum of criminal activity.
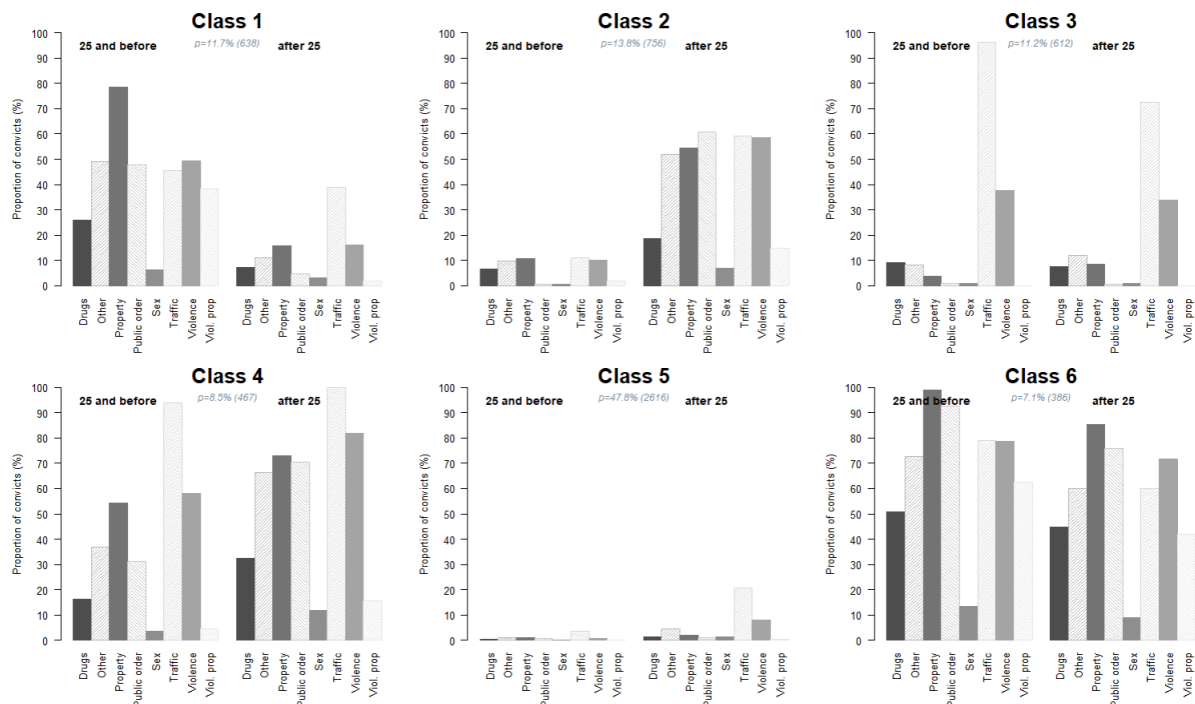
Figure 15 – Proportion of convicts per offense category and lifetime period for each class

The criminal pattern in Class 4 (8.5% of the sample) is characterized by an increasing tendency to offend both in terms of variety (crime mix index : 3.0 in the juvenile period vs 4.5 in the adult period) and intensity (median number of convictions: 5.2 and 9.8 respectively). Class 4 is also characterised by the highest proportion of convicts for traffic offense. We labelled the members of this class *versatile criminals with growing criminal activity.* Class 5 is, by far, the largest class in our sample (47.6%). The Class 5 members are peculiar in the sense that their criminal careers are practically inexistent before the homicide. In fact, for 68.1% of the members of this class, there is simply no prior conviction. In general, the likelihood of being convicted in the juvenile period is virtually nil. With the exception of traffic offense (20.8%) and to, a lesser extent, violence (8.0%), the probability of conviction remains extremely low in the adult period. Given the absence and the low severity of prior convictions, if any, we labelled Class 5 as the class of the *Primo-criminals*. At the other end of the spectrum in comparison with Class 5, we find the homicide offenders of Class 6. These offenders (7.1% of the sample) have a particularly wide range of offenses (crime mix index: 6.0) for which they are highly likely to be convicted in both the juvenile and adult periods (mean number of convictions : 8.9 and 8.2 respectively). We labelled the homicide offenders of Class 6 as *highly active persistent and versatile delinquents*.

Interestingly, these six classes differed not only in terms of criminal patterns but also in terms of the age at which their members committed their homicide (e.g., the median age for the Class 4 is 37 whereas that of Class 1 is 27), and the age at which they began their criminal careers (median onset age ranging from 18 in the Classes 1 and 6 to 31 in the Class 5). Another interesting finding is that the gender of the homicide offenders was not evenly distributed across the six classes. In particular, women seemed very unlikely to belong to classes with high criminal activity, especially in the juvenile period. To illustrate this point, 77.9% of women were assigned to Class 5, but only 0.6 % and 1.4% to Classes 6 and 4 respectively.

In conclusion, our results clearly demonstrated that homicide offenders are not a homogenous group. The latent class analysis revealed the existence of six distinct groups whose criminal careers differed in terms of crime variety, career continuity and intensity of the criminal activity. Future studies are planned to investigate whether these classes also differ in terms of recidivism, on the one hand, and socio-economic characteristics on the other.

More details relating to the method and the results can be found in an article to be submitted in the near future (Vande Velde et al., *in prep*).

4.4. Studying the feasibility of a recidivism module

The study of the feasibility of a "recidivism monitor" was based firstly on the scientific literature on recidivism in criminology (see point 2 - State of art) and secondly on all the documentation collected as part of this project concerning the Central Criminal Record database (see point 4.1. Documenting The Criminal Justice Databases). This study led to the creation of a prototypical version of such a monitor, using data from the Central Criminal Record.

In the absence of a generally accepted definition in the scientific literature, recidivism is defined in the context of this project as: "the act of committing a new offence, for which there is a judgment (final judicial decision), after a first judgment". A discussion of the construction of this definition, as well as the circumscription of the concepts derived from it, can be found in Huynen, Mine et al. (2024).

The Central Criminal Record's Office has an administrative database, and fulfills its mission of producing extracts at the request of citizens or authorized authorities. However, a number of obstacles have been encountered when it comes to the scientific use of the data to monitor recidivism, some of which have been overcome, some of which are in the process of being resolved, while others will remain unresolved, at least as far as historical data is concerned. All the problems encountered and the solutions implemented can be consulted in Huynen, Mine et al. (2024).

The principles for selecting and organizing data from the Central Criminal Record are explained in detail in Huynen, Mine et al. (2024). In brief, the files, bulletins and decisions taken into account must be either "active" or "legally deleted". A bulletin must mention at least one dated offence (failing which, the judgment date is taken into account) and at least one sentence and measure (thus excluding administrative decisions). For each file, the first conviction of a calendar year is taken as the reference. From this point onwards, the first criminal act committed and for which a final judgment has been handed down constitutes a "recidivist act", making the offender a repeat offender.

As mentioned in the ad hoc methodological section (cf. point 3 - Studying the feasibility of a recidivism module), the production of this prototype was entrusted to the Business Intelligence department of the FPS Justice (Federal Public Service Justice), through a set of specifications explaining the selection and calculation rules. These specifications can be found in Appendix 3 of Huynen, Mine et al. (2024).

The prototype provides the four main measures of recidivism, i.e. its prevalence (the recidivism "rate"), the number of antecedents and recidivisms, and the measure of the time to recidivism. These four statistics are available for the years 1995-2017, and can be selected and broken down by gender, age group, nationality, offence type and sentence type. The interface's main products are statistical

tables. For example, Table 1 below shows recidivism rates (in years, from zero to 9 years) for the years 2000, 2005, 2010, 2015 and 2016, based on fictitious data and a sub-group of the population.

Table 23 – Recidivism prevalence

☰	v1.1 - Evolution des cohortes (Délai année)										

Aperçu des données   Définitions   Info Groupe de faits   Info Groupe de peines

1995 to 2021
1995 ——●——●—— 2021

Cohorte ▾   Délai récidive (an) ▾   Catégorie d'âge ▾   Sexe ▾   Nationalité ▾   Pays juridiction ▾

‹   Apercu général   Par Sexe   Par Catégorie d'âge   Par Groupe de faits   Par Groupe de peines   ›

Cohorte ▲	2000		2005		2010		2015		2016	
Délai récidive (an) ▲	Base	Taux de récidive	Base	Taux de récidive	Base	Taux de récidive	Base	Taux de récidive	Base	Taux de récidive
0	138,400	16.3%	179,861	18.8%	197,740	18.5%	185,588	19.2%	185,288	19.1%
1	137,613	25.1%	179,024	28.4%	196,819	27.2%	184,690	28.5%	184,433	27.6%
2	136,803	31.7%	178,091	35.1%	195,842	32.9%	183,751	34.4%	183,546	32.3%
3	136,034	36.9%	177,100	39.9%	194,891	37.6%	182,804	38.0%	182,633	35.9%
4	135,253	40.4%	176,076	43.6%	193,802	41.1%	181,898	40.8%	181,704	38.5%
5	134,437	43.1%	175,001	46.5%	192,719	44.1%	180,904	42.9%	180,692	39.8%
6	133,604	45.4%	173,927	48.6%	191,622	46.5%	179,842	43.9%	179,586	40.0%
7	132,774	47.6%	172,797	50.3%	190,511	48.5%	178,772	44.0%	178,733	40.0%
8	131,828	49.3%	171,643	52.0%	189,393	49.9%	177,817	44.1%	.	.
9	130,811	50.9%	170,452	53.5%	188,221	51.1%	.	.	.	.

The graphs are few in number because, to be legible, they would require the construction of "ad-hoc" queries. The main graphs show the evolution over time of the four above-mentioned statistics (in total and/or on sub-group(s)).

5. RECOMMENDATIONS

The IIHA project unfolded over more than two years amidst the emerging attention for the study of recidivism, criminal careers and criminal justice databases at the National Institute of Criminalistics and Criminology (NICC). Accomplishing the research effort through IIHA led to reinforce some of our past observations (see Mine & Vanneste, 2011); it also led to new realizations.

In a nutshell, **we advocate for a culture of more and better scientific evidence that can serve to inform criminal justice policies and practices in Belgium, which implies a smooth and safe access to the Criminal Justice Administration System (CJAS) data, and data of better quality (in terms of internal, horizontal, vertical, and contextual integration).** Explanations about what this means, and concrete recommendations follow.

5.1. Background

5.1.1. A culture of more and better scientific evidence to inform criminal justice policies

The IIHA project intended to provide **national recidivism base rates** through the development of a recidivism monitor prototype (Objective 4) as a **crucial first step** in providing descriptive measures on the phenomenon of interest: recidivism. Such a step had to be accomplished before studying and explaining the phenomenon (i.e., relating it to other variables), and before studying the impact of criminal justice policies and interventions (including punishments) on the phenomenon.

Such step is indeed necessary to **conduct in the future informed evaluations** of the criminal justice policies, to assess their effectiveness in tackling crime (be it through general or specific deterrence, incapacitation, or rehabilitation). Policy making that is not based on appropriate evidence is not only potentially damaging to the concerned population but could be also detrimental to the whole Belgian society (e.g., in terms of allocating financial and human resources to the wrong interventions or practices).

Unfortunately, in Belgium, there is **lack of culture of evaluation of the criminal justice policies** (e.g., measures, interventions). This evaluation could be carried out in terms of **processes** (i.e., how things work in the system purportedly affected by the decision-making process), **effects** (i.e., what unforeseen consequences of the decision-making can be observed), and **impact** (i.e., what can we measure on the specific indicators we have chosen in advance, once the policies have been applied).

Developing a culture of informed criminal justice policies, including evaluations of decisions-making, requires the **sensitization of the policymakers as well as of the professionals** working in the entire chain of the criminal justice system (police, prosecution, courts, prisons, etc.). Such sensitization could have numerous consequences.

It could lead to encouraging **specific research** that explicitly aims at evaluating the consequences of decisions in terms of processes, effects, or impact. It could lead to **decisions that are much more grounded in scientific empirical reasoning**. Next, such an approach to decision-making could be a **source of motivation and support** for policymakers and professionals involved (i.e., by helping them align their policymaking and/or professional practices with sound scientific findings and thus making

their policy-making and/or professional practices more empirically anchored, which helps to provide legitimacy to professionals and their decision-making).

Moreover, a culture of more attention for sound empirical evidence also requires more attention for the collection and use of empirical data. Such data would help the improvement and testing of existing practices and policies and could serve as a springboard to **test new ideas**. For example, based on discussions with data producers across the criminal justice chain, we observed, inter alia, that criminal justice databases of all kinds include hardly any data about the victim, the number of victims, the age of victims, and the relationship between a victim and the offender. Including such information in criminal justice databases would help to deepen scientific analyses that draw on criminal justice databases and have the potential to inform policymakers and professionals about the relationship between victims and offenders in a range of specific offense types.

Also, it remains a point of concern that many criminal justice databases seem to be suffering from difficulties in the possibilities to link with other criminal justice and non-criminal justice databases, in a large part because of the absence of a shared unique personal identifier. The prospect of **linking databases within and beyond criminal justice** opens a horizon towards a more profound scientific analysis of criminal justice policies.

In the **IIHA project, we have taken several steps towards more and better empirical evidence b**y preserving and exploiting knowledge from the State collections, whether through documenting two of its criminal justice databases (Objective 1), preserving and facilitating the exploitation of data extracted from these two databases at the person level (Objective 2), performing statistical analyses to address criminological issues (Objective 3) and developing a dynamic system producing reliable statistics on recidivism (Objective 4).

A culture of more and better scientific evidence to inform criminal justice policies needs **data of good quality**. The data that are produced at different levels of the criminal justice system (i.e., police, judicial and sentence enforcement levels), require four types of integration regarding both the data that are being produced, and the multiple processes that lead to these data (processing concerned with collecting, classifying, recording, controlling, correcting, documenting, etc.): 1) internal integration, 2) horizontal integration, 3) vertical integration and 4) contextual integration. These four types of data integration are addressed in the sections below.

5.1.2. Better data at the Internal Level

Internal integration aims to identify and eliminate contradictions and inconsistencies in data collection practices. They concern the practices of the staff members working in a particular service (e.g., in the Central Criminal Record). All encoders should **encode data in the same way**, by following the same procedures. They should be **sensitized to the importance of homogeneity** in encoding practices between staff members. Even if the act of encoding is repetitive and laborious, it should be stressed that a good quality encoding serves not only the purpose of the operations of the service that employs them, but also the other goals of the society at large (e.g., scientific aims). The homogeneity of the work of the staff members can be supported by an **adequate recording manual, appropriate quality control procedures and IT processes** that reduce the probability of human errors or variability.

As a way of illustrating the previous point, during the realization of the tasks of the IIHA project that involved the processing of data (Objective 2, Objective 3 and especially Objective 4), it appears that some data was not as clean or complete as one might have wished. For instance, the RRN was not always recorded although it could have been in some cases, and for some of them, the RRN was not coherent (it did not respect the format given to RRNs) or did not match with the rest of the information (e.g., did not match with the gender of the person, perhaps because the person has undergone a sex change). In general, the recording of data should be enforced when it is possible, and appropriate checks and constraints on encoding should be imposed wherever possible.

To take another example, some addresses of birth and addresses of residence were present in text fields. Although it is probably the easiest and most natural way of encoding an address, this also allows a variety of forms (e.g., in terms of spelling or language) during encoding. It would be beneficial if the input of this kind of information was replaced, or accompanied, by information falling under pre-defined items. For example, addresses should be chosen from a pre-defined list of addresses.

Although nowadays such types of advice may seem obvious in matters of IT and user interfaces, they were likely not broadly applied in the FPS Justice as late as 2014, as can be seen with the data of SIDIS. It may still be the case in some parts of the FPS Justice. Whether this is actually the case or not deserves to be assessed.

5.1.3. Better data at the Horizontal Level

Horizontal integration consists, within each phase of the criminal justice administration system, in establishing **channels of communication** between the departments responsible for dealing with data of a similar nature, so that a common methodology for collecting, recording, and processing data can be applied.

For instance, regarding detention, the encoders of all prisons should encode the information in a similar way. Likewise, regarding prosecution, the encoders of all prosecutors' offices should encode the information in a similar way, etc.

For SIDIS we have observed that discrepancies or variability in the encoding of different information types such as the infractions (e.g., both generic and detailed codes are used).

Nonetheless, there are efforts within the DG EPI to promote uniform registration (e.g., internal control within the application, visa by registry officer per prison, formerly service inspection registry at the central level).

5.1.4. Better data at the Vertical Level

Vertical integration seeks to reconstruct the unity of the data across the different phases of the criminal justice administration system, thus making accessible to quantitative analysis the functioning of the system in its entirety, i.e., the flow of people and cases through it, and the trajectories they follow.

The development of the Integrated Historical Database (IHD) consisted in creating a database system (Objective 2) that would store and preserve the data on the one hand, and recognize which persons are concerned by which records on the other hand.

The task of establishing which records in a database belongs to what person, within a database or across different databases is not a trivial task in the absence of a unique personal identifier.

First, reliable personal data is not always available to establish these links between records. For instance, the national Register number (RRN) that is assigned to Belgian nationals and residents is not always found in the data, and therefore cannot be solely relied upon as a unique personal identifier. In other words, the systematic use of the RRN is lacking in the data.

Second, even when the RRN is available to make a link between two records, the data might be wrongly assigned to a personal record or can contain a mistake. These events may occur in the data because there haven't always been automatic controls of the data (e.g., for CJCS, such controls exist only since 2011). In other words, the reliability of the RRN in the data might sometimes be problematic.

For these two reasons, a **specific methodology had to be developed for the IHD**. With such a system, it is possible to make further progress in studying the recidivism and criminal careers while envisioning different scenarios (i.e., different methodological choices). However, although some of these choices will be considered safer than others, and that some serious rationales have been applied in developing the overall procedure, it certainly could be improved in many ways. For instance, different weighting schemes could be tested to weight evidence (some of them being inferred by probabilistic means), some errors (e.g., incorrect RRN) found in the dataset could be corrected to further enrich the data, and a more sophisticated definition of person records similarity could be designed (e.g., by using some new variables such as variables pertaining to the criminal careers). These new avenues will be the subject of **future research**.

Finally, what is true of the unique identification of people, is true of other aspects of the data. For instance, infractions are both coded by the prosecution offices (in MaCH) and the Central Criminal Record (in CJCS). However, the two systems (MaCH and CJCS) have been developed independently, which makes the mapping between the two sets of codes difficult. Like for horizontal integration, the ability of both services to map the codes of one another could be facilitated by ensuring that **channels of communication** are opened between them.

5.1.5. Better data at the Contextual Level

Contextual integration aims at situating data relating to the Criminal Justice Administration System (CJAS) in a more global context that gives them a more accurate meaning.

A key contextual variable is time. As databases tend to evolve over time both in content and functions, there is a potential for a problematic change in the nature of the data. A variable can be either completely modified or partially modified (e.g., only changed at a particular point in time). The later type of modification introduces a break in a time series which might be particularly damaging to data interpretation if it is not known to the data analyst. All changes to the data should be **thoroughly documented** so that future users are duly informed of the process that led to generate the data.

Whether drastic changes concern the data or not, the reasons or causes responsible for the values of the variable (other phenomena or events of interest, i.e., the characteristics of the world at a specific point in time that led to encode these values) are of direct interest to analysts or researchers. This is the issue of **third variables**, i.e., **additional variables that may have an explanatory value** to explain the variable of interest. For instance, if the variable of interest is the rate of recidivism of an individual, we might be interested to have access to other variables such as demographic data, as well as more qualitative information relating to the development of criminal activities and the societal response. To have access to the proper variables, implies that the data is **collected in the first place, maintained over time or archived, and can then be accessed in an efficient and secure manner** (see the next section)**.**

In the IIHA project, while assembling documentation to describe the content and functions of two criminal justice databases for criminological purposes (Objective 1), we realized even more that the information available to complete this task was sparse, cluttered, and sometimes even incomplete.

It was especially true for SIDIS-greffe (SIDIS), because it was an older system that had been discontinued in 2014, i.e., about six years before the official start of the project. No single overarching document about SIDIS existed before we produced the one that we were able to come up with. The existing documentation was spread across different documents and was sometimes incomplete. Moreover, a large quantity of documentation was apparently lost.

For CJCS, the situation was not quite as problematic because this system is still being used and under active development by the Federal Public Service (FPS) Justice. Moreover, some documentation on the database structure existed in the ICT department of the FPS Justice. However, it was geared more towards the technical understanding of the database and its operations rather than scientific, criminological, purposes. Furthermore, while analysing the CJCS data (Objective 2, Objective 3, and especially Objective 4), it became apparent that the CJCS system had undergone some transformations over time, which would affect trends in the data and therefore their interpretation. Finally, other changes were in the process of being made or had to be performed (e.g., regarding the information about erased records), the nature of some of them being unclear at the time of the research.

The process of collecting the information, by finding the proper experts with knowledge of the systems, soliciting their help, and engaging with them in order to elucidate certain points, was an effortful one.

In general, considering other research activities involving the recording systems of the Criminal Justice Administration System (CJAS) such as the FAR project, it is likely that documentation on these systems, the information they store, and the underlying encoding practices, is often scattered and incomplete. The reasons for this include the heavy workload and demands placed on staff, limited human capacity, and the rapid pace of technological change, all of which are not conducive to the documentation of this digital heritage. The 'institutional memory' is all the more fragile in that expertise is usually held by one or two people and dies out when they leave the institution or are assigned to different missions. Existing documentation may be lost or misplaced. However, documenting the structure of these recording systems and their contents, as well as archiving the documentation, means ensuring that the data remains accessible and usable, and thus preserving a certain continuity for future research.

Regarding Objective 1, we have decided to store the documents we produced in this project in the Social Sciences and Digital Humanities Archive (SODHA)[28] of the Belgian State Archives (e.g., reports, metadata). The IHD and the extracts on which the analyses are based cannot be made available to the public, as the NICC does not own these data.

Beyond the life of distinct research projects (such as the IIHA project), it is advisable to implement a policy of systematic archiving and documentation of recording systems, their structure and evolution, to ensure the transmission of knowledge in the event of personnel changes. The preservation of collective memory must be guaranteed by more than just the goodwill of a few individuals.

5.1.6. A Secure and Smooth access to data

Facilitating access to data for scientific research (while ensuring proportionality and pragmatism) is obviously essential to the realization of scientific research. However, carrying out criminological studies using the Criminal Justice Administration System (CJAS) databases remains difficult, not only because of the complexity of the registration systems, but also because of the procedures required to obtain access to the data. To guarantee the feasibility and sustainability of such studies, a **structural solution** is required, which must remain pragmatic and proportionate to the societal objectives pursued by scientific research. **Legal provision** should be made for the long-term exploitation of criminal justice databases for scientific, statistical, and historical purposes, the use of which is currently almost exclusively envisaged in terms of administrative or operational purposes.

Once such sensitive data have been obtained, they must be **safeguarded and exploited in a modern way**, that offers important guarantees regarding the safety of the data and make it available to researchers in an efficient manner. Although steps have been taken towards safeguarding and exploiting data through developing the IHD (Objective 2), this system has not been designed to facilitate the access of data to a team of researchers in a secure way. To address these issues a **genuine research infrastructure** is necessary. Such an objective will be pursued in the DOT project.

The development of the research infrastructure should serve as an opportunity to develop a **culture of (digital) security** within the FPS Justice, given the sensitivity of the data processed. This involves not only staff training, but also the choice of tools used (e.g., open-source software vs. multinational software).

5.2. List of recommendations

The recommendations listed below are numbered such as to follow the motivations expressed in the previous sections.

1. In the interest of applying relevant policies, the public authorities should engage in a significant effort of **encouraging a culture of more and better scientific evidence to inform criminal justice policies**, which could be done by:
    1.1. **sensitizing** the policymakers and the professional working in the criminal justice system,
    1.2. **funding** research that aims at evaluating the consequences of the decisions,

---

[28] https://www.sodha.be/

1.3. supporting, and collaborating with, researchers by helping with the collection of **new data** in the criminal justice records systems,

1.4. **draft new legislations** that stimulate a culture of more and better scientific evidence to inform criminal justice policies by explicitly including the evaluation of such policies, by facilitating the access of data for scientific purposes, and by extending the purpose of databases given the future treatment of the data for scientific endeavours.

2. In the interest of **reducing data variability and inconsistencies between encoders** within the administration, the FPS Justice should take measures that support **internal integration**, by:

    2.1. **developing and promoting policies** that encourage uniformity of encoding (including with the help of encoding manuals and other devices),

    2.2. **enhancing the value of encoders and sensitizing them through specific training** to the importance of encoding consistency for operational, scientific, historical, and statistical purposes, which may have a substantial impact on society,

    2.3. developing **IT systems** that reduce human errors and variability.

3. In the interest of **reducing data variability and inconsistencies between services** within the administration, the FPS Justice should take more measures that support **horizontal integration**, by:

    3.1. **developing and promoting policies** that encourage uniformity of encoding (including with the help of encoding manuals and other devices),

    3.2. establishing **channels of communication** between services.

4. In the interest of **connecting the different parts of the Criminal Justice Administration System** (CJAS), the FPS Justice should take measures that support **vertical integration**, by:

    4.1. **enforcing the use of unique personal identifiers** such as the RRN and the APFIS numbers to recognize the persons that are being processed by the different parts of the CJAS,

    4.2. establishing **channels of communication** between services (e.g., prosecution offices and the Central Criminal Records),

    4.3. **stimulate research and development in IT systems** aimed at reconciling data (e.g., automated mapping between the MaCH and CJCS system, the automatic assignment and checking of unique personal identifiers).

5. In the interest of **situating data of the CJAS in a more global context** that gives more accurate meaning, the FPS Justice should take measures that support **contextual integration**, by:

    5.1. **developing and promoting policies and relevant tools aimed at documenting and archiving data**,

    5.2. **training existing staff** to sensitize them to the importance of documenting and archiving or **hiring staff dedicated to that effort** (knowledge managers, librarians, documentalists, writers, archivists, technical documentation specialists, etc.),

    5.3. encouraging its services to reinforce their collaboration with the **State Archive of Belgium** and consider the **SODHA project** as a repository for some of their data and documents.

6. In the interest of working with data and conducting original research, the FPS Justice and the public authorities should encourage a secure and smooth access to data, by:

    6.1. **facilitating access** to data for scientific research,

    6.2. funding the development of **research infrastructure**,

    6.3. stimulating a **culture of digital security** within the administration.

6. DISSEMINATION AND VALORISATION

6.1. Presentations

1) Mine, B. (2021). La récidive et les carrières criminelles des personnes condamnées pour terrorisme en Belgique. Présentation donnée dans le cadre du cours de criminologie clinique de Luc Robert à l'Université de Mons.

2) Mine, B. (2021). La récidive et les carrières criminelles des personnes condamnées pour terrorisme en Belgique. Présentation donnée dans le cadre du groupe de travail « prisons et exécution des peines » de la Task Force nationale.

3) Maes, E., Mine, B., & Robert, L. (2022). La récidive, Formation des aumôniers de prison francophones. Manoir d'Anjou, Woluwe Saint Pierre. Organisation par la DG EPI, SPF Justice.

4) Sharief, S. (2022). Recidive bij veroordeelden voor terrorisme: Een overlevingsduuranalyse gebaseerd op het Belgisch Centraal Strafregister, NVC, Leiden.

5) Jeuniaux, P., Mine, B., Maes, E., Robert, L. & Geudens, T. (2022). Constructing a graph-based Integrated Historical Database (IHD) to study Recidivism and Criminal Careers. Presentation on 23 September 2022 at the 22nd Annual Conference of the European Society of Criminology, in Málaga, Spain. Organization: Universidad de Málaga.

6) Maes, E., Mine, B., Robert, L., Jeuniaux, P., Huynen, P. & Sarief, S. (2022). Perspectives on national reconviction statistics in Belgium. Presentation on 24 September 2022 at the 22nd Annual Conference of the European Society of Criminology, in Málaga, Spain. Organization: Universidad de Málaga.

7) Mine, B. (2022). La récidive et les carrières criminelles. Presentation on 6 October 2022, in Brussels. Organization: Institut de Formation judiciaire.

8) Robert, L. (2022). Recidive en Criminele Carrières. Presentation on 13 October 2022, in Brussels. Organization: Instituut voor Gerechtelijke Opleiding.

9) Robert, L., Wijkman, M.M.D.S., Blokland, A. & Jeuniaux, P. (2022). Criminal Trajectories of Women Convicted of Sexual Offenses in Belgium: A Nationalwide Study. Presentation at the Annual Meeting of the American Society of Criminology, Atlanta, 16-19 November.

10) Jeuniaux, P. & Mine, B., Detry, I. (2023). The development of a graph-based integrated historical database to study the trajectories of radicals in Belgium. Presentation on 26 January 2023. Second Annual Colloquium of the European Society of Criminology, Working Group on Radicalization, Extremism, and Terrorism (WG-EXTREME) in association with the Vrije Universiteit Amsterdam.

11) Mine, B. (2023). La récidive et les carrières criminelles. Presentation on 18 April 2023, in Brussels. Organization: FIDEX (Les Midis de la FIDEX).

12) Vande Velde, M., Mine, B., Maes, E., Jeuniaux, P. & Robert, L (2023). Differentiating homicide offenders based on their criminal careers : A latent class analysis. Presentation on 7 September 2023 at the 23nd Annual Conference of the European Society of Criminology, in Firenze, Italy. Organization: Università degli Studi di Firenze.

13) Schils, E., Vande Velde, M., Robert, L., Maes, E. & Mine, B (2023). Empirical typologies of homicide offenders: a scoping review. Presentation on 7 September 2023 at the 23nd

Annual Conference of the European Society of Criminology, in Firenze, Italy. Organization: Università degli Studi di Firenze.

14) Huynen, P., Mine, B., Maes, E., Robert, L. & Jeuniaux, P (2023). New perspectives on national reconviction statistics in Belgium: towards a recidivism monitor. Presentation on 7 September 2023 at the 23nd Annual Conference of the European Society of Criminology, in Firenze, Italy. Organization: Università degli Studi di Firenze.

15) Mine, B. & Maes, E. (2023). Documentation du CJCS et SIDIS-Greffe | Documentatie van CJCS en SIDIS-Griffie. Talk given on 14 November 2023 at the Study day « Recidivism and criminal careers: the digital heritage of criminal justice at the service of justice professionals », in Brussels, Belgium. Organization: NICC-INCC & Belgian Judicial Training Institute.

16) Jeuniaux, P. (2023). Développement et exploitation d'une base de donnée historique intégrée au moyen de graphes | Ontwikkeling en werking van een geïntegreerde historische database met behulp van grafen. Talk given on 14 November 2023 at the Study day « Recidivism and criminal careers: the digital heritage of criminal justice at the service of justice professionals », in Brussels, Belgium. Organization: NICC-INCC & Belgian Judicial Training Institute.

17) Robert, L. & Vande Velde, M. (2023). Résultats d'analyses sur la récidive et les carrières criminelles | Resultaten van analyses van recidive en criminele carrières. Talk given on 14 November 2023 at the Study day « Recidivism and criminal careers: the digital heritage of criminal justice at the service of justice professionals », in Brussels, Belgium. Organization: NICC-INCC & Belgian Judicial Training Institute.

18) Huynen, P. (2023). Moniteur de la récidive |Recidivemonitor. Talk given on 14 November 2023 at the Study day « Recidivism and criminal careers: the digital heritage of criminal justice at the service of justice professionals », in Brussels, Belgium. Organization: NICC-INCC & Belgian Judicial Training Institute.

19) Jeuniaux, P. (2024). Integrated Historical Database (IHD) : usage et potentiel d'une base de données en graphe. Présentation du 30 janvier 2024 au séminaire interne de la DO criminologie de l'INCC.

20) Schils, E., Robert, L., Mine, B. & Maes, E. (2024). *Typologies empiriques des auteurs d'homicide : scoping review*. Presentation on 22 May 2024 at the 18nd Conference of the Association internationale des criminologues de langue française, in Liège, Belgium. Organization: Université de Liège (ULg).

21) Jeuniaux, P., Mine, B., Detry, I. & Remacle, C. (2024). L'appréhension des parcours pénaux des personnes condamnées pour terrorisme en Belgique au travers des bases de données de la justice pénale. Presentation on 22 May 2024 at the 18nd Conference of the Association internationale des criminologues de langue française, in Liège, Belgium. Organization: Université de Liège (ULg).

22) Mine, B., Vande Velde, M & Jeuniaux, P. (2024). Analyse multivariée du délai de récidive des personnes condamnées pour terrorisme en Belgique. Presentation on 22 May 2024 at the 18nd Conference of the Association internationale des criminologues de langue française, in Liège, Belgium. Organization: Université de Liège (ULg).

23) Mine, B., Jeuniaux, P. & Detry, I. (2024). 'Crime-terror nexus' ou l'objectivation des parcours pénaux des personnes considérées comme radicales par les autorités publiques belges. Presentation on 22 May 2024 at the 18nd Conference of the Association

internationale des criminologues de langue française, in Liège, Belgium. Organization: Université de Liège (ULg).

24) Robert, L. & Jeuniaux, P. (2024). La durée de la carrière criminelle (officielle) des délinquants sexuels. Presentation on 22 May 2024 at the 18nd Conference of the Association internationale des criminologues de langue française, in Liège, Belgium. Organization: Université de Liège (ULg).

25) Vande Velde, M., Mine, B., Maes, E. & Robert, L. Différencier les auteurs d'homicide sur base de leurs carrières criminelles : une analyse de classes latentes (LCA). Presentation on 22 May 2024 at the 18nd Conference of the Association internationale des criminologues de langue française, in Liège, Belgium. Organization: Université de Liège (ULg).

6.2. Events

1) Recidivism and criminal careers: the digital heritage of criminal justice at the service of justice professionals. Study day on 14th November 2023, in Brussels. Organization : NICC & IGO/IFJ.

6.3. Publications

1) Huynen, Philippe, Patrick Jeuniaux, Benjamin Mine, Eric Maes, et Luc Robert. *La base de données du Casier judiciaire central*. Rapport de recherche. Bruxelles: Institut National de Criminalistique et de Criminologie, n°58, 2024.

2) Maes, Eric, Benjamin Mine, Patrick Jeuniaux, Shanty Sarief, Philippe Huynen, et Luc Robert. *SIDIS-Griffie databank*. Brussel: Nationaal Instituut voor Criminalistiek en Criminologie, n°59, 2024.

3) Jeuniaux, Patrick, Benjamin Mine, Luc Robert, Eric Maes, Michaël Vande Velde. *Le développement d'une base de données historique intégrée pour l'étude de la récidive et des carrières criminelles*. Bruxelles: Institut National de Criminalistique et de Criminologie, n°60, 2024.

4) Huynen, Philippe, Benjamin Mine, Eric Maes, Patrick Jeuniaux, et Luc Robert. *Vers un moniteur belge de la récidive : jalons pour le développement d'un prototype basé sur le Casier judiciaire central*. Rapport de recherche. Bruxelles: Institut National de Criminalistique et de Criminologie, n°61, 2024.

5) Mine, Benjamin, Michael Vande Velde, Patrick Jeuniaux, Eric Maes, et Luc Robert. « Recidivism among convicts for terrorism: A survival analysis based on Belgian Central Criminal Record » (Submitted).

6) Schils, Elodie, Luc Robert, Benjamin Mine, Michaël Vande Velde, et Eric Maes. « Empirical Typologies of Homicide Offenders: A Scoping Review » (Submitted).

7) Vande Velde, Michael, Benjamin Mine, Eric Maes, Elodie Schils, et Luc Robert. « Differentiating homicide offenders based on their criminal careers : A latent class analysis » (Submitted).

7. ACKNOWLEDGEMENTS

8. REFERENCES

Albrecht, H.-J., & Jehle, J.-M. (Eds.). (2014). *National reconviction statistics and studies in Europe: = Nationale Rückfallstatistiken und -untersuchungen in Europa*. Univ.-Verl. Göttingen.

Andersen, S. N., & Skardhamar, T. (2017). Pick a Number: Mapping Recidivism Measures and Their Consequences. *Crime & Delinquency*, *63*(5), 613–635. https://doi.org/10.1177/0011128715570629

Blokland, A. (2018). Taking a criminal career approach to sexual offending. In P. Lussier & E. Beauregard (Eds.), *Sex Offending. A Criminological Perspective* (pp. 141–157). Routledge.

Blokland, A., & Lussier, P. (Eds.). (2015). *Sex Offenders: A Criminal Career Approach* (1st ed.). Wiley. https://doi.org/10.1002/9781118314630

Blokland, A., & Van Os, R. (2010). Life Span Offending Trajectories of Convicted Dutch Women. *International Criminal Justice Review*, *20*(2), 169–187. https://doi.org/10.1177/1057567710368938

Blumstein, A. (2016). From Incapacitation to Criminal Careers. *Journal of Research in Crime and Delinquency*, *53*(3), 291–305. https://doi.org/10.1177/0022427815622020

Blumstein, A., & Cohen, J. (1987). Characterizing Criminal Careers. *Science*, *237*(4818), 985–991. https://doi.org/10.1126/science.237.4818.985

Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, *34*(2), 187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

Daems, T., Maes, E., & Robert, L. (2013). Crime, criminal justice and criminology in Belgium. *European Journal of Criminology*, *10*(2), 237–254. https://doi.org/10.1177/1477370812464467

Daeninck, P., Jonckheere, A., Deltenre, S., & Maes, E. (2005). *Onderzoek inzake de voorlopige hechtenis. Analyse van de juridische mogelijkheden om de toepassing van de voorlopige hechtenis te verminderen. Recherche sur la détention préventive. Analyse des moyens juridiques susceptibles de réduire la détention préventive—Rapport 13* (Rapports INCC/NICC; Rapport Final 13; p. 367). Institut national de criminalistique et de criminologie; INCC/NICC Bib. http://nicc.fgov.be/de/node/140

De Blander, R., Robert, L., Mincke, C., Maes, E., & Mine, B. (2019). *Étude de faisabilité d'un moniteur de la récidive/Haalbaarheidsstudie betreffende een recidivemonitor* (Rapports INCC/NICC; Rapport de recherche/Eindrapport 42; p. 44). NICC; INCC/NICC Bib. https://nicc.fgov.be/upload/publicaties/rapport_42-recidivemonitor-20210504.pdf

DeLisi, M., & Piquero, A. R. (2011). New frontiers in criminal careers research, 2000–2011: A state-of-the-art review. *Journal of Criminal Justice*, *39*(4), 289–301. https://doi.org/10.1016/j.jcrimjus.2011.05.001

Detry, I., Mine, B., & Jeuniaux, P. (2021). *La radicalisation au prisme des bases de données* (Rapport de Recherche 47; Collection Des Rapports et Notes de Recherche, p. 76). Institut National de Criminalistique et de Criminologie, Direction Opérationnelle de Criminologie.

Farrington, D. P. (2019). The Duration of Criminal Careers: How Many Offenders Do Not Desist up to Age 61? *Journal of Developmental and Life-Course Criminology*, *5*(1), 4–21. https://doi.org/10.1007/s40865-018-0098-5

Fazel, S., & Wolf, A. (2015). A Systematic Review of Criminal Recidivism Rates Worldwide: Current Difficulties and Recommendations for Best Practice. *PLOS ONE*, *10*(6), e0130390. https://doi.org/10.1371/journal.pone.0130390

Fox, B. H., & Farrington, D. P. (2012). Creating Burglary Profiles Using Latent Class Analysis: A New Approach to Offender Profiling. *Criminal Justice and Behavior*, *39*(12), 1582–1611. https://doi.org/10.1177/0093854812457921

Gendreau, P., Little, T., & Goggin, C. (1996). Meta-Analysis of the Predictors of Adult Offender Recidivism: What Works! *Criminology*, *34*(4), 575–608. https://doi.org/10.1111/j.1745-9125.1996.tb01220.x

Huynen, P., Jeuniaux, P., Mine, B., Maes, E., & Robert, L. (2024). *La base de données du Casier judiciaire central* [Rapport de recherche]. Institut National de Criminalistique et de Criminologie.

Huynen, P., Mine, B., Maes, E., Jeuniaux, P., & Robert, L. (2024). *Vers un moniteur belge de la récidive: Jalons pour le développement d'un prototype basé sur le Casier judiciaire central* [Rapport de recherche]. Institut National de Criminalistique et de Criminologie.

Jeuniaux, P., De Moor, S., Robert, L., Renard, B., Stappers, C., & Vanvooren, V. (2018). Reconstruction and study of offending trajectories through forensic evidence – an illustration using a forensic DNA database. In Q. Rossy, D. Décary-Hétu, O. Delémont, & M. Mulone (Eds.), *The Routledge International Handbook of Forensisc Intelligence and Criminology* (pp. 177–190). Oxford Routledge.

Jeuniaux, P., Mine, B., & Detry, I. (2022). *Le développement d'une base de données intégrée pour l'étude des trajectoires pénales des radicaux* (Rapport de recherche 53; Issue 53, p. 234). Institut National de Criminalistique et de Criminologie. https://incc.fgov.be/le-developpement-dune-base-de-donnees-integree-pour-letude-des-trajectoires-penales-des-radicaux

Jeuniaux, P., Mine, B., Robert, L., Maes, E., & Vande Velde, M. (2024). *Le développement d'une base de données historique intégrée pour l'étude de la récidive et des carrières criminelles* (60). NICC / INCC.

Khoshnood, A., Ohlsson, H., Sundquist, J., & Sundquist, K. (2023). Firearm-Related Homicide in Sweden: A Latent Class Analysis of Suspected Offenders. *Crime & Delinquency*, *69*(1), 88–104. https://doi.org/10.1177/00111287211046519

Linzer, D. A., & Lewis, J. B. (2011). poLCA : An R Package for Polytomous Variable Latent Class Analysis. *Journal of Statistical Software*, *42*(10). https://doi.org/10.18637/jss.v042.i10

MacLeod, J. F., Grove, P. G., & Farrington, D. P. (2012). *Explaining Criminal Careers: Implications for Justice Policy* (1st ed.). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199697243.001.0001

Maes, E., Mine, B., Jeuniaux, P., Sarief, S., Huynen, P., & Robert, L. (2024). *SIDIS-Griffie databank*. Nationaal Instituut voor Criminalistiek en Criminologie.

Maes, E., Robert, L., & Mine, B. (2021). État des lieux des travaux criminologiques sur la récidive en Belgique. In B. Mine (Ed.), *La récidive et les carrières criminelles en Belgique* (Politeia, pp. 55–88).

Maes, E., & Tange, C. (2014). Langgestrafte veroordeelden in de SURB-wachtkamer voor voorwaardelijke invrijheidstelling. 'En attendant Godot'? In K. Beyens, T. Daems, & E. Maes (Eds.), *Exit gevangenis? Werking van de strafuitvoeringsrechtbanken en de wet op de externe rechtspositie van veroordeelden tot een vrijheidsstraf [reeks Panopticon Libri, nr. 6]* (pp. 97–131). Maklu.

Maltz, M. D. (1984). *Recidivism*. Academic Press.

Mine, B., Jeuniaux, P., & Detry, I. (2021). La récidive et les carrières criminelles des personnes condamnées pour terrorisme en Belgique. In B. Mine (Ed.), *La récidive et les carrières criminelles en Belgique* (pp. 117–160). Politeia, Bruxelles.

Mine, B., Robert, L., & Maes, E. (2015a). *Recidive na een rechterlijke beslissing. Nationale cijfers op basis van het Centraal Strafregister/La récidive après une décision judiciaire. Des chiffres nationaux sur la base du Casier judiciaire central* (Onuitgegeven Onderzoeksrapport 38). NICC/INCC. https://nicc.fgov.be/upload/publicaties/rapport_38_0.pdf

Mine, B., Robert, L., & Maes, L. (2015b). Soulever un coin du voile sur la récidive en Belgique. La prévalence de la récidive à partir des données du Casier judiciaire central. *Revue de Droit Pénal et de Criminologie*, *6*, 620–650.

Mine, B., & Vanneste, C. (2011). *Recherche relative aux conditions de faisabilité d'une articulation des bases de données statistiques sous la forme d'un datawarehouse* (Rapport de recherche 25; Rapport et notes de recherche, p. 220). Institut national de criminalistique et de criminologie. https://nicc.fgov.be/upload/publicaties/rapport_25_0.pdf

Robert, L. (2020). Dark sides and black holes: A study of criminological research utilization in two sex offender policies. In T. Daems & S. Pleysier (Eds.), *Criminology and Democratic Politics* (pp. 206–222). Routledge (Advances in Criminology).

Robert, L., De Blander, R., Detry, I., Maes, E., Mine, B., & Vanneste, C. (2019). Recidivism Research at the NICC. In C. Mincke, D. Brutyn, D. Burssens, A. Lemonne, E. Maes, B. Renard, & L. Robert (Eds.), *20 years of Criminology at the NICC. A scientific journey and its perspectives* (pp. 165–198). Gompel & Svacina.

Robert, L., & Maes, E. (2012). *Wederopsluiting na vrijlating uit de gevangenis* (Rapports INCC/NICC; onderzoeksrapport 27; p. 151). Nationaal Instituut voor Criminalistiek en Criminologie; INCC/NICC Bib. https://nicc.fgov.be/upload/publicaties/rapport_27.pdf

Robert, L., Maes, E., Blokland, A., & Wermink, H. (2017). 'Virtual' versus 'real' prison: Which is best? Comparing the re-incarceration rates after electronic monitoring and imprisonment in Belgium. In A. Blokland & V. van der Geest (Eds.), *The Routledge International Handbook of Life-Course Criminology* (pp. 417–435). Routledge.

Robert, L., Spaan, P., Blokland, A., Maes, E., Pauwels, L., Blom, M., & Wartna, B. S. J. (2018). Hoe "anders" zijn de criminele carrières van seksuele delinquenten? Een vergelijking van seksuele met niet-seksuele delinquenten op basis van nationale veroordelingsdata in België en Nederland. *Panopticon. Tijdschrift voor strafrecht, criminologie en forensisch welzijnswerk*, *39*(6), 519–539.

Robert, P. (1977). Les statistiques criminelles et la recherche. Réflexions conceptuelles. *Déviance et Société*, *1*(1), 3–27.

Schils, E., Robert, L., Mine, B., Vande Velde, M., & Maes, E. (submitted). *Empirical Typologies of Homicide Offenders: A Scoping Review*.

Spaan, P., Blokland, A., De Blander, R., Robert, L., Maes, E., Blom, M., & Wartna, B. (2020). Differentiating Individuals Convicted of Sexual Offenses: A Two-Country Latent Class Analysis. *Sexual Abuse*, *32*(4), 423–451. https://doi.org/10.1177/1079063219893370

Thijssen, G., Masthoff, E., Sijtsema, J., & Bogaerts, S. (2023). Understanding violent extremism: Socio-demographic, criminal and psychopathological background characteristics of detainees residing in Dutch terrorism wings. *Criminology & Criminal Justice*, *23*(2), 290–308. https://doi.org/10.1177/17488958211049019

Vaughn, M. G., DeLisi, M., Beaver, K. M., & Howard, M. O. (2009). Multiple murder and criminal careers: A latent class analysis of multiple homicide offenders. *Forensic Science International*, *183*(1–3), 67–73. https://doi.org/10.1016/j.forsciint.2008.10.014

Wartna, B. S. J. (2009). *In de oude fout: Over het meten van recidive en het vaststellen van het succes van strafrechtelijke interventies*. Boom Juridische Uitgevers.

Wartna, B. S. J., & Nijssen, L. T. J. (2006). *National studies on recidivism. An inventory of large-scale recidivism research in 33 European countries. Fact sheet, nr. 2006-11*. WODC.

Wartna, B. S. J., Tollenaar, N., & Essers, A. A. M. (2005). *Door na de gevangenis. Een cijfermatig overzicht van de strafrechtelijke recidive onder ex-gedetineerden. Onderzoek en beleid 228*. WODC.

Wijkman, M., Bijleveld, C., & Hendriks, J. (2011). Female sex offenders: Specialists, generalists and once-only offenders. *Journal of Sexual Aggression*, *17*(1), 34–45. https://doi.org/10.1080/13552600.2010.540679

Willems, M. (2007). *Quelle validité et quelle fiabilité pour les statistiques pénales du Service de la politique criminelle* (p. 50) [Mémoire de stage]. Ministère de la Fonction publique (I.F.A.) – Ministère de la Justice (S.P.C.).

Yukhnenko, D., Sridhar, S., & Fazel, S. (2019). A systematic review of criminal recidivism rates worldwide: 3-year update. *Wellcome Open Research*, *4*, 28. https://doi.org/10.12688/wellcomeopenres.14970.2

Zara, G., & Farrington, D. P. (2016). *Criminal recidivism: Explanation, prediction and prevention*. Routledge.